Year: 2022

Phrase-level and edge marking in Drehu

Torres, Catalina ; Fletcher, Janet

Abstract: This study investigates prosodic correlates of phrasing in Drehu, an Oceanic language from New Caledonia. The analysis is concerned with the demarcation of prosodic levels in the language, namely the Accentual Phrase and the Intonation Phrase. First impressionistic descriptions of Drehu state there is fixed word initial stress, however recent experimental evidence does not support this claim. Instead, it has been suggested that Drehu could be an edge-marking language which relies on right boundary marking. To determine whether the patterns recorded in the literature are borne out, the phonological and phonetic realisation of post-lexical word level prosody is investigated. An experiment was conducted to examine the extent to which fundamental frequency (F0) and duration contribute to boundary marking in Drehu. The results show that F0 cues mark the right boundary of two prosodic levels, the AP and IP, and that the strength of the boundary is related to its level in the prosodic hierarchy. Preboundary lengthening also cues IP boundaries but not AP boundaries.

Open Library of Humanities

Phrase-level and edge marking in Drehu

Catalina Torres, University of Zurich, ARC Centre of Excellence for the Dynamics of Language, University of Melbourne, Affolternstrasse 56, 8050, Zurich, CH, catalina.torres@ivs.uzh.ch

Janet Fletcher, ARC Centre of Excellence for the Dynamics of Language, University of Melbourne, Parkville, 3010, Melbourne, AU, j.fletcher@unimelb.edu.au

This study investigates prosodic correlates of phrasing in Drehu, an Oceanic language from New Caledonia. The analysis is concerned with the demarcation of prosodic levels in the language, namely the Accentual Phrase and the Intonation Phrase. First impressionistic descriptions of Drehu state there is fixed word initial stress, however recent experimental evidence does not support this claim. Instead, it has been suggested that Drehu could be an edge-marking language which relies on right boundary marking. To determine whether the patterns recorded in the literature are borne out, the phonological and phonetic realisation of post-lexical word level prosody is investigated. An experiment was conducted to examine the extent to which fundamental frequency (F0) and duration contribute to boundary marking in Drehu. The results show that F0 cues mark the right boundary of two prosodic levels, the AP and IP, and that the strength of the boundary is related to its level in the prosodic hierarchy. Preboundary lengthening also cues IP boundaries but not AP boundaries.

1 Introduction

In intonational phonology, it is by now established that head-marking and edge-marking languages differ in how (or where) they encode not only prominent syllables but also prosodic boundaries. Edge-marking languages are known for relying on phrasing and phrase-based acoustic cues to mark these two parameters (Jun 1998; Jun & Jiang 2019; Kügler & Calhoun 2020). Within autosegmental-metrical phonology (AM) (Pierrehumbert 1980; Ladd 2008) a number of studies have been devoted to the study of phrasing as a process that involves the modulation of fundamental frequency (F0) over larger prosodic constituents and the utterance (Pierrehumbert 1980; Liberman & Pierrehumbert 1984; Pierrehumbert & Beckman 1988; Ladd 1990; Prieto, Shih & Nibert 1996; Truckenbrodt 2002; D'Imperio & Michelas 2014). However, detailed studies on prosodic phrasing that have included experimental and statistical data analysis have focussed on a small number, of mostly, European languages. This study seeks to investigate acoustic data from an understudied Oceanic language and evaluate how this contributes to our knowledge of prosodic phrasing. More precisely, the aim of this study is to propose a prosodic hierarchy for Drehu and to investigate a range of phonetic and phonological cues that mark different levels of constituency within the hierarchy. Specifically, a range of durational and intonational cues are examined, extending earliery analyses in Torres, Fletcher & Wigglesworth (2018) and Torres & Fletcher (2020).

Drehu is a Southern Oceanic language spoken in New Caledonia (Crowley, Lynch & Ross 2011). According to the 2014 census (ISEE 2014), there were around 9600 inhabitants in Lifou, the island where the language originates from. Two descriptive grammars have documented the language (Tryon 1968; Moyse-Faurie 1983), however there is limited work on its prosodic system and in particular very little research on its post-lexical prosodic structure. In the following sections, different aspects of the intonational system will be analysed, contributing to the study of prosody in Drehu and of the Oceanic languages more broadly.

1.1 Prosodic marking

Phrase-level prosody serves different functions such as marking boundaries between phrases (demarcating) but also marks important elements within phrases (highlighting) (Kaland & Baumann 2020). The study of prosodic prominence marking aims at identifying the phonetic and phonological elements used in language when rendering a portion of the speech stream more salient. To further narrow down the concept of prosodic prominence in this study, we follow Cangemi & Baumann (2020) who note that phonetically prominent syllables are not only produced with higher force and perceived with more ease. Importantly, prosodically prominent syllables are also prominent in the sense that they organise their environment and that they provide a structure to their context. Languages of a head-marking type such as English or German usually rely on the use of distinctive *pitch accents* on a lexically stressed syllable to mark

a prominent syllable and word (Kügler & Calhoun 2020). Languages that use word order to mark focus, such as Italian and Spanish, can also use intonation to mark focus on a stressed syllable (Face & D'Imperio 2005). Languages of an edge-marking type such as Korean or Japanese, that lack lexical stress, rely on manipulations of phrasing and the grouping of words in order to highlight a part of speech. This means that in these languages prosodic boundaries are either enhanced or additionally inserted in order to mark important speech elements (Baltazani & Jun 1999; Jun 2011; 2014a; Athanasopoulou & Vogel 2016; Lee 2017; Kügler & Calhoun 2020).

Next to intonational marking of post-lexical prominence, the use of durational cues is also commonly attested (Kaland & Baumann 2020). In languages with head-marking, accented syllables show longer duration values than non-accented ones (Kügler 2008; Hanssen, Peters & Gussenhoven 2008). Similarly, languages with edge-marking show increased syllable duration at the boundary of prosodic constituents (Athanasopoulou & Vogel 2016; Jun & Fougeron 2002). Variation of duration, in *pre-pausal* or *pre-boundary lengthening* of syllables and/or segments has been doubly associated with a demarcating function of prosodic constituency and a highlighting function of prominence marking in French, an edge-marking language (Jun & Fougeron 2002; Tabain 2003; Michelas & D'Imperio 2012; Smith, Erickson & Savariaux 2019). Crosslinguistic observations have found that not only the syllable, but also a word in Intonation Phrase final position is lengthened (Seifart et al. 2021). Additionally, the degree of juncture between words, meaning the insertion of pauses in pre- or post-position can accompany the prominent element in the speech stream (Gordon 2008).

In this study we are interested in examining the phonetic and phonological realisation of words in different frame sentences to identify prosodic phrasing patterns in Drehu. Preliminary instrumental studies on Drehu suggest that post lexical prominence is marked at the right edge (Torres, Fletcher & Wigglesworth 2018; Torres & Fletcher 2020). This study seeks to build on these preliminary analyses to investigate the extent to which fundamental frequency (F0) patterns, preboundary lengthening and pausing contribute to the demarcation of different degrees of prosodic phrasing above the level of the word.

1.2 Oceanic languages

Within the Oceanic subgroup, Drehu has been classified as a language of the Southern Melanesian linkage which belongs to the Loyalty Islands family (Crowley, Lynch & Ross 2011). Until recently, only a small number of acoustic studies have investigated languages spoken in New Caledonia (Maddieson & Anderson 1994; Gordon & Maddieson 2004; Monnin & Loevenbruck 2010). Not surprisingly, the number of empirical studies devoted to the investigation of prosody is rather small, although there is at least one exploratory study on the perception of word level prominence in Nengone (Konyi 1996). Drehu represents an interesting case in the study of word level prosody since early records report a rather unusual pattern relative to other Oceanic languages. The first

impressionistic description reports word initial stress (Lenormand 1954) which was later supported by Tryon (1968). Arguably, an areal description of word accent patterns or stress typology in the Austronesian language family and therein classified Oceanic subgroup proves to be rather difficult. The literature shows that only a fraction of these languages has been documented and that there is an imbalance between the subgroups studied (Grimes 2000; Hulst, Goedemans & Zanten 2010). From the estimated 1236 languages in the Austronesian family, only 10% were found in the database (Hulst, Goedemans & Zanten 2010). Their survey shows that from 493 Oceanic languages, only data for around 6% (31 languages) was available. Despite the relatively small sample size, certain tendencies could be established, for instance that from the 117 languages evaluated, in 89% of the cases, the stress domain is located at the right edge of the word (104 languages). Further, it was established that for 75% stress is restricted to the penultimate or ultimate syllable (88 languages). Initial stress was found in only four languages of the data set, all of which belong to the Oceanic subgroup, two of them are from the Loyalty Islands, one of them is Drehu. Thus, earlier surveys suggest that word-initial stress is uncommon in Austronesian languages compared to final or penultimate stress. Interestingly, the remaining 21 languages for which the stress position is not bound to the penultimate or final syllable of the word all have a unique stress type. Most of these are Oceanic languages (Hulst, Goedemans & Zanten 2010).

1.3 Drehu

Drehu is the indigenous language with the largest number of speakers in New Caledonia, and it is increasingly also spoken in the urban region of Noumea (Vernaudon 2015; Dotte, Geneix-Rabault & Vandeputte 2017). New Caledonia is a collectivity of France[1] and the education system on the island follows the French metropolitan model. This means that apart from optional Drehu language classes all other subjects are taught in French and pupils follow the metropolitan school curriculum. Field work observations suggest that today almost all speakers, especially younger generations, are bilingual in French and Drehu (Torres, Fletcher & Wigglesworth 2020).

Drehu allows for SVO and VOS word order and no particular pattern has yet been identified as the canonical one (Moyse-Faurie 1983). There is no verbal inflexion and grammatical markers are placed in pre-verbal position, as shown in in example (1) where the aspect marker *kola* indicates a progressive action (Moyse-Faurie 1998).

(1)   kola meköl la föe
        kola mekəl la fəe
        DUR sleep ART woman
        'The woman is sleeping'

---

[1] After the Noumea accord from 1998 New Caledonia became *une collectivité d'outre-mer à statut particulier* hence a collectivity with a *special* status different to its old one as overseas territory.

Regarding the phonology of Drehu, it has been established that there is a 30 consonant system including stops, nasals, fricatives, laterals, and approximants. Further, there are fourteen vowels and this inventory includes a contrast between short and long vowels. The language's phonotactics allow a syllabic structure with the combinations (C)V(C) and (C)VV(C). Additionally, there is a strong preference for CV syllabification which does not allow consonant clusters word internally. Descriptions also note that in word final position, a syllable ending with a coda consonant will carry a supplementary epenthetic vowel /e/ if the following word starts with a consonant (Lenormand 1954; Tryon 1968; Moyse-Faurie 1983). Sam (2009) explains that this epenthetic vowel is optional in writing and Torres & Fletcher (2020) report a high degree of variability in the realisation of word final epenthesis. As mentioned above, some accounts offer an informal description of prominence and classify Drehu as a stress language with a word prosodic system. According to Lenormand (1954), primary stress (accent d'intensité), always falls on the first syllable of a word, pëkö [ˈpɛkə] 'none, there is nothing', fifikë [ˈfifikɛ] 'toy'. In case a prefix is added to a word, the stress pattern remains and main stress shifts to the inserted first syllable e.g.: malan [ˈmalan] 'to fall' vs. amalan [ˈamalan] 'CAUSATIVE-fall'. Compound words were described similarly, meaning that stress always shifts to the first syllable of the word. Further, Tryon (1968) proposed secondary stress in polysyllabic words, with it always falling on the third syllable [ˈamaˌlan] 'CAUSATIVE-fall'. Finally, stress is described as not being weight sensitive having a demarcative function, marking out word edges and it is not used to mark lexical distinctions. Although not much has been noted about intonation in Drehu, Tryon (1968) describes sentence final intonation as being characterised by a fall in pitch and a low tone towards the end of what is referred to here as the Intonation Phrase.

1.3.1 Previous phonetic descriptions of Drehu prosody

Recent phonetic investigations on word level prosody and tonal marking in Drehu (Torres, Fletcher & Wigglesworth 2018; Torres & Fletcher 2020) have not been able to find acoustic evidence for word initial primary stress (Lenormand 1954) or secondary stress on the third syllable, reported by Tryon (1968). Instead, these studies suggest that a prominent syllable is marking the right edge of the prosodic word. Torres & Fletcher (2020) found that noun phrases are mostly realised with a rising tonal pattern, demarcating the prosodic word with a low tone (L) to the left and a high tone (H) to the right edge. The evidence shows that the initial low tone is consistently aligned to the left edge of the word. More precisely, it aligns between a function word and a content word, thus suggesting it is a phrase initial tone. Additionally, it was found that a word final H tone is aligned to the last full syllable of the word. This study also reports that tonal targets have similar terminal F0 values at normal and fast speech rate, showing no evidence of steeper rises when the speech tempo was increased, contrary to predictions of a strict segmental anchoring hypothesis (Arvaniti, Ladd & Mennen 1998).

A preliminary study on words embedded in sentence frames examined words realised in non-utterance final position (Torres, Fletcher & Wigglesworth 2018). This study found that there is a preference to make the right edge prosodically prominent through means of a word final H tone and wider pitch excursions to the right edge. In contrast to expectations set by previous descriptions (Lenormand 1954; Tryon 1968), no evidence was found for an acoustically salient word initial syllable, but instead a small increase in duration was reported for word final vowels.

1.4 Research aims

Although many Oceanic languages remain understudied, first descriptions and ongoing work on Drehu make it possible to formulate hypotheses about the phonological structure of the language and investigate aspects relative to prosodic phrasing and intonational phonology. More precisely, in this study, the aim is to investigate how phrasing is prosodically realised and determine whether tonal evidence for different higher levels of the prosodic hierarchy can be found. For this purpose, in this paper, we take into account previous descriptions (Torres & Fletcher 2020) and propose a reanalysis of the data in Torres, Fletcher & Wigglesworth (2018) together with additional data not analysed before.

Since the intonational description in Torres, Fletcher & Wigglesworth (2018) focused on words in sentence initial and medial positions the question remains how words in sentence final position are realised. Therefore, in our experiment, the aim is to further examine the realisation of words embedded in sentence frames, including more data from words realised in sentence final position, taking into account words in pre-pausal and post-pausal position, and reporting additional data from one more speaker.

Figure 1: Proposed Drehu prosodic hierarchy including two levels: the AP and the IP. '*La melimala ka za*'. The turtle dove is beautiful.

Starting with the prosodic word as the lowest tonally marked constituent, we investigate whether there is evidence for two prosodic levels, the Accentual Phrase (AP) and the Intonation Phrase (IP). **Figure 1** shows our proposal for the prosodic hierarchy for Drehu, suggesting two tonally defined levels, using an autosgemental-metrical framework (after Ladd 2008; Jun 2014b). It shows that a function word and a content word are grouped together as an AP. Building on the analyses presented in Torres & Fletcher (2020), the dotted lines indicate the double association of the phrase initial L tone. Tones in parentheses indicate that these are optionally realised, most likely in polysyllabic words and presumably for rhythmic reasons. The AP final H tone is marked with the letter a which indicates this is the prominently marked syllable. The utterance is realised as one IP, delimited by a final tone that can be H or L and is marked with a percentage sign. Finally, it is of interest to determine whether durational cues contribute to the marking of constituent boundaries in the proposed prosodic structure.

2 Experiment

As mentioned above, recent observations for Drehu suggest prominent syllables are marked towards the right edge of the prosodic word (Torres, Fletcher & Wigglesworth 2018; Torres & Fletcher 2020). These phonetic studies did not include an intonational description of words realised in utterance final position. Since Torres, Fletcher & Wigglesworth (2018) did not consider effects of pauses inserted in pre or post-position, realisations of this type are now also examined. Considering that edge-marking languages rely on cues associated with boundaries (Kügler & Calhoun 2020), pitch movements towards the right boundary are likely to contribute to prosodic phrase demarcation in Drehu. For target items realised with a rising tonal movement, we anticipate that the insertion of a pause will not displace the previously attested H tone at the right edge of a word or unit larger than the word. Instead, as known from edge-marking languages, such as French (D'Imperio & Michelas 2014), it is expected that boundary and prominence marking will coincide on the last full syllable of a word. In case there is a preference to prominently mark the right edge of the word, it is further hypothesised that pitch span expansion, between the preceding L tone and the right-most H tone, will be greater when the H tone is placed at the right edge of higher level prosodic units than at the right edge of lower level units.

Additionally, the study will examine word and syllable-level durational patterns in order to determine whether position of the word in a sentence has an influence on syllable and word duration. Since for languages with edge-prominence marking increased duration has been detected at higher level prosodic boundaries, we also predict that the effect of lengthening not only affects the vowel but the entire syllable and thus that the word final syllable will show greater duration in comparison to other syllables in the word, especially in pre-pausal position.

2.1 Hypotheses

In this experiment, the aim is to see whether there is phonetic evidence to support the model of Drehu prosodic structure as shown in **Figure 1**. In case of APs and IPs being present in Drehu speech, it is expected to find more phonological processes of lending prominence and acoustic correlates accompanying phonological processes. Therefore, it is expected that increased syllable duration and particular tonal patterns and movements at phrase boundaries, as well as increased phonetic prominence at larger phrase breaks (longer word or syllable duration and more extreme F0 excursions) will be observed. In this description of phrasing we build on previous observations (Torres, Fletcher & Wigglesworth 2018; Torres & Fletcher 2020), and take into account findings reported in the literature on phonetic and tonal cues to phrase boundaries and correlates of different levels of prosodic phrasing (1.1). The following hypotheses are proposed:

H1 Major intonational boundary in sentence final position hypothesis :
It is expected that utterance final tonal movements will demarcate a higher prosodic level. Thus, words produced in sentence final position will carry a tonally marked prosodic boundary demarcating the right edge of the intonational phrase.

H2 Intonational right-edge marking hypothesis:
In words carrying rising tonal movements, it is expected that the pitch span of a rise produced towards the right edge will be greater than that of a rise found word initially. Thus, a high tone will demarcate the right boundary of the accentual phrase. As known from phrasal languages, we expect that when co-occurring with a pause, the right-boundary pitch span expansion will be greater than that of rises not co-occurring with a pause. In the latter case the high tone will demarcate an IP.

H3 Sentence final duration of words hypothesis:
The position in which a word is realised in the sentence will affect overall word duration. This hypothesis predicts that words realised at the end of a sentence will be the longest, assuming that sentence-final position equates with the highest level of phrasing in the prosodic hierarchy.

H4 Duration of word final syllables hypothesis:
Word-final syllables are longer than preceding syllables. Further, it is hypothesised that this effect will increase in pre-pausal contexts.

H5 Simultaneous prominence and boundary marking hypothesis:
The insertion of a pause will not affect the previously noted preference to mark the right edge of the prosodic word with a H tone. Since it is assumed that words in pre-pausal position will

be at the edge of the highest level prosodic unit in our hierarchy, we expect that effects of prominence and boundary marking will coincide on the right most full syllable. This will lead to longer final syllables and more extreme pitch excursions to the right.

2.2 Participants

Five adult female speakers (age 29–47) were recorded in Lifou. All participants are bilinguals of French and Drehu and responded to an adapted version of the Bilingual Language Profile (BLP) asking about demographic data and linguistic preferences (Gertken, Amengual & Birdsong 2014). The questionnaire was administered on line and each participant responded from their workplace where the necessary internet connection was available. This questionnaire provides useful insights into linguistic practices of the participants and a score on language dominance. All reported they acquired French and Drehu in their childhood, two started learning French from birth and one reports acquiring Drehu from the age of seven. All participants were schooled in French and have varying degrees of education, ranging from a primary school diploma to the equivalent of a bachelor's degree (Licence) for a participant 36 years old. Four reported having received school instruction in Drehu (1 to 10 years) and all were comfortable reading in this language. Further, participants reported they work in the local community in professions that require them to speak in the two languages (e.g. librarian, secretary, language documentation). The score obtained from the BLP suggests that the older four speakers (age 36–47) are dominant in Drehu while the youngest speaker (age 29) is slightly more dominant in French.

2.3 Materials and Design

A word insertion paradigm was chosen for which three carrier phrases and 55 token words were used (Jun & Fletcher 2014). As shown in Examples (2a), (2b) and (2c), tokens were inserted in three positions: sentence Initial, Medial, and Final. The target words consisted of 1, 2, 3, and 4 syllables. Due to variability in the realisation of a word final epenthetic vowel two disyllabic tokens could also be produced with one less syllable (as monosyllables), the same was true for two trisyllables, and one quadrisyllable. Target items had variable syllabic structure (V, CV, VC, CVV, or CVC in final position) although most were CV-syllables since this is the preferred syllable structure in word non-final position in Drehu (see appendix A). The target tokens contained short vowels or diphthongs. The token words were checked for comprehensibility and the carrier phrases were developed with help of a linguist who is a native speaker of Drehu.

(2) a. __ la ëjen qene drehu
         __ la ɛðen ʌene d͡ʒehu
         __ ART name language drehu
         '__ is the word in Drehu'

b. Ame la __ tre ka lolo
Ame la __ t͡ʃe ka lolo
PRS1 ART __ PRS2 STAT beautiful
'This __ is beautiful/good'

c. Ngöne la qene drehu kola hape __
ŋœne la ʌene d͡ʒehu kola hape __
in ART language drehu DUR say __
'In Drehu we say __ '

2.4 Procedure

In this experiment each speaker had to produce 55 tokens in three different frames (55 × 3 × 5 = 825). For this purpose visual stimuli and written material were used. In the first step, the token was given in written form. Thereafter, the frame (written) and the token (image) were provided and the participant had to produce the correct frame with the expected token. The order in which the frames appeared was randomised. The prompting of every token aimed at eliciting highlighted words similar to how a new referent would be introduced in speech. Slides in Power Point were used to present the material and participants had a test trial prior to recordings.

2.5 Recordings

The first author who is a fluent speaker of French carried out the recording of the experiment during a fieldwork trip in Lifou. Participants were recorded in a quiet room in a library or at a community centre. Speakers had time to familiarise themselves with the procedure and were then recorded with a Zoom H6 Handy recorder, and a head mounted-microphone, at a sampling rate of 48 kHz and 16-bit depth. For further processing the recordings were down sampled to 44.1 kHz.

2.5.1 Analysis procedures

The recordings were manually transcribed and then force aligned in WebMAUS, using a grapheme to phoneme conversion, with a parameter model based on SAMPA (Kisler, Reichel & Schiel 2017). This alignment process provided a textgrid for each utterance in which all phonemes were segmented and marked with boundaries. Since automatic forced alignment does not detect phonological processes in Drehu, and is otherwise also prone to alignment errors, manual correction was required. All utterances were thereafter visually inspected in Praat 6.0.48 (Boersma & Weenink 2017) and the segmentation of phones in target tokens was corrected when necessary. During the correction of segmental alignment, special attention was paid to the setting of phone boundaries. A boundary was set between vowels and nasals, laterals or approximants at the point where sudden changes in both amplitude and formant structure occurred. In case the

change in formant structure was gradual, the segment boundaries were marked at the midpoint of the transition from vowel to liquid or approximant. For obstruents, the start of closure was marked as the onset and the start of high amplitude periodicity was marked as the onset of the next vowel. Fricatives were marked at the onset and offset of high amplitude aperiodicity (Harrington 2010). Pauses that were inserted prior or after the target tokens were also identified and marked. Pauses were visually identified in the spectrogram based on the acoustic signal, when no speech was uttered. In ambiguous cases when a pause could be realised prior to a stop closure, it was decided that the stop closure would be marked at 50 ms before the release of the burst, unless otherwise indicated by the presence of formants or an F0 trace.

Contrary to earlier descriptions in the literature (Lenormand 1954; Moyse-Faurie 1983), and similar to Torres & Fletcher (2020), variability regarding a strict CV syllabification in word final syllables was found. Words where we would predict epenthesis to occur were often realised without the epenthetic vowel ending on a coda consonant instead. This caused some variability in the number of syllables contained in token words. For example, a disyllabic word such as *trenge* [t͡ʃe.ŋe] 'basket', could instead be realised as the monosyllabic word [t͡ʃeŋ]. There were five words that showed this variability and a slight preference for epenthesis realisation was noted since 54.7% of the tokens were realised with a final vowel.
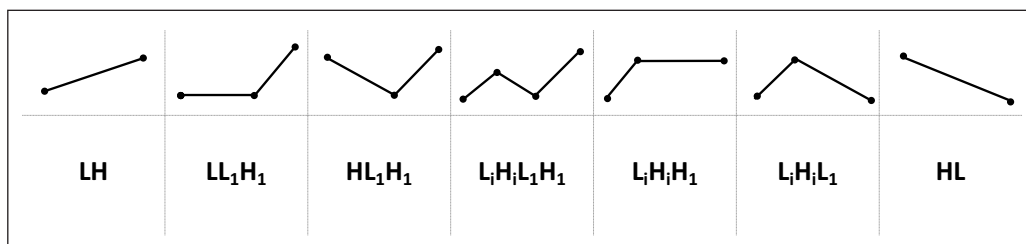

Figure 2: Schematic representation of seven tonal patterns identified in the Drehu corpus.

For query purposes within emu-R (Winkelmann et al. 2017), the following parameters were marked: position in carrier phrase (Initial, Medial, Final), position of the syllable, and pauses inserted prior or after the target token. Additionally, an AM-style annotation (Beckman & Hirschberg 1994; Pitrelli, Beckman & Hirschberg 1994; Welby 2006; Ladd 2008) was used to mark tonal targets in the items of interest. The F0 contour was visually inspected in Praat and the tones were marked according to fundamental frequency minima and maxima points that were found. The location of L was defined as the F0 minimum, and mostly occurred around the first syllable of the test word. High tones were also visually identified and marked with H. When more than one rising tone was observed, the tones in the initial rise were marked with an additional letter (L_iH_i) and every subsequent tone in a token was also marked with an additional number (L_1H_1). In case there was a plateau, valley or several points of equal F0 value, the tonal

target was marked at the onset of the tone. Attention was given to contexts when the word initial consonant was an obstruent or a fricative, and tonal targets were marked to avoid the marking of micro-prosodic perturbations as tones. In some instances, especially for words in sentence final position, it was difficult to determine a tonal target due to strong glottalisation or a very breathy voice quality. In these cases, it was preferred not to assign any value to the final tone target.

Although there was variation in the intonational configuration of patterns, four tonal movements that could be associated with phrasal prosodic boundaries, were identified and analysed. Three types of rises and one falling tone were observed in the data which, for the sake of clarity, are referred to as: the left-boundary rise (LiHi), the across-token rise (LH), the right-boundary rise (L_1H_1), and falling tone (HL). These movements were sometimes combined on a single word, see **Figure 2**. A small proportion of target items produced with disfluencies or noise (e.g. coughing) were discarded, leaving 91% of the data in our corpus. A total of 751 utterances, that were mostly evenly distributed amongst speakers, were collected and used for analysis.

Figure 3 shows the notation used for the data and how target tokens, position, syllables, phones, and tones were labelled. A hierarchical database was constructed using the EMU Speech Database Management System (Winkelmann, Harrington & Jänsch 2017). It included five tiers for the tones, phonemic segments, syllables, words, and target token position. The acoustic and durational characteristics produced in the target words were extracted and analysed using the `emuR` package in `R` (R Core Team 2017; Winkelmann et al. 2017). First, the tonal patterns were noted, then duration values for syllables, and F0 values of low and high tones were extracted. For further analysis, the values obtained in Hz for tonal targets were converted into semitones with a benchmark of 100 (Nolan 2003). Since we were interested in the magnitude of rises, we measured pitch span expansion as the difference between two consecutive tonal targets (T1-T2) in semitones. Following Campbell (1999), in order to examine variation related to a prosodic order, durational values obtained for syllables were normalised to z-scores per individual speaker. Expressing duration on a scale from long to short, with the average being the centre point, the z-score transform is known for factoring out predictable contextual differences due to inherent length of segments. Thus, predictable effects related to longer duration of segments can be factored out by calculating the averages separately for each syllable. This makes it possible to consider lengthening as a result of prosodic effects.

2.6 Statistical analyses

Data were analysed using linear mixed effects models to investigate (i) effects related to F0, and (ii) durational patterns of words and syllables. Statistical analyses were carried out in `R` (R Core Team 2017) with help of the statistics package `lme4` and `emmeans` (Bates, Mächler & Walker 2019; Lenth & Herve 2019). Values were fitted into a theoretically motivated and maximally specified linear mixed effects model and different models were tested out to investigate specific

factors of interest. In agreement with Matuschek et al. (2017), we refrained from adding random slopes to the models as this affects statistical power of small data sets, such as the ones used in this study. An automatic backward model selection of the linear mixed models was performed. Additionally, estimated marginal means were obtained for factor comparisons and p-values were adjusted with the Tukey method.

2.7 Results
2.7.1 Tonal patterns

First, collected data, tonal patterns and observations regarding the insertion of pauses will be presented. Sentences that contained misspellings or hesitations in the critical target word were discarded from further inspection. As mentioned above, due to strong glottalisation, devoicing or a breathy voice quality, not all tones could be identified, meaning that some words could not be tonally marked, especially in sentence final position. In the corpus, tonal patterns were identified for 606 target words, in three positions: sentence Initial (227), Medial (205), and Final (174). **Figure 2** shows a schematic representation of the seven different tonal patterns found. In comparison to Torres, Fletcher & Wigglesworth (2018), one additional pattern was identified ($L_iH_iH_1$), although this represents a low frequency pattern (1.8%).

   Figure 4 shows the occurrences of tonal patterns identified in our corpus according to position in the carrier phrase. **Figure 3** (top panel) shows the pitch trace of a LH rising pattern, the most frequent in the corpus, found in 46.8% of the entire data set. This rising contour also represents the across-token rise (LH). The LL_1H_1 pattern represents the second most frequent rising pattern, occurring in 12% of the cases. Note that the F0 trajectory of LH in **Figure 3** (top panel) differs from that in LL_1H_1 in **Figure 3** (mid panel). While LH displays a gradually ascending F0 curve, the pattern of LL_1H_1 is characterised by a low plateau, sometimes with a slight dip, which then rises to a peak at the right edge of the prosodic word (and of the Intonational Phrase in this example). The LL_1H_1 tonal pattern illustrated in **Figure 3** (mid panel) contains the right-boundary rise (L_1H_1). Three other patterns that ended on a high peak at the right edge of token words were identified: HLH_1 (7.8%), $L_iH_1L_1H_1$ (4.6%), and $L_iH_iH_1$ (1.8%). Additionally, two tonal patterns that ended on a low target were also identified, HL (26%) and $L_iH_iL_1$ (1%). These two patterns carry a falling tone (HL) to the right edge. As can be expected, most of the patterns ending on a low tone occurred when the token word was in sentence Final position. However, the pattern HL was also identified in sentence Initial and Medial positions a total of 12 times, and in 5 cases it was followed by a pause.

   From the 174 patterns identified in sentence Final position 84% were realised with a HL pattern. However, note that the LH rising pattern was also identified in sentence Final position, occurring in 13.7% of the data. Additionally, the $L_iH_iL_1$ and LL_1H_1 were identified, each in 1.15% of tokens in sentence Final position.

It was noted that participants sometimes inserted pauses, as illustrated in **Figure 3** (mid panel), where a pause was inserted after the target item, and **Figure 3** (bottom panel) with a pause occurring prior to the target item. The frequency of the insertion of pauses was quantified and it was found that in 22.5% of the data a pause was inserted prior or after the target word. Sentence Initial target words were realised in pre-pausal position in 6.8% of the cases. Sentence Final target items were in post-pausal position in 15.3%. Interestingly, sentence Medial target words were accompanied by a pause in 44.5% of the cases, in 92.4% the target word was realised in pre-pausal position and only in 7.6% in post-pausal position.
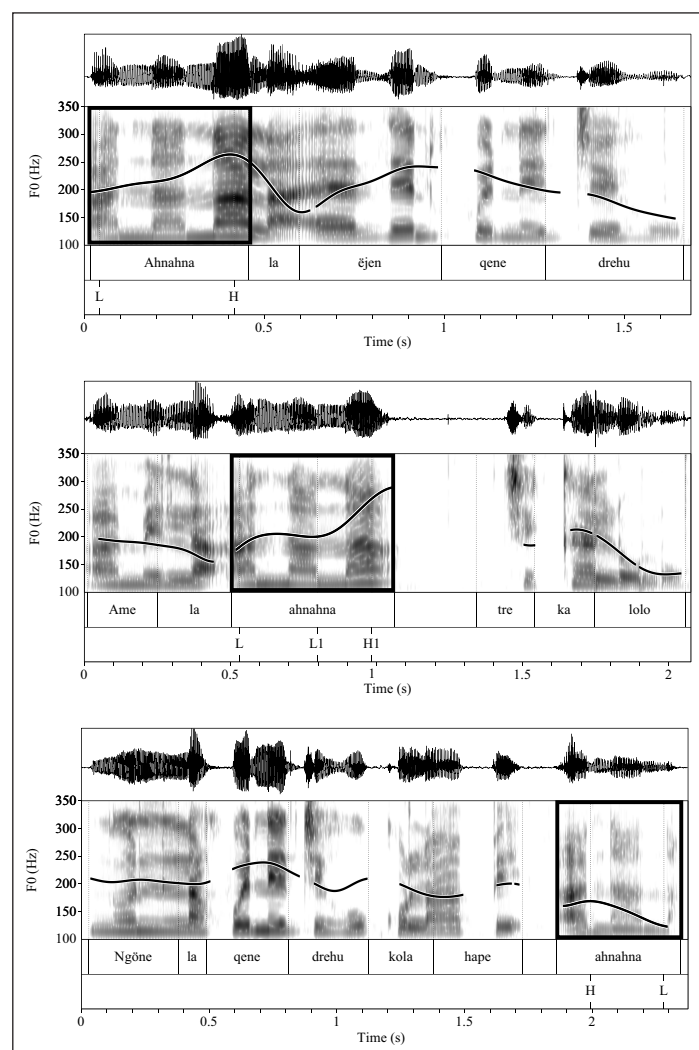

Figure 3: Speech waveform, spectrogram, and F0 trace of target token *ahnahna* 'gift' highlighted in a black box, from top to bottom, inserted in sentence Initial, Medial (with following pause), and Final (after a pause) position. All examples were produced by the same female speaker.
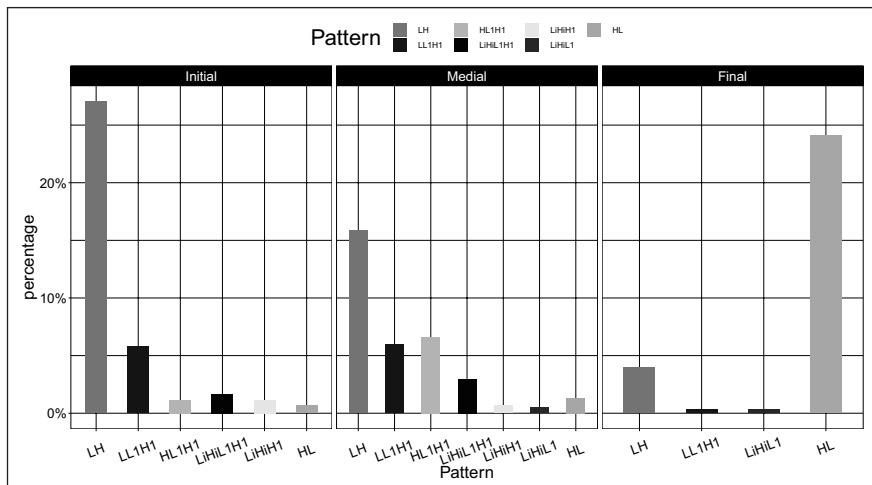
Figure 4: Bar plots summarising the tonal patterns identified in the corpus according to position in the carrier phrase. There were three positions Initial, Medial, and Final. The plot includes pre- and post-pausal tokens.

2.7.2 Tonal movements

In addition to examining the distribution of tonal patterns in the corpus, in this experiment the magnitude of tonal movements was examined. This was done in order to verify whether greater expansion could be signalling a higher prosodic level, such as the IP. For this purpose, the magnitude of pitch span in rising and falling tones was calculated as T1-T2. With this procedure, rising tones receive a pitch span measurement with a positive value, while falling tones a negative one. **Table 1** shows the mean values for pitch span expansion between tonal targets in three positions. These values indicate that in absence of a pause, tonal movements in sentence Final position appear to be greater in magnitude than in sentence Initial and Medial positions.

Tonal movement	Initial	Pre-Pause	Medial	Pre-Pause	Final	Post-Pause
left-boundary L_iH_i	2.6 (1)	2.8 (2)	3.1 (2.2)	2.6 (1.6)		
across-token LH	4.7 (2.9)	6.8 (0.9)	3.9 (2.3)	5.6 (2.8)	8.7 (3)	
right-boundary L_1H_1	4.4 (1.7)	7.7 (3.8)	5.2 (2.5)	6.3 (3.1)		
Falling tone HL	−1.8 (0.3)	−9.4 (2.6)	−5.2 (2.4)	−3.9 (0.2)	−6 (2.9)	−5.4 (2.1)

Table 1: Values for mean pitch span in semitones for tonal movements in three positions, sentence Initial, Medial, and Final. Measurements for tokens in pre-pausal position are given for sentence Initial and Medial. For tokens in post-pausal position, measurements are given when sentence Final. Standard deviation is given in parentheses.

As mentioned earlier, target items in sentence Final position were mostly realised with a HL pattern, although also a LH pattern was reported, and only in less than 3% of the data $L_iH_iL_1$ and LL_1H_1 were observed in this position, see right panel in **Figure 4**.

In the following, the aim is to examine the use of F0 in demarcating prosodic boundaries. Therefore, we re-examine tonal patterns to see which ones a) demarcate the right edge of accentual phrases, and b) higher level tonally-marked units like intonation phrases. Additionally, we examine the F0 realisations of these tonal movements in more detail including pitch span (i.e. the distance between T1 and T2).

For instance, it could be that in the HL and $L_iH_iL_1$ patterns the word final L is always associated with a major prosodic boundary. **Figure 5** shows the magnitude of pitch span (T1-T2) of four tonal movements observed in three positions (see also **Table 1**). Note that the falling tone (HL), was a low frequency pattern in Initial (4 occurrences by 3 speakers) and Medial (8 occurrences by 3 speakers) positions, while it was more frequent in Final position (146 occurrences). Therefore, mean values should be treated with caution. As shown in **Figure 5**, the L tone in HL in sentence Final position, is lower than in other positions. This L tone is thus associated with final lowering, an additional lowering relative to other Ls earlier in the phrase. Moreover, the deeper falls (around – 6 st) could be additionally affected by declination. Taking into consideration measurements in **Table 1**, this shows that both tonal movements, the across-token (LH, 8.7 st) and falling tone (HL, –9.4 st) in sentence Final and pre-pausal position are realised at the extremes of the tonal range of speakers.
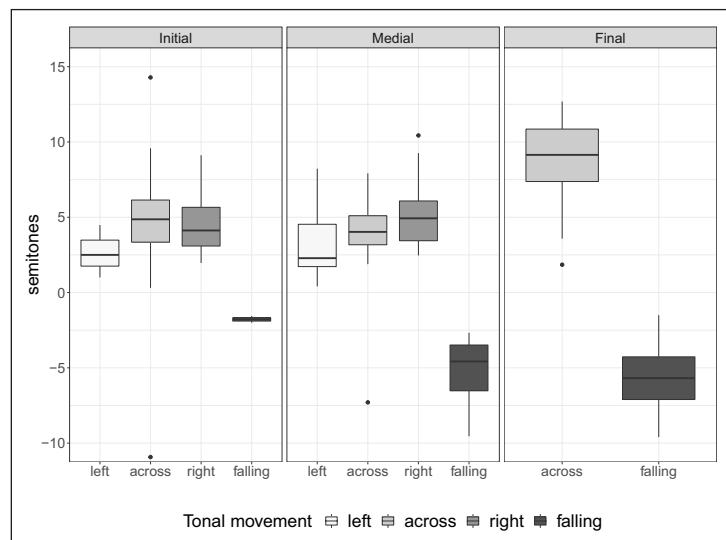

Figure 5: Box plots show magnitude of pitch span in semitones for tonal movements in sentence Initial and Medial position and (right-edge tonal movements) in Final position. Pre-pausal and post-pausal items excluded.

Table 1 and **Figure 5** show that sentence position and movement type are important factors in determining pitch movement magnitude so a linear mixed effects model was employed to investigate whether the position in the sentence had an effect on the magnitude of the tonal movement. The model had POSITION (Initial, Medial, Final) and TONAL MOVEMENT (left-boundary, across-token, right-boundary, falling tone) as fixed effects, together with participant and token as random intercepts, no random slopes were included. From the initial 606 annotated intonation patterns 134 co-occurred with a pause, and tokens in pre-pausal and post-pausal positions are excluded from this analysis. In 70 cases, due to pitch track errors, we could not obtain a value for one of the two tones required to measure pitch span, leaving thus 421 data points for further analysis. The reader is reminded that in e.g., $L_iH_iL_1H_1$, two rising tones are present. There were 19 such instances for which two tonal movements (left-boundary and right-boundary) were included. In the case of tri-tonal patterns, of the type $L_iH_iL_1$, where one tone could be associated with a rising and a falling tonal movement at the same time, both tonal movements were included. Since we did not want to discard tonal movements *a priori*, it was decided that both the left-boundary and the falling tone would be considered. Additionally, for the pattern $L_iH_iH_1$, the left-boundary and for LL_1H_1 the right-boundary tonal movements were taken into account, as we did not expect any strong differences between two consecutive low and high tones, and these were low frequency patterns.

Results of the linear mixed effects model are in Appendix B.1. After performing a backward model selection, POSITION and TONAL MOVEMENT were retained as fixed factors, together with random intercepts for participant. This model was used for a post hoc analysis which examined an interaction between the two retained fixed factors.

Selected results obtained from the Tukey corrected post hoc tests are summarised in **Table 2**. They show that there is an interaction between position in the sentence and tonal movement but crucially that for the across-token rise there is a significant difference in sentence Final position when compared to Initial and Medial positions. This shows that the magnitude of pitch span is substantially greater when in sentence Final position. As expected, no significant effect was found when comparing the across-token rise in sentence Initial and Medial positions. However, no significant difference is found when the Final falling tone is compared to Initial and Medial positions. As stated before, the HL pattern was a low frequency pattern in Initial and Medial positions (7 occurrences in absence of a pause). Considering that the model included 11 data points for falling tones, we take this result with caution as additional data points would be necessary to contribute to a more definitive result.

Arguably, for sentence Final tokens the position in which they were realised is already indicative of their right boundary being at a higher prosodic level, such as the Intonation Phrase. Given that tones in sentence Final position are associated with phrasal boundary marking, demarcating a higher prosodic level, it should be no surprise when larger differences are measured

in this position. Considering the negative values for the falling tone in sentence Initial and Medial position, we take HL and $L_iH_iL_1$ to be patterns whose right boundary tone (L) can be associated with a boundary of a higher prosodic level, namely the IP. When occurring in utterance final position, the notation L% and H% should be used to indicate they are IP-boundary tones.

Factor	Contrast	Estimate	SE	t-value	*p*-value
RISE TYPE	Position				
across-token	Initial, Final	–4.1	0.6	–6.5	<.0001
across-token	Initial, Medial	0.6	0.4	1.3	=0.9 (n.s)
across-token	Medial, Final	–4.6	0.7	–6.4	<.0001
falling tone	Initial, Final	4.9	1.8	2.7	=0.2 (n.s)
falling tone	Initial, Medial	3.8	2	2	=0.7 (n.s)
falling tone	Medial, Final	1.1	0.9	1.2	=0.9 (n.s)

Table 2: Selected results of Tukey corrected factor comparisons between rise type and position in the sentence.

2.7.3 Rising tones and pauses

Pre-pausal and post-pausal rising tones in sentence Initial and Medial positions were of interest, since the insertion of a pause could be signalling a higher prosodic level, namely the IP. Therefore, the pitch span between the low and the high tones associated with these two contexts was compared. As mentioned in section 2.5.1, three types of rises within the word were established, the left-boundary rise (LiHi), across-token rise (LH), and right-boundary rise (L_1H_1). Box plots in **Figure 6** show the differences in pitch span for the three rises when realised without a pause and in pre-pausal position.
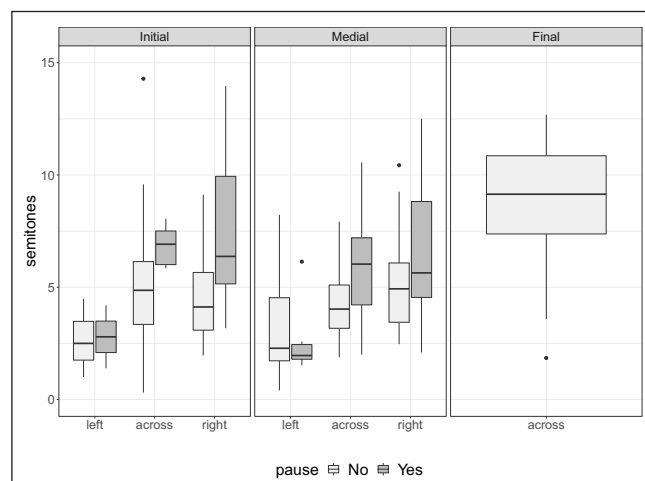

Figure 6: Box plots show magnitude of pitch span in semitones for three types of rises found in Initial and Medial positions. Tokens realised with a following pause are in dark grey.

To investigate factors influencing the magnitude of the excursion a linear mixed effects model was employed. The model included POSITION (Initial, Medial, Final), RISE TYPE (left-boundary, across-token, right-boundary), PRE-PAUSE (yes, no), and POST-PAUSE (yes, no) as fixed factors as well as participant, and token as random factors, no random slopes were included.[2] The model included 401 observations, from which 96 were produced in pre-pausal and 5 in post-pausal position. Additionally, for 22 words two rises were included, and the remainder were taken from 357 patterns with only one rise.

The results of the linear mixed effects model are in Appendix B.2. They show that rises of post-pausal words were not significantly affected by the pause. Thus this factor was excluded from the subsequent model. After performing a backward model selection, the factors POSITION, RISE TYPE, and PRE-PAUSE were retained together with random intercepts for participant and token.

Selected results for the Tukey corrected post hoc test are summarised in **Table 3**. Results are averaged over the levels of position. The analysis reveals that there was a significant effect of pause on rise type, and that the pitch span expansion was greater when tokens were followed by a pause. This effect was significant for the across rise and for the left-boundary rise which decreased when in pre-pausal position, see also **Table 1**. Additionally, compared to the left rise (L_iH_i), the magnitude of the excursion of the right boundary rise (L_1H_1) is significantly greater. Moreover, there is a significant difference between the left rise (L_iH_i) and the across-token rise (LH), with the latter also being significantly greater. Interestingly, there is no significant difference in the size of excursion for the across-token and right-boundary comparison. These results suggest that the across-token and the right-boundary rises show similar strength and contribute to demarcating the right boundary of the AP or the IP when in pre-pausal position.

Factor	Contrast	Estimate	SE	t-value	*p*-value
RISE TYPE	Pre-pausal				
across-token	no, yes	–1.2	0.31	–3.8	= .002
right-boundary	no, yes	–1.2	0.31	–3.8	= .002
PRE-PAUSAL	Rise Type				
no	left-boundary vs. across-token	–1.9	0.4	–4.3	= .0003
no	left-boundary vs. right-boundary	–2.2	0.4	–5.1	< .0001
no	right-boundary vs. across-token	–3.2	0.3	–1.1	= .8 (n.s)

Table 3: Selected results of Tukey corrected factor comparisons for rises realised in pre-pausal position and without a pause.

---

[2] A similar model that excluded data points from sentence Final position was also tested. The overall results did not change. It was decided to include data points in sentence Final position to determine whether the insertion of a pause had a comparable effect on the right-most tonal movement, as that of the right boundary in sentence Final position.

2.7.4 Word duration

An additional aim of this experiment was to determine whether the manipulation of position in the sentence had an effect on duration patterns of the word and the constituent syllables. **Table 4** provides a summary of word length in number of syllables per word, and token numbers in three different sentence positions. **Figure 7** shows the duration of words with one, two, three or four syllables in three different positions in ms. As expected the acoustic duration of words increases incrementally depending on word length in syllables.

Interestingly, words with identical syllable numbers are longer the further to the right (or later in the sentence) they are produced, especially in sentence Final position. A linear mixed effects model was employed to investigate the effect of position on token duration. From the 751 collected utterances, 166 target words co-occurred with a pause and were excluded from this analysis, leaving 585 observations. The model included POSITION (Initial, Medial, Final) and SYLLABLES (monosyllable, disyllable, trisyllable, quadrisyllable) as fixed factors together with the random factors participant and token, no random slopes were included. Results of the linear mixed effects model are summarised in the Appendix B.2.

Syllables/word	Initial	Medial	Final
one	20	19	21
two	87	94	46
tree	102	98	82
four	41	49	82

Table 4: Number of syllables per word in three different positions, sentence Initial, Medial, and Final. Table takes into account all 751 collected target words.
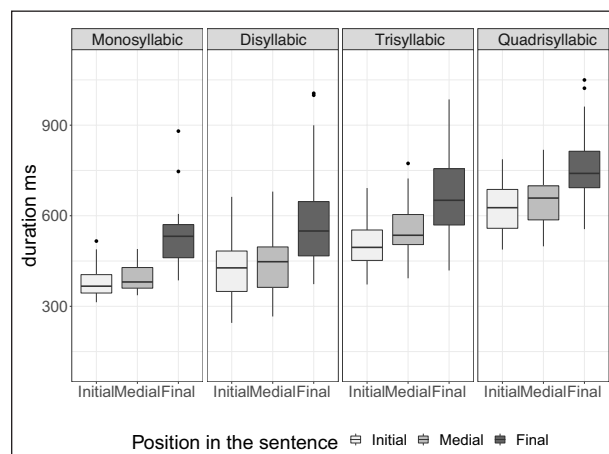

Figure 7: Box plots show duration values in ms of words with one to four syllables in three different positions in the sentence: Initial, Medial, and Final. Words in pre- and post-pausal position were excluded.

After performing the backward model selection, the model retained the fixed factors POSITION and SYLLABLES with random intercepts for participant and token. Selected results of the post hoc analysis are summarised in **Table 5**, they indicate that the position in which, e.g., disyllabic words were produced, had an effect on their duration, and that words produced in sentence Medial and Final positions are significantly longer than words in sentence Initial position. Not surprisingly, as word length increases in terms of number of syllables, acoustic duration also increases. However, there is no significant difference when monosyllables and disyllables are compared in sentence Initial, Medial, and Final positions.

Factor	Contrast	Estimate	SE	t-value	*p*-value
SYLLABLES	Position				
disyl.	Initial vs. Medial	−26.5	7.6	−3.5	= .03
disyl.	Initial vs. Final	−147	6.9	−21.3	< .0001
disyl.	Medial vs. Final	−120.7	7.9	−15.1	< .0001
POSITION	Syllables				
Initial	monosyl. vs. disyl.	8	26	0.3	= 1 (n.s)
Initial	monosyl. vs. trisyl.	−108.3	27.6	−3.9	= .007
Initial	monosyl. vs. quadrisyl.	−217.4	30.1	−7.1	< .0001

Table 5: Selected results of Tukey corrected factor comparisons for word duration.

2.7.5 Syllable duration in the word

Apart from identifying whether the manipulation of position of the word had an effect on its duration, the aim was to examine whether there is evidence for an acoustically more salient syllable in the word. For this purpose, syllabic duration in polysyllabic words was examined. A linear mixed effects model was employed to investigate syllable duration of polysyllabic words, including observations from 1267 syllables. To allow for better comparison and avoid confounding effects, this analysis was restricted to syllables with a CV-structure, taken from 524 words (191 disyllables, 220 trisyllables, 113 quadrisyllables). Tokens in pre- or post-pausal position were excluded. Normalised duration values were fitted to a model that included the POSITION IN SENTENCE (Initial, Medial, Final) POSITION OF THE SYLLABLE (first, second, third, final) and SYLLABLE NUMBER PER WORD (disyllable, trisyllable, quadrisyllable) as fixed factors, together with participant and token as random factors, no random slopes were included. Results of the linear mixed effects model are in Apendix B.3.

Table 6 shows the mean duration of CV-syllables in words with varying syllable number produced in the three positions included in the experiment. Note that for disyllabic and trisyllabic words mean duration for syllables with an epenthetic vowel is shorter than for other final

syllables. Although in this data set only 23 words were realised with epenthesis it was decided to exclude these items from the following statistical investigation. Note that mean duration values in **Table 6** indicate that syllables in polysyllabic words are shortened.

Syllable position	disyllabic	trisyllabic	quadrisyllabic
first	209 (47)	173 (47)	150 (39)
second	–	172 (39)	153 (26)
third	–	–	170 (34)
final	229 (73)	215 (61)	195 (52)
epenthesis	184 (18)	192 (70)	217 (45)

Table 6: Values for mean syllable duration in ms in disyllabic, trisyllabic, and quadrisyllabic words, in three positions, CV structure, pre- and post-pausal tokens excluded. Epenthesis refers to the word final syllable carrying an epenthetic vowel. Standard deviations are given in parentheses.

After performing backward model selection, the model retained the fixed factors POSITION IN SENTENCE, POSITION OF THE SYLLABLE, SYLLABLE NUMBER PER WORD and token, as random factor. Of particular interest was the evaluation of a possible interaction between the number of syllables in the word and the position of the syllable in the word. Additionally, an interaction between the position in the sentence and the position of the syllable in the word was examined. Selected results of the Tukey corrected post hoc analyses are summarised in **Table 7**.

Factor	Contrast	Estimate	SE	t-value	*p*-value
SYLLABLE NUMBER	Position syllable				
disyl.	first vs. final	–0.2	0.1	–2.9	= .1 (n.s)
trisyl.	second vs. final	–0.4	0.1	–7.5	= <.0001
quadrisyl.	third vs. final	–0.3	0.1	–4.9	= <.0001
POSITION SYLLABLE	Position sentence				
final	Initial vs. Medial	–0.2	0.1	–3.1	= .08 (n.s)
final	Initial vs. Final	–0.7	0.1	–14.1	<.0001
final	Medial vs. Final	–0.5	0.1	–8.9	<.0001

Table 7: Results of Tukey corrected post hoc analysis for syllable duration (z-scores) of polysyllabic words in three positions in the sentence.

The analysis shows that the duration of a syllable is influenced by its position in the word and the position of the word in the sentence. There is a statistically significant interaction for the

position of the syllable in the word and the number of syllables per word in polysyllabic words with three and four syllables. However, this result was not the same in words with only two syllables. Similarly, there is a significant interaction for the position of the word in the sentence and the position of the syllable in the word, showing that word final syllables are significantly longer in sentence Final position when compared to Initial and Medial positions. In contrast, no significant difference is found for word final syllables in sentence Initial and Medial positions.

2.7.6 Additional effects on syllable duration

Similar to the analysis of the effect of pauses on rises, it is of interest to determine how syllable duration was affected by the insertion of a pause. **Figure 8** visualises CV-syllable duration of disyllabic, trisyllabic, and quadrisyllabic words produced without (no) and with a pause (yes). The box plots indicate that when produced with a pause, the target word was longer but specially that the final syllable appears to be lengthened.
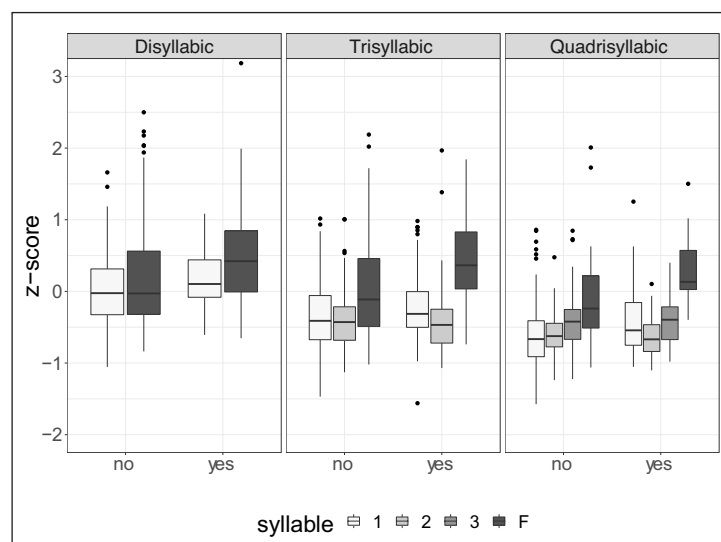

Figure 8: Box plot of normalised syllable duration in words containing two, three, and four syllables. Produced without (no) and with a pause (yes). Items from three positions in the sentence (Initial, Medial, Final) are included.

To further investigate syllable duration and a possible interaction with the insertion of a pause a linear mixed effects model was employed. The model included measurements of 1665 syllables, taken from 678 words. In this sample 38 words were in post-pausal, 116 in pre-pausal position, and 524 did not co-occur with a pause. Based on the previous sets of analyses we tested the fixed factors POSITION OF SYLLABLE (first, second, third, final), POSITION IN SENTENCE (Initial, Medial, Final), PRE-PAUSE (yes, no), POST-PAUSE (yes, no) and token as random intercept, no random slopes were included. The results for the model are in Appendix B.5. After performing

backward model selection, the model retained the factors POSITION OF SYLLABLE, POSITION IN SENTENCE, PRE-PAUSE, and token as random intercept. Selected results for the Tukey corrected factor comparisons between Pre-Pause and Position in the sentence as well as Pre-Pause and Position of syllable in word are summarised in **Table 8**.

Factor	Contrast	Estimate	SE	t-value	*p*-value
SENTENCE	Pause				
Initial	yes vs. no	–0.2	0.04	–5.6	<.0001
PAUSE	Position				
yes	Initial vs. Medial	–0.1	0.03	–2.5	=.1 (n.s)
SYLLABLE	Pause				
final	yes vs. no	–0.4	0.1	–7.1	<.0001
PAUSE	Syllable				
yes	third vs. final	–0.5	0.1	–4.1	<.0001
yes	first vs. third	–0.1	0.1	–0.5	=.9 (n.s)

Table 8: Selected results of Tukey corrected factor comparisons. Dependent measure is normalised syllable duration as z-score.

The analysis reveals that the insertion of a pause affects tokens in the same position and that when a pause is inserted, there is lengthening. Note that the previously established difference between duration values of words realised in sentence Initial and Medial positions, is again confirmed for the normalised duration of syllables. The results further indicate that the final syllable was significantly lengthened, when a pause followed, in comparison to when no pause was inserted. Additional evidence comes from the comparison between the third and the final syllable, whereby the final syllable is significantly lengthened when produced prior to a pause. There is no effect of pause on non-final syllable duration, i.e. non-final syllables in pre-pausal tokens do not lengthen. Taking into account the previous intonational analyses, it appears to be the case that sentence Final and pre-pausal word-final syllables that carry an IP-final boundary tone (L% and H%) are also significantly longer, thus providing non-tonal evidence of a higher level prosodic boundary.

3 Discussion

In this paper the intonational realisation and duration patterns of Drehu words and phrases was investigated. In our corpus, evidence was found for seven intonational patterns as well as right-boundary marking associated with two prosodic levels, the AP and the IP, as proposed in **Figure 1**. As predicted, the MAJOR INTONATIONAL BOUNDARY IN SENTENCE FINAL POSITION HYPOTHESIS

was confirmed. An examination of the extent and magnitude of the tonal movements at the right edge showed that these were stronger in sentence Final position. It was found that words in sentence Final position were realised with a final L% or H% tone demarcating the IP. Tokens that were realised in sentence Initial and Medial position generally displayed a rising tonal movement that culminated on a H tone at the right edge. Right-boundary low tones in sentence Initial and Medial positions were also observed. This low tone appeared to be shallow, and presumably it marks an IP right boundary. However, due to the low number of tones in this position it would be beneficial to gather more data to confirm this interpretation.

Pitch span expansion and the influence of an inserted pause on the three rising tones that were found in the data were examined. The aim was to evaluate whether there is additional tonal marking that marks the right edge of a higher intonational constituent. First, it was confirmed that in absence of a pause, the pitch excursion between an L and H tone is greater for across-token and right-boundary rises than for left-boundary rises. This means that words in Initial and Medial positions show intonational right-boundary marking in form of a high tonal target. Phonetic right-boundary marking in these positions is associated with an AP boundary (see **Figure 1**). The left-boundary rise appears to be an optional word-initial tone that presumably serves a rhythmic function. Additionally, it was established that pre-pausal words with the across-token and the right-boundary rise showed greater pitch span expansions than when no pause was inserted. This is in agreement with the INTONATIONAL RIGHT-EDGE MARKING HYPOTHESIS which examines whether the right-edge is acoustically marked in Drehu and shows that a H tone is associated with an AP, while a H% tone marks the IP boundary. Arguably, the word initial left-boundary rise could have been perceived and interpreted as the word initial stress previously reported in the literature (Lenormand 1954; Tryon 1968).

Regarding duration patterns, it was established that the position in the sentence has an influence on word and syllable duration. Not only was the SENTENCE FINAL DURATION OF WORDS HYPOTHESIS confirmed, it was also found that words in sentence Medial position are longer than words in Initial position. Note however that this comparison did not reach statistical significance, which may be due to the small sample size of the data. Future studies with more data could help resolve this issue. Additionally, when examining syllable duration of polysyllabic words, it is found that the word final syllable is longer than the second and third syllables in trisyllabic and quadrisyllabic words.

However, in disyllabic words, no statistically significant difference between the first and the final syllable could be established. Neither, did we find a significant difference for word final syllables when the words were non-final in the sentence. This indicates that word length (in number of syllables in the word) and the position of the word in the sentence have an influence on duration patterns. Hence, there is polysyllabic shortening or syllable compression in longer words so these syllables are shorter. This means that when a word contains more syllables, the final syllable

is longer than preceding ones and additionally there is a lengthening effect in sentence Final position. Note that no word-final syllable lengthening was observed in Initial of Medial position.

Further, the DURATION OF WORD FINAL SYLLABLES HYPOTHESIS was confirmed, since longer duration of word final syllables was established. However, word final lengthening is strongly influenced by the position of the word in the sentence and is also motivated by the insertion of a pause, which taken together with the intonational cues, suggests a higher level of prosodic phrasing. Durational differences of word final syllables were the greatest in pre-pausal position. This suggests that final lengthening effects are prosodically motivated and the result of cumulative influences. In words with three or four syllables there is overall syllable shortening and the final syllable appears to be lengthened in pre-pausal position. Finally, in agreement with the SIMULTANEOUS PROMINENCE AND BOUNDARY MARKING HYPOTHESIS, it could be established that the insertion of a pause which simultaneously inserts an IP-boundary, significantly affected final syllable duration and the size of excursion at the right boundary and the last full syllable of target tokens.

4 Conclusion

This study examined the realisation of words in frame sentences and provides an analysis of prosodic phrasing in Drehu, a Southern Melanesian language from the Loyalty Islands. As noted, the rare stress pattern previously recorded in the literature makes Drehu a particularly interesting case for the study of prosody and its acoustic correlates, especially since more recent phonetic investigations found no acoustic evidence for word intial stress (Torres, Fletcher & Wigglesworth 2018; Torres & Fletcher 2020). This study shows evidence for post-lexical right-edge marking in Drehu and represents a contribution to current knowledge in the field of intonational phonology and the autosegmental-metrical approach. It is found that Drehu shows similarities to other edge-marking languages, like French and Korean, while no evidence is found in favour of a stress pattern similar to the one known from e.g., English (Hyman 2014).

The present work shows that Drehu relies on manipulations of F0 and duration at the right edge of constituents as can be expected from an edge-marking language. The evidence collected shows that prosodic constituents are tonally demarcated at the right edge and the strength of this boundary is related to the level in the prosodic hierarchy where it is placed. More precisely, based on the magnitude of tonal movements, we find evidence for two separate levels, the Accentual Phrase and the Intonation Phrase. Regarding the role of acoustic duration, it is found that syllable lengthening is used when demarcating the highest prosodic level, the IP but not the AP. However, durational patterns resemble those of Korean more than those in French, in that there is a lack of AP-final syllable lengthening. (Jun 1998; Jun & Fougeron 2002).

The analysis of tonal movements showed that there is a preference to prominently mark the right boundary of the word with a H tone, in sentence Initial and Medial positions. Words

in Final position usually ended on a H or L boundary tone, and the latter could be associated with final lowering. Results from our experiment showed that from three different types of rises found (left-boundary L_iH_i, across-token LH, and right-boundary L_1H_1 rise), the magnitude of the pitch excursion between a L and H tone is greater for across-token and right-boundary rises than for left-boundary rises. This is in line with previous observations in Torres, Fletcher & Wigglesworth (2018), and confirms right-edge marking with a final high tone which in most instances is preceded by a low tone that is not always word initial. The strength of prosodic cues at the right boundary is indicative of the prosodic level they demarcate. The Accentual Phrase is marked with a right-most H tone, while the Intonation Phrase is marked with a either a H% or L% boundary tones and a following pause. This shows that Drehu, similar to other edge-marking languages, like French and Korean, relies on intonational right-edge marking (Jun 1998; Welby 2006).

The fairly variable left-boundary rise does not occur in every rendition and could presumably serve a rhythmic function. Considering that the insertion of a pause did not change the tonal configuration of right-boundary L_1H_1 rising tones and that the strength of the rise is augmented, we interpret these results as the enhancement of the right-boundary. Thus, our results suggest a simultaneous realisation of highlighting and demarcating at the right edge. Note that no significant results were found for the less frequent post-pausal words.

Not only is the right edge marked through a rising tone that mostly culminates as a peak, there is also a prosodically motivated increased duration of the word final syllable in sentence Final words. Moreover, the insertion of a pause after the word also suggests that this strategy is used to further strengthen the right boundary which is accompanied by a more extreme pitch excursion and longer duration of the final syllable. Note however, that increased final syllable duration resulted from combined effects related to word length (in syllables), the position of the word in the sentence, and the insertion of a pause. This indicates that the word final syllable in Drehu is not always significantly lengthened, suggesting that it is not *per se* a lexically stressed syllable (Gordon & Roettger 2017). Although our results are not conclusive in this regard, we find a lack of evidence for (word initial) stress in Drehu, and suggest that there is post-lexical right edge prominence marking which could be phrasal.

To better understand how duration operates, it would be of interest to test the processing of acoustic duration cues in a perception experiment. Moreover, it might be helpful to instrumentally examine further Oceanic languages which reportedly display rare stress patterns. Finally, this study emphasises the importance of carrying out instrumental studies in the field which can help broaden our understanding of prosody as an organisational structure for speech production and prosodic typology in general.

Abbreviations

ART = article, DUR = durative, PRS = presentative

Additional file

The additional file for this article can be found as follows:

- **Appendices.** Appendix A and B. DOI: https://doi.org/10.16995/glossa.5845.s1

Acknowledgements

Funding information

Competing interests

The authors have no competing interests to declare.

References

Arvaniti, Amalia & Ladd, D. Robert & Mennen, Ineke. 1998. Stability of tonal alignment: the case of Greek prenuclear accents. *Journal of Phonetics* 26(1). 3–25. DOI: https://doi.org/10.1006/jpho.1997.0063

Athanasopoulou, Angeliki & Vogel, Irene. 2016. The Acoustic Manifestation of Prominence in Stressless Languages. In *Proceedings of Interspeech*, 82–86. DOI: https://doi.org/10.21437/Interspeech.2016-1424

Baltazani, Mary & Jun, Sun-Ah. 1999. Focus and topic intonation in Greek. In *Proceedings of the 14 th International Congress of Phonetic Sciences*.

Bates, Douglas & Mächler, Martin & Walker, Steven. 2019. *Linear Mixed-Effects Models using 'Eigen' and S4*. Version 1.1-21.

Beckman, Mary E. & Hirschberg, Julia. 1994. The ToBI annotation conventions. *Ohio State University*.

Boersma, Paul & Weenink, David. 2017. *Praat: doing phonetics by computer (Version 6.0.26) [Computer program]*.

Campbell, Nick. 1999. A Study of Japanese Speech Timing from the Syllable Perspective. *Journal of the Phonetic Society of Japan* 3(2). (< Feature Articles> Contrasting English and Japanese Phonetics), 29–39.

Cangemi, Francesco & Baumann, Stefan. 2020. Integrating phonetics and phonology in the study of linguistic prominence. *Journal of Phonetics* 81. 100993. DOI: https://doi.org/10.1016/j.wocn.2020.100993

Crowley, Terry & Lynch, John & Ross, Malcolm. 2011. *The Oceanic languages*. Routledge.

D'Imperio, Mariapaola & Michelas, Amandine. 2014. Pitch scaling and the internal structuring of the Intonation Phrase in French. *Phonology* 31(1). 95–122. DOI: https://doi.org/10.1017/S0952675714000049

Dotte, Anne-Laure & Geneix-Rabault, Stéphanie & Vandeputte, Leslie. 2017. Nouméa at the Crossroad of New Caledonian Multilingualism. *Amerasia Journal* 43(1). 13–32. DOI: https://doi.org/10.17953/aj.43.1.13-32

Face, Timothy L. & D'Imperio, Mariapaola. 2005. Reconsidering a focal typology: Evidence from Spanish and Italian. *Rivista di Linguistica* 17(2). 271–289.

Gertken, Libby M. & Amengual, Mark & Birdsong, David. 2014. Assessing language dominance with the Bilingual Language Profile. *Measuring L2 proficiency: Perspectives from SLA*. 208–225. DOI: https://doi.org/10.21832/9781783092291-014

Gordon, Matthew. 2008. The intonational realization of contrastive focus in Chickasaw. In *Topic and focus*, 69–82. Springer. DOI: https://doi.org/10.1007/978-1-4020-4796-1_4

Gordon, Matthew & Maddieson, Ian. 2004. The phonetics of Paicî vowels. *Oceanic linguistics* 43(2). 296–310. DOI: https://doi.org/10.1353/ol.2005.0004

Gordon, Matthew & Roettger, Timo. 2017. Acoustic correlates of word stress: A crosslinguistic survey. *Linguistics Vanguard* 3(1). DOI: https://doi.org/10.1515/lingvan-2017-0007

Grimes, Barbara F. 2000. *Ethnologue*. SIL International. http://www.sil.%20org/ethnologue/.

Hanssen, J. E. G. & Peters, Jörg & Gussenhoven, Carlos. 2008. Prosodic effects of focus in Dutch declaratives. In *Proceedings of speech prosody*, 609–612. Campinas, Brazil: Editora RG/CNPq.

Harrington, Jonathan. 2010. Acoustic phonetics. In William, J. Hardcastle & Laver, John & Gibbon, Fiona E. (eds.), *The handbook of phonetic sciences* 2. 81–129. Wiley-Blackwell. DOI: https://doi.org/10.1002/9781444317251.ch3

Hulst, Harry van der & Goedemans, Rob & van Zanten, Ellen. 2010. *A survey of word accentual patterns in the languages of the world*. Walter de Gruyter. DOI: https://doi.org/10.1515/9783110198966

Hyman, Larry M. 2014. Do all languages have word accent. In van der Hulst, Harry (ed.), *Word stress: theoretical and typological issues*, 56–82. Cambridge: Cambridge University Press. DOI: https://doi.org/10.1017/CBO9781139600408.004

ISEE. 2014. *Recensement général de la population.* I. S. E. E. Institut de la Statistique et des Études Économiques (ed.). [Online]. Nouméa. https://www.insee.fr/fr/statistiques/1560282.

Jun, Sun-Ah. 1998. The accentual phrase in the Korean prosodic hierarchy. *Phonology* 15(2). 189–226. DOI: https://doi.org/10.1017/S0952675798003571

Jun, Sun-Ah. 2011. Prosodic markings of complex NP focus, syntax, and the pre-/postfocus string. In *Proceedings of the 28th west coast conference on formal linguistics,* 214–230.

Jun, Sun-Ah. 2014a. *Prosodic typology II: The phonology of intonation and phrasing.* Jun, Sun-Ah (ed.). Oxford: Oxford University Press. DOI: https://doi.org/10.1093/acprof:oso/9780199567300.001.0001

Jun, Sun-Ah. 2014b. Prosodic typology: By prominence type, word prosody, and macrorhythm. In Jun, Sun-Ah (ed.), *Prosodic typology II: The phonology of intonation and phrasing,* 520–539. Oxford University Press. DOI: https://doi.org/10.1093/acprof:oso/9780199567300.003.0017

Jun, Sun-Ah & Fougeron, Cécile. 2002. Realizations of accentual phrase in French intonation. *Probus* 14(1). 147–172. DOI: https://doi.org/10.1515/prbs.2002.002

Jun, Sun-Ah & Fletcher, Janet. 2014. Methodology of studying intonation: From data collection to data analysis. In *Prosodic typology II: The phonology of intonation and phrasing,* Chap. 16, 493–519. Oxford University Press. DOI: https://doi.org/10.1093/acprof:oso/9780199567300.003.0016

Jun, Sun-Ah & Jiang, Xiannu. 2019. Differences in prosodic phrasing in marking syntax vs. focus: Data from Yanbian Korean. *The Linguistic Review* 36(1). 117–150. DOI: https://doi.org/10.1515/tlr-2018-2009

Kaland, Constantijn & Baumann, Stefan. 2020. Demarcating and highlighting in Papuan Malay phrase prosody. *The Journal of the Acoustical Society of America* 147(4). 2974–2988. DOI: https://doi.org/10.1121/10.0001008

Kisler, Thomas & Reichel, Uwe & Schiel, Florian. 2017. Multilingual processing of speech via web services. *Computer Speech & Language* 45. 326–347. DOI: https://doi.org/10.1016/j.csl.2017.01.005

Konyi, Wassissi. 1996. L'accent du mot en Nengone. In *Oceanic studies: proceedings of the first international conference on oceanic linguistics* (C-133), 243–247. Pacific Linguistics.

Kügler, Frank. 2008. The role of duration as a phonetic correlate of focus. In *Proceedings of speech prosody,* 591–594.

Kügler, Frank & Calhoun, Sasha. 2020. Prosodic encoding of information structure: A typological perspective. In *The oxford handbook of language prosody*. Oxford University Press. DOI: https://doi.org/10.1093/oxfordhb/9780198832232.013.30

Ladd, D. Robert. 1990. Metrical representation of pitch register. *Papers in laboratory phonology* 1. 35–57. DOI: https://doi.org/10.1017/CBO9780511627736.003

Ladd, D. Robert. 2008. *Intonational phonology*. Cambridge University Press. DOI: https://doi.org/10.1017/CBO9780511808814

Lee, Yong-cheol. 2017. Prosodic focus in seoul korean and south kyungsang korean. *Linguistic Research* 34(1). 133–161. DOI: https://doi.org/10.17250/khisli.34.1.201703.005

Lenormand, Maurice-H. 1954. La phonologie du mot en lifou (iÎ es loyalty). *Journal de la Société des Océanistes* 10(10). 91–109. DOI: https://doi.org/10.3406/jso.1954.1812

Lenth, R. & Herve, M. 2019. *Emmeans: Estimated marginal means, aka least-square means. R package version 1.1. 2.*

Liberman, Mark & Pierrehumbert, Janet. 1984. Intonational invariance under changes in pitch range and length. In *Language Sound Structure: Studies in Phonology Presented to Morris Halle,* Chap. 10, 157–233. MIT Press.

Maddieson, Ian & Anderson, Victoria Balboa. 1994. Phonetic structures of iaai. *UCLA Working Papers in Phonetics* 87. 163–182.

Matuschek, Hannes & Kliegl, Reinhold & Vasishth, Shravan & Baayen, Harald & Bates, Douglas. 2017. Balancing Type I error and power in linear mixed models. *Journal of memory and language* 94. 305–315. DOI: https://doi.org/10.1016/j.jml.2017.01.001

Michelas, Amandine & D'Imperio, Mariapaola. 2012. When syntax meets prosody: Tonal and duration variability in French Accentual Phrases. *Journal of Phonetics* 40(6). 816–829. DOI: https://doi.org/10.1016/j.wocn.2012.08.004

Monnin, Julia & Loevenbruck, Hélène. 2010. Language-specific influence on phoneme development: French and Drehu data. In *11th annual conference of the international speech communication association 2010 (interspeech 2010)*, 1882–1885. DOI: https://doi.org/10.21437/Interspeech.2010-543

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

Moyse-Faurie, Claire. 1998. Relations actancielles et aspect en drehu et en xârâcùù. *Actances* 9. 135–145.

Nolan, Francis. 2003. Intonational equivalence: an experimental evaluation of pitch scales. In *Proceedings of the 15th international congress of phonetic sciences, barcelona*, vol. 39.

Pierrehumbert, Janet B. 1980. *The phonology and phonetics of English intonation.* Massachusetts Institute of Technology dissertation.

Pierrehumbert, Janet B. & Beckman, Mary E. 1988. *Japanese tone structure.* Vol. 15 (Linguistic Inquiry monograph). MIT Press.

Pitrelli, John F. & Beckman, Mary E. & Hirschberg, Julia. 1994. Evaluation of prosodic transcription labeling reliability in the ToBI framework. In *Third international conference on spoken language processing.*

Prieto, Pilar & Shih, Chilin & Nibert, Holly. 1996. Pitch downtrend in Spanish. *Journal of Phonetics* 24(4). 445–473. DOI: https://doi.org/10.1006/jpho.1996.0024

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

Sam, Léonard Drilë. 2009. *Dictionnaire drehu-français.* Centre de Documentation Pédagogique de Nouvelle Calédonie; revised edition (2009).

Seifart, Frank & Strunk, Jan & Danielsen, Swintha & Hartmann, Iren & Pakendorf, Brigitte & Wichmann, Søren & Witzlack-Makarevich, Alena & Himmelmann, Nikolaus P. & Bickel, Balthasar.

2021. The extent and degree of utterance-final word lengthening in spontaneous speech from 10 languages. *Linguistics Vanguard* 7(1). DOI: https://doi.org/10.1515/lingvan-2019-0063

Smith, Caroline L. & Erickson, Donna & Savariaux, Christophe. 2019. Articulatory and acoustic correlates of prominence in French: Comparing L1 and L2 speakers. *Journal of Phonetics* 77. 100938. DOI: https://doi.org/10.1016/j.wocn.2019.100938

Tabain, Marija. 2003. Effects of prosodic boundary on /aC/ sequences: acoustic results. *The Journal of the Acoustical Society of America* 113(1). 516–531. DOI: https://doi.org/10.1121/1.1523390

Torres, Catalina & Fletcher, Janet. 2020. The alignment of F0 tonal targets under changes in speech rate in Drehu. *The Journal of the Acoustical Society of America* 147(4). 2947–2958. DOI: https://doi.org/10.1121/10.0001006

Torres, Catalina & Fletcher, Janet & Wigglesworth, Gillian. 2018. Investigating word prominence in Drehu. In *Proceedings of the 17th australasian international speech science and technology conference*, 141–144.

Torres, Catalina & Fletcher, Janet & Wigglesworth, Gillian. 2020. Fundamental frequency and regional variation in Lifou French. *Language and Speech*. DOI: https://doi.org/10.1177/0023830920952497

Truckenbrodt, Hubert. 2002. Upstep and embedded register levels. *Phonology* 19(1). 77–120. DOI: https://doi.org/10.1017/S095267570200427X

Tryon, Darrell T. 1968. *Dehu grammar*. Australian National University.

Vernaudon, Jacques. 2015. Linguistic Ideologies: Teaching Oceanic Languages in French Polynesia and New Caledonia. *The Contemporary Pacific* 27(2). 433–462. DOI: https://doi.org/10.1353/cp.2015.0048

Welby, Pauline. 2006. French intonational structure: Evidence from tonal alignment. *Journal of Phonetics* 34(3). 343–371. DOI: https://doi.org/10.1016/j.wocn.2005.09.001

Winkelmann, Raphael & Harrington, Jonathan & Jänsch, Klaus. 2017. Emu-sdms: advanced speech database management and analysis in r. *Computer Speech & Language* 45. 392–410. DOI: https://doi.org/10.1016/j.csl.2017.01.002

Winkelmann, Raphael & Jaensch, Klaus & Cassidy, Steve & Harrington, Jonathan. 2017. *emuR: Main Package of the EMU Speech Database Management System*. R package version 0.2.3.