The alignment of F0 tonal targets under changes in speech rate in Drehu

Catalina Torres and Janet Fletcher

The University of Melbourne, Centre of Excellence for the Dynamics of Language[a)]

This study investigates fundamental frequency alignment to segmental landmarks in Drehu, an Oceanic language. We present a production experiment that aimed to evaluate the marking of prosodic prominence, within the autosemental-metrical phonology, since stress and prominence system of the language has not been phonetically investigated. A rate manipulation paradigm was chosen to test the *segmental anchoring hypothesis*, namely to see whether prominence lending tonal movements exhibit a constant slope due to rate manipulation and whether tonal targets can be associated to *segmental anchoring* points in the speech stream. We find, that a rising tonal movement, between a word initial low (L), and a word final high (H) tone, is the most frequent tonal pattern. The word initial L tone seeks to align with the left edge of the word whereas the H tone, at the right edge, seeks to anchor to the last full syllable. In fast speech, tonal targets are produced closer together but the slope remains constant in both speech rates. It is shown that high tones seek to anchor to the word-final syllable, yet not to any specific segment which suggests a weak version of the segmental anchoring hypothesis applies.

[a)] catalina.torres@unimelb.edu.au

I. INTRODUCTION

A. Tonal alignment

Within autosegmental-metrical phonology (Ladd, 2008; Pierrehumbert, 1980), the study of text-to-tune alignment has shown for a number of languages, that a series of phonetic factors and language specific phonological characteristics come into play in determining how the speech stream is organized. One important concept that has received considerable attention is the *segmental anchoring hypothesis*, a term that refers to how tonal targets, realised through fundamental frequency, unfold in time while being coordinated with the segmental string (Arvaniti *et al.*, 1998; Ladd *et al.*, 1999, 2009). According to Arvaniti *et al.* (1998), this hypothesis predicts that a pitch accent is aligned or *anchored* to a segmentally defined position, and the duration and rise time of the pitch movement are determined by the segmental composition of the accented word. Strictly speaking, the segmental anchoring hypothesis predicts anchors, of low and high tones, to be closer together, and the rises to be shorter and steeper as the speech rate increases. Arvaniti et al.'s results for Modern Greek and the concluding theoretical assumptions of their study were in agreement with similar observations that had been previously made for Stockholm and Malmö Swedish, Mexican Spanish, and English[1] (Bruce, 1977; Prieto *et al.*, 1995; Silverman and Pierrehumbert, 1990). It is important to bear in mind that these studies were concerned with the realisation of prenuclear accents and also that the languages investigated shared one important phonological property, namely lexical prominence. To date a number of studies have demonstrated that language specific phonological patterns such as syllabic structure can have an influence on the anchoring points of *F0* maxima. In an experiment testing tone realisation in CV and CVC-nasal syllables, it was found that regardless of syllable structure, the tone always aligned at a similar point in the syllable in Mandarin (Xu, 1998). However, peaks of accents are not always realised in the same position but can be aligned variably. An example of variable alignment comes from Japanese, for which it was found that the peak alignment can occur earlier, when the first mora was accented in a two-mora syllable, like CVV or CVN, than in words with a one-mora syllable, like CV (Ishihara, 2003). A similar peak alignment has been observed in Dutch where peaks of prenuclear accents are aligned at the end of a vowel if the syllable contains a long vowel, but during the following consonant if the accented syllable contains a short vowel (Ladd *et al.*, 2000). Also studies on Peninsular Spanish and French, have cast some doubt on the strict segmental anchoring hypothesis since they have shown that accented syllables containing a coda consonant can be aligned differently when compared to those with no coda consonant (Prieto and Torreira, 2007; Welby and Lœvenbruck, 2006). In view of variable anchoring points for peaks in French, where final maxima could be anchored to either the vowel or the final nasal consonant, in a CVC-nasal syllable, Welby and Lœvenbruck (2006) propose the notion of *segmental anchorage*. In contrast to a strict segmental anchoring, this term refers to a region, like the syllable (and not exclusively a segment), to which a tone can be anchored. Moreover, it has been observed that alignment patterns can differ between varieties of the same language. For example, it could be established that the location of *F0* minima varies between Northern and Southern German (Atterer and Ladd, 2004), just like it was found for two varieties of English (Scottish Standard English and Southern British English) although differences did not reach a statistically significant level in the latter study (Ladd *et al.*, 2009). Anchoring points of *F0* minima and maxima offer a way to examine how the speech stream is organised, especially regarding the marking of prominence. In addition, the modulation of *F0* over larger constituents, accompanied by other markers, like pause, are considered to be indicators distinguishing between different prosodic levels in a number of languages (D'Imperio and Michelas, 2014, for French), (Truckenbrodt, 2002, for German), (Pierrehumbert and Beckman, 1988, for Tokyo Japanese). As Ladd *et al.* (2009) note, there is a considerable number of questions as to how language specific phonological factors condition the alignment of accentual peaks. In the present study, the goal is to describe prominence lending tonal targets and their position relative to the segmental string. The aim is to verify whether the *strict segmental anchoring* or a *segmental anchorage hypotheses* are best suited for a description of Drehu text to tune alignment. With this, we wish to find out whether tonal targets also contribute to the marking of prosodic constituency beyond the prosodic word. Hence, our analysis also considers effects related to the marking of higher levels of prosodic constituency.

B. Drehu

Drehu is an Oceanic language from Lifou, New Caledonia, and the language with the largest number of speakers in the archipelago (Vernaudon, 2015). According to the 2009 census (ISEE, 2009), Lifou counts around 8600 inhabitants from which approximately 5500 are Drehu speakers older than 14 years old. Protestant missionaries developed the first writing system (Magulué, 2011), which is in its majority still used for Drehu language class in primary, high school, and religious Sunday school. New Caledonia is a French overseas territory and the education system on the island follows the French metropolitan model. This means that apart from the optional Drehu language class all other subjects are taught in French. Today almost all speakers, especially younger generations, are bilinguals of French and Drehu.

Within the Oceanic languages, Drehu has been classified as a language of the Southern Melanesian linkage which belongs to the Loyalty Islands family (Crowley *et al.*, 2013). Drehu, Iaai, and Nengone are closely related languages (Loyalty Islands) and were described as having similar phonological systems. They share several properties, like a rich vowel system that includes a length distinction, a lack of weight sensitivity regarding stress, which in turn is described as demarcative and fixed. Similarly, in Drehu, Iaai, and Nengone the rhythm type was analysed as trochaic (Tryon, 1967, 1968a,b; Tryon and Dubois, 1969). However, phonetic studies further investigating the acoustics of rhythm or prosody in these three languages are under-represented. 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 V, VC, CV, CVC and VV, VVC, CVV, CVVC. It has been established that there is a strong preference for syllabification that does not allow for consonant clusters in the word. Descriptions also note that in word final position, a syllable ending with a coda consonant may receive a supplementary epenthetic vowel /e/ if the following word starts with a consonant (Lenormand, 1954; Moyse-Faurie, 1983; Tryon, 1968a). Some accounts offer an informal description of prominence and classify Drehu as a stress language with a word prosodic system. According to Lenormand (1954), stress (accent d'intensité), always falls on the first syllable of a word, pëkö [ˈpɛkʌ] (*none, there is nothing*), fifikë [ˈfifikɛ] (*toy*). In word derivation, when words obtain a prefix, the stress pattern remains and main stress shifts to the inserted first syllable e.g.: malan [ˈmalan] (*to fall*) vs. amalan [ˈamalan] (CAUS-*fall*). Compound words are described similarly, meaning that stress always shifts to the first syllable of the word. Further, Tryon (1968a) proposes secondary stress in polysyllabic words, with it always falling on the third syllable [ˈamaˌlan] (CAUS-*fall*). Finally, stress is described as not being weight sensitive and is said to have a demarcative function, which marks out word edges. Although not much has been noted about intonation in Drehu, Tryon (1968a) describes sentence final intonation as being characterised by a fall in pitch to a low tone towards the end of what we refer to here as the Intonation Phrase.

More recent studies have shown that Drehu might be undergoing phonetic changes since it was described in the fifties and early eighties. The once attested voiced and voiceless retroflex stops /ɖ/ and /ʈ/ are now realised as voiced and voiceless alveolar affricates /d͡ʒ/ and /t͡ʃ/ (Monnin, 2010). Moreover, a recent phonetic investigation on word prominence in Drehu, suggests that a phrasal rather than strictly lexical prominence marking is observed (Torres *et al.*, 2018). This study investigated words in informational focus and found that there is a preference to make the right edge prosodically prominent through longer duration and an *F0* peak on the final syllable. Further, the most frequent intonation pattern observed was a rising pattern that spanned over the word in focus. Since this study investigated words in informational focus the question remains if there is indeed word initial lexical stress or if there is rather a phrasal prominence marking system. The rapidly changing linguistic landscape, where the people of Lifou have turned into a bilingual community, might be influencing some of these phonetic changes. This hypothesis needs however to be evaluated which is why more phonetic research is needed. With this study of synchronic tune to text alignment in Drehu, we wish to add to the body of research on Melanesian languages of New Caledonia, and the phonetics of an under-documented language.

II. THE CURRENT STUDY

In this experiment, we sought to investigate the *F0* alignment of prominent tonal targets in Drehu. Further, we evaluate how modulation in scaling of *F0* relates to prosodic constituency. According to intonational phonology (Ladd, 2008), tonal events are defined through relative contrast from one tone to another. This contrast is perceived in relative tonal height and differs depending on the speaker's tone range. *High* tones are denoted with an (H) and *low* tones with an (L). Tonal events can be simple monotonal targets (H and L) but they can also be complex and bitonal (HL) falling or (LH) rising tones. Based on previous findings (Torres *et al.*, 2018), we hypothesise that there will be a preference to tonally mark the left edge and the right edge of a prosodic constituent respectively with a low tone on the initial syllable, and a high tonal target on the word final syllable. Since no phonetic evidence has been found to support a word-initial stress hypothesis, an examination of tonal alignment should provide further evidence either in favour or against the proposed phrasal prominence for Drehu. We are interested in investigating if the predicted rising tonal pattern prevails under non-focal conditions and seek to determine location and timing patterns relative to segmental landmarks. Moreover, we are interested in investigating the magnitude of rise excursion and whether under manipulation of speech rate, slope is constant. This will be important in determining whether prominent tones in Drehu are best described with the *strict segmental anchoring* or the *segmental anchorage* hypotheses. Based on previous findings for Japanese and Dutch (Ishihara, 2003; Ladd *et al.*, 2000), we further predict that a long vowel in the final syllable will be the anchor of the H tone, which should be found towards the end of the long vowel. Similarly, it is expected that for word final CVC syllables, the constituent final H tone won't be placed on the vowel but on the following sonorant coda. Additionally, it is expected that the H tone will be placed somewhere on the last full syllable and not on the syllable containing the epenthetic vowel.

III. METHODOLOGY

To investigate the placement of prominent tonal targets, a set of sentences was prepared with a native speaker of Drehu who is also a language worker from the *académie des langues kanak*. In order to test if the observed rising pattern prevails and the constituent initial L and constituent final H tones consistently demarcate prosodic boundaries a set of stimuli was prepared. In the experimental procedure, the number of syllables and syllable structure are manipulated. If our hypothesis is correct, in Drehu there is a preference to prominently mark the right edge by means of a peak, and there is no weight sensitive material that could attract the peak away from the right edge. In this case, a contrast in vowel length in the initial syllable of a disyllabic word (CVV.CV) would not have an impact on the rising pattern, and constituents would always show a peak at the right edge. Further, to examine the stability of tonal targets under time pressure, a rate manipulation paradigm was chosen.

This made it possible to investigate tonal alignment and the excursion size in *F0* between low and high targets at two self-selected speech rates, normal and fast.

A. Speech materials

The experiment used fifteen target words that were all nouns, preceded by either one or two monosyllabic function words and were embedded in sentences which were read aloud at a self selected normal and fast speech rate. The target tokens were inserted in two types of sentences, one in which the item of interest was the first Noun Phrase (NP) and another in which it represented a second NP that was preceded by a previous NP:

(1) Ame **[la maamu]**, tre, hna sile hnei itre qatr

PRS.1 ART bogeyman PRS.2 PST invent A PL old

'The bogeyman was invented by the old'

(2) Ame [la satana] **[me la maamu]**, tre, hna sile hnei itre qatr

PRS.1 ART devil and ART bogeyman PRS.2 PST invent A PL old

'The devil and the bogeyman were invented by the old'

This manipulation of position allowed us to verify if the predicted rising pattern would be constant despite changes in its placement within the utterance. The selected items of interest always started on either a vowel, a nasal consonant, or a voiced labio-velar approximant and mostly contained sonorant consonants or approximants in order to ensure a continuous *F0* curve. Attention was given that when this was not the case, the segments in the target tokens were voiced, with the exception of the word *emexemin* /emexemin/ ("echo") where the velar fricative isn't. As Table I shows, the items of interest could be grouped into two categories, according to whether they were short or long. There were nine disyllabic target tokens and six others with three, four and five syllables with varying syllable structure. The target token was embedded in a sentence. It was never placed at the beginning or end of the sentence in order to avoid any confounding effects clearly related to prosodic boundary marking at a higher level, like the Intonation Phrase. The instructions of the language informant were always followed and only phrases that were judged to sound natural and were in agreement with the current Drehu punctuation were used in the experiment. It should be noted that there is variability in the occurrence of the epenthetic vowel. Speakers varied in cases where the final syllable ended on epenthesis, and some times omitted the vowel, producing so a shorter word ending on a coda consonant. We also observed that tokens starting with a vowel (unexpectedly) tended to be produced with strong glottalization or a glottal stop.

B. Speakers

Thirteen speakers of Drehu were recorded, in a quiet room, at Lifou's high school Lycé Polyvalent des Îles. Participants provided basic demographic information (age, tribe) and responded to a sociolinguistic questionnaire that also included questions about language use. Seven were male, all of them identified as Kanak[2], and all reported being raised in a bilingual or multilingual environment, typical of the region.

TABLE I. Syllabic structure of token words. When in parenthesis (V) the word final vowel can be realised as epenthetic /e/ or be fully omitted.

Type	Count	Syllabic structure
Disyllabic	3	CV.CV
(short)	2	CVV.CV
	1	CVV.C(V)
	1	V.CVVC
	2	CV.CVC
Polysyllabic	1	CV.CV.C(V)
(long)	2	CV.CV.CV
	2	CV.CV.CV.CV
	1	CV.CV.CV.CV.C(V)

C. Recording and analysis procedures

1. Recording

The first author who is a fluent speaker of French carried out the recording of the experiment during a field work trip in Lifou. This study is part of a larger project which investigates prosody in Lifou French and Drehu. The participants had previously taken part in a similar experiment in French which took place at least one week prior to the Drehu

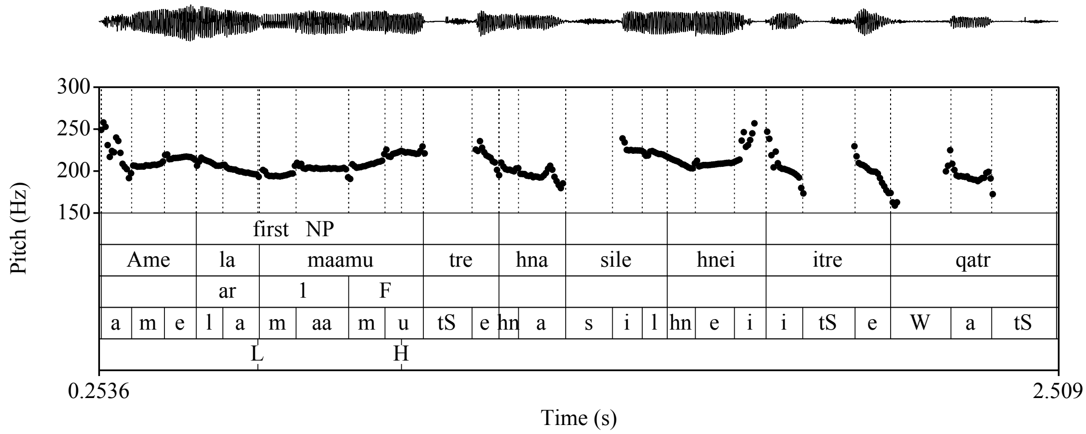


FIG. 1.

F0 trace and waveform of the utterance *Ame la maamu, tre, hna sile itre qatr*, "The devil and the bogeyman were invented by the old". Token word is *maamu* 'bogeyman', which is preceded by one monosyllabic function word, the definite article *la*, and represents the first noun in the phrase. Produced by a female speaker at fast speech rate.

experiment. This ensured that the participants understood the experimental procedure. The material for the current experiment was presented on a slide show written in Drehu and which included no written French. Speakers had time to familiarise them selves with the sentences and read these then twice at a self selected normal and a fast speech rate. They were recorded at a sampling rate of 48kHz and 24-bit depth, using a Zoom H6 Handy recorder, and a head mounted-microphone. The sentence materials were randomised. Participants were asked to read the sentences as naturally as possible, and were asked to repeat any misread sentences. It should be noted that participants seemed more at ease during the French recordings and that reading aloud in Drehu appeared to be unusual leading so to more hesitations and mispronunciations. Data that contained hesitations or pauses within the target tokens were discarded and only tokens that were produced in a prosodic constituent larger than the word were kept for analysis. The data analysed correspond to 85% of the corpus.

2. Analysis procedures

The first author manually transcribed and then force aligned the recorded utterances in WebMAUS, using a grapheme to phoneme conversion, with a parameter model based on SAMPA (Kisler *et al.*, 2017). This alignment process provided a textgrid for each utterance in which all phonemes were segmented and marked with boundaries. Since automatic forced aligment does not detect phonological processes in Drehu, manual correction was required. All utterances were thereafter visually inspected in Praat 6.0.48 (Boersma and Weenink, 2017) and the segmentation of target tokens was corrected when necessary. Additionally, a broad ToBI-style annotation (Beckman and Hirschberg, 1994; Pitrelli *et al.*, 1994) was used to mark *F0* minima and maxima points in the items of interest. The tonal targets were marked with L for low and H for high tones. Every subsequent tone identified in a token was also marked with an additional number (e.g. LHL_1H_1).

During the correction of segmental alignment, special attention was paid to the setting of phone boundaries. Following Ladd *et al.* (2000), 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. Pauses and strong glottalization, in form of creaky voice or a glottal stop, that were inserted prior or after the target tokens were also identified and marked.

Figure 1 shows the notation used for the data and how target tokens (including function words preceding it), position, syllables, phones, and tones were labelled. A hierarchical database was constructed using the EMU Speech Database Management System (Winkelmann *et al.*, 2017a). 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 queried and analysed using the emuR package in R (R Core Team, 2017; Winkelmann *et al.*, 2017b). The *F0* contour was visually inspected in Praat and the two tonal targets were marked according to fundamental frequency minima and maxima points that were found. The location of L was defined as the *F0* minimum, or elbow of the prosodic constituent and mostly occurred around the first segment of the test word. In some instances, especially due to strong glottalization, when the pitch curve was not continuous, the lowest point indicating an elbow was chosen. *F0* maxima points were also visually identified and marked with H. In some polysyllabic words, there appeared additional H and L tones which were also marked. Following Ladd *et al.* (2000), in case consonants caused perturbations on the *F0* track (for instance, nasal and laterals provoking slight dips and rises in the *F0* contour at onsets and offsets) the annotator compensated for confounding effects by ignoring *F0* points that were clearly perturbations, and marking the next highest or lowest point as measurement point. In case there was a plateau, a valley or two points of equal *F0* value, the tonal target was marked at its very beginning.

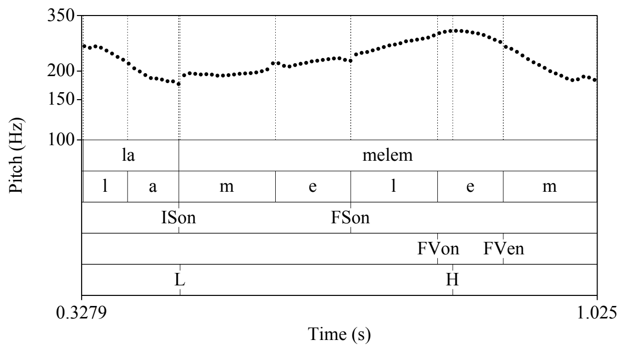


FIG. 2.

Measurement points used for the alignment of the low and high tones. Token word taken from the utterance: *Ame la melem, tre, ka mingöming*, Engl. The full moon night is beautiful. Token word is *melem* 'Full moon night', it is preceded by one monosyllabic function word, the definite article *la*.

Figure 2 illustrates the selected points for measurements. The segmental point chosen to measure the alignment of the low tone was the onset of the initial syllable (ISon). Points for the high tone were the onset of the token word last syllable (FSon); the onset of the vowel in the final syllable (FVon); the end of the same vowel (FVen). Based on these segmental landmarks, the following measurements for tone alignment were calculated:

1. Alignment L : L minus ISon in ms; negative values indicate alignment of L before ISon ;

2. Alignment H : H minus FSon in ms; negative values indicate alignment of H before FSon ;

3. Peak in FV : H minus FVon (in ms) divided by DurFV

Additionally, measurements were taken for the alignment of *F0* events relative to segmental events and the excursion size of rises between *F0* minima and maxima points. Finally, duration values were queried for the vowel of the last syllable (DurFV). To investigate if segmental anchoring was constant under rate manipulation a value for slope was calculated. The formula used divides the *F0* excursion size of a rise by its rise time. More precisely, the slope was calculated as $m = y^b - y^a / x^b - x^a$, where y^a is the *F0* value of the L tone, y^b is the *F0* value of the H tone, x^a is the realisation in time of the L tone, and x^b the time of the H tone, for both the fast and normal speech rates (Welby, 2006).

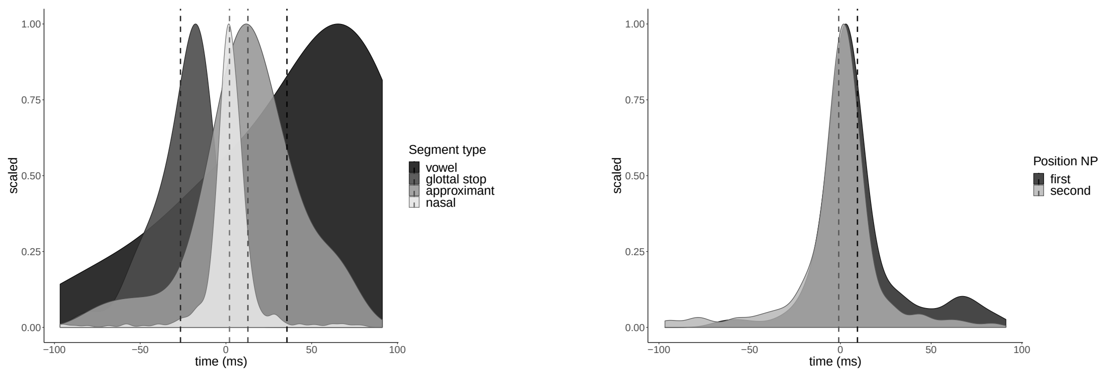
D. Statistical analyses

Data were analysed using linear mixed effects models investigating each of the following parameters of interest : (i) the alignment of the low tone, (ii) the duration of segments (iii) the alignment of the high tone, as well as (iv) the rise time, and (v) the size of the excursion in fundamental frequency. Statistical analyses were carried out in R (R Core Team, 2017) with help of the statistics package lme4 (Bates *et al.*, 2015). An automatic backward model selection of fixed parts of the linear mixed model was performed in all cases. Further, estimated marginal means were obtained for factor comparisons and p-values were adjusted with the Tukey method. Additionally, correlation tests were used and when needed Wilcoxon signed-rank tests were used for direct comparisons of the slope between rates.

IV. RESULTS

It was found that the majority of tokens (93%) were realised with a LH rising pattern and only a small set (7%) were produced with a LHL_1H_1 pattern. Overall, the initial L tone showed greater stability and tokens with both patterns (LH and LHL_1H_1) were kept for statistical analysis. The LHL_1H_1 pattern was more often found in the normal speech rate and occurred on words with three or more syllables. Usually, the initial rise contained a peak (mean *F0* 200 Hz) that was followed by a rise to a peak with a higher pitch (mean *F0* 220 Hz). The initial rise is not the rule and is variable which is why its function cannot be determined yet. Since an influence from the initial peak on the second peak could not be ruled out, we restricted our analysis of the alignment of final peaks on tokens realised with a LH pattern. Further, it was found that if there was an insertion of a pause, this predominantly occurred after the second NP, namely in 211 instances.

As expected, all token words were realised with a H tone towards the end of the prosodic constituent, on the final syllable of the token word, apart from the tokens containing an epenthetic vowel. As previously noted, the realisation of the epenthetic vowel was highly variable although it was more often realised, namely in 62% of cases. The words containing an epenthetic vowel confirmed our predictions in 85% of the cases where the peak was either realised on the penultimate syllable or the epenthetic vowel was deleted so that the peak was realised on the last full syllable of the word. Additionally, the variation in syllabic structure



(a) by segment type (b) by position of the noun

FIG. 3. Latency to initial L tone to onset of the syllable initial segment as smoothed density plot with mean values according to (a) segment where the tone is placed; (b) position of the noun in the phrase. A value lower than zero indicates that the low tone was placed prior to the onset of token word.

in disyllabic words did not have an influence on the rising pattern (LH), which was always produced in the same way, despite syllabic variation (CVV.CV, vs. CV.CVC, vs. CV.CV). In the following, the alignment of the L and H tones as well as the analyses on rise time, excursion, and slope will be presented.

A. Alignment of L

A low tone was always found in the vicinity of the first segment of the token words, regardless of tonal pattern, which is why all low tones were kept for the analysis. A linear mixed effects model that investigated the alignment of the low tone (in ms) relative to the onset of the first syllable (ISon) was employed (n=664). The automatic backward model selection kept the factors POSITION NP (first or second) and SEGMENT TYPE (vowel, nasal, approximant, glottal stop) as fixed factors, together with participants and items as random intercepts. The factor speech rate was not retained in the model. Figures in 3 show the estimated alignment and mean values of the low tone by (a) segment type and (b) position of the noun. Figure 3(a) shows that the low tone is realised at the left boundary (ISon) when the initial segment is a nasal consonant or an approximant and further to the right when it is a vowel. In cases with a glottal stop, which occurred in 36 items, the L tone was often displaced left from ISon. Figure 3(b) shows a rather consistent alignment of L to ISon. Note that in case the token word was in the first NP then the L tone was realised slightly further to the right from ISon.

Statistical analyses reveal that there was a significant interaction between segment type and position. Table II summarises the statistical results of the differences found between segment types in the same position in the phrase. This shows that the glottal stop which was produced before vowels has a significant effect on where the low tone was placed. Arguably, the production of a glottal stop does not offer a good anchor point for any tone which is why in Figure 3(a) we can see that the low tone was often displaced, and pushed to the left, past the word boundary, towards the function word preceding the item of interest.

We were further interested in determining which factor could best predict the alignment of the low tone and hypothesised the duration of the token initial segment could be a factor. To determine whether the difference in alignment of the low tone could be due to the duration of the segment to which it is anchored we first examined the duration of word initial segments. To test this, it was important to first evaluate whether there was a significant difference in the duration of the word initial segments which were grouped according to

TABLE II. Summary of the interaction between position of noun in phrase and segment type for the alignment of the L tone. Segments are only compared in the same position.

Segment type	comparison	Est. SE	df	p
Approx.	Glot. stop	36.03 ± 7.22	23.4	<.001
Nasal	Glot. stop	26.29 ± 6.35	29.08	<.005
Vowel	Glot. stop	51.73 ± 7.43	246.02	<.0001
Nasal	Vowel	25.02 ± 6.54	21.7	<.01

segment type. Table III summarises the alignment of the L tone, per segment type, in the two positions of the experiment, and the mean duration of token initial segments, per segment type, for two speech rates. A linear mixed effects model investigating the duration of the word initial segment included RATE (normal fast), and SEGMENT TYPE (approximant, nasal, vowel, glottal stop) as fixed factors, together with participant, and token as random intercepts (n=664). As can be expected, we find that that there was a significant effect of rate on duration for each segment (Est. 11.63 ± 1.72 ms, $p = .0001$). Under the same speech rate conditions there was a significant difference of duration for the segment types vowel and nasal (Est. 35.12 ± 10.46 ms, $p = .035$) but not for other segments.

Since the low tone cannot be anchored to a glottal stop, tokens that were produced with a word initial glottal stop were excluded from the following analysis. To investigate the alignment of L relative to ISon we employed a model that used RATE (normal, fast), POSITION NP (first, second) and SEGMENT TYPE (vowel, nasal, approximant) as fixed

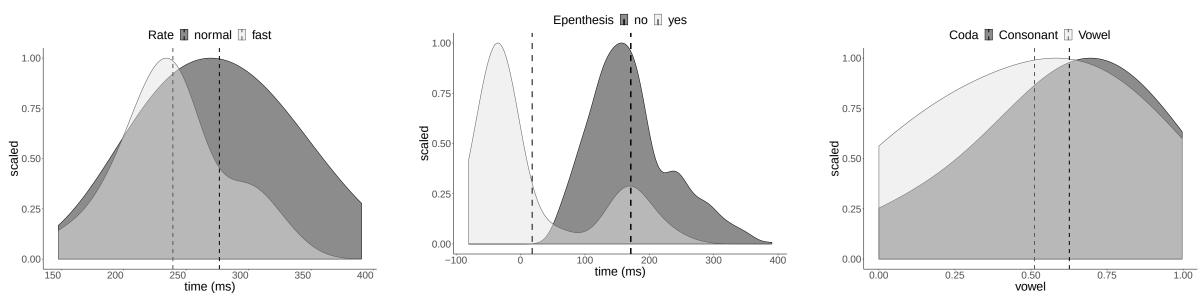
TABLE III. Mean distance of L tone to ISon (in ms), in two positions (first NP, second NP), both speech rates included. Mean duration (in ms) of token initial segments, at two speech rates (normal, fast). Standard deviation in parenthesis.

L to ISon alignment		
Segment	First NP	Second NP
Approximant	15.6 (32.5)	4.43 (47.65)
Nasal	5.1 (14.98)	-2.28 (21.78)
Vowel	58.81 (57.06)	-10.95 (65.43)
Glot. stop	-17.39 (33.15)	-38.26(34.62)
Duration		
Segment	Normal	Fast
Approximant	119.43 (34.58)	99.67 (26.9)
Nasal	110.74 (30.43)	100.81 (25)
Vowel	75.22 (20.74)	73.7 (22.05)
Glot. stop	97.68 (29.73)	91.11 (21.86)

factors, together with participants and items as random intercepts (n=628). Results show that the rate manipulation had no significant effect on where the low tone was placed (Est. 0.9 $\pm$ 2.3 ms, p = .7 (ns.)) but that position (Est. 11.51 $\pm$ 2.28 ms, p = .0001) and segment type (Est. 13.30 ± 5.76 ms, p = .04) had a significant effect. The Tukey corrected post hoc test revealed that there was a significant interaction between segment type and position, for nasals and vowels (Est. 25.19 ± 6.86 ms, p = .02). Hence, it is found that the length of the segment where the tone seeks to anchor has an influence. As Figure 3(a) shows, the low tone is realised closer to ISon when the initial segment is a nasal consonant and further away from ISon when it is a vowel. In case the token word was in the first NP and the initial segment was a vowel then the L tone was realised to the right from ISon, at around 52.81 ms. If it was realised on a vowel in the second NP, then the L tone was placed around 10.95 ms prior to ISon (see 3(b)). As table III shows, this difference was smaller between the two positions for other segments. In sum, there are small, yet consistent differences, as to where the initial low tone can be anchored depending on the position of the token in the utterance.

Additionally, we wanted to test if the location of the low tone was influenced by differences in duration of the initial segment which are related to modifications in speech rate. If we assume that duration of the word initial segment has an influence on where the low tone will be placed and a longer duration (at normal speech rate) of the segment provides more time for the low tone to be anchored further away from the left edge (ISon) and more towards the syllable nucleus, we should observe a positive correlation. If instead the correlation is negative, this indicates that the tone is seeking to align with the left edge of the syllable. To test our hypothesis two Pearson correlation tests were performed, one for tokens at fast, and another for tokens at normal speech rate. Results show that the location of the low tone does not correlate with the duration of the initial segment under the influence of speech rate modifications. For the fast rate, there is a small positive correlation (r=0.02, n=334, $p = 0.77$ (n.s)); for the normal speech rate there is also a small positive correlation (r= 0.02, n=289, p=0.74 (n.s)). The two tests failed in showing a significant effect of correlation between the duration of the initial segment under rate manipulation and the time of alignment of the low tone. In view of this result, it seems fair to discard the hypothesis of speech rate influencing the anchor point of the low tone.

B. Alignment of H



(a) by speech rate

(b) by realisation of epenthesis

(c) by final segment

FIG. 4. Latency in ms to final H tone to onset of the final syllable (FSon) as smoothed density plot with mean values according to: (a) speech rate in monosyllabic tokens; (b) tokens with or without epenthesis. A value lower than zero indicates that the peak was placed prior to the onset of the final syllable. (c) Proportional latency to final H tone to onset of the vowel carrying the peak. The vowel midpoint is at 0.5.

Figures in 4 show estimated density plots with mean values for differences in the alignment on the peak related to (a) a modification of speech rate; (b) final epenthesis realisation. To investigate the alignment of the final peak relative to the onset of the final syllable (FSon) a model was used which included RATE (fast, normal), FINAL SEGMENT (consonant or vowel), EPENTHESIS (yes, no), and WORD LENGTH (short, long) as fixed factors together with participant, token and pause as random intercepts (n=624). Results show that the rate manipulation always had a significant effect on the location of the H tonal target (Est. 12.6 ± 2.92 ms, p <.0001). Also the factors final segment (Est. 37.15 ± 17.14 ms, p <.04), epenthesis (Est. 63.67 ± 21.10 ms, p <.005), and word length (Est. 63.59 ± 15.05 ms, p <.0001) showed a significant effect.

We had observed that some tokens tended to be realised with final glottalization of segments or were followed by a full glottal stop. This was marked in our corpus so that we could investigate if glottalization had an effect on where the peak was placed. Additionally, we were interested in determining the position of the peak within the syllable and found that in 91% of the cases the H peak was realised on a vowel. Further, 35% of the target tokens were produced ending on a coda consonant, however in 81% of these cases the tonal target was also found on a vowel. In fact, words ending on an open syllable or on coda consonant show similar placement of the peak. For syllables ending on a vowel the mean peak location is proportionally at two thirds (0.64) of the syllable and for syllables containing a coda the mean peak location is also at around two thirds (0.64) of the final syllable. Figure 4(c) shows a density plot with the estimated placement of the peak relative to the vowel, depending on whether there is a coda consonant or not on the last syllable. A linear mixed effects model investigated the position of the peak proportionally to the vowel carrying it (n=568). The model included RATE (fast, normal), FINAL SEGMENT (consonant, vowel), GLOTTALIZATION (yes, no), and VOWEL LENGTH (long, short) as fixed factors together with participant, token, as well pause as random intercepts. Results show that the realisation of a final coda had an influence on where the peak was situated on the vowel (Est. 0.112 ± 0.0414, p <.01). The peak was realised earlier in the vowel, when this was the last segment, and later in the vowel, when there was a coda consonant. No other factors, such as the vowel glottalization (p=0.14) or vowel length (p=0.58) showed a significant influence on where the peak was placed in the vowel.

C. Rise time

Rise time was calculated as the difference in ms between the H tone and the preceding L tone. To investigate the influence of rate manipulation on rise time a linear mixed effects model was employed. The model included RATE (fast, normal), WORD LENGTH (short, long), SEX (male, female), and POSITION NP (first, second) as fixed factors as well as participant, token, and pause as random factors. Table IV summarises the results showing that the rise time between the low tone and the high tone was affected by all four factors, but most importantly, by the modification of speech rate. As expected, in the first NP, the rise time was longer at normal speech rate (mean 419.42 ms), than at fast speech rate (mean 396.99 ms). This difference was greater for the second NP for which the mean of the rise time was 442.09 ms at normal speech rate and 395.47 ms at fast speech rate. Overall the rise time of female participants was greater (mean 470.08 ms, at normal speech rate; 427.27 ms, at fast speech rate) than that of the male participants (mean 397.43 ms at normal speech rate; 369.68 ms at fast speech rate). Not surprisingly, longer words displayed greater rise time than short words (mean 538.69 ms vs. 384.61 ms at normal speech rate).

TABLE IV. Comparison and interaction between normal and fast rate with the listed fixed factors.

Factor	Est. SE	df	p
Rate	40.55±5.46	578.18	<.0001
Position	14.60±5.52	579.08	<.008
Sex	65.9 ±26.07	12.23	<.02
Word length	56.52±27.51	144.08	<.04

D. Excursion

The size of excursion was calculated as the difference in Hz between the H tone and the preceding L tone. A linear mixed effects model investigated the effect of the rate manipulation on the size of excursion. As the automatic backward model selection of fixed parts of the model was performed, all factors with the exception of pause PAUSE (yes, no) were eliminated. Results show that there is a significant effect of the pause on the size of excursion (Est. 52.5 ± 11.8 Hz, p <.001). For female speakers, tokens at normal speech rate had an excursion size (mean 46.16 Hz) which increased in case there was a following pause (mean 57.61 Hz). This effect was stronger at fast speech rate, where the excursion size was again greater if a pause followed (mean 60.25 Hz) then when it did not (mean 40.15 Hz). These results indicate that the insertion of a pause after the target token had an influence on the excursion size and that the rate manipulation marginally affected excursion size.

E. Slope

In order to see whether rate influenced the slope of the rise or whether the slope was constant across rates, a paired two-sample Wilcoxon signed-rank test was performed by speaker. The comparison included 252 token words. Results suggest that the distribution is symmetric and the slope is constant (p=.001; V = 12106). However, as the examination of the excursion between the two tones has shown the effect of pause was the only significant factor influencing its size, two additional tests comparing tokens realised with and without a following pause were conducted separately. The number of comparable tokens realised with no pause at the right boundary only left 59 comparisons possible. The result of the test does not support the constant slope hypothesis (p=.08; V = 657). This result could be due to the low number of tokens included. The number of comparisons for tokens followed by a pause was higher (n=122) and the result in this case is in favour of a constant slope (p=.01; V = 2775).

V. DISCUSSION

In this study we sought to examine tonal targets in Drehu and investigate how they align to segmental landmarks. In order to evaluate whether tonal targets are produced to delimit prosodic constituency, token words with varying syllable number, and phonotactic structure (at the end of the word) were included. The results of this study show that in Drehu there is a consistent demarcation of the prosodic word, with a low tonal target at the beginning, and a high tone at the end of target tokens. A variation in syllabic structure of disyllabic words showed that long vowels do not attract additional pitch movements, demonstrating that in Drehu there is no weight sensitivity attracting an *F0* peak. The investigation of the alignment of the low tone revealed that this target is aligned at the left edge of content words in all cases, which means the realisation of a LH or LHL_1H_1 tonal pattern did not have an influence on the position of the tone. There were cases in which tokens beginning on a vowel were realised with a glottal stop as initial segment. When this happened, the tone could not be anchored to this first segment and the *F0* minimum was pushed into the monosyllabic function word that preceded it. Further, it was observed that the factor segment type had an influence on where the low tone was realised. Token-initial nasals differed significantly in duration when compared to vowels and this had an effect on the position of the low tone. The L tone was realised at the onset of the initial syllable when the first segment was a nasal and further away from the the same point when it was a vowel. However, these differences were rather small and the alignment of the L tone was consistently towards the left edge of the token word. The stability in the alignment of the L tone supports the hypothesis of the L being a phrase initial tone which demarcates a constituent like the prosodic word.

Our investigation of the final peak suggests that this tone is prominently marking the last full syllable. The alignment of the high tone confirmed our hypothesis that a long vowel would be the anchor if in word final position or a monosyllabic word. In contrast, our hypothesis about words ending on CVC-nasal syllables could not be confirmed since in the majority of instances (85%) the peak was reached in the vowel of the same syllable. However, it was found that the peak aligns at a proportionally similar distance from the onset of the final syllable which suggests that the peak is aligned in relation to the syllable boundary, similarly to what has been found in Mandarin (Xu, 1998). This result strengthens a hypothesis of the alignment being anchored relative to the onset of the final syllable. An investigation of vowels carrying *F0* peaks showed a significant effect originating from the coda consonant, where the peak was realised later in the vowel, than in syllables with no coda showing that alignment occurs earlier in syllables with CVC but later in CV syllables. Thus, a certain degree of variability is found in the alignment of H which is why we cannot conclude that this tone has one particular segment to which it will always be anchored. These results show that the *segmental anchorage hypothesis* is better suited to account for the observed alignment of the final peak in Drehu. In this approach the anchorage point refers to a region, like the syllable, to which a tone can be anchored, as appears to be the case in these Drehu data. Further examination of the segmental anchoring hypothesis shows that the rate manipulation had a significant effect on rise time but a less strong impact on excursion size. Recall, a strict segmental anchoring hypothesis predicts anchors to be closer together, and the rises to be shorter and steeper as the speech rate increases. An investigation of the rise time, which was calculated as the difference in ms between the high tone an the low tone, showed a consistent effect of rate. Moreover, we could establish a significant effect of word length, sex, and position. This indicates increased speaking rate resulted in tonal targets being realised closer together, as the segmental anchoring hypothesis would predict. We did not find the same effect of rate on the excursion size of *F0*, which was calculated as the difference in Hz between the high and the low tones. Instead, we found a significant effect conditioned by the insertion of a pause after the target token. A follow up analysis that compared the slope of the same tokens by participant and rate suggests the slope is mostly constant although some variability is found. Arguably, the insertion of pause could have and effect and confound the measurement of slope. Therefore, separate statistical tests were conducted for tokens that were produced with or without a following pause. The result for the slope of tokens realised with a subsequent pause suggest a symmetric distribution but this was not confirmed for tokens without following pause. Since the test for tokens with a subsequent pause included considerably more items (122) than the other (59) the result could be due to a lack of statistical power but a clear conclusion cannot be determined at this point. It was observed that the pause mostly occurred after the second NP which could be due to prosodic phrasing since two NPs represent a larger constituent. With regard to the significant effect of pause on excursion size, it is proposed that the pause participates in marking a higher level in the prosodic hierarchy. As known from research on prosodic constituency of other languages, the insertion of a pause, as well as significant manipulations of *F0* prior to the pause, are considered strong indicators of a higher level prosodic constituent. We therefore propose that the Intonation Phrase in Drehu does not only end on a phrase final low tone but can also be realised with a high boundary tone.

VI. CONCLUSION

The present study finds evidence for a stable initial low tone and a more variable final high tone demarcating the prosodic word in Drehu. We propose that tonal alignment is better described with the segmental anchorage hypothesis, since we find that the word final prominence seeks to anchor with respect to the left boundary of the syllable, but that there is variability regarding the segment to which the peak will be anchored. Moreover, it was found that *F0* minima seek to align towards the left edge of nouns and that neither a variation of tonal pattern, nor a rate manipulation had an effect on this. The small variability in the alignment of this tone could be understood a indicator of an initial phrase tone. The final high tone showed variability under the pressure of rate manipulation. The shortening of rise time in fast speech rate indicates segmental anchors are closer together. However, the excursion size of the rise was not affected in the same way, meaning that the rises were not steeper. Instead, it seems that speakers seek to reach the same *F0* level despite time pressure. Interestingly, it was found that the insertion of a pause had an effect on the excursion size which could be related to the demarcation of a higher prosodic level, namely the Intonation Phrase. In view of these results, it is proposed that Drehu shows a phrasal prominence marking system with an initial L tone and a prominence lending H tone at the right edge of the prosodic word. Contrary to earlier impressionistic accounts (Lenormand, 1954; Tryon, 1968a), this study on tonal alignment supports the claim that a phrasal rather than strictly lexical prominence marking is found in Drehu (Torres *et al.*, 2018). Arguably, tonal patterns outlined in this experiment and alignment of tunes to segmental landmarks resemble those of the contact language French although this requires further investigation. This study has investigated short noun phrases containing a function word and a noun. It would be of interest to test larger prosodic constituents and examine if this would lead to a grouping of a higher level prosodic constituent. Further, to better identify how the peak is anchored in different segmental contexts, more data including a larger set of words ending on CVC, and a larger set of consonants at the onset of the syllable, would be of interest in a future study. Based on the data we cannot conclude that the attested use of *F0* is the result of language contact. Further studies examining Drehu prominence marking and use of information structure are therefore planned.

APPENDIX: SPEECH MATERIALS

(1a) Kola drenge la emexemine ngöne la hlapa / DUR hear ART echo in ART field / "The echo can be heard in the field". (1b) Kola drenge ngöne la hlapa la emexemine me kola ulil / DUR hear in ART field ART echo and DUR noise / "In the field the echo and the noise can be heard".

(2a) Ame la maja, tre, ka a nyipi ewekë troa e i / PRS.1 ART bait PRS.2 STAT PRS important to catch fish / "Then the bait is important to catch fish". (2b) Ame la thiitre me la maja, tre, ka a nyipi ewekë troa e i / PRS.1 ART trap and ART bait PRS.2 STAT PRS important to catch fish / "Then the trap and the bait are important to catch fish".

(3a) Ame la mamala, tre, hna xupe ngöne la uma / PRS.1 ART window PRS.2 PST build in ART house / "Then the window is part of the house". (3b) Ame la qënelö me la mamala, tre, hna xupe ngöne la uma / PRS.1 ART door and ART window PRS.2 PST build in ART house / "Then the door and the window are part of the house".

(4a) Ame la maamu, tre, hna sile hnei itre qatr/ PRS.1 ART bogeyman PRS.2 PST invent A PL old / "The bogeyman was invented by the old". (4b) Ame la satana me la maamu, tre, hna sile hnei itre qatr / PRS.1 ART devil and ART bogeyman PRS.2 PST invent A PL old / "The devil and the bogeyman were invented by the old".

(5a) Thaa tro kö a nue la wamine falawa hune la laulau/ NEG FUT emphasis PRS leave ART crumb bread on ART table/ “No (bread) crumb should be left on the table”. (5b) Thaa tro kö a nue la laese me la wamine falawa hune la laulau/ NEG FUT emphasis PRS leave ART rice and crumb bread on ART table/ “No rice and (bread) crumb should be left on the table”.

(6a) Ame la imaag, tre, e kuhu ngöne gadran/ PRS.1 ART mango tree PRS.2 down there in garden / “Then the mango tree is down there in the garden”. (6b) Ame la ipiic me la imaag, tre, e kuhu ngöne gadran/ PRS.1 ART orange tree ART mango tree PRS.2 down there in garden / “Then the orange tree and the mango tree are down there in the garden”.

(7a) Ame la maage, ka lapa hunei sinöe/ PRS.1 ART mango PRS.2 STAT be on tree / “Then the mango is on the tree”. (7b) Ame la waco me maage, ka lapa hunei sinöe/ PRS.1 ART bird and ART mango PRS.2 STAT be on tree / “Then the bird and the mango are on the tree”.

(8a) Ame la mani, tre, ka catrehnine la ijine hnötr/ PRS.1 ART rain PRS.2 STAT strong ART period cold / “Then the rain is very strong this winter”. (8b) Ame la enyi me la mani, tre, ka catrehnine la ijine hnötr/ PRS.1 ART wind and ART rain PRS.2 STAT strong ART period cold / “Then the wind and the rain are very strong this winter”.

(9a) Ame la mano, tre, ketre götrane la ngönetrei/ PRS.1 ART chest PRS.2 INDF part ART body / “Then the chest is a part of the body”. (9b) Ame la qëmeke me mano, tre, ketre götrane la ngönetrei/ PRS.1 ART face and chest PRS.2 INDF part ART body / “Then the chest is a part of the body”.

(10a) Ame la maano, tre, hna hamëne kowe la joxu/ PRS.1 ART fabric PRS.2 PST give to ART chief / “Then the fabric was given to the chief”. (10b) Ame la itre koko me la maano, tre, hna hamëne kowe la joxu/ PRS.1 ART PL yam and ART fabric PRS.2 PST give to ART chief / “Then the yams and the fabric were given to the chief”.

(11a) Ame la melem, tre, ka mingöming/ PRS.1 ART full moon night PRS.2 STAT beautiful / “Then the full moon night is beautiful”. (11b) Thaupune la jidri ne melem, tre, ka mingöming/ calm ART night of full moon night PRS.2 STAT beautiful / “The calm of the full moon night is beautiful”.

(12a) Ame la melimala, ka za/ PRS.1 ART green turtledove STAT beautiful / “Then the green turtledove is beautiful”. (12b) Ame la fenifene me la melimala, ka za/ PRS.1 ART butterfly and ART green turtledove STAT beautiful / “Then the butterfly and the green turtledove are beautiful”.

(13a) Ame la munun, tre, ka mele ngöne la hnagejë/ PRS.1 ART fish picot PRS.2 STAT be in ART sea/ “Then the picot fish is in the sea”. (13b) Ame la eötre me la munun, tre, ka mele ngöne la hnagejë/ PRS.1 ART fish picot PRS.2 STAT be in ART sea/ “Then the shark and the picot fish are in the sea”.

(14a) Ame la wahnawa, tre, ka hnyapa la itre wen/ PRS.1 ART banana PRS.2 STAT sweet ART PL fruit/ “Then the banana is a sweet fruit”. (14b) Ame la guafa me la wahnawa, tre, ka hnyapa la itre wen/ PRS.1 ART guava ART banana PRS.2 STAT sweet ART PL fruit/ “Then then guava and the banana are sweet fruits”.

(15a) Ame la wanangongo, tre, ie ka lapa ngöne la huuca/ PRS.1 ART fish dawa PRS.2 STAT live in ART reef/ “Then the dawa fish lives in the reef”. (15b) Ame la nunuce me wanangongo, tre, ie ka lapa ngöne la huuca/ PRS.1 ART pufferfish ART fish dawa PRS.2 STAT live in ART reef/ "Then the pufferfish and the dawa fish live in the reef".

List of abbreviations: ART, article; DUR, durative; FUT, futur; INDF, indefinite; NEG, negation; PL, Plural; PRS, present; PRS.1, presentative1; PRS.2, presentative2; PST, past; STAT, stative.

[1] Although the authors do not specify which dialect of English is spoken by the participants in their study, it seems fair to assume it is one American variety.

[2] Indigenous people from New Caledonia self identify as Kanak (Vernaudon, 2015).

Arvaniti, A., Ladd, D. R., and Mennen, I. (**1998**). "Stability of tonal alignment: the case of greek prenuclear accents," Journal of phonetics **26**(1), 3–25.

Atterer, M., and Ladd, D. R. (**2004**). "On the phonetics and phonology of 'segmental anchoring' of F0: evidence from German," Journal of Phonetics **32**(2), 177–197.

Bates, D., Mächler, M., Bolker, B., and Walker, S. (**2015**). "Fitting linear mixed-effects models using lme4," Journal of Statistical Software **67**(1), 1–48, doi: 10.18637/jss.v067.i01.

Beckman, M. E., and Hirschberg, J. (**1994**). "The ToBI annotation conventions," Ohio State University .

Boersma, P., and Weenink, D. (**2017**). "Praat: doing phonetics by computer (version 6.0.26)[computer program]. Retrieved November 2, 2017" .

Bruce, G. (**1977**). “Swedish accents in sentence perspective,” Working Papers in Linguistics **12**(1).

Crowley, T., Lynch, J., and Ross, M. (**2013**). *The oceanic languages* (Routledge).

D’Imperio, M., and Michelas, A. (**2014**). “Pitch scaling and the internal structuring of the intonation phrase in french,” Phonology **31**(1), 95–122.

ISEE (**2009**). “Recensement général de la population” Institut de la Statistique et des Études Économiques.

Ishihara, T. (**2003**). “A phonological effect on tonal alignment in tokyo japanese,” in *15th ICPhS*, pp. 615–618.

Kisler, T., Reichel, U., and Schiel, F. (**2017**). “Multilingual processing of speech via web services,” Computer Speech & Language **45**, 326–347.

Ladd, D. R. (**2008**). *Intonational phonology* (Cambridge University Press).

Ladd, D. R., Faulkner, D., Faulkner, H., and Schepman, A. (**1999**). “Constant “segmental anchoring” of F0 movements under changes in speech rate,” The Journal of the Acoustical Society of America **106**(3), 1543–1554.

Ladd, D. R., Mennen, I., and Schepman, A. (**2000**). “Phonological conditioning of peak alignment in rising pitch accents in dutch,” The Journal of the Acoustical Society of America **107**(5), 2685–2696.

Ladd, D. R., Schepman, A., White, L., Quarmby, L. M., and Stackhouse, R. (**2009**). “Structural and dialectal effects on pitch peak alignment in two varieties of british english,” Journal of Phonetics **37**(2), 145–161.

Lenormand, M.-H. (**1954**). "La phonologie du mot en lifou (îles loyalty)," Journal de la Société des Océanistes **10**(10), 91–109.

Magulué, P. F. (**2011**). "Les teachers du pacifique au xixe siècle ou l'émergence d'une nouvelle élite océanienne entre tradition et modernité," Histoire et missions chrétiennes (4), 139–156.

Monnin, J. (**2010**). "Influence de la langue ambiante sur l'acquisition phonologique: une comparaison du français et du drehu," Ph.D. thesis, Université de Grenoble; Université de Nouvelle Calédonie.

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

Pierrehumbert, J., and Beckman, M. (**1988**). *Japanese tone structure (Linguistic Inquiry monograph 15)* (Cambridge, MA: MIT Press).

Pierrehumbert, J. B. (**1980**). "The phonology and phonetics of english intonation," Ph.D. thesis, Massachusetts Institute of Technology.

Pitrelli, J. F., Beckman, M. E., and Hirschberg, J. (**1994**). "Evaluation of prosodic transcription labeling reliability in the tobi framework," in *Third International Conference on Spoken Language Processing*.

Prieto, P., and Torreira, F. (**2007**). "The segmental anchoring hypothesis revisited: Syllable structure and speech rate effects on peak timing in spanish," Journal of Phonetics **35**(4), 473–500.

Prieto, P., Van Santen, J., and Hirschberg, J. (**1995**). "Tonal alignment patterns in Spanish," Journal of Phonetics **23**(4), 429–451.

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

Silverman, K., and Pierrehumbert, J. (**1990**). "The timing of prenuclear high accents in English," Papers in laboratory phonology I 72–106.

Torres, C., Fletcher, J., and Wigglesworth, G. (**2018**). "Investigating word prominence in drehu," in *Proc. 17th Speech Science and Technoloy conference*.

Truckenbrodt, H. (**2002**). "Upstep and embedded register levels," Phonology **19**(1), 77–120.

Tryon, D. T. (**1967**). *Nengone Grammar*, **6** of *Pacific Linguistics, Series B* (Australian National University, Canberra).

Tryon, D. T. (**1968**a). *Dehu grammar* (Australian National University).

Tryon, D. T. (**1968**b). *Iai Grammar*, **8** of *Pacific Linguistics, Series B* (Australian National University, Canberra).

Tryon, D. T., and Dubois, M.-J. (**1969**). "Nengone dictionary. part i, nengone-english," in *Pacific Linguistics*, **Series C-Books, No. 9** (ERIC).

Vernaudon, J. (**2015**). "Linguistic Ideologies: Teaching Oceanic Languages in French Polynesia and New Caledonia," The Contemporary Pacific **27**(2), 433–462.

Welby, P. (**2006**). "French intonational structure: Evidence from tonal alignment," Journal of Phonetics **34**(3), 343–371.

Welby, P., and Lœvenbruck, H. (**2006**). "Anchored down in anchorage: Syllable structure and segmental anchoring in French," Italian Journal of Linguistics/Rivista di linguistica **18**, 74–124.

Winkelmann, R., Harrington, J., and Jänsch, K. (**2017**a). "Emu-sdms: Advanced speech database management and analysis in r," Computer Speech & Language **45**, 392–410.

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

Xu, Y. (**1998**). "Consistency of tone-syllable alignment across different syllable structures and speaking rates," Phonetica **55**(4), 179–203.