

# 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<sup>a)</sup>*

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

---

<sup>a)</sup> [catalina.torres@unimelb.edu.au](mailto:catalina.torres@unimelb.edu.au)

## 16 I. INTRODUCTION

### 17 A. Tonal alignment

18 Within autosegmental-metrical phonology (Ladd, 2008; Pierrehumbert, 1980), the study  
19 of text-to-tune alignment has shown for a number of languages, that a series of phonetic  
20 factors and language specific phonological characteristics come into play in determining  
21 how the speech stream is organized. One important concept that has received considerable  
22 attention is the *segmental anchoring hypothesis*, a term that refers to how tonal targets,  
23 realised through fundamental frequency, unfold in time while being coordinated with the  
24 segmental string (Arvaniti *et al.*, 1998; Ladd *et al.*, 1999, 2009). According to Arvaniti *et al.*  
25 (1998), this hypothesis predicts that a pitch accent is aligned or *anchored* to a segmentally  
26 defined position, and the duration and rise time of the pitch movement are determined by  
27 the segmental composition of the accented word. Strictly speaking, the segmental anchoring  
28 hypothesis predicts anchors, of low and high tones, to be closer together, and the rises to be  
29 shorter and steeper as the speech rate increases. Arvaniti *et al.*'s results for Modern Greek  
30 and the concluding theoretical assumptions of their study were in agreement with similar  
31 observations that had been previously made for Stockholm and Malmö Swedish, Mexican  
32 Spanish, and English<sup>1</sup> (Bruce, 1977; Prieto *et al.*, 1995; Silverman and Pierrehumbert, 1990).  
33 It is important to bear in mind that these studies were concerned with the realisation of  
34 prenuclear accents and also that the languages investigated shared one important phonolog-  
35 ical property, namely lexical prominence. To date a number of studies have demonstrated  
36 that language specific phonological patterns such as syllabic structure can have an influence

37 on the anchoring points of  $F_0$  maxima. In an experiment testing tone realisation in CV  
38 and CVC-nasal syllables, it was found that regardless of syllable structure, the tone always  
39 aligned at a similar point in the syllable in Mandarin (Xu, 1998). However, peaks of accents  
40 are not always realised in the same position but can be aligned variably. An example of  
41 variable alignment comes from Japanese, for which it was found that the peak alignment can  
42 occur earlier, when the first mora was accented in a two-mora syllable, like CVV or CVN,  
43 than in words with a one-mora syllable, like CV (Ishihara, 2003). A similar peak alignment  
44 has been observed in Dutch where peaks of prenuclear accents are aligned at the end of a  
45 vowel if the syllable contains a long vowel, but during the following consonant if the accented  
46 syllable contains a short vowel (Ladd *et al.*, 2000). Also studies on Peninsular Spanish and  
47 French, have cast some doubt on the strict segmental anchoring hypothesis since they have  
48 shown that accented syllables containing a coda consonant can be aligned differently when  
49 compared to those with no coda consonant (Prieto and Torreira, 2007; Welby and Løeven-  
50 bruck, 2006). In view of variable anchoring points for peaks in French, where final maxima  
51 could be anchored to either the vowel or the final nasal consonant, in a CVC-nasal syllable,  
52 Welby and Løevenbruck (2006) propose the notion of *segmental anchorage*. In contrast to a  
53 strict segmental anchoring, this term refers to a region, like the syllable (and not exclusively  
54 a segment), to which a tone can be anchored. Moreover, it has been observed that align-  
55 ment patterns can differ between varieties of the same language. For example, it could be  
56 established that the location of  $F_0$  minima varies between Northern and Southern German  
57 (Atterer and Ladd, 2004), just like it was found for two varieties of English (Scottish Stan-  
58 dard English and Southern British English) although differences did not reach a statistically

59 significant level in the latter study (Ladd *et al.*, 2009). Anchoring points of  $F0$  minima and  
60 maxima offer a way to examine how the speech stream is organised, especially regarding  
61 the marking of prominence. In addition, the modulation of  $F0$  over larger constituents,  
62 accompanied by other markers, like pause, are considered to be indicators distinguishing be-  
63 tween different prosodic levels in a number of languages (D’Imperio and Michelas, 2014, for  
64 French), (Truckenbrodt, 2002, for German), (Pierrehumbert and Beckman, 1988, for Tokyo  
65 Japanese). As Ladd *et al.* (2009) note, there is a considerable number of questions as to  
66 how language specific phonological factors condition the alignment of accentual peaks. In  
67 the present study, the goal is to describe prominence lending tonal targets and their position  
68 relative to the segmental string. The aim is to verify whether the *strict segmental anchoring*  
69 or a *segmental anchorage hypotheses* are best suited for a description of Drehu text to tune  
70 alignment. With this, we wish to find out whether tonal targets also contribute to the mark-  
71 ing of prosodic constituency beyond the prosodic word. Hence, our analysis also considers  
72 effects related to the marking of higher levels of prosodic constituency.

## 73 B. Drehu

74 Drehu is an Oceanic language from Lifou, New Caledonia, and the language with the  
75 largest number of speakers in the archipelago (Vernaudon, 2015). According to the 2009  
76 census (ISEE, 2009), Lifou counts around 8600 inhabitants from which approximately 5500  
77 are Drehu speakers older than 14 years old. Protestant missionaries developed the first  
78 writing system (Magulué, 2011), which is in its majority still used for Drehu language class  
79 in primary, high school, and religious Sunday school. New Caledonia is a French overseas

80 territory and the education system on the island follows the French metropolitan model.  
81 This means that apart from the optional Drehu language class all other subjects are taught  
82 in French. Today almost all speakers, especially younger generations, are bilinguals of French  
83 and Drehu.

84 Within the Oceanic languages, Drehu has been classified as a language of the Southern  
85 Melanesian linkage which belongs to the Loyalty Islands family ([Crowley \*et al.\*, 2013](#)). Drehu,  
86 Iaii, and Nengone are closely related languages (Loyalty Islands) and were described as  
87 having similar phonological systems. They share several properties, like a rich vowel system  
88 that includes a length distinction, a lack of weight sensitivity regarding stress, which in  
89 turn is described as demarcative and fixed. Similarly, in Drehu, Iaii, and Nengone the  
90 rhythm type was analysed as trochaic ([Tryon, 1967, 1968a,b](#); [Tryon and Dubois, 1969](#)).  
91 However, phonetic studies further investigating the acoustics of rhythm or prosody in these  
92 three languages are under-represented. Regarding the phonology of Drehu, it has been  
93 established that there is a 30 consonant system including stops, nasals, fricatives, laterals,  
94 and approximants. Further, there are fourteen vowels and this inventory includes a contrast  
95 between short and long vowels. The language's phonotactics allow a syllabic structure with  
96 the combinations V, VC, CV, CVC and VV, VVC, CVV, CVVC. It has been established that  
97 there is a strong preference for syllabification that does not allow for consonant clusters in  
98 the word. Descriptions also note that in word final position, a syllable ending with a coda  
99 consonant may receive a supplementary epenthetic vowel /e/ if the following word starts  
100 with a consonant ([Lenormand, 1954](#); [Moyse-Faurie, 1983](#); [Tryon, 1968a](#)). Some accounts  
101 offer an informal description of prominence and classify Drehu as a stress language with a

102 word prosodic system. According to [Lenormand \(1954\)](#), stress (accent d'intensité), always  
103 falls on the first syllable of a word, pëkö ['pɛkɔ] (*none, there is nothing*), fifikë ['fifikɛ] (*toy*).  
104 In word derivation, when words obtain a prefix, the stress pattern remains and main stress  
105 shifts to the inserted first syllable e.g.: malan ['malan] (*to fall*) vs. amalan ['amalan] (*CAUS-*  
106 *fall*). Compound words are described similarly, meaning that stress always shifts to the first  
107 syllable of the word. Further, [Tryon \(1968a\)](#) proposes secondary stress in polysyllabic words,  
108 with it always falling on the third syllable ['ama,lan] (*CAUS.fall*). Finally, stress is described  
109 as not being weight sensitive and is said to have a demarcative function, which marks out  
110 word edges. Although not much has been noted about intonation in Drehu, [Tryon \(1968a\)](#)  
111 describes sentence final intonation as being characterised by a fall in pitch to a low tone  
112 towards the end of what we refer to here as the Intonation Phrase.

113 More recent studies have shown that Drehu might be undergoing phonetic changes since  
114 it was described in the fifties and early eighties. The once attested voiced and voiceless  
115 retroflex stops /d/ and /t/ are now realised as voiced and voiceless alveolar affricates / $\hat{d}$ / and / $\hat{t}$ / ([Monnin, 2010](#)). Moreover, a recent phonetic investigation on word prominence in  
116 Drehu, suggests that a phrasal rather than strictly lexical prominence marking is observed  
117 ([Torres et al., 2018](#)). This study investigated words in informational focus and found that  
118 there is a preference to make the right edge prosodically prominent through longer duration  
119 and an *F0* peak on the final syllable. Further, the most frequent intonation pattern observed  
120 was a rising pattern that spanned over the word in focus. Since this study investigated words  
121 in informational focus the question remains if there is indeed word initial lexical stress or  
122 if there is rather a phrasal prominence marking system. The rapidly changing linguistic  
123

124 landscape, where the people of Lifou have turned into a bilingual community, might be  
125 influencing some of these phonetic changes. This hypothesis needs however to be evaluated  
126 which is why more phonetic research is needed. With this study of synchronic tune to text  
127 alignment in Drehu, we wish to add to the body of research on Melanesian languages of New  
128 Caledonia, and the phonetics of an under-documented language.

## 129 II. THE CURRENT STUDY

130 In this experiment, we sought to investigate the  $F0$  alignment of prominent tonal tar-  
131 gets in Drehu. Further, we evaluate how modulation in scaling of  $F0$  relates to prosodic  
132 constituency. According to intonational phonology (Ladd, 2008), tonal events are defined  
133 through relative contrast from one tone to another. This contrast is perceived in relative  
134 tonal height and differs depending on the speaker’s tone range. *High* tones are denoted with  
135 an (H) and *low* tones with an (L). Tonal events can be simple monotonal targets (H and L)  
136 but they can also be complex and bitonal (HL) falling or (LH) rising tones. Based on pre-  
137 vious findings (Torres *et al.*, 2018), we hypothesise that there will be a preference to tonally  
138 mark the left edge and the right edge of a prosodic constituent respectively with a low tone  
139 on the initial syllable, and a high tonal target on the word final syllable. Since no phonetic  
140 evidence has been found to support a word-initial stress hypothesis, an examination of tonal  
141 alignment should provide further evidence either in favour or against the proposed phrasal  
142 prominence for Drehu. We are interested in investigating if the predicted rising tonal pat-  
143 tern prevails under non-focal conditions and seek to determine location and timing patterns  
144 relative to segmental landmarks. Moreover, we are interested in investigating the magnitude

145 of rise excursion and whether under manipulation of speech rate, slope is constant. This  
146 will be important in determining whether prominent tones in Drehu are best described with  
147 the *strict segmental anchoring* or the *segmental anchorage* hypotheses. Based on previous  
148 findings for Japanese and Dutch (Ishihara, 2003; Ladd *et al.*, 2000), we further predict that  
149 a long vowel in the final syllable will be the anchor of the H tone, which should be found  
150 towards the end of the long vowel. Similarly, it is expected that for word final CVC sylla-  
151 bles, the constituent final H tone won't be placed on the vowel but on the following sonorant  
152 coda. Additionally, it is expected that the H tone will be placed somewhere on the last full  
153 syllable and not on the syllable containing the epenthetic vowel.

### 154 III. METHODOLOGY

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

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

### 168 A. Speech materials

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

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

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

176 ‘The bogeyman was invented by the old’

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

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

179 ‘The devil and the bogeyman were invented by the old’

180 This manipulation of position allowed us to verify if the predicted rising pattern would be  
181 constant despite changes in its placement within the utterance. The selected items of interest  
182 always started on either a vowel, a nasal consonant, or a voiced labio-velar approximant and  
183 mostly contained sonorant consonants or approximants in order to ensure a continuous  $F0$   
184 curve. Attention was given that when this was not the case, the segments in the target tokens  
185 were voiced, with the exception of the word *emexemin* /emexemin/ (“echo”) where the velar

186 fricative isn't. As Table I shows, the items of interest could be grouped into two categories,  
187 according to whether they were short or long. There were nine disyllabic target tokens and  
188 six others with three, four and five syllables with varying syllable structure. The target token  
189 was embedded in a sentence. It was never placed at the beginning or end of the sentence  
190 in order to avoid any confounding effects clearly related to prosodic boundary marking at  
191 a higher level, like the Intonation Phrase. The instructions of the language informant were  
192 always followed and only phrases that were judged to sound natural and were in agreement  
193 with the current Drehu punctuation were used in the experiment. It should be noted that  
194 there is variability in the occurrence of the epenthetic vowel. Speakers varied in cases where  
195 the final syllable ended on epenthesis, and some times omitted the vowel, producing so a  
196 shorter word ending on a coda consonant. We also observed that tokens starting with a  
197 vowel (unexpectedly) tended to be produced with strong glottalization or a glottal stop.

## 198 **B. Speakers**

199 Thirteen speakers of Drehu were recorded, in a quiet room, at Lifou's high school Lycé  
200 Polyvalent des Îles. Participants provided basic demographic information (age, tribe) and  
201 responded to a sociolinguistic questionnaire that also included questions about language  
202 use. Seven were male, all of them identified as Kanak<sup>2</sup>, and all reported being raised in a  
203 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)

## 204 C. Recording and analysis procedures

### 205 1. *Recording*

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

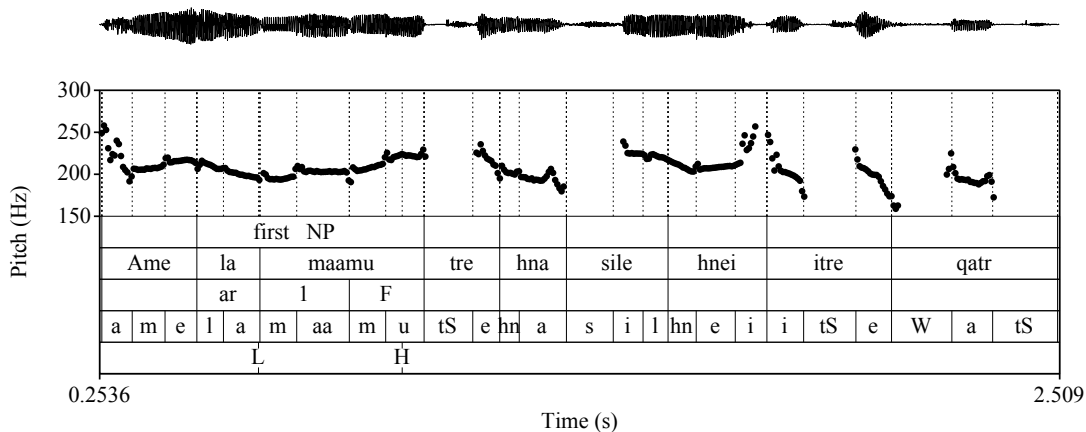


FIG. 1.

*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.

210 experiment. This ensured that the participants understood the experimental procedure.  
 211 The material for the current experiment was presented on a slide show written in Drehu  
 212 and which included no written French. Speakers had time to familiarise them selves with  
 213 the sentences and read these then twice at a self selected normal and a fast speech rate.  
 214 They were recorded at a sampling rate of 48kHz and 24-bit depth, using a Zoom H6 Handy  
 215 recorder, and a head mounted-microphone. The sentence materials were randomised. Par-  
 216 ticipants were asked to read the sentences as naturally as possible, and were asked to repeat  
 217 any misread sentences. It should be noted that participants seemed more at ease during the  
 218 French recordings and that reading aloud in Drehu appeared to be unusual leading so to

219 more hesitations and mispronunciations. Data that contained hesitations or pauses within  
220 the target tokens were discarded and only tokens that were produced in a prosodic con-  
221 stituent larger than the word were kept for analysis. The data analysed correspond to 85%  
222 of the corpus.

## 223 2. *Analysis procedures*

224 The first author manually transcribed and then force aligned the recorded utterances in  
225 WebMAUS, using a grapheme to phoneme conversion, with a parameter model based on  
226 SAMPA (Kisler *et al.*, 2017). This alignment process provided a textgrid for each utterance  
227 in which all phonemes were segmented and marked with boundaries. Since automatic forced  
228 alignment does not detect phonological processes in Drehu, manual correction was required.  
229 All utterances were thereafter visually inspected in Praat 6.0.48 (Boersma and Weenink,  
230 2017) and the segmentation of target tokens was corrected when necessary. Additionally, a  
231 broad ToBI-style annotation (Beckman and Hirschberg, 1994; Pitrelli *et al.*, 1994) was used  
232 to mark *F0* minima and maxima points in the items of interest. The tonal targets were  
233 marked with L for low and H for high tones. Every subsequent tone identified in a token  
234 was also marked with an additional number (e.g. LHL<sub>1</sub>H<sub>1</sub>).

235 During the correction of segmental alignment, special attention was paid to the setting  
236 of phone boundaries. Following Ladd *et al.* (2000), a boundary was set between vowels  
237 and nasals, laterals or approximants at the point where sudden changes in both amplitude  
238 and formant structure occurred. In case the change in formant structure was gradual, the  
239 segment boundaries were marked at the midpoint of the transition from vowel to liquid or

240 approximant. For obstruents, the start of closure was marked as the onset and the start of  
241 high amplitude periodicity was marked as the onset of the next vowel. Fricatives were marked  
242 at the onset and offset of high amplitude aperiodicity. Pauses and strong glottalization, in  
243 form of creaky voice or a glottal stop, that were inserted prior or after the target tokens  
244 were also identified and marked.

245 Figure 1 shows the notation used for the data and how target tokens (including function  
246 words preceding it), position, syllables, phones, and tones were labelled. A hierarchical  
247 database was constructed using the EMU Speech Database Management System ([Winkel-](#)  
248 [mann et al., 2017a](#)). It included five tiers for the tones, phonemic segments, syllables,  
249 words, and target token position. The acoustic and durational characteristics produced in  
250 the target words were queried and analysed using the emuR package in R ([R Core Team,](#)  
251 [2017](#); [Winkelmann et al., 2017b](#)). The  $F0$  contour was visually inspected in Praat and the  
252 two tonal targets were marked according to fundamental frequency minima and maxima  
253 points that were found. The location of L was defined as the  $F0$  minimum, or elbow of  
254 the prosodic constituent and mostly occurred around the first segment of the test word. In  
255 some instances, especially due to strong glottalization, when the pitch curve was not contin-  
256 uous, the lowest point indicating an elbow was chosen.  $F0$  maxima points were also visually  
257 identified and marked with H. In some polysyllabic words, there appeared additional H and  
258 L tones which were also marked. Following [Ladd et al. \(2000\)](#), in case consonants caused  
259 perturbations on the  $F0$  track (for instance, nasal and laterals provoking slight dips and  
260 rises in the  $F0$  contour at onsets and offsets) the annotator compensated for confounding  
261 effects by ignoring  $F0$  points that were clearly perturbations, and marking the next highest

262 or lowest point as measurement point. In case there was a plateau, a valley or two points of  
 263 equal  $F0$  value, the tonal target was marked at its very beginning.

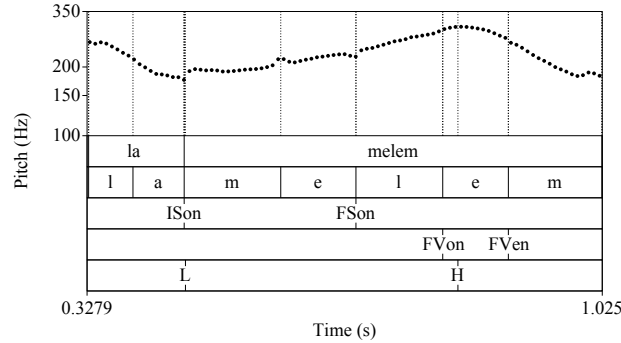


FIG. 2.

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*.

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

- 269 1. Alignment L : L minus ISon in ms; negative values indicate alignment of L before ISon  
 270 ;

271 2. Alignment H : H minus F<sub>Son</sub> in ms; negative values indicate alignment of H before  
272 F<sub>Son</sub> ;

273 3. Peak in FV : H minus F<sub>Von</sub> (in ms) divided by DurFV

274 Additionally, measurements were taken for the alignment of *F0* events relative to segmen-  
275 tal events and the excursion size of rises between *F0* minima and maxima points. Finally,  
276 duration values were queried for the vowel of the last syllable (DurFV). To investigate if  
277 segmental anchoring was constant under rate manipulation a value for slope was calculated.  
278 The formula used divides the *F0* excursion size of a rise by its rise time. More precisely, the  
279 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  
280 the *F0* value of the H tone,  $x^a$  is the realisation in time of the L tone, and  $x^b$  the time of  
281 the H tone, for both the fast and normal speech rates (Welby, 2006).

#### 282 D. Statistical analyses

283 Data were analysed using linear mixed effects models investigating each of the following  
284 parameters of interest : (i) the alignment of the low tone, (ii) the duration of segments (iii)  
285 the alignment of the high tone, as well as (iv) the rise time, and (v) the size of the excursion  
286 in fundamental frequency. Statistical analyses were carried out in R (R Core Team, 2017)  
287 with help of the statistics package lme4 (Bates *et al.*, 2015). An automatic backward model  
288 selection of fixed parts of the linear mixed model was performed in all cases. Further,  
289 estimated marginal means were obtained for factor comparisons and p-values were adjusted

290 with the Tukey method. Additionally, correlation tests were used and when needed Wilcoxon  
291 signed-rank tests were used for direct comparisons of the slope between rates.

#### 292 IV. RESULTS

293 It was found that the majority of tokens (93%) were realised with a LH rising pattern  
294 and only a small set (7%) were produced with a LHL<sub>1</sub>H<sub>1</sub> pattern. Overall, the initial L  
295 tone showed greater stability and tokens with both patterns (LH and LHL<sub>1</sub>H<sub>1</sub>) were kept  
296 for statistical analysis. The LHL<sub>1</sub>H<sub>1</sub> pattern was more often found in the normal speech  
297 rate and occurred on words with three or more syllables. Usually, the initial rise contained  
298 a peak (mean *F0* 200 Hz) that was followed by a rise to a peak with a higher pitch (mean  
299 *F0* 220 Hz). The initial rise is not the rule and is variable which is why its function cannot  
300 be determined yet. Since an influence from the initial peak on the second peak could not  
301 be ruled out, we restricted our analysis of the alignment of final peaks on tokens realised  
302 with a LH pattern. Further, it was found that if there was an insertion of a pause, this  
303 predominantly occurred after the second NP, namely in 211 instances.

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

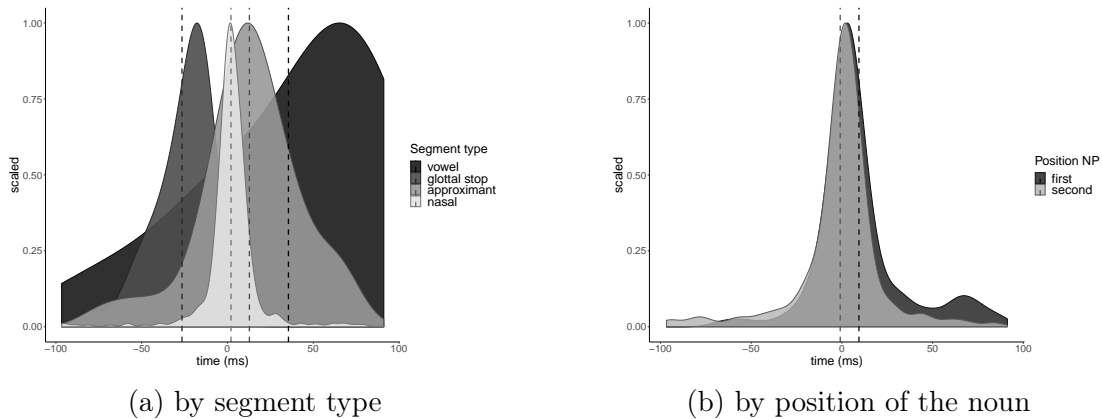


FIG. 3. Latency to initial L tone to onset of the syllable initial segment as smoothed density plot with mean values according to (a) segment where the tone is placed; (b) position of the noun in the phrase. A value lower than zero indicates that the low tone was placed prior to the onset of token word.

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

### 315 A. Alignment of L

316 A low tone was always found in the vicinity of the first segment of the token words,  
 317 regardless of tonal pattern, which is why all low tones were kept for the analysis. A linear  
 318 mixed effects model that investigated the alignment of the low tone (in ms) relative to the  
 319 onset of the first syllable (ISon) was employed (n=664). The automatic backward model  
 320 selection kept the factors POSITION NP (first or second) and SEGMENT TYPE (vowel, nasal,

321 approximant, glottal stop) as fixed factors, together with participants and items as random  
322 intercepts. The factor speech rate was not retained in the model. Figures in 3 show the  
323 estimated alignment and mean values of the low tone by (a) segment type and (b) position  
324 of the noun. Figure 3(a) shows that the low tone is realised at the left boundary (ISon)  
325 when the initial segment is a nasal consonant or an approximant and further to the right  
326 when it is a vowel. In cases with a glottal stop, which occurred in 36 items, the L tone was  
327 often displaced left from ISon. Figure 3(b) shows a rather consistent alignment of L to ISon.  
328 Note that in case the token word was in the first NP then the L tone was realised slightly  
329 further to the right from ISon.

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

337 We were further interested in determining which factor could best predict the alignment  
338 of the low tone and hypothesised the duration of the token initial segment could be a  
339 factor. To determine whether the difference in alignment of the low tone could be due to the  
340 duration of the segment to which it is anchored we first examined the duration of word initial  
341 segments. To test this, it was important to first evaluate whether there was a significant  
342 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	<i>p</i>
Approx.	Glott. stop	36.03 ± 7.22	23.4	<.001
Nasal	Glott. stop	26.29 ± 6.35	29.08	<.005
Vowel	Glott. stop	51.73 ± 7.43	246.02	<.0001
Nasal	Vowel	25.02 ± 6.54	21.7	<.01

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

352 Since the low tone cannot be anchored to a glottal stop, tokens that were produced  
 353 with a word initial glottal stop were excluded from the following analysis. To investi-  
 354 gate the alignment of L relative to ISON we employed a model that used RATE (normal,  
 355 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.

<b>L to ISON alignment</b>		
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)
Glott. stop	-17.39 (33.15)	-38.26(34.62)
<b>Duration</b>		
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)
Glott. stop	97.68 (29.73)	91.11 (21.86)

356 factors, together with participants and items as random intercepts (n=628). Results show  
 357 that the rate manipulation had no significant effect on where the low tone was placed  
 358 (Est.  $0.9 \pm 2.3$  ms,  $p = .7$  (ns.)) but that position (Est.  $11.51 \pm 2.28$  ms,  $p = .0001$ ) and

359 segment type (Est.  $13.30 \pm 5.76$  ms,  $p = .04$ ) had a significant effect. The Tukey corrected  
360 post hoc test revealed that there was a significant interaction between segment type and  
361 position, for nasals and vowels (Est.  $25.19 \pm 6.86$  ms,  $p = .02$ ). Hence, it is found that  
362 the length of the segment where the tone seeks to anchor has an influence. As Figure 3(a)  
363 shows, the low tone is realised closer to ISON when the initial segment is a nasal consonant  
364 and further away from ISON when it is a vowel. In case the token word was in the first NP  
365 and the initial segment was a vowel then the L tone was realised to the right from ISON, at  
366 around 52.81 ms. If it was realised on a vowel in the second NP, then the L tone was placed  
367 around 10.95 ms prior to ISON (see 3(b)). As table III shows, this difference was smaller  
368 between the two positions for other segments. In sum, there are small, yet consistent dif-  
369 ferences, as to where the initial low tone can be anchored depending on the position of the  
370 token in the utterance.

371       Additionally, we wanted to test if the location of the low tone was influenced by differences  
372 in duration of the initial segment which are related to modifications in speech rate. If we  
373 assume that duration of the word initial segment has an influence on where the low tone will  
374 be placed and a longer duration (at normal speech rate) of the segment provides more time  
375 for the low tone to be anchored further away from the left edge (ISON) and more towards  
376 the syllable nucleus, we should observe a positive correlation. If instead the correlation is  
377 negative, this indicates that the tone is seeking to align with the left edge of the syllable.  
378 To test our hypothesis two Pearson correlation tests were performed, one for tokens at fast,  
379 and another for tokens at normal speech rate. Results show that the location of the low tone  
380 does not correlate with the duration of the initial segment under the influence of speech rate

381 modifications. For the fast rate, there is a small positive correlation ( $r=0.02$ ,  $n=334$ ,  $p = 0.77$   
 382 (n.s)); for the normal speech rate there is also a small positive correlation ( $r= 0.02$ ,  $n=289$ ,  
 383  $p=0.74$  (n.s)). The two tests failed in showing a significant effect of correlation between the  
 384 duration of the initial segment under rate manipulation and the time of alignment of the low  
 385 tone. In view of this result, it seems fair to discard the hypothesis of speech rate influencing  
 386 the anchor point of the low tone.

387 **B. Alignment of H**

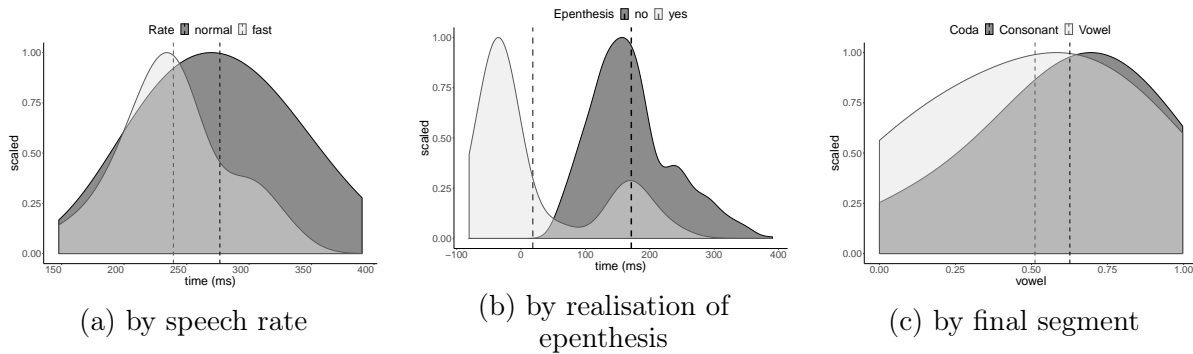


FIG. 4. Latency in ms to final H tone to onset of the final syllable (FSon) as smoothed density plot with mean values according to: (a) speech rate in monosyllabic tokens; (b) tokens with or without epenthesis. A value lower than zero indicates that the peak was placed prior to the onset of the final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

388 Figures in 4 show estimated density plots with mean values for differences in the  
 389 alignment on the peak related to (a) a modification of speech rate; (b) final epenthe-  
 390 sis realisation. To investigate the alignment of the final peak relative to the onset of

391 the final syllable (F<sub>Son</sub>) a model was used which included RATE (fast, normal), FINAL  
392 SEGMENT (consonant or vowel), EPENTHESIS (yes, no), and WORD LENGTH (short, long) as  
393 fixed factors together with participant, token and pause as random intercepts (n=624).  
394 Results show that the rate manipulation always had a significant effect on the location  
395 of the H tonal target (Est.  $12.6 \pm 2.92$  ms,  $p < .0001$ ). Also the factors final segment  
396 (Est.  $37.15 \pm 17.14$  ms,  $p < .04$ ), epenthesis (Est.  $63.67 \pm 21.10$  ms,  $p < .005$ ), and word  
397 length (Est.  $63.59 \pm 15.05$  ms,  $p < .0001$ ) showed a significant effect.

398 We had observed that some tokens tended to be realised with final glottalization of  
399 segments or were followed by a full glottal stop. This was marked in our corpus so that we  
400 could investigate if glottalization had an effect on where the peak was placed. Additionally,  
401 we were interested in determining the position of the peak within the syllable and found that  
402 in 91% of the cases the H peak was realised on a vowel. Further, 35% of the target tokens  
403 were produced ending on a coda consonant, however in 81% of these cases the tonal target  
404 was also found on a vowel. In fact, words ending on an open syllable or on coda consonant  
405 show similar placement of the peak. For syllables ending on a vowel the mean peak location  
406 is proportionally at two thirds (0.64) of the syllable and for syllables containing a coda the  
407 mean peak location is also at around two thirds (0.64) of the final syllable. Figure 4(c) shows  
408 a density plot with the estimated placement of the peak relative to the vowel, depending  
409 on whether there is a coda consonant or not on the last syllable. A linear mixed effects  
410 model investigated the position of the peak proportionally to the vowel carrying it (n=568).  
411 The model included RATE (fast, normal), FINAL SEGMENT (consonant, vowel), GLOTTALIZATION  
412 (yes, no), and VOWEL LENGTH (long, short) as fixed factors together with participant, token,

413 as well pause as random intercepts. Results show that the realisation of a final coda had  
414 an influence on where the peak was situated on the vowel (Est.  $0.112 \pm 0.0414$ ,  $p < .01$ ).  
415 The peak was realised earlier in the vowel, when this was the last segment, and later in the  
416 vowel, when there was a coda consonant. No other factors, such as the vowel glottalization  
417 ( $p=0.14$ ) or vowel length ( $p=0.58$ ) showed a significant influence on where the peak was  
418 placed in the vowel.

### 419 C. Rise time

420 Rise time was calculated as the difference in ms between the H tone and the preceding  
421 L tone. To investigate the influence of rate manipulation on rise time a linear mixed effects  
422 model was employed. The model included RATE (fast, normal), WORD LENGTH (short, long),  
423 SEX (male, female), and POSITION NP (first, second) as fixed factors as well as participant,  
424 token, and pause as random factors. Table IV summarises the results showing that the  
425 rise time between the low tone and the high tone was affected by all four factors, but most  
426 importantly, by the modification of speech rate. As expected, in the first NP, the rise time  
427 was longer at normal speech rate (mean 419.42 ms), than at fast speech rate (mean 396.99  
428 ms). This difference was greater for the second NP for which the mean of the rise time was  
429 442.09 ms at normal speech rate and 395.47 ms at fast speech rate. Overall the rise time  
430 of female participants was greater (mean 470.08 ms, at normal speech rate; 427.27 ms, at  
431 fast speech rate) than that of the male participants (mean 397.43 ms at normal speech rate;  
432 369.68 ms at fast speech rate). Not surprisingly, longer words displayed greater rise time  
433 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	<i>p</i>
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

434 **D. Excursion**

435 The size of excursion was calculated as the difference in Hz between the H tone and the  
436 preceding L tone. A linear mixed effects model investigated the effect of the rate manipula-  
437 tion on the size of excursion. As the automatic backward model selection of fixed parts of  
438 the model was performed, all factors with the exception of pause PAUSE (yes, no) were elim-  
439 inated. Results show that there is a significant effect of the pause on the size of excursion  
440 (Est.  $52.5 \pm 11.8$  Hz,  $p < .001$ ). For female speakers, tokens at normal speech rate had an  
441 excursion size (mean 46.16 Hz) which increased in case there was a following pause (mean  
442 57.61 Hz). This effect was stronger at fast speech rate, where the excursion size was again  
443 greater if a pause followed (mean 60.25 Hz) then when it did not (mean 40.15 Hz). These  
444 results indicate that the insertion of a pause after the target token had an influence on the  
445 excursion size and that the rate manipulation marginally affected excursion size.

## 446 E. Slope

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

## 459 V. DISCUSSION

460 In this study we sought to examine tonal targets in Drehu and investigate how they align  
461 to segmental landmarks. In order to evaluate whether tonal targets are produced to delimit  
462 prosodic constituency, token words with varying syllable number, and phonotactic structure  
463 (at the end of the word) were included. The results of this study show that in Drehu there  
464 is a consistent demarcation of the prosodic word, with a low tonal target at the beginning,  
465 and a high tone at the end of target tokens. A variation in syllabic structure of disyllabic

466 words showed that long vowels do not attract additional pitch movements, demonstrating  
467 that in Drehu there is no weight sensitivity attracting an *F0* peak. The investigation of the  
468 alignment of the low tone revealed that this target is aligned at the left edge of content words  
469 in all cases, which means the realisation of a LH or LHL<sub>1</sub>H<sub>1</sub> tonal pattern did not have an  
470 influence on the position of the tone. There were cases in which tokens beginning on a vowel  
471 were realised with a glottal stop as initial segment. When this happened, the tone could not  
472 be anchored to this first segment and the *F0* minimum was pushed into the monosyllabic  
473 function word that preceded it. Further, it was observed that the factor segment type had  
474 an influence on where the low tone was realised. Token-initial nasals differed significantly  
475 in duration when compared to vowels and this had an effect on the position of the low tone.  
476 The L tone was realised at the onset of the initial syllable when the first segment was a nasal  
477 and further away from the the same point when it was a vowel. However, these differences  
478 were rather small and the alignment of the L tone was consistently towards the left edge of  
479 the token word. The stability in the alignment of the L tone supports the hypothesis of the  
480 L being a phrase initial tone which demarcates a constituent like the prosodic word.

481 Our investigation of the final peak suggests that this tone is prominently marking the  
482 last full syllable. The alignment of the high tone confirmed our hypothesis that a long  
483 vowel would be the anchor if in word final position or a monosyllabic word. In contrast,  
484 our hypothesis about words ending on CVC-nasal syllables could not be confirmed since  
485 in the majority of instances (85%) the peak was reached in the vowel of the same syllable.  
486 However, it was found that the peak aligns at a proportionally similar distance from the  
487 onset of the final syllable which suggests that the peak is aligned in relation to the syllable

488 boundary, similarly to what has been found in Mandarin (Xu, 1998). This result strengthens  
489 a hypothesis of the alignment being anchored relative to the onset of the final syllable. An  
490 investigation of vowels carrying  $F0$  peaks showed a significant effect originating from the  
491 coda consonant, where the peak was realised later in the vowel, than in syllables with no  
492 coda showing that alignment occurs earlier in syllables with CVC but later in CV syllables.  
493 Thus, a certain degree of variability is found in the alignment of H which is why we cannot  
494 conclude that this tone has one particular segment to which it will always be anchored.  
495 These results show that the *segmental anchorage hypothesis* is better suited to account for  
496 the observed alignment of the final peak in Drehu. In this approach the anchorage point  
497 refers to a region, like the syllable, to which a tone can be anchored, as appears to be the  
498 case in these Drehu data. Further examination of the segmental anchoring hypothesis shows  
499 that the rate manipulation had a significant effect on rise time but a less strong impact  
500 on excursion size. Recall, a strict segmental anchoring hypothesis predicts anchors to be  
501 closer together, and the rises to be shorter and steeper as the speech rate increases. An  
502 investigation of the rise time, which was calculated as the difference in ms between the high  
503 tone and the low tone, showed a consistent effect of rate. Moreover, we could establish a  
504 significant effect of word length, sex, and position. This indicates increased speaking rate  
505 resulted in tonal targets being realised closer together, as the segmental anchoring hypothesis  
506 would predict. We did not find the same effect of rate on the excursion size of  $F0$ , which  
507 was calculated as the difference in Hz between the high and the low tones. Instead, we  
508 found a significant effect conditioned by the insertion of a pause after the target token.  
509 A follow up analysis that compared the slope of the same tokens by participant and rate

510 suggests the slope is mostly constant although some variability is found. Arguably, the  
511 insertion of pause could have an effect and confound the measurement of slope. Therefore,  
512 separate statistical tests were conducted for tokens that were produced with or without a  
513 following pause. The result for the slope of tokens realised with a subsequent pause suggest a  
514 symmetric distribution but this was not confirmed for tokens without following pause. Since  
515 the test for tokens with a subsequent pause included considerably more items (122) than  
516 the other (59) the result could be due to a lack of statistical power but a clear conclusion  
517 cannot be determined at this point. It was observed that the pause mostly occurred after  
518 the second NP which could be due to prosodic phrasing since two NPs represent a larger  
519 constituent. With regard to the significant effect of pause on excursion size, it is proposed  
520 that the pause participates in marking a higher level in the prosodic hierarchy. As known  
521 from research on prosodic constituency of other languages, the insertion of a pause, as well  
522 as significant manipulations of  $F_0$  prior to the pause, are considered strong indicators of a  
523 higher level prosodic constituent. We therefore propose that the Intonation Phrase in Drehu  
524 does not only end on a phrase final low tone but can also be realised with a high boundary  
525 tone.

## 526 VI. CONCLUSION

527 The present study finds evidence for a stable initial low tone and a more variable final  
528 high tone demarcating the prosodic word in Drehu. We propose that tonal alignment is  
529 better described with the segmental anchorage hypothesis, since we find that the word final  
530 prominence seeks to anchor with respect to the left boundary of the syllable, but that there is

531 variability regarding the segment to which the peak will be anchored. Moreover, it was found  
532 that *F0* minima seek to align towards the left edge of nouns and that neither a variation  
533 of tonal pattern, nor a rate manipulation had an effect on this. The small variability in  
534 the alignment of this tone could be understood as an indicator of an initial phrase tone. The  
535 final high tone showed variability under the pressure of rate manipulation. The shortening  
536 of rise time in fast speech rate indicates segmental anchors are closer together. However,  
537 the excursion size of the rise was not affected in the same way, meaning that the rises were  
538 not steeper. Instead, it seems that speakers seek to reach the same *F0* level despite time  
539 pressure. Interestingly, it was found that the insertion of a pause had an effect on the  
540 excursion size which could be related to the demarcation of a higher prosodic level, namely  
541 the Intonation Phrase. In view of these results, it is proposed that Drehu shows a phrasal  
542 prominence marking system with an initial L tone and a prominence lending H tone at the  
543 right edge of the prosodic word. Contrary to earlier impressionistic accounts ([Lenormand,](#)  
544 [1954](#); [Tryon, 1968a](#)), this study on tonal alignment supports the claim that a phrasal rather  
545 than strictly lexical prominence marking is found in Drehu ([Torres et al., 2018](#)). Arguably,  
546 tonal patterns outlined in this experiment and alignment of tones to segmental landmarks  
547 resemble those of the contact language French although this requires further investigation.  
548 This study has investigated short noun phrases containing a function word and a noun. It  
549 would be of interest to test larger prosodic constituents and examine if this would lead to a  
550 grouping of a higher level prosodic constituent. Further, to better identify how the peak is  
551 anchored in different segmental contexts, more data including a larger set of words ending  
552 on CVC, and a larger set of consonants at the onset of the syllable, would be of interest in

553 a future study. Based on the data we cannot conclude that the attested use of *F0* is the  
554 result of language contact. Further studies examining Drehu prominence marking and use  
555 of information structure are therefore planned.

## 556 APPENDIX: SPEECH MATERIALS

557 (1a) Kola drengē la emexemine ngöne la hlapa / DUR hear ART echo in ART field / “The echo  
558 can be heard in the field”. (1b) Kola drengē ngöne la hlapa la emexemine me kola ulil  
559 / DUR hear in ART field ART echo and DUR noise / “In the field the echo and the noise can  
560 be heard”.

561 (2a) Ame la maja, tre, ka a nyipi ewekē troa e i / PRS.1 ART bait PRS.2 STAT PRS important to  
562 catch fish / “Then the bait is important to catch fish”. (2b) Ame la thiitre me la maja,  
563 tre, ka a nyipi ewekē troa e i / PRS.1 ART trap and ART bait PRS.2 STAT PRS important to  
564 catch fish / “Then the trap and the bait are important to catch fish”.

565 (3a) Ame la mamala, tre, hna xupe ngöne la uma / PRS.1 ART window PRS.2 PST build in ART  
566 house / “Then the window is part of the house”. (3b) Ame la qēnelö me la mamala, tre,  
567 hna xupe ngöne la uma / PRS.1 ART door and ART window PRS.2 PST build in ART house  
568 / “Then the door and the window are part of the house”.

569 (4a) Ame la maamu, tre, hna sile hnei itre qatr / PRS.1 ART bogeyman PRS.2 PST invent A PL  
570 old / “The bogeyman was invented by the old”. (4b) Ame la satana me la maamu, tre,  
571 hna sile hnei itre qatr / PRS.1 ART devil and ART bogeyman PRS.2 PST invent A PL old /  
572 “The devil and the bogeyman were invented by the old”.

573 (5a) Thaa tro kö a nue la wamine falawa hune la laulau/ NEG FUT emphasis PRS leave ART  
574 crumb bread on ART table/ “No (bread) crumb should be left on the table”. (5b) Thaa tro  
575 kö a nue la laese me la wamine falawa hune la laulau/ NEG FUT emphasis PRS leave ART  
576 rice and crumb bread on ART table/ “No rice and (bread) crumb should be left on the table”.

577 (6a) Ame la imaag, tre, e kuhu ngöne gadran/ PRS.1 ART mango tree PRS.2 down there in garden  
578 / “Then the mango tree is down there in the garden”. (6b) Ame la ipiic me la imaag, tre,  
579 e kuhu ngöne gadran/ PRS.1 ART orange tree ART mango tree PRS.2 down there in garden /  
580 “Then the orange tree and the mango tree are down there in the garden”.

581 (7a) Ame la maage, ka lapa hunei sinöe/ PRS.1 ART mango PRS.2 STAT be on tree / “Then the  
582 mango is on the tree”. (7b) Ame la waco me maage, ka lapa hunei sinöe/ PRS.1 ART bird  
583 and ART mango PRS.2 STAT be on tree / “Then the bird and the mango are on the tree”.

584 (8a) Ame la mani, tre, ka catrehnine la ijine hnötr/ PRS.1 ART rain PRS.2 STAT strong ART  
585 period cold / “Then the rain is very strong this winter”. (8b) Ame la enyi me la mani, tre,  
586 ka catrehnine la ijine hnötr/ PRS.1 ART wind and ART rain PRS.2 STAT strong ART period  
587 cold / “Then the wind and the rain are very strong this winter”.

588 (9a) Ame la mano, tre, ketre götrane la ngönetrei/ PRS.1 ART chest PRS.2 INDF part ART body  
589 / “Then the chest is a part of the body”. (9b) Ame la qëmeke me mano, tre, ketre götrane  
590 la ngönetrei/ PRS.1 ART face and chest PRS.2 INDF part ART body / “Then the chest is a  
591 part of the body”.

592 (10a) Ame la maano, tre, hna hamëne kowe la joxu/ PRS.1 ART fabric PRS.2 PST give to ART  
593 chief / “Then the fabric was given to the chief”. (10b) Ame la itre koko me la maano, tre,

594 hna hamëne kowe la joxu/ PRS.1 ART PL yam and ART fabric PRS.2 PST give to ART chief /  
595 “Then the yams and the fabric were given to the chief”.

596 (11a) Ame la melem, tre, ka mingöming/ PRS.1 ART full moon night PRS.2 STAT beautiful / “Then  
597 the full moon night is beautiful”. (11b) Thaupune la jidri ne melem, tre, ka mingöming/  
598 calm ART night of full moon night PRS.2 STAT beautiful / “The calm of the full moon night is  
599 beautiful”.

600 (12a) Ame la melimala, ka za/ PRS.1 ART green turtledove STAT beautiful / “Then the green  
601 turtledove is beautiful”. (12b) Ame la fenifene me la melimala, ka za/ PRS.1 ART butterfly  
602 and ART green turtledove STAT beautiful / “Then the butterfly and the green turtledove are  
603 beautiful”.

604 (13a) Ame la munun, tre, ka mele ngöne la hnagejë/ PRS.1 ART fish picot PRS.2 STAT be in ART  
605 sea/ “Then the picot fish is in the sea”. (13b) Ame la eötre me la munun, tre, ka mele  
606 ngöne la hnagejë/ PRS.1 ART fish picot PRS.2 STAT be in ART sea/ “Then the shark and the  
607 picot fish are in the sea”.

608 (14a) Ame la wahnawa, tre, ka hnyapa la itre wen/ PRS.1 ART banana PRS.2 STAT sweet ART  
609 PL fruit/ “Then the banana is a sweet fruit”. (14b) Ame la guafa me la wahnawa, tre, ka  
610 hnyapa la itre wen/ PRS.1 ART guava ART banana PRS.2 STAT sweet ART PL fruit/ “Then  
611 then guava and the banana are sweet fruits”.

612 (15a) Ame la wanangongo, tre, ie ka lapa ngöne la huuca/ PRS.1 ART fish dawa PRS.2 STAT live  
613 in ART reef/ “Then the dawa fish lives in the reef”. (15b) Ame la nunuce me wanangongo,

614 tre, ie ka lapa ngöne la huuca/ PRS.1 ART pufferfish ART fish dawa PRS.2 STAT live in ART  
615 reef/ “Then the pufferfish and the dawa fish live in the reef”.

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

619 <sup>1</sup>Although the authors do not specify which dialect of English is spoken by the participants in their study,  
620 it seems fair to assume it is one American variety.

621 <sup>2</sup>Indigenous people from New Caledonia self identify as Kanak (Vernaudon, 2015).

622

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

625 Atterer, M., and Ladd, D. R. (2004). “On the phonetics and phonology of ‘segmental an-  
626 choring’ of F0: evidence from German,” *Journal of Phonetics* **32**(2), 177–197.

627 Bates, D., Mächler, M., Bolker, B., and Walker, S. (2015). “Fitting linear mixed-effects  
628 models using lme4,” *Journal of Statistical Software* **67**(1), 1–48, doi: [10.18637/jss.v067.](https://doi.org/10.18637/jss.v067.i01)  
629 [i01](https://doi.org/10.18637/jss.v067.i01).

630 Beckman, M. E., and Hirschberg, J. (1994). “The ToBI annotation conventions,” Ohio State  
631 University .

632 Boersma, P., and Weenink, D. (2017). “Praat: doing phonetics by computer (version  
633 6.0.26)[computer program]. Retrieved November 2, 2017” .

634 Bruce, G. (1977). “Swedish accents in sentence perspective,” Working Papers in Linguistics  
635 12(1).

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

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

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

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

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

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

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

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

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

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

657 Magulué, P. F. (2011). “Les teachers du pacifique au xixe siècle ou l’émergence d’une nou-  
658 velle élite océanienne entre tradition et modernité,” *Histoire et missions chrétiennes* (4),  
659 139–156.

660 Monnin, J. (2010). “Influence de la langue ambiante sur l’acquisition phonologique: une  
661 comparaison du français et du drehu,” Ph.D. thesis, Université de Grenoble; Université de  
662 Nouvelle Calédonie.

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

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

667 Pierrehumbert, J. B. (1980). “The phonology and phonetics of english intonation,” Ph.D.  
668 thesis, Massachusetts Institute of Technology.

669 Pitrelli, J. F., Beckman, M. E., and Hirschberg, J. (1994). “Evaluation of prosodic tran-  
670 scription labeling reliability in the tobi framework,” in *Third International Conference on*  
671 *Spoken Language Processing*.

672 Prieto, P., and Torreira, F. (2007). “The segmental anchoring hypothesis revisited: Syllable  
673 structure and speech rate effects on peak timing in spanish,” *Journal of Phonetics* **35**(4),  
674 473–500.

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

677 R Core Team (2017). *R: A Language and Environment for Statistical Computing*, R Foun-  
678 dation for Statistical Computing, Vienna, Austria, <https://www.R-project.org/>.

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

681 Torres, C., Fletcher, J., and Wigglesworth, G. (2018). “Investigating word prominence in  
682 drehu,” in *Proc. 17th Speech Science and Technology conference*.

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

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

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

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

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

691 Vernaudon, J. (2015). “Linguistic Ideologies: Teaching Oceanic Languages in French Poly-  
692 nesia and New Caledonia,” *The Contemporary Pacific* **27**(2), 433–462.

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

695 Welby, P., and Løevenbruck, H. (2006). “Anchored down in anchorage: Syllable structure  
696 and segmental anchoring in French,” *Italian Journal of Linguistics/Rivista di linguistica*  
697 **18**, 74–124.

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

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

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