Pitch range modulations in an edge-marking language

Catalina Torres\textsuperscript{1,2}

\textsuperscript{1}University of Zurich
\textsuperscript{2}ARC Centre of Excellence for the Dynamics of Language
catalina.torres@ivs.uzh.ch

Abstract

Prosodic phrasing is a topic that has received considerable attention over the last decades. However, most research has dealt with well studied (mostly European) languages, and quantitative production studies of under-resourced languages are under-represented. To better inform the field of intonational phonology, more data from a more diverse set of languages is needed.

This study investigates pitch range modulations in Drehu, an Oceanic language from New Caledonia. Recent experimental work suggests Drehu is edge-marking and the right-edge is prosodically salient. In this study, the phonological and phonetic realisation of prosodic boundary marking is investigated.

To determine whether pitch range modulations contribute to phrasing, the intonational marking of noun phrases of different sizes is analysed. An experiment was conducted to examine the extent to which fundamental frequency (F0) modulations contribute to the signalling of right-boundaries and if these are associated with the marking of different prosodic levels. The results show evidence for pitch range adjustments between a phrase initial low tone and a phrase final high tone depending on the position in the noun phrase. These modulations show a blocking of downstep and suggest pitch range adjustments could be indicative of an intermediate phrase (ip) level.

Index Terms: F0 range, Drehu, New Caledonia, prosodic phrasing, downstep, span narrowing

1. Introduction

It is commonly accepted that intonational cues serve to demarcate prosodic constituency. Part of the set of intonational cues that can serve to demarcate boundaries are clear boundary movements, other junctural cues like pre-boundary lengthening, but also pitch range adjustments can be a cue to prosodic constituency. According to Ladd [1, 2], speaking range is defined in terms of speaker specific parameters which include a baseline and a default initial register setting which can be thought of as a top and bottom line. Linguistically meaningful modulations of F0 are typically detected within a narrower portion of the range. The term tonal space refers to an interval where high (H) tones are at the top and low (L) tones towards the bottom level of a band of F0 values. Variations in the tonal space serve as phonological representation of prosodic constituency [3] and can interact with other linguistic domains, such as syntactic structure, in edge-marking languages like French or Japanese [4, 5, 6, 7].

The modulations of tonal space, especially of H tonal targets, are said to be influenced by downstep, a phenomenon initially described for sub-Saharan African tone languages, such as Efik (Benue-Congo, Nigeria) [8, 9]. A similar downtrend has also been documented for other languages, such as English, German, Japanese or French [10, 6, 3, 4]. For the purpose of this study, the focus will be on downstep in relation to ‘non-tonal’ languages, thus languages that do not base lexical or morphological distinctions purely on tone. Downstep is described as a process by which lowering and/or narrowing of the tonal space interval is phonologically triggered through language-specific conditions. As a consequence, a frequently observed trend is that subsequent H peaks are realised with a lower F0 value than phrase-initial ones. Note that downstep is different from declination, a gradual lowering of F0 values that is not phonologically constrained, and occurs from left to right, meaning, from the beginning of the utterance to its end [6]. Thus, the domain of declination is the utterance whereas downstep can be reset utterance internally and this manipulation can be associated with phrasal constituency.

Ladd [2] argues that intonational target levels are controlled by the speaker. Based on the position of intonational tone targets in a hierarchical prosodic structure, the phonetic realisation parameters can be specified locally for any accentual pitch target which ultimately is linked with phrasing. Figure 1 shows a schematic representation of two F0 contours, one displaying continuous downstep (solid line), while the other shows downstep blocking (dashed line). For example, it has been shown that in Hexagonal and Lifou French, an intermediate phrase (ip) boundary is triggered when there is a major syntactic break between a noun and a verb phrase, leading to a blocking of downstep [4, 11, 12]. Similarly, studies about Tokyo Japanese have shown that a syntactic (left-)boundary blocks downstep [13, 14]. Perception studies have shown that when F0 peaks are at the same height, English and Dutch listeners perceive peaks occurring later in a phrase as being more prominent [15, 16]. Another study [17] showed that in French a second, equally high F0 peak was more often associated with a higher prosodic level (ip), and that a second lower F0 peak, was more often associated with a lower prosodic level, the accentual phrase (AP).

The aim of this study is to examine the tonal space and evaluate whether modulations of F0, such as the blocking of downstep and span narrowing play a role in demarcating prosodic constituency in Drehu. Additionally, it is also of interest to evaluate whether F0 contributes to the demarcation of different potential prosodic levels.

Figure 1: Schematic representation of F0 contour with four peaks. The solid black line shows continuous downstep and span narrowing. The dashed line shows the blocking of downstep at the third peak. Top and bottom levels are marked with solid grey lines.
1.1. Drehu

Drehu is a Southern Oceanic language [18] spoken in New Caledonia, a collectivity of France 1. Recent phonetic investigations show evidence for post-lexical right edge-marking in the language [19, 20]. At the word prosody level several intonational patterns have been identified [20], however the LH rising pattern appears to be the most frequent. One study [19] shows that the word initial low tone and word final high tone most commonly demarcate the prosodic word. It also shows that the initial low tone is consistently aligned to the left edge of the word. More precisely, the low tone aligns between a function word and a following content word, thus suggesting it is a phrase initial tone. The word final H tone aligns to the last full syllable of the word, making it the most prominent syllable.

2. Research questions

This study examines whether noun phrases (NP) of different lengths will be associated with downstep and span narrowing blocking effects. It is hypothesised that, similar to French or Japanese, where syntactic breaks block recurring downstep, in Drehu, the prosodic realisation of a NP will affect the tonal space. If syntactic structure had no effect on subsequent peaks, these should display downstep and the span in the tonal space become narrower. The Drehu data selected for this analysis contains regular L and H tones demarcating prosodic words. In case of a continuous narrowing of the tonal space both tonal targets can be affected. Thus, span expansion between these two tones represents the relevant measure for this examination. Based on previous findings in intonational phonology, and the Drehu prosody, the following hypotheses are postulated:

H1 DOWNSTEP BLOCKING HYPOTHESIS:
The peak in the right most NP will be at the same height or higher than that of preceding peaks, in a preceding NP.

H2 SPAN NARROWING BLOCKING HYPOTHESIS:
The pitch span between a low and a high tone of the right-most NP will show the same or greater expansion than a preceding rising tone, in a preceding NP.

3. Methods

3.1. Participants

Thirteen speakers of Drehu from a local high school in Lifou participated in this study. The speakers recruited are pupils of the optional Drehu language or Kanak culture class and 14 to 19 years old. All participants provided basic demographic information (age, tribe) and responded to an adapted version of the BLP [21], suitable for younger speakers. Seven were male and all reported being raised in a bilingual or multilingual environment, typical of the region. The BLP indicates that seven were dominant in Drehu, two in French, and four were close to being balanced bilinguals.

3.2. Materials

The experiment used a set of 25 utterances realised with the topic marker aame at the onset of the sentence (13 speakers x 25 sentences x 2 repetitions = 650). Data used in this study are taken from the corpus used in [19]. There were two lengths depending on number of noun phrases: one including a Single NP (1NP x 13) and another larger NP with a First NP and Second NP (2NP x 12):

(1) Ame la mano, tre, ketre götrane la ngönetrei

(2) Ame la qêmeke me mano, tre, ketre götrane

3.3. Procedure

Figure 2 illustrates an utterance containing two NPs and the selected points for measurements of F0. Since one aim is to examine the scaling of tonal targets associated with high tones, every utterance initial peak was marked with hi (for initial high) and the subsequent peaks were marked with H. Additionally, the goal is to investigate whether a narrowing or expanding of the range could be observed and therefore, all utterance-initial low tones were marked with li and every subsequent low tone with L. Pitch span was calculated as the difference between a H tone and a corresponding L tone contained in the same NP. The same was done for the utterance initial low and high tones. F0 values in Hz were normalised to semitones, benchmark 100 [22]. Recordings were carried out at a self-selected normal and fast speech rate. For more details on experimental procedure, recordings, transcription, processing of data and creation of the data base see [19].

3.4. Statistical analysis

To investigate the effects related to F0, data were analysed using linear mixed effects models. Statistical analyses were carried out in R [23] with help of the statistics package lme4 and emmeans [24, 25]. Values were fitted into a theoretically motivated model to investigate specific factors of interest. In agreement with [26], no random slopes are added to the models as this affects statistical power of small data sets, such as the one used in this study. Estimated marginal means were obtained for factor comparisons and p-values were adjusted with the Tukey method.

4. Results

In the corpus, rising tones were the most frequent. In fact, in 96.5% of the cases, utterances started with a rising tonal movement directly on the topic marker, as shown in Figure 2. The
following LH rising tones were obtained for noun phrases: 270 in Single NP, 248 in First NP, 222 in Second NP, and 523 in the topic marker (TOP). The insertion of pauses was quantified, revealing that Single NP was in pre-pausal position in 59.6%, First NP in 30.8%, and Second NP in 72.9% of the cases. A pause was never inserted after the topic marker. Since the focus lies on pitch modulations that are not affected by the insertion of a pause, the following analysis excludes all pre-pausal items.

Figure 2: Oscillogram, spectrogram, and F0 trace of the utterance Ame la eötre me la munun, tre, ka mele ngöne la hnagejë. “The shark and the picot fish live in the sea”. This utterance contains two NPs and was produced by a female speaker.

### 4.1. Downstep blocking of H peaks

To examine whether there is downstep blocking of peaks in NPs measurements for H tones were investigated. The model included POSITION (TOP, Single NP, First NP, Second NP), SEX (male, female), and SPEECH RATE (fast, normal) as fixed factors together with participant and words as random factors. This model included 612 observations, from 110 utterances with 1NP, and 188 with 2NP.

<table>
<thead>
<tr>
<th>Length</th>
<th>Contrast</th>
<th>Est.</th>
<th>SE</th>
<th>t-Value</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>1NP</td>
<td>Single TOP</td>
<td>-0.5</td>
<td>0.2</td>
<td>-2.6</td>
<td>0.07 (n.s.)</td>
</tr>
<tr>
<td>2NP</td>
<td>TOP-First</td>
<td>-0.7</td>
<td>0.2</td>
<td>-3.7</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>2NP</td>
<td>TOP-Second</td>
<td>-0.4</td>
<td>0.2</td>
<td>-1.8</td>
<td>0.02 (n.s.)</td>
</tr>
<tr>
<td>2NP</td>
<td>First-Second</td>
<td>0.2</td>
<td>0.2</td>
<td>1.8</td>
<td>0.07 (n.s.)</td>
</tr>
<tr>
<td>1NP-2NP</td>
<td>Single-Second</td>
<td>0.1</td>
<td>0.2</td>
<td>0.3</td>
<td>0.1 (n.s.)</td>
</tr>
<tr>
<td>1NP-2NP</td>
<td>Single-First</td>
<td>-0.2</td>
<td>0.2</td>
<td>-1.1</td>
<td>0.09 (n.s.)</td>
</tr>
</tbody>
</table>

Table 1: Relevant results of Tukey corrected factor pairwise comparisons for pitch peaks according to position, for female speakers.

Figure 3 shows values for H peaks in NPs of two lengths. As predicted, for 1NP, values for the rightmost peaks of NPs are higher than those in the preceding peak in the topic marker. Additionally, for 2NP the peak in the First NP is higher than the peak of the preceding topic marker. However, the peak of the Second NP is not systematically higher. Results of the statistical analysis show that there is a significant difference for the factors sex (Est. -5.9, ± 1.8 semitones, $p < 0.01$) and speech rate (Est. -0.4, ± 0.2 semitones, $p = 0.02$). Results of the Tukey corrected pairwise comparisons for female speakers are summarised in Table 1. Statistical results indicate that peaks in the First NP are slightly higher (estimate 0.2 semitones) than peaks in the Second NP, in 2NP.

Figure 3: Notched box plots of H peaks in semitones for all speakers. The left panel summarises values in stimuli with one NP, the right panel summarises values in stimuli with two NPs. Measurements for males are in dark grey.

### 4.2. Span narrowing blocking

Prosodic phrasing could be realised in relation to the tonal range and not restricted to high tones. To examine whether there is a narrowing process that equally involves low and high tonal targets, it was important to consider the role of pitch span. Figure 4 shows the pitch span expansion in utterances in 1NP and 2NP. It shows that in 1NP the span is greater for Single NP than for TOP, and in 2NP the pitch span of Second NP is greater than First NP and TOP.

To investigate whether the position of the NP in the utterance had an effect on the magnitude of pitch span a statistical model was employed. The model included POSITION (TOP, Single NP, First NP, Second NP), SEX (male, female) and SPEECH RATE (normal, fast) as fixed factors and participant plus word as random factors. This model included 612 observations. No statistically significant difference could be established for the factors sex and speech rate. Results of the Tukey
Table 2: Relevant results of Tukey corrected factor comparisons for pitch peaks according to position, at normal speech rate.

<table>
<thead>
<tr>
<th>Length</th>
<th>Contrast</th>
<th>Est.</th>
<th>SE</th>
<th>t-value</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>1NP</td>
<td>Single</td>
<td>-1.3</td>
<td>0.3</td>
<td>-4.4</td>
<td>=.001</td>
</tr>
<tr>
<td>1NP</td>
<td>First</td>
<td>-0.6</td>
<td>0.3</td>
<td>-2.3</td>
<td>=.2 (n.s)</td>
</tr>
<tr>
<td>2NP</td>
<td>Single</td>
<td>-1.3</td>
<td>0.3</td>
<td>-4.3</td>
<td>=.001</td>
</tr>
<tr>
<td>2NP</td>
<td>First</td>
<td>-0.7</td>
<td>0.2</td>
<td>-3</td>
<td>=.02</td>
</tr>
<tr>
<td>2NP</td>
<td>Second</td>
<td>-0.7</td>
<td>0.2</td>
<td>-3</td>
<td>=.02</td>
</tr>
<tr>
<td>1NP-2NP</td>
<td>Single</td>
<td>0.4</td>
<td>0.2</td>
<td>2.1</td>
<td>=.1 (n.s)</td>
</tr>
<tr>
<td>1NP-2NP</td>
<td>Single</td>
<td>-0.3</td>
<td>0.3</td>
<td>-1.1</td>
<td>=.8 (n.s)</td>
</tr>
</tbody>
</table>

corrected factor comparisons are summarised in Table 2. The results of the statistical analyses confirm there is a significant difference between TOP and Single NP in 1NP. Importantly, evidence is found that the pitch span of Second NP is significantly greater than that of the preceding TOP and First NP in 2NP. Hence, evidence is found that for both lengths the pitch span of the right most NP is significantly greater than the span in preceding rising tones in the speech.

5. Discussion

This study investigates whether, similar to French or Japanese, syntax relates to prosodic phrasing leading to a blocking of a downstep trend and span narrowing in Drehu. Modulations in the tonal space were examined in two different sets of sentences containing NPs of different sizes. The aim was to evaluate whether the height of pitch peaks and pitch span expansion relate to the grouping of phrases as prosodic constituents. A blocking effect was found at the right boundary of the NP. Additionally, it was found that the measure for span narrowing provides the more robust result. First, the aim was to determine whether effects of downstep blocking could be established on H tonal targets. Values for pitch peaks were investigated. A statistically significant difference was found between female and male speakers. Moreover, it was found that H peaks did not consistently increase or decrease across successive right-edge peaks suggesting that there was no consistent blocking effect of downstep. However, this result could be due to diverging trends in the scaling of peaks between female and male speakers. Thus although a trend confirming the DOWNSTEP BLOCKING HYPOTHESIS was found, it was not consistent across all speakers.

Second, the aim was to test whether there was span narrowing. In agreement with predictions, it could be established that the pitch excursion at the right most rising tone in 2NP was significantly greater than that of preceding rises (on the topic marker and the First NP). Similarly, the right-most peak of the NP was larger than that of a preceding rise in the topic marker in 1NP. Since the right-most rise shows no span narrowing relative to the preceding one, these results confirm the SPAN NARROWING BLOCKING HYPOTHESIS. This investigation showed that in 2NP the First NP has a narrower span than the Second NP, meaning that a narrowing of pitch expansion is stopped at the right edge of the Second NP. Since in noun phrases the span narrowing is blocked, this process shows similarities to other edge-marking languages like Japanese and French. An evaluation of tonal annotation revealed that in the NPs in the corpus, the utterance initial particle *mae* is mostly realised with a rising tone. It was determined that the particle is realised with a weak rise and is presumably not really accented. Arguably, the sentence initial rise serves a rhythmic function originating from the regular LH rising tones across an utterance in Drehu. This could be indicative of Drehu showing a strong macro-rhythm [27].

The results of this study suggest that examining modulations in the tonal space, which involve minima and maxima points, provides a more informative analysis that could yield more robust results. An examination of the entire tonal space allows for an investigation of tonal targets in relation to each other without isolating single tonal events. Arguably, examining exclusively H tones restricts the analysis to the top level of a band of F0 values, providing less information about the overall modulations. This indicates that measuring the distance between low and high tones could be particularly relevant for a language with regular LH tones such as Drehu. Additionally, it could be established that the expansion related to the right boundary of the Single and Second NP was similar. A prosodic hierarchical approach can explain these discrepancies which arise from a mediation between two levels, the accentual phrase and the intermediate phrase. The difference between 1NP and 2NP was the number of NPs that constituted them, however both were syntactically NPs. Taking into consideration that in Drehu there is a preference to mark the right edge, it shouldn’t be surprising that the Single and Second NP are actually delimiting the same prosodic level to the right. Hence, the expanded pitch excursion can be accounted for through the existence of an intermediate phrase level. Conversely, the pitch span in the First NP is narrower than in the Second NP because it is demarcating a lower prosodic level, embedded in the ip, namely the AP. Taken together the results of this study provide preliminary tonal evidence for an intermediate phrase H-tone at the right edge. However, further investigation of this tone and prosodic boundary marking in the Drehu is required. For example, it is of interest to examine whether the same evidence can be found in speech styles other than controlled reading, including studies using a wider variety of genres.

6. Conclusions

This study examined modulations of the tonal space in relation to the grouping of noun phrases in declarative sentences. It could be established that NPs are generally realised as one prosodic constituent. Two possible mechanisms involved in prosodic phrasing were investigated, the blocking of downstep of H peaks, and the blocking of span narrowing of LH rising tones. The analysis found evidence for a blocking of downstep and span narrowing, with the latter showing the more robust results. This supports the idea that to better understand modulations of the tonal space (in languages that display regular rising tones) it is helpful to consider F0 minima and maxima together. The tonal evidence found could be indicative of an ip level boundary in Drehu. However, since this study only examined a limited number of utterances more data including varying syntactic and information structure should be taken into consideration in future work.

7. Acknowledgements

Special thank to the participants who made this study possible. This research was conducted with support from the ARC Centre of Excellence for the Dynamics of Language (Project ID: CE140100041).
8. References