Tonal scaling and alignment in Javanese

Fabian Schubo\textsuperscript{1}

\textsuperscript{1}University of Stuttgart, Germany
fabian.schuboe@ifla.uni-stuttgart.de

Abstract

The present study provides basic insights into the tonal properties of Javanese, an Austronesian language that is mainly spoken in the central and eastern parts of Java, Indonesia. This language is particularly interesting with regard to prosody because it lacks lexical stress and lexical tone. A pilot production study was conducted with a single speaker, who produced a total of 96 sentences with SVO word order. The distribution of tonal events suggests that the sentences are subdivided into two prosodic phrases, one containing the subject and another one containing the remaining material. Both phrases involved a local pitch rise on the final syllable. In sentence-final position, the pitch peak was followed by a fall in case a sonorant coda consonant was present, but not if such a consonant was absent. Based on these findings, a preliminary account on the tonal patterns of Javanese is discussed, addressing several discrepancies found in the literature.

Index Terms: prosody, intonation, tonal alignment, tonal scaling, accentual phrase, Banyumasan, Basa Ngapak, Javanese

1. Introduction

This paper presents results from a pilot study on the sentence prosody of Javanese, an Austronesian language that is mainly spoken in the central and eastern parts of Java, Indonesia. Javanese is one of the largest languages in terms of native speakers (68.3 million according \cite{1}), but has not been studied in detail with regard to its prosodic system. It has been found that Javanese lacks lexical stress and lexical tone (e.g. \cite{2,3,4}) and thus organizes its tonal structure entirely with reference to higher prosodic domains (e.g. \cite{5}). Languages of this type have been termed ‘phrase languages’ in a recent system of prosodic typology \cite{6,7}. Many Austronesian languages seem to organize their tonal structure on the phrase level, but only few of these languages have been studied so far (see e.g. \cite{3,8,9}). The present study investigates Banyumasan, also referred to as Basa Ngapak, a Western Javanese variety that is spoken in the western part of Central Java and a few adjacent areas.

An early description of Javanese prosody is provided in Uhlenbeck (1941) \cite{10}. Based on auditory perception, he analyzed different types of sentences and found that they are subdivided into shorter prosodic phrases involving a pitch rise towards their end whose peak is aligned with the boundary. The height of the peak is lowered from one phrase to the next so that the rise is smallest in the sentence-final phrase. The phrases in Uhlenbeck’s analysis often matched syntactic constituents comprising collocations of several words.

The first qualitative description based on acoustic data is given by Stoel (2006) \cite{5}, who analyzed a total of nine utterances from a single speaker of Banyumasan. The productions included declarative and interrogative sentences that varied in length and structure and in most cases contained a narrow focus. Stoel \cite{5} assumed that sentences are subdivided into accentual phrases (APs), which mostly correspond to syntactic phrases. Based on a qualitative analysis of pitch tracks, he observed that, in non-final APs, the pitch contour continuously rises from left to right and reaches a peak at the right edge. This contour is henceforth referred to as ‘gradual-rise pattern’. Stoel \cite{5} assumed this contour as the default pattern for APs in non-final position, which he modeled by means of an L boundary tone at the left and an H boundary tone at the right edge of the AP. The final AP of the focused part in a declarative sentence showed a falling pattern, which was modeled as an HL boundary tone at the right edge AP. The peak of the H target of this tone occurred earlier than the peak in the gradual-rise pattern and was mostly aligned with the penultimate syllable (only one instance showed alignment with the final syllable). This pattern is referred to as ‘final-fall pattern’ in the following. The pitch tracks presented in Stoel \cite{5} show a steep rise immediately before the fall so that the AP pattern may as well be interpreted as comprising a rising-falling contour. The rise is clearly different from the one in the gradual-rise pattern, as its level is elevated mostly in the second half of the AP. In interrogatives and non-final declaratives, Stoel found a rising contour at the right edge of the last AP (without a following fall), which he modeled as an LH boundary tone. This tone appears to indicate question intonation or a continuation rise. Note that the patterns identified in Stoel’s analysis are different from those described by Uhlenbeck \cite{10}.

Rahyono (2007) \cite{11} reports on a production study on the variety of Javanese spoken in the palace of Yogyakarta, a Central Javanese variety of the formal register (krama). The study involved the elicitation of a declarative, an interrogative, and an imperative sentence with SVO word order. Four subjects produced the three sentences four times. The results showed that all productions involved a local rise at the end of the subject and a more complex movement at the end of the object, suggesting a sequence of two APs. The rise at the end of the first AP differs from the gradual-rise pattern in that it starts relatively late in the AP and shows a rapid increase in pitch whose peak is reached near the right AP edge. This contour is referred to as ‘late-rise pattern’ in the following. Note that this pattern matches Uhlenbeck’s \cite{10} description, but was not found by Stoel \cite{5} in a non-final AP. As for the sentence-final contour, Rahyono \cite{11} found that the final pitch contour of statements involved a rising-falling pattern on the final syllable (which was a closed syllable with a sonorant coda consonant). This pattern appears to be the same as the final-fall pattern in Stoel’s \cite{5} data (but with peak alignment on the final syllable). The imperative also showed a rising-falling pattern, but the rise occurred on the penultimate syllable and the fall followed on the final syllable. The interrogative involved a rising-falling-rising pattern that started already on the antepenultimate syllable.

In sum, prior studies on Javanese sentence prosody employed auditory perception \cite{10} or qualitative inspections of pitch contours based on very limited data sets \cite{5,11}. The observations in \cite{5,10,11} regarding the tonal patterns of APs and
their alignment properties in declarative sentences involve the following discrepancies:

- The default contour for non-final APs is assumed to be the gradual-rise pattern in [5], but the late-rise pattern in [10,11].
- The default contour for final APs is the final-fall pattern in [5,11], but the late-rise pattern in [10].
- The peak of the final-fall pattern is aligned (in most cases) with the penultimate syllable in [5], but with the final syllable in [11].

2. Research questions and hypotheses

The research questions addressed in this study are as follows:

1. What is the default tonal pattern for (a) non-final and (b) final APs in Javanese declarative sentences?

With regard to non-final APs, it is tested if the rise leading to the peak at the right AP edge begins at the left AP edge (gradual-rise pattern) or at a later point (late-rise pattern). With regard to final APs, it is tested if the peak at the right edge is followed by a fall (final-fall pattern) or if it constitutes the sentence-final target (late-rise pattern).

2. To which syllables or segments are the tonal targets aligned in (a) non-final and (b) final APs?

This question is addressed by identifying the locations of the given high (H) and low (L) turning points of the pitch contour. If the gradual-rise pattern is attested in non-final APs, it is expected that the L is aligned with the AP-initial and the H with the AP-final syllable. In case of the late-rise pattern, it is expected that the L is aligned with the beginning and the H with the end of the AP-final syllable (as in [11]). As for final APs, the following hypotheses are motivated based on the literature: In case the final-fall pattern is attested, the H is aligned either to the penultimate syllable (as in [5]) or to the final syllable (as in [11]) and the following L is aligned to the end of the syllable. In case the late-rise pattern is attested in this position, the H is predicted to be aligned with the end of the final syllable.

3. Methods

The production study involved the elicitation of read speech in a laboratory setting. The materials consisted of 24 declarative sentences with SVO word order. The verb was preceded by a particle expressing progressive aspect (agi). Examples are given in (3). Twelve of the sentences had a trisyllabic subject, as in (3a), and twelve of the sentences had a disyllabic subject, as in (3b). The subject and object always comprised a single noun. The subject was always a proper name. In some cases, as in (3a), and twelve of the sentences had a disyllabic subject, as in (3b), which contains a front vowel (not a schwa). Voiceless obstruents were avoided in order to minimize consonantal distortions in the pitch contour. The verb contained two to three syllables. Each verb was used in two different sentences, one with a trisyllabic subject and one with a disyllabic subject and a trisyllabic object.

4. Results

Figure 1 presents the results for F0 scaling of the tonal targets identified in the productions. T1-T3 (yellow) are the targets identified on the subject. The relative scaling pattern suggests that the contour remains flat in the first part and rises in the second part of the subject. T4 (grey) constitutes the first target after the subject, which was located on the verb. This target is considerably lower than the preceding ones. T5-T7 (green) are the targets identified on the object. The sequence of targets suggests that the contour does not move significantly from the verb (T4) to the first part of the object (T5). After that, a rise follows, reaching its peak at T6. The final target (T7) was only present in 43 out of the 96 productions. In the remaining 53 productions, T6 constituted the final target (see below for details).

The overall scaling pattern is compatible with the analysis given in Figure 2, which assumes two APs, one containing the subject (T1-T3 in Figure 1) and another one containing the remaining part of the sentences (T4-T6/T7 in Figure 1). The first AP can be assumed to involve an LLH sequence of tonal targets whereas the second AP can be assumed to involve either an LLH or an LLLH sequence. The pitch range of the second AP appears to be lowered in relation to the one of the first AP. The final L (T7 in Figure 1) is considerably lower than all preceding targets, possibly reaching the bottom of the speaker’s register. In the following, we will focus on the L and H targets involved in the rising and rising-falling parts of the contour.
Figure 1: $F_0$ scaling (Hz) of the turning points; the yellow points are on the subject, the grey point is on the verb, the green points are on the object (T1-T6: n=96; T7: n=43)

Figure 2: Stylized $F_0$ contour with annotation of AP structure (the dashed line indicates the variable presence of the final fall)

Figure 3: Duration (ms) from the beginning of the subject to the L preceding the rise (n=96)

Table 1: Distribution of final LHL and final LH contours by structure of the final syllable (coda obstruents are voiceless)

<table>
<thead>
<tr>
<th></th>
<th>LHL</th>
<th>LH</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>CVC sonorant</td>
<td>43</td>
<td>5</td>
<td>48</td>
</tr>
<tr>
<td>CVC obstruent</td>
<td>0</td>
<td>16</td>
<td>16</td>
</tr>
<tr>
<td>CV</td>
<td>0</td>
<td>32</td>
<td>32</td>
</tr>
</tbody>
</table>

Figure 4 presents the alignment results for the LH contour at the end of the object (part of the data in T5 and T6 in Figure 1). The bottom plot shows the alignment of the L with reference to the beginning of the final syllable and the top plot shows the alignment of the H with reference to the end of the vowel (zero axes). The drops suggest that the L is aligned near the beginning of the syllable and the H is aligned at the end of the vowel.

Figure 5 presents the alignment results for the LH contour at the end of the object (n=53): bottom: L with reference to the beginning of the syllable; top: H with reference to the end of the vowel; dark grey areas indicate mean duration of CV syllables
Figure 6 presents the alignment results for the LHL contour at the end of the object (part of the data in T5 and T6 and all of the data in T7 in Figure 1), which occurred only in CVC\textsubscript{consonant} syllables. The top plot shows the alignment of the H with reference to the end of the vowel, the middle plot shows the alignment of the preceding L, and the bottom plot shows the alignment of the final L, each with reference to the respective edge of the syllable (zero axes). The plots suggest that the first L is aligned near the beginning of the syllable, the H is aligned at the end of the vowel, and the final L is aligned at the end of the syllable.

5. Discussion

The results showed that Javanese SVO sentences of the type at hand involve a rising contour on the subject-final and a rising or rising-falling contour on the object-final syllable. In line with prior observations [5,10,11], the location of these tonal events suggest that the sentences were subdivided into two APs, one containing the subject and one containing the rest of the sentence. This assumption is also supported by the apparent pitch range lowering effect on the second AP.

For the non-final AP, an LLH sequence of tonal targets was identified. The beginning of the rise (i.e., the second L in the sequence) occurred later in trisyllabic than in disyllabic subjects. This suggests that the late-rise pattern constitutes the default tonal contour of non-final APs, which is in line with the observations in [10,11], but not with the one in [5]. The rise occurred on the AP-final syllable, starting at the beginning of the left syllable edge and reaching its peak at the end of the vowel. This is in line with the contour shown in [11].

The final AP involved an LLH or an LHL sequence of tonal targets. Notwithstanding the final fall, both of these sequences behaved in a similar way: The initial L occurred on the verb and the rise reached from near the beginning of the AP-final syllable (second L; pitch elbow) to the end of the vowel (H). The LLH sequence suggests the presence of a late-rise pattern also on the final AP, as described in [10].

The variability with regard to the presence of the final fall was unexpected and not controlled for. A closer examination of the data suggested that a final fall is realized only in case the final syllable contains a sonorant coda consonant. This fall reached a considerably lower pitch level than all preceding targets. The presence and alignment properties of the LHLH pattern is in line with the contour presented in [11], which was also obtained from a production with a sonorant coda consonant in final position. The finding that a final fall in declarative sentences is phonologically conditioned reconciles the opposing observations in [5,11] and [10].

The finding that an LLH sequence occurs in final and non-final APs, in both cases showing similar alignment properties, gives rise to the assumption that the late-rise pattern constitutes the default tonal pattern in both kinds of APs. The final fall reaches the bottom of the register and might thus be related to an L\% boundary tone associated with the edge of an intonation phrase. Furthermore, it might be that the L\% is only realized if sonorant material is available after the vowel, but is otherwise truncated.

The patterns observed here differ in some respects from those observed in [5], who investigated the same variety of Javanese as in the present study. First, the default tonal contour for non-final APs assumed in [5] was the gradual-rise pattern, modeled as a sequence of an L at the left and an H at right edge. The pitch tracks presented in [5] however contain very few of these instances, which, at least in some cases, also allow for different interpretations regarding the tonal targets (e.g., Figure 6 in [5]). It is also possible that the gradual-rise pattern constitutes a reduced variant of the late-rise pattern in which the second L is omitted. Models accounting for the tonal variants of non-final APs by means of omission of one or more tones of a default contour have been proposed for other ‘phrase languages’ (e.g., [16] for French, [17] for Korean). It might well be that a similar strategy applies to Javanese. Certainly, further research is required in this regard.

The observations from the present study also differ from those in [5] with regard to the alignment of the H preceding a final fall. In the present study, the H was consistently aligned with the final syllable whereas, in [5], it was aligned mostly with the penultimate syllable. Note, however, that all instances of APs ending in a final fall in [5] were realized on a narrowly focused constituent. Thus, the earlier alignment of the peak might be related to the expression of focus. The rise-fall pattern shown for imperatives in [11] also involves an earlier aligned peak than the pattern for declaratives. This difference might be related to emphasis, which might have been stronger on the imperative. Further research should address the relation between tonal alignment and pragmatic meaning.

The assumptions proposed in the present study must be regarded as preliminary, as they are based on production data from a single speaker. Future research should be based on larger data sets containing productions from several speakers. Furthermore, Javanese phrasal prosody should be addressed with regard to speech perception. As a tonal rise appears to be aligned near the right edge of an AP, it is likely that listeners employ them for speech segmentation, for example identifying a word boundary at the location of the pitch peak. The present study was meant to provide a starting point for further investigations on Javanese prosody.

6. Acknowledgments

Many thanks to the Ngapak speaker who participated in this study. I am also grateful for the comments I received from two anonymous reviewers and the participants of the 1st International Conference on Tone and Intonation, held in December 2021 at the University of Southern Denmark.

7. References


