Effects of Consonant Type on Tone Realization in Luganda

Scott Myers

Department of Linguistics, University of Texas at Austin, U.S.A.

s.myers@austin.utexas.edu

Abstract

Since the f0 realization of tone categories requires voicing, interruptions of voicing disrupt that realization. Speakers have been shown to truncate the f0 excursion or compress its time-course in order to fit it into voiced sonorant intervals shortened by encroaching voiceless consonants [5–8]. In this study, the effect of consonants on f0 trajectories was investigated in Luganda (Bantu, Uganda).

Speakers produced Luganda sentences including a C1-V1-C2-V2 interval, in which V1 was a short vowel with a high tone, and C1 and C2 varied in manner (sonorant, voiced obstruent, voiceless obstruent) and quantity (singleton, geminate). It was found that when C2 was voiceless, speakers sped up the f0 rise up to the voiceless interval, and slowed down the f0 fall during the voiceless interval. These adjustments had the effect of concentrating the f0 movements for the high tone in the voiceless intervals before and after the voiceless interval. Voiceless C1, on the other hand, had no effect on the timing of the f0 trajectory.

Index Terms: Luganda, lexical tone, phonation, f0 timing

1. Introduction

Fundamental frequency (f0) is a property of voiced intervals in speech, yet such intervals are constantly interrupted by voiceless consonants and pauses, and by the low-amplitude voicing of voiced obstruents [18]. These intervening voiceless or weakly voiced intervals shorten the voiced intervals necessary for the f0 trajectories realizing tone categories. Speakers adjust the f0 trajectory to fit it into the reduced available voiced interval, either through truncation of the f0 excursion, or through compression of the excursion through an increase in the rate of change [5–8].

The present study addresses the effect of intervening voiced and voiceless obstruents on the f0 trajectory in Luganda, a Bantu tone language spoken in Uganda. Luganda has three contrasting tone categories: high, falling and low [17]. Previous acoustic phonetic work on Luganda has shown that a high-tone syllable has an f0 rise from the syllable onset to a peak near the end of the vowel, with a subsequent f0 fall mainly in the following syllable [12,13]. The falling tone is the same, but with an f0 peak earlier in the syllable. Up to now, this research has been limited to contexts in which the neighboring consonants were sonorants, and so has left unresolved how voiced and voiceless obstruents affect f0 trajectories in the language.

2. Methods

Participants

19 adult native speakers of Luganda participated (6 female, 13 male). They were from all over the Central region of Uganda, where Luganda is spoken.

Materials

There were two sets of test sentences: one for Experiment 1 and one for Experiment 2. In both sets, the test syllable, the second syllable in the sentence, had a high tone and a short vowel. The syllables on either side of the test syllable were low-toned. The sentences were declarative statements.

In the Experiment 1 set, the test syllable began with an initial sonorant consonant C1, and the vowel was followed by a consonant C2. In half of the test items, C2 was a singleton and in the other half it was geminate. Within each of these classes, a third of the items had a sonorant nasal stop, a third had a voiceless obstruent stop, and a third had a voiceless obstruent stop. There were thus 6 classes of C2: 2 quantity classes and 3 manner classes.

In the Experiment 2 set, C1 belonged to one of the 6 classes defined for C2 in Experiment 1, and C2 was a singleton sonorant consonant.

In each set, there were 120 sentences, consisting of 20 sentences in each of the 6 conditions.

Procedure

The participants read the test sentences aloud from a self-paced PowerPoint presentation on a computer screen, with one sentence presented at a time. They were instructed to read each sentence as a separate statement, and to pause before beginning the next one. They were also told that they could repeat the sentence until they felt that they had it right. The sentences were presented in randomized order, interspersed with sentences belonging to other experiments.

The productions were recorded using a solid-state recorder with a head-mounted microphone, at a sampling rate of 44.1 kHz.

Measurements

Within the C1-V1-C2-V2 measurement interval, the duration of the closure intervals of C1 and C2, and of the vowel intervals V1 and V2 were measured in Praat [3].

F0 measurements in the C1-V1-C2-V2 interval were made using a Praat script. F0 value and timepoint were obtained for the landmarks Min1 (the f0 minimum at the beginning of the test syllable), Max1 (the f0 maximum at the end of the initial rise), and Min2 (the f0 minimum following the maximum). In addition, to accommodate those items in which the f0 trajectory...
was interrupted by a voiceless interval, there was a \( \text{Max2} \) measurement (the \( f_0 \) maximum at the onset of voicing in V2).

Statistical analyses were done using mixed-model regression models in R [2, 11, 15]. The fixed effects in Experiment 1 were C2 quantity (Singleton/Geminate) and C2 manner (Sonorant/Voiced Obstruent/Voiceless Obstruent). The default levels were Sonorant for C2 manner and Singleton for C2 quantity. In Experiment 2 it was the same groups for C1. The random effects in both experiments were Speaker and Item.

3. Results

Experiment 1: The effects of C2

In Experiment 1, the 19 participants produced a total of 2280 sentences. Of these 93 were excluded from the study: 90 due to speaker errors, and 3 because the measurement interval was too reduced for segmentation. This left 2187 items for analysis.

The mean closure duration for singletons was 52 ms and for geminates was 117 ms. The mean voiceless interval was 102 ms in voiceless obstruent C2 (135 ms in voiceless geminates) and 9 ms in voiced obstruent C2. There were no voiceless intervals in the sonorant C2 condition.

Figures 1-3 present sample spectrograms and pitch tracks for test words with geminate C2 consonants of the three different manner classes. In Figure 1, with a sonorant C2, \( f_0 \) rises from the beginning of C1 to a point just after the end of V1 and falls mainly during the C2 closure interval. The trajectory is similar with the voiced obstruent C2 in Figure 2. But in Figure 3, with a voiceless C2, the peak is a bit before the end of V1 and the visible fall is concentrated in V2.

Figure 1: Sonorant C2 [lánmal] alannamye “he stretched”

Figure 2: Voiced C2 [müdddo] omuddo “grass”

Figure 3: Voiceless C2: [müttô] omutto “pillow”

Figure 4: Mean normalized \( f_0 \) (z) by time from C1 onset (ms), broken down by C2 quantity and manner

Figure 4 presents the mean normalized \( f_0 \) trajectory in the measurement interval for each of the 6 C2 conditions. In all conditions, \( f_0 \) rises from the onset of C1 to a peak about 100 ms after that. It is in the subsequent fall that the conditions diverge. In the voiceless C2 condition (the third panel) the gap in the series of points represents the voiceless interval.

The initial \( f_0 \) maximum (Max1) was similar across conditions, with no significant differences according to C2 Quantity or Manner. The only significant effect was the interaction Quantity: Geminate * Manner: Voiced Obstruent (\( t_{44} = 2.3, p = 0.03 \)).

However, there were differences across the conditions in the timing of the \( f_0 \) trajectory. This can be seen in Figure 5, which presents peak delay (the duration of the interval from the onset of C1 to the Max1 peak) by the duration of the C1-V1 interval. The diagonal line represents \( x = y \), so points on that line are exactly at the offset of V1. It can be seen that all points are close to that, but there are more points above that line in the Sonorant C2 condition, and in particular the Sonorant Geminate C2 condition (the red points in the top panel).
Figure 5: Peak delay (ms) by C1-V1 duration (ms), broken down by C2 quantity and manner

Relative peak delay is peak delay divided by the duration of the C1-V1 interval. It represents the proportion of the test syllable at which the f0 peak occurred. Mean relative peak delay in sonorant C2 was 1.05, indicating that on average the f0 peak in this condition was slightly after the end of V1, in the closure interval of C2. For geminate sonorant C2, the mean relative peak delay was 1.14. For obstruent C2, however, mean relative peak delay was lower: 0.88 for voiced obstruent C2 and 0.90 for voiceless obstruent C2, indicating that in these conditions the peak was on average before the end of V1. The f0 peak occurs later in syllable in those syllable types with a longer voiced sonorant interval.

The effects of C2 are clearest in the velocity of the f0 trajectory. To accommodate the voiceless interval, this trajectory is divided into three intervals: the Initial Rise (Min1 to Max1), the Early Fall (Max1 to Max2), and the Final Fall (Max2 to Min2). The end of the Early Fall interval is the onset of voicing in V2, which for voiceless C2 is the end of the voiceless interval. The mean rate of f0 change is given in Tables 1–3 for each of these three intervals in each C2 manner condition.

Table 1: Mixed-model analysis of f0 change rate in the Initial Rise interval

<table>
<thead>
<tr>
<th></th>
<th>β</th>
<th>d.f.</th>
<th>t</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>0.31</td>
<td>23.3</td>
<td>9.2</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>C2: Voiced</td>
<td>-0.00</td>
<td>46.5</td>
<td>-0.2</td>
<td>0.90</td>
</tr>
<tr>
<td>C2: Voiceless</td>
<td>0.07</td>
<td>46.2</td>
<td>4.1</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>C2: Geminate</td>
<td>0.02</td>
<td>53.6</td>
<td>0.9</td>
<td>0.35</td>
</tr>
<tr>
<td>C2:Voiced *</td>
<td>0.09</td>
<td>43.0</td>
<td>3.3</td>
<td>0.002</td>
</tr>
<tr>
<td>C2:Geminate</td>
<td>0.04</td>
<td>43.3</td>
<td>1.6</td>
<td>0.12</td>
</tr>
</tbody>
</table>

Table 2: Mixed-model analysis of f0 change rate in the Early Fall interval

<table>
<thead>
<tr>
<th></th>
<th>β</th>
<th>d.f.</th>
<th>t</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>0.19</td>
<td>26.5</td>
<td>6.3</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>C2: Voiced</td>
<td>0.03</td>
<td>47.4</td>
<td>1.8</td>
<td>0.08</td>
</tr>
<tr>
<td>C2: Voiceless</td>
<td>-0.13</td>
<td>46.9</td>
<td>-6.8</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>C2: Geminate</td>
<td>0.05</td>
<td>46.1</td>
<td>2.0</td>
<td>0.049</td>
</tr>
<tr>
<td>C2: Voiced * C2: Geminate</td>
<td>-0.03</td>
<td>42.9</td>
<td>-1.0</td>
<td>0.34</td>
</tr>
<tr>
<td>C2: Voiceless * C2: Geminate</td>
<td>0.03</td>
<td>43.3</td>
<td>1.0</td>
<td>0.30</td>
</tr>
</tbody>
</table>

Table 3: Mixed-model analysis of f0 change rate in the Final Fall interval

<table>
<thead>
<tr>
<th></th>
<th>β</th>
<th>d.f.</th>
<th>t</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>0.35</td>
<td>27.6</td>
<td>10.3</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>C2: Voiced</td>
<td>-0.09</td>
<td>44.1</td>
<td>-4.0</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>C2: Voiceless</td>
<td>0.09</td>
<td>43.6</td>
<td>4.0</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>C2: Geminate</td>
<td>-0.05</td>
<td>47.4</td>
<td>-1.6</td>
<td>0.13</td>
</tr>
<tr>
<td>C2: Voiced * C2: Geminate</td>
<td>0.03</td>
<td>40.2</td>
<td>0.8</td>
<td>0.45</td>
</tr>
<tr>
<td>C2: Voiceless * C2: Geminate</td>
<td>-0.07</td>
<td>40.6</td>
<td>-1.9</td>
<td>0.07</td>
</tr>
</tbody>
</table>

If the f0 trajectory in the voiceless obstruent C2 condition was the same as in the sonorant C2 condition, the f0 peak and most of the subsequent fall would coincide with the voiceless interval. But this is not what we see. The initial rise is faster in the voiceless C2 condition, putting the f0 peak well before the onset of voicelessness. The fall is slowed during the voiceless interval, leaving the bulk of the f0 movement for the sonorant voiced interval of V2.

Experiment 2: The effects of C1

In experiment 2, items differed in the quantity and manner of the consonant C1 preceding the high-toned V1. 2280 sentences were recorded, and 75 were excluded from the analysis (73 due to speaker error and 2 due to segmentation issues). This left 2205 tokens for analysis.

Figure 6 gives a mean normalized f0 trajectory for the measurement interval, broken down by C1 quantity and manner. The most striking aspect of these graphs is that the peak with geminate C1 is considerably later than that for singleton C1. This is because C1 is part of the C1-V1 interval represented in the x-axis, and geminate C1 is longer (mean closure duration 129 ms) than singleton C1 (mean closure duration 51 ms).

In Figure 7, f0 peak delay is plotted against this C1-V1 duration, with the solid line marking x=y, i.e. f0 peaks at the end of C1. It can be seen that the f0 peak is close to that line.
across C1-V1 durations and across conditions. There is no clear difference among the C1 quantity and manner classes as far as their grouping relative to the $x=y$ line.

Mean relative peak delay is very similar across the manner classes: Sonorant (0.95), Voiced Obstruent (0.91) and Voiceless Obstruent (0.85). The mean also does not vary much according to C1 quantity: Singleton (0.89) and Geminate (0.92). In a mixed-model regression analysis of relative peak delay, none of the experimental C1 factors had any significant effect. Unlike the case with C2, then, there is no evidence in this case of an adjustment of the f0 trajectory to fit the available voiced sonorant interval.

![Figure 6: Mean normalized f0 (z) by time from C1 onset (ms), broken down by C1 quantity and manner](image1)

![Figure 7: Peak delay (ms) by C1-V1 duration (ms), broken down by C1 quantity and manner](image2)

**4. Discussion**

Experiment 1 provided evidence that the voicing of a postvocalic consonant affects the timing of the f0 trajectory of a high tone in Luganda. With a voiceless C2, the initial f0 rise is faster, resulting in an earlier f0 peak, and the f0 fall is slower during the voiceless interval and faster after that. These adjustments have the effect of maximizing the proportion of the whole f0 trajectory that occurs in the voiced sonorant intervals before and after the voiceless interval.

These adjustments can be described neither as truncation nor compression [5–8]. The extent of the f0 excursion is similar across conditions, and the rate of f0 change in each interval of the trajectory depends on the interval’s position relative to the voiceless interval. But these adjustments are similar to truncation and compression in that they result in fitting the f0 trajectory to the available voiced interval. Other such listener-oriented adjustments in articulatory timing include the timing of f0 landmarks to avoid intervals of creaky voicing [14, 16], or the timing of the peak of a laryngeal gesture in a stop to coincide with the closure release [9]. The partial suspension of the f0 fall during the voiceless interval is similar to the resumption of the f0 topline after a parenthetical [10].

Perception studies [1, 4] have demonstrated that in just this configuration of a rise-fall f0 trajectory interrupted at the maximum by a voiceless interval, listeners perceive a pitch corresponding to the maximum value in the remaining voiced interval, without any extrapolation or interpolation through the intervening blank period. Because listeners do not fill in f0 movements obscured by voicelessness, speakers have to ensure that the relevant f0 information is available in the voiced intervals.

It is interesting that voiceless consonants in the prevocalic C1 position do not trigger the same adjustments in f0 timing. One interpretation of this would be that the interval of the f0 trajectory that is compromised by a voiceless C1, the initial part of the rise, is less important to the perception of high tone than the interval affected by a voiceless C2, which is the interval including the peak and the subsequent beginning of the fall.

**5. Acknowledgements**

The author would like to thank Dr. Sadaah Namyalo, Sam and Rose Museke, Anatole Kiriggwajjo, and Paul Bbosa for their invaluable help in designing and running this experiment, and the participants for sharing their knowledge of the language. This research was supported by a fellowship from the Harold C. and Alice T. Nowlin Regents Professorship in Liberal Arts, from the College of Liberal Arts of the University of Texas at Austin.

**6. References**


