**Extreme Tonal Depressor Effects in Khoisan: Evidence from Tsua**

Timothy K. Mathes

Department of Linguistics, New York University, USA
tkm237@nyu.edu

**Abstract**

This paper reports on the interaction between Fundamental Frequency (F0), voicing and aspiration in Tsua, a Central Khoisan click language spoken in Botswana. Tsua lacks the tonal mobility of spreading and shifting found in many neighboring Bantu languages. Nevertheless, original field research in 2012 and 2013 revealed that Tsua has an extreme tonal depressor effect that lowers a post-consonantal, root-initial H-tone’s production by ~50 Hz or more. Tonal depression can be triggered by voiced stops, aspirated stops or the glottal fricative /h/, an effect that has been missed by the small amount of previous research on Tsua. However, this inquiry also found that 19.2% of the root-initial H-tones that are expected to be depressed by the triggering environment are produced without depression, perhaps providing a reason as to why the effect has been overlooked. Cognates with the Central Khoisan languages G|ui and Kua were investigated to formulate a plausible explanation. It was determined that tonal depression exceptions with root-initial voiced stops in Tsua always correspond with G|ui and Kua cognates that have root-initial nasals or nasalized clicks, consonants which do not trigger depression in Tsua. These results provide insight into diachronic sound change that has influenced the production of Tsua lexical tone.

**Index Terms:** tone, production, depressors, consonant-tone interaction, Khoisan

1. Introduction

Tonal depression was first described in Beach [1] for the South African Nguni language Xhosa. Beach divided initial consonants into two classes according to their tonal affinities, with one set of consonants associated with the Low tone class. Subsequently, Doke [5] chronicled the phenomenon of a group of segments followed by Low tones in the Nguni language Zulu. It was not until Lanham [10] when the term ‘depressor’ was introduced to the study of Nguni languages, which specified a consonant class that triggered a lowering effect on all tones except Low level tones. Furthermore, Rycroft [12] claimed to have found lexically and grammatically conditioned cases of tonal depression in the Nguni language Swati.

Although the set of pitch depressors in Nguni languages consists of phonologically voiced obstruants such as bh, d, g, v, z, dl, lh and the nasal compounds occurring with voiced obstruants, e.g., mb, nd, etc. [17], Traill et al. [16] wrote that the set of depressors in Zulu was phonetically heterogeneous, claiming /bh dh gh/ were phonetically voiceless unaspirated stops and /v z dl/ phonetically voiced fricatives that caused tonal depression. Schachter [13] and Downing [6] provided additional evidence of voiced and voiceless consonants triggering tonal depression in the languages of the Shona and Nguni groups. What is particularly striking about the depressor consonants in Nguni languages is they often induce exaggerated pitch lowering: one female subject in Traill et al.’s [16] study had an onset F0 difference of ~70 Hz between High and depressed High tones, with the rest of the subjects having a 30-60 Hz differential. Traill [17] also reported the onset of a depressed High tone as about 70 Hz lower than a non-depressed High tone in a study of Swati. All of the aforesaid Nguni languages are of Bantu stock.

There is evidence that aspiration may trigger tonal depression in Asian tone languages. In a paper investigating the phonetic realization of the contrastively aspirated affricates /tsʰ/ and /dzʰ/ in Nepali, Clements and Khatiwada [4] found the F0 of /tsʰa/ lower than that of /tsa/ by an average of more than 20 Hz for two speakers. The F0 of /dzʰv/ was an average of 13 Hz lower for one speaker and 20 Hz for a second compared to /dzv/. In a preliminary report on the Leng-shiu-jiang dialect of Chinese, Caicai [2] calculated a significantly lower F0 for two out of three tone groups following aspirated obstruents produced by six speakers (three male, three female). Xu and Xu [20] studied Mandarin Chinese to clarify the effect of consonant aspiration on the following F0. There were five factors involved in the experiment: consonant, lexical tone, syllable position, tonal context and carrier sentence. The main effects of consonant aspiration and lexical tone, as well as their interaction, were all significant, with a mean difference of about 20 Hz in the expected direction for unaspirated versus aspirated consonants. Moreover, the magnitude of the effect of consonant aspiration on F0 varied according to the preceding tone. When the preceding tone was High or Rising, the F0 onset was much higher for /ta/ than for /tʰa/. In addition, the effect of aspiration on F0 was greater for the Rising and Low Tones than for the High and Falling Tones.

What makes the current inquiry intriguing is the extreme tonal depression of ~50 Hz or more on root-initial High tones (henceforth H-tones) in the Khoisan language Tsua. Original field research in 2012 and 2013 with three native Tsua consultants in Botswana revealed that Tsua has a pitch depressor effect on a post-consonantal tone, which may be triggered by three activities of the larynx: (i) when the plain stop or click consonant preceding the tone is voiced (e.g., the voiced alveolar stop /d/ or the voiced dental click /ɡd/; (ii) when the plain or click consonant is aspirated (e.g., the voiceless aspirated velar stop /kʰ/ or the voiceless aspirated lateral click /l/; or (iii) when the glottal fricative /h/ is the root-initial consonant. This paper provides acoustic evidence of Tsua tonal depression and discusses the implications of the findings, which are based on ~1,300 elicited lexical items.

2. The Tsua language

2.1. Classification

Tsua is a critically-endangered Khoisan language classified by Güldemann [8] and Güldemann and Vossen [7] as a Kalahari Khoe East language of the Central Khoisan branch and part of the Tsxwa subgroup. The Kalahari Khoe East languages occupy the eastern part of Botswana, southeast of the Central Kalahari Game Reserve and eastward towards the Western Sandveld of the Central District. These languages are spoken mainly around the Makgadikgadi Salt Pans, as well as towards
the Shashe river, Serule and Mabesekwa areas in eastern Botswana, Serowe and Shoshong.

2.2. Depressor consonants

Tsua has a large consonant inventory in line with previous studies of other Khoisan languages. The non-depressor, non-click consonants are: p, t, ts, c, k, q, ʔ, t’, ts’, t’s, c’, k’, q’, q’', t”, ts”, t’s”, c”, k”, q”, h, are depressors. The click consonants presented in this paper follow Ladefoged and Trill [9] by using ‘click type’ and ‘accompaniment’ as guides for transcribing the clicks. The click type is determined by the release of the anterior closure and the accompaniment consists of: (i) voicing, aspiration and ejection associated with the posterior closure; (ii) posterior closure place of articulation; (iii) the oro-nasal setting; and (iv) the plosion or friction in the uvula or glottis following the release of the anterior closure. Tsua has four click types symbolized as: [i] dental, [i] palatal, [!] alveolar, and [!l] lateral, with voicing and nasality transcribed before the click symbol. For example, [g] represents the voiced dental click, [q] is the voiceless dental click and [ŋ] is the nasalized dental click. The non-depressor click consonants are: |, ǂ, ǃ, ǁ, |ʔ, ǂʔ, ǃʔ, ǁʔ, |χ, ǁχ, |χ’, ǃχ’, ǁχ’, |q, ǂq, ǃq, ǁq, while click and [ŋ|] is the nasalized dental click. The non-depressor effect that has been overlooked, an oversight that has led to an incomplete understanding of Tsua lexical tone. Nevertheless, while Tsua lacks the tonal mobility found in many African languages, it does have a tonal depressor effect which is about 80 Hz. Speaker M’s fall is normally from around her tokens. Speaker B produces her H-L contour with a fall quickly from the H-tone to the L-tone target. For speaker S, F0 falls very rapidly from the H-tone to the L-tone target. What is notable about the melodies exemplified in (1) is the F0 contours unfold over time in a way that is different from Niger-Congo and Nilo-Saharan tone languages, especially with respect to level tones. For example, the M-M melody’s F0 rises, dips and then rises again towards the end of the second Mid tone, i.e., it has a double rise. Each melody has its own distinctive F0 curve shape, and it is the comparisons of these shapes that were used as one criteria for grouping and transcribing tonal melodies in citation form. A near-minimal CVV sextuplet in (2) provides supporting evidence for the melodies listed in (1).

Table 1. Vossen’s five tone classes in the Tsha subgroup.

<table>
<thead>
<tr>
<th>Tone Class</th>
<th>Sequence</th>
<th>Verb</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>H-H</td>
<td>əoə 'shoot'</td>
</tr>
<tr>
<td>2</td>
<td>L-L</td>
<td>ɡaː ʔaː ʔaː 'spread a hide'</td>
</tr>
<tr>
<td>3</td>
<td>H-L</td>
<td>pii 'suck'</td>
</tr>
<tr>
<td>4</td>
<td>L-H (I)</td>
<td>tshô  'dig'</td>
</tr>
<tr>
<td>5</td>
<td>L-H (II)</td>
<td>xún 'grind'</td>
</tr>
</tbody>
</table>

Positing tonal classes 4 and 5 as L-H (I) and L-H (II), respectively, obscures the Tsua depressor effect triggered by aspiration in tshau 'dig,' which can be observed when instrumental and comparative data from the closely-related language Kua are taken into consideration. In the case of Kua, tshau 'dig' is analyzed as H-M in Chebanne and Collins [3], while Tsua tshao 'dig' as DH-M (Depressed High-Mid) in my analysis. I analyze Xun 'grind' as L-H, which is in agreement with Vossen.

2.4. Tone melodies

Lexical items in Tsua are bimoraic, with tones linking to each mora, resulting in tonal melodies of the form [Tₗ-Tₑ]. Furthermore, I am positing three underlying tones High, Mid and Low: /H/, /M/, and /L/. Nasal consonants can bear tone. Six non-depressed tonal melodies have been observed, as seen in (1).
2.6. Tone melody H-M

The initial H-tone for the H-M melody is produced at around 200 Hz for speaker S — F0 subsequently drops to a little over 150 Hz to reach the Mid-tone target. Speakers B and M produce their initial H-tones starting at ~225 Hz, with their respective contours falling to around 160 Hz (see Fig. 2). (4) presents a subset of the Kua, Tsua H-M cognates.

Kua H-M and Tsua H-M cognates

<table>
<thead>
<tr>
<th>Kua H-M</th>
<th>Tsua H-M</th>
<th>Gloss</th>
</tr>
</thead>
<tbody>
<tr>
<td>lae</td>
<td>lae</td>
<td>'to chew'</td>
</tr>
<tr>
<td>ʔee</td>
<td>ʔee</td>
<td>'fire; firewood'</td>
</tr>
<tr>
<td>ϱa</td>
<td>ϱa</td>
<td>'meat'</td>
</tr>
<tr>
<td>ǀqore</td>
<td>ǀqore</td>
<td>'claw; fingernail'</td>
</tr>
<tr>
<td>tsoro</td>
<td>tsoro</td>
<td>'husk; shell'</td>
</tr>
<tr>
<td>ϱam</td>
<td>ϱam</td>
<td>'lion'</td>
</tr>
<tr>
<td>saa</td>
<td>saa</td>
<td>'to heal; to rest'</td>
</tr>
<tr>
<td>ϱan</td>
<td>ϱan</td>
<td>'to sew'</td>
</tr>
</tbody>
</table>

The DH-M contour has an F0 decrease in the realization of the first part of the initial tone as seen in Fig. 4. F0 then rises before it falls to the M-tone target. The fall from the F0 peak to the F0 offset is usually 40-50 Hz for all three consultants. (6) lists a subset of the Kua H-M and Tsua DH-M cognates.

Kua H-M and Tsua DH-M cognates

<table>
<thead>
<tr>
<th>Kua H-M</th>
<th>Tsua DH-M</th>
<th>Gloss</th>
</tr>
</thead>
<tbody>
<tr>
<td>kʰæc</td>
<td>kʰæc</td>
<td>'to stab'</td>
</tr>
<tr>
<td>dam</td>
<td>dam</td>
<td>'tongue'</td>
</tr>
<tr>
<td>glam</td>
<td>glam</td>
<td>'thorn'</td>
</tr>
<tr>
<td>tsʰaa</td>
<td>tsʰaa</td>
<td>'water'</td>
</tr>
<tr>
<td>cʰuro</td>
<td>tʰoro</td>
<td>'to pluck'</td>
</tr>
<tr>
<td>gui</td>
<td>gui</td>
<td>'rope; trapping string'</td>
</tr>
<tr>
<td>tsʰau</td>
<td>tsʰau</td>
<td>'to dig'</td>
</tr>
</tbody>
</table>

2.7. Evidence for tonal depression

We now turn to evidence for tonal depression in Tsua by observing the F0 contours and reviewing comparative data from Kua, which does not have tonal depression. The two depressed contours in question will henceforth be labeled as DH-L (Depressed Low-Low) and DH-M (Depressed High-Mid). The lexicon has numerous examples of DH-L and DH-M with voiced stops, aspirated stops or the glottal fricative /h/ as triggers. For all three consultants, there is an F0 decrease of 50 Hz or more in the initial H-tone realization of DH-L, which rises towards the H-tone target before it falls sharply to the L-tone target, a contour that is similar to the H-L melody after the initial pitch lowering is realized. The fall from the F0 peak to the F0 offset is 80-100 Hz. (5) presents a list of Kua H-L and Tsua DH-L cognates with illustrative spectrograms and F0 traces in Figure 3.

Kua H-L and Tsua DH-L cognates

<table>
<thead>
<tr>
<th>Kua H-L</th>
<th>Tsua DH-L</th>
<th>Gloss</th>
</tr>
</thead>
<tbody>
<tr>
<td>ǀqoo</td>
<td>ǀqoo</td>
<td>'aardvark'</td>
</tr>
<tr>
<td>ǀʔana</td>
<td>ǀʔana</td>
<td>'split wood'</td>
</tr>
<tr>
<td>ǀqaa</td>
<td>ǀqaa</td>
<td>'Silver tree'</td>
</tr>
<tr>
<td>kʰan</td>
<td>kʰan</td>
<td>'to crawl'</td>
</tr>
<tr>
<td>gam</td>
<td>gam</td>
<td>'to throw'</td>
</tr>
<tr>
<td>ǀpaba</td>
<td>ǀpaba</td>
<td>'to stumble,umble'</td>
</tr>
<tr>
<td>ǀpobe</td>
<td>ǀpobe</td>
<td>'to creep toward'</td>
</tr>
<tr>
<td>ǀqʰam</td>
<td>ǀqʰam</td>
<td>'spiderweb silk'</td>
</tr>
</tbody>
</table>

Kua H-M and Tsua DH-M cognates

<table>
<thead>
<tr>
<th>Kua H-M</th>
<th>Tsua DH-M</th>
<th>Gloss</th>
</tr>
</thead>
<tbody>
<tr>
<td>kʰæc</td>
<td>kʰæc</td>
<td>'to stab'</td>
</tr>
<tr>
<td>dam</td>
<td>dam</td>
<td>'tongue'</td>
</tr>
<tr>
<td>glam</td>
<td>glam</td>
<td>'thorn'</td>
</tr>
<tr>
<td>tsʰaa</td>
<td>tsʰaa</td>
<td>'water'</td>
</tr>
<tr>
<td>cʰuro</td>
<td>tʰoro</td>
<td>'to pluck'</td>
</tr>
<tr>
<td>gui</td>
<td>gui</td>
<td>'rope; trapping string'</td>
</tr>
<tr>
<td>tsʰau</td>
<td>tsʰau</td>
<td>'to dig'</td>
</tr>
</tbody>
</table>
3. Discussion

It has been argued here that positing tonal melodies L-H (I) and L-H (II) misses the Tsua tonal depressor effect on root-initial H-tones, an oversight that becomes apparent when instrumental and comparative data are taken into consideration. A factor that renders the effect easy to overlook are exceptions in which the initial H-tone is not depressed. Table 2 lists the exceptions count in the first row; the second row labeled ‘environments’ considers the total number of roots where tonal depression is expected to occur, including the exceptions. Tonal depression exceptions for H-tones constitute 19.2% of the data, a rather large percentage.

Table 2. Tonal depression exceptions by triggering category.

<table>
<thead>
<tr>
<th>G</th>
<th>ui</th>
<th>Kua</th>
<th>Tsua</th>
<th>Gloss</th>
</tr>
</thead>
<tbody>
<tr>
<td>voiced stops</td>
<td>/h/</td>
<td>Total</td>
<td></td>
<td></td>
</tr>
<tr>
<td>exceptions</td>
<td>33</td>
<td>4</td>
<td>4</td>
<td>41</td>
</tr>
<tr>
<td>environments</td>
<td>112</td>
<td>88</td>
<td>14</td>
<td>214</td>
</tr>
</tbody>
</table>

It is informative to consider why so many exceptions occur by inspecting cognates between the Kalahari Khoe West language G|ui, along with Kua and Tsua, which gives us a sense of the click erosion observations (i.e. click replacement and click loss) as described in Traill [14], [15] and Traill and Vossen [18]. In the present case of G|ui, Kua and Tsua, I posit that click erosion for the exceptions comes in three forms: (i) nasalization loss; (ii) click and nasalization loss; and (iii) click and nasalization loss plus fronting. Table 3 lists a subset of the exceptions sorted by click erosion type. All G|ui data come from Nakagawa [11] while the Kua data come from Chebanne and Collins [3] and original field research in 2013.

Table 3. Tonal depression exceptions with comparative data.

| G|ui | Kua | Tsua | Gloss |
|---|---|---|---|
| ҭǃUi | ҭǃUi | ҭǃUi | ‘oil’ |
| ҭǁOt | ҭǁOt | ҭǁOt | ‘moon’ |
| 𦰡ǃRo | SPATH | .PathVariable | ‘chameleon’ |
| agascar | .HashSet | HashSet | ‘to hide’ |
| .DropTable | Queryable | Queryable | ‘backbone’ |
| ҭǁOt | ҭǁOt | ҭǁOt | ‘hut’ |

My interpretation is Tsua’s diachronic click erosion involved nasalization loss and in some cases the root-initial consonants underwent fronting. These roots did not originally have depressed tones since nasalized clicks and nasal non-clicks do not trigger tonal depression in Tsua. What remains uncertain is why tonal depression did not affect these roots after the sound changes anyway as a synchronic rule. One explanation is the consonant-tone interactions were firmly established by the time click erosion commenced. Therefore, the roots in question retained their original H-tones after click erosion, even though the tones were eventually preceded by voiced stops as seen in the Table 3 cases. What may be unusual here is the apparently frozen-in-time nature of the extreme tonal depressor effect, since it appears to be triggered by voiced stops that did not originate as nasalized clicks or plain nasal consonants. This is a potentially fruitful area of future inquiry on the nature of tonal change.

There are three predictions made by my interpretation: (i) root-initial voiced clicks in G|ui and Kua should correspond with depressed Tsua H-tones, whether the clicks have been eroded or not; (ii) depressed H-tones in Tsua should never correspond to root-initial nasals in Kua and G|ui; and (iii) any lexical items entering the grammar after the Tsua sound changes would not be subject to a depressor effect. Evidence supporting the first prediction is displayed in Table 4. The examples show the cognates between Giui, Kua and Tsua where the voiced alveolar click has eroded but tonal depression is still triggered in the Tsua correspondent.

Table 4. Giui, Kua and Tsua cognates where tonal depression is triggered in Tsua despite click erosion.

| G|ui | Kua | Tsua | Gloss |
|---|---|---|---|
| ҭǁUi (H-M) | ƣUi (H-M) | ƣUi (DH-M) | ‘steenbok’ |
| g’oo (H-L) | g’oo (H-L) | g’oo (DH-L) | ‘aardvark’ |
| ƣUi (H-M) | ƣUi (H-M) | ƣUi (DH-M) | ‘rope’ |
| g’um (H-L) | g’um (H-L) | g’um (DH-L) | ‘to blow’ |
| gluu (H-M) | gluu (H-M) | gluu (DH-M) | ‘chess’ |

With respect to the second prediction, there are no instances of Tsua depressed H-tones following voiced stops with cognate nasal correspondences in Kua or Giui in my data set. There is evidence to support the third prediction as well: recent Setswana loanwords are not subject to tonal depression. For instance, the Tsua H-M-L [dèbèdu] ‘dewlap; chicken jowl’ and H-M [diā] ‘to delay’, which should have word-initial depressed H-tones but do not, are likely the Setswana loans lebedu ‘dewlap’ and diīwa ‘to delay’, respectively. At least 1 of the 4 exceptions for the glottal fricative /h/ category as listed in Table 2 is a loanword: the H-L [huma] ‘to become rich’ is from Setswana huma. Thus, by appealing to sound change processes many of the exceptions can be plausibly accounted for.

4. Conclusions

The results presented in this paper give us crucial pieces of information that have not been reported in the Khoisan literature. For instance, click erosion has consequences not just on the Tsua clicks themselves but also on tone. Vis-à-vis click replacement and click loss, Traill and Vossen [18] states: ‘...tone plays no part in the phonetic processes we are dealing with.’ While the statement may be generally true, these processes obscure the Tsua tonal system, as exceptions to tonal depression are best explained by looking at comparative data, taking into account click erosion. Future research on Khoisan tone would greatly benefit from this approach.

5. Acknowledgements

Many thanks to Chris Collins, Andy Chebanne and Gillian Gallagher for their advice and support in this ongoing dissertation research project. A special thanks goes to the Tsua and Kua consultants for their patience and wisdom during my numerous elicitation sessions.

6. References


