A spectral moments analysis of Tsua and Juǀ’hoan alveolar and palatal click consonants

Timothy K. Mathes                        Alex de Voogt
Assured Information Security, Inc.        Drew University

Tsua is a critically endangered language of the Eastern Kalahari Khoe languages where the historical processes of click replacement and click loss are relatively common. An acoustic analysis of Tsua’s remaining alveolar and palatal click consonants that have not undergone click loss reveals that on three statistical spectral moments — center of gravity, skew, and kurtosis — the Tsua alveolar and palatal click bursts do not markedly differ from each other. T-tests reveal that the differences are not statistically significant. Moreover, the click burst duration and intensity differences are not statistically significant either. In contrast, an analysis of the alveolar and palatal clicks in Juǀ’hoan, which does not exhibit click loss, shows that these clicks are significantly different on all the acoustic measures. These findings may help shed light on the diachronic process of click loss in Tsua.

Keywords: Khoisan, click consonants, spectral moments, acoustic phonetics, click loss, statistical analysis

1. Introduction

In their study of southern African click languages, Traill and Vossen (1997) showed that the phenomena of click replacement (a click replaced by another click) and click loss (a click replaced by a non-click) have implications for phonetic and phonological theory as well as our understanding of clicks historically. They compared spectra and waveforms of weakened and unweakened /ǃ/ and /ǂ/ from !Gana, !Gui and !Xóõ. Their analysis of acoustic and articulatory factors identified a more powerful click influx for the unweakened clicks in !Xóõ compared to the weakened clicks. The weakened clicks were the result of articulatory undershoot, according to their analysis. In the same year, Traill (1997:115) argued that clicks add “a salient acoustic signal” to speech and are an enhanced version of stops, a point he illustrated with the patterns of click replacement in Khoe languages.

In the current discussion, Tsua presents an example of the Eastern Kalahari Khoe languages where click replacement and click loss are relatively common (Vossen 2013, Fehn 2020a). Click loss tends to be found the most on the periphery of the Kalahari Basin linguistic area (Sands 2020), an area where Tsua is spoken. For the Khoe languages, there is a diachronic, click loss pattern in which alveolar clicks /ǃ/ become non-click velar consonants and palatal clicks /ǂ/ become non-click palatal consonants. These so-called abrupt clicks, /ǃ/ and /ǂ/, are commonly replaced or lost in Khoe.

* The authors wish to thank Chris Collins and Andy Chebanne for making the Tsua research possible. We would also like to thank the staff and management of Jack’s Camp, specifically Ralph Bousfield, Eugene Khumalo, and Chabba Selei. We are grateful to Richard Butler and Elizabeth DeGaetano for creating the opportunities for this work. This research would not have been possible without the insight, enthusiasm, and patience of the Tsua and Juǀ’hoan people. This paper has also benefited from comments by two anonymous reviewers.
with the alveolar click type being lost before the palatal click type (see Fehn 2020a for an overview). The results of a preliminary acoustic analysis of these clicks are reported here in comparison to a dataset from Juǀ’hoan to further our understanding of click loss in Khoe.

Table 1: Tsua click consonant inventory \((n = 34)\), based on Mathes and Chebanne (2013) and Mathes (2015, 2020).

<table>
<thead>
<tr>
<th></th>
<th>Dental</th>
<th>Palatal</th>
<th>Alveolar</th>
<th>Lateral</th>
</tr>
</thead>
<tbody>
<tr>
<td>voiceless</td>
<td>l</td>
<td>ǂ</td>
<td>!</td>
<td>l</td>
</tr>
<tr>
<td>voiced</td>
<td>g</td>
<td>l</td>
<td>g</td>
<td>l</td>
</tr>
<tr>
<td>aspirated</td>
<td>l’h</td>
<td>ǂ’h</td>
<td>l’h</td>
<td></td>
</tr>
<tr>
<td>ejective</td>
<td>l’h</td>
<td>ǂ’h</td>
<td>l’h</td>
<td></td>
</tr>
<tr>
<td>glottalized</td>
<td>l’h</td>
<td>ǂ’h</td>
<td>l’h</td>
<td></td>
</tr>
<tr>
<td>clusters</td>
<td>l</td>
<td>x</td>
<td>!q</td>
<td>x’</td>
</tr>
<tr>
<td></td>
<td>l</td>
<td>q</td>
<td>!q</td>
<td>l</td>
</tr>
<tr>
<td></td>
<td>l</td>
<td>g</td>
<td>ǂ</td>
<td>g</td>
</tr>
<tr>
<td></td>
<td>l</td>
<td>q’h</td>
<td>ǂ</td>
<td>q’h</td>
</tr>
<tr>
<td>nasal</td>
<td>Ĺ</td>
<td>l</td>
<td>Ĺ</td>
<td>l</td>
</tr>
</tbody>
</table>

2. Background

Tsua is classified as an Eastern Kalahari Khoe language in the Tshwa subgroup (Güldemann and Vossen 2000; Güldemann 2006). The Eastern Kalahari Khoe languages are spoken in the eastern part of Botswana, southeast of the Central Kalahari Game Reserve. Tsua has not been studied extensively, as the Tshwa subgroup languages have not attracted the same interest as other languages (Andersson and Janson 1997). Moreover, the Kalahari Khoe East speech communities display signs of widespread language shift to Setswana and Ikalanga (Chebanne 2008, 2014).

We have opted to follow Ladefoged and Traill (1994) by using the terms “click type” and “accompaniment” to describe the Tsua and Juǀ’hoan clicks. Tsua and Juǀ’hoan have four click types symbolized as: dental /ǀ/, palatal /ǂ/, alveolar /ǃ/, and lateral /‖/, with voicing and nasality transcribed before the click symbol. For instance, /gǀ/ represents the voiced dental click and /ŋǀ/ is the nasalized dental click. A ‘k’ before the click symbol to indicate voicelessness is not used in the transcriptions to avoid redundancy.
The Tsua click consonant inventory is presented in Table 1. The cluster analysis for some of the clicks in Table 1 follows Nakagawa (2006) and references therein.1 The click consonant inventory gaps in Table 1 reflect the extent of Tsua click loss. Some of the alveolar and palatal clicks as well as certain accompaniments are superseded by non-clicks. The unattested alveolar and palatal clicks in Tsua are: /ɡǃ, ǃh, ǂ', ǃ', ǂχ, ǃχ, ǂqχ', ǂɢ, ǃɢ, and ǃqχ/.

Jul’hoan — also referred to as Ju’hoansi, Zu’l’hōasi, !Xu, among others (see Lewis et al. 2015 for an overview) — is one of the best-described click languages, with a grammar and dictionary by

---

1 Bennett (2020) argues that cluster analyses of clicks are problematic in part because of their inability to predict which clicks a language has, based on its non-click inventory and vice-versa. We have limited our study to Tsua palatal and alveolar clicks that do not form a cluster and leave the unit vs. cluster analysis to ongoing work on Tsua phonology.
Snyman (1970, 1975), expanded dictionaries and a teaching grammar by Dickens (1994, 2005, 2009), and recent, extensive phonetic and phonological studies by Miller-Ockhuizen (e.g., 2003, 2013). It is part of the Kx’a family following a proposal in Heine and Honken (2010) and is spoken mainly in northwestern Botswana and Tsumkwe in Namibia (Chebanne 2008).

The Juǀ’hoan click consonant inventory is presented in Table 2. Each click type has the full set of accompaniments except the voiced aspirated lateral click (see Snyman 1975: 88, 125). The interested reader is referred to Miller-Ockhuizen (2003, 2013) for a more detailed phonetic description and analysis. For our purposes here, there is no evidence to suggest that click replacement or click loss has occurred in Juǀ’hoan. However, it should be noted that Juǀ’hoan ! is a reflex of the Proto-Ju retroflex click *‼!, which is an instance of click replacement (Fehn 2020b).

3. Click Replacement and Click Loss

The earliest unpublished vocabulary of an eastern dialect of the Khoe languages was collected by the explorer David Livingstone as he traveled through the Ngwato territory via Serowe and north of the Makarikari Pan [Makadigadi Salt Pans] around 1851 (Maingard 1961). A few of the relevant examples demonstrate the language under investigation by Livingstone had lost clicks in a pattern typical of Khoe languages (Traill and Vossen 1997). Moreover, Dornan (1911) supplies an abbreviated description of the language spoken by the Hiechware in eastern Botswana, with examples showing evidence of click loss.

There is evidence of click stability in other languages, e.g., Nama, Juǀ’hoan, and !Xóô (Beach 1938, Snyman 1975, and Traill 1985, respectively). It was not until Traill (1986) that the systematic nature of click replacement and click loss was carefully studied. He found an implicational relationship in which alveolar and palatal clicks were most often lost to their velar and palatal non-click oral or nasal stop counterparts, with loss of the palatal clicks implying loss of the alveolar clicks. Indeed, there are Khoe varieties that lose at least some /ǃ/ clicks while preserving all /ǂ/ clicks, or varieties that lose both, but never a language that has lost /ǂ/ while keeping /ǃ/. Citing Dressler (1972), Traill (1986) agrees with the observation pointing to language attrition as a factor in click replacement and/or click loss, most often occurring in communities facing economic or social challenges. Contrarily, Sands et al. (2007) found little evidence of lexical attrition despite a language shift from Nǀu to Afrikaans by its few remaining speakers. It could be the case that shift and attrition are distinct processes, although they often co-occur.

Traill (1986) observes that certain Khoe languages but not all lose clicks to non-clicks, and when the phenomenon occurs, /ǂ/ and /ǃ/ clicks may be lost to cognate palatal and velar stops, respectively. Further details of all the click loss patterns can be found in Traill and Vossen (1997) and Fehn (2020a). Nama, Naro, and Giui are conservative in that they preserve all their clicks as phonetically and phonologically stable elements at one end of the spectrum, but Shua, Teti, and Tyire are non-conservative varieties that lose both alveolar and palatal clicks at the other end of the spectrum. Click loss operates on a continuum in terms of the degree to which a language has undergone a particular change, i.e., the degree to which it has spread through the lexicon. In Tsua, this sound change has not completely spread through the lexicon.

4. Related Work

The acoustic differences between alveolar and palatal clicks have been studied for several click languages. Ladefoged and Traill (1994) report the mean duration of alveolar click bursts to be 12
ms, almost twice as long as the 6.6 ms for palatal click bursts in !Xóõ. In addition, they found that clicks can be generally divided into three classes depending on the distribution of energy within the auditory spectrum: palatal clicks have more energy above 14 Bark (about 2.5 kHz) while alveolar clicks have more energy below this frequency. This confirms that alveolar and palatal clicks have been measurably distinct using the duration of the click bursts and the distribution of energy in the auditory spectrum.

Fulop et al. (2003) show that two measures of duration for the closure and release phases of clicks in Yeyi can distinguish all four click articulations. This was possible to a lesser extent using the anterior burst power spectra, “as parameterized using the first four spectral moments” where interspeaker variation affected the anterior burst spectra. In Miller and Shah’s (2009) study of the acoustics of Mangetti Dune !Xung clicks it is shown that burst duration, amplitude and rise time are correlated. Here the interspeaker variation is found in the acoustics of the palatal click that is expected to be correlated to the anterior constriction release dynamics. In an earlier study, Miller, Brugman and Sands (2007) provide data on COG and resonances of two central peaks in acoustic spectra of Nǀuu clicks. They found that these values do not differ for the velar and uvular click articulations, the main topic of their study. These studies suggest that other measures to distinguish alveolar and palatal clicks show either more interspeaker variation or have not been applied specifically to the distinction of these two clicks.

Exter (2011) focuses on modeling specifically the contrasting alveolar and palatal click types. He argues that there are fundamentally different overall tongue configurations between these two types of clicks with a generally higher posture and shallower click cavity for the palatal click compared to the alveolar click type. The tongue shapes influence the source spectra as well as the spectral properties of these clicks. Exter does not mention interspeaker variation but provides a model that transcends language differences. He tested his model with Nǀuu data that largely fit his ideas.

The focus of the current study is on the targets of click loss in Tsua, i.e., the extant alveolar and palatal clicks. Several acoustic-phonetic properties of these clicks are investigated. Additionally, the characteristics of alveolar and palatal clicks are reported for the click language Jul’hoan as an acoustic point of reference, as Jul’hoan does not exhibit click loss.

5. Methods

Recordings of Tsua words, phrases, and sentences were collected in Botswana from 2012–2013. Three consultants from the village of Moralane in eastern Botswana provided multiple utterances of selected words in isolation. The consultants were 3 elderly women who spoke each word a minimum of three times. They were recorded with a Zoom H4n digital recorder set to a sampling rate of 44.1 kHz at 16-bit resolution.

Recordings of Jul’hoan phrases were collected in Botswana in 2014. Four consultants from the village of !Xaeua near the Botswana-Namibian border provided multiple utterances of selected words within sentential contexts. The consultants include a senior woman, a young woman, a senior man, and a young man. Each utterance was collected a minimum of three times and recorded with a “shotgun” hyper-cardioid condenser microphone using a digital recorder set to 44.1 kHz, 16 bits.

Spectral moments have been used to quantify spectral shape features in speech analysis (Jakobson et al. 1963; Forrest et al. 1988; Johnson 1993). The current analysis incorporates COG, skew, and kurtosis to further understand clicks that are the target of click loss in Tsua. COG is a
spectrum’s weighted mean: if the spectrum has more high-frequency energy than low, the mean will be high, and vice versa if a spectrum has more low-frequency energy. Skew is a measure of the spectral shape above the COG versus below the COG. For example, if there is more energy above the COG than below, the spectrum has a positive skew. Kurtosis is a measure of how the spectral shape around the COG is different from a Gaussian shape, i.e., the “peakedness” of the spectrum.

To compare the three spectral moments properties of Tsua alveolar and palatal clicks, the spectra of the combined anterior + posterior click bursts were analyzed. Three repetitions per speaker of the lexical items produced in isolation in (1a)–(1b) were segmented by burst onset and offset, with the moments calculated in Praat 5.3.7 (see Boersma and Weenink 2014 for details on spectral moments calculation). Juǀ’hoan alveolar and palatal clicks were compared to each other as a point of reference (1c)–(1d). There were three Tsua speakers and four Juǀ’hoan speakers as mentioned earlier. It was not possible to completely control for segmental and syllabic structure since the data were not elicited with a spectral study in mind. Figure 1 provides waveforms and spectrograms of Tsua !ari ‘yellow’ and ǂare ‘turn and look’ for a visual comparison of the Tsua alveolar and palatal clicks.

Figure 1: Waveforms and spectrograms of Tsua !ari ‘yellow’ (top) and ǂare ‘turn and look’ (bottom).
(1) a. **Tsua alveolar clicks**
   Ṕao ‘tall; long’
   Ṕari ‘yellow’
   Ṕʔuu ‘white’

b. **Tsua palatal clicks**
   Ṕea ‘valley’
   Ṕim ‘tobacco roll; a load wrapped in hide’
   Ṕare ‘turn and look’

c. **Juǀ’hoan alveolar clicks**
   Ṕae ‘hunt’
   Ṕoo ‘older brother’
   Ṕui ‘older sister’

d. **Juǀ’hoan palatal clicks**
   Ṕae ‘look down at’
   Ṕ’in ‘think’
   Ṕ’ari ‘forget’

Figure 2: Spectral moments plot of Tsua palatal (Tǂ), Tsua alveolar (Tǃ), Juǀ’hoan palatal (Jǂ) and Juǀ’hoan alveolar (Jǃ) click bursts. Center of gravity (COG), skew and kurtosis means are compared in this acoustic spectra 3-D plot.

6. **Results**

The spectral moments’ means are plotted by language and click type in Figure 2. Interestingly, the Tsua palatal (Tǂ) and alveolar (Tǃ) click burst means cluster together in terms of COG, skew and kurtosis. Welch’s t-tests reveal that the Tsua alveolar and palatal click burst differences are not statistically significant. Tǂ COG ($M = 6437$ Hz, $SD = 1036$) is not significantly different from Tǃ
COG ($M = 5853$ Hz, $SD = 2035$), $t(39) = -1.33$, $p = .192$, n.s. Moreover, Tǂ skew ($M = .35$, $SD = .66$) is not significantly different from Tǃ skew ($M = .44$, $SD = 1.11$), $t(42) = 0.338$, $p = .737$. Tǂ kurtosis ($M = .13$, $SD = 1.17$) is not significantly different from Tǃ kurtosis ($M = .15$, $SD = 2.88$), $t(34) = 0.038$, $p = .970$.

The Juǀ’hoan palatal (Jǂ) and alveolar (Jǃ) click bursts diverge spectrally from each other in the plot. Jǂ COG ($M = 5001$ Hz, $SD = 1173$) has a significantly higher mean compared to Jǃ COG ($M = 1376$ Hz, $SD = 623$), $t(51) = -16.192$, $p < .001$. Jǂ skew ($M = 1.1$, $SD = .96$) has a significantly lower mean compared to Jǃ skew ($M = 6.5$, $SD = 2.93$), $t(43) = 10.471$, $p < .001$. Jǂ kurtosis ($M = 4.6$, $SD = 5$) has a significantly lower mean compared to Jǃ kurtosis ($M = 63.4$, $SD = 54$), $t(36) = 6.498$, $p < .001$.

Two further acoustic properties, click burst duration and intensity, were measured to check for potential differences (Figure 3a-b). For Tsua, Tǂ duration ($M = 12$ ms, $SD = 7$) is not significantly different from Tǃ duration ($M = 15$ ms, $SD = 6$), $t(51) = 1.44$, $p = .156$. In addition, Tǂ intensity ($M = 69$ dB, $SD = 7.93$) is not significantly different from Tǃ intensity ($M = 72$ dB, $SD = 6.54$), $t(50) = 1.45$, $p = .153$. Regarding Juǀ’hoan, Jǂ duration ($M = 21$ ms, $SD = 8$) is significantly longer than Jǃ duration ($M = 14$ ms, $SD = 4$), $t(48) = -4.4$, $p < .001$. The longest duration and largest degree of variability is found for the Juǀ’hoan palatal click. Jǂ intensity ($M = 75$ dB, $SD = 4.39$) is significantly lower than Jǃ intensity ($M = 84$ dB, $SD = 3.34$), $t(63) = 9.74$, $p < .001$. In sum, for the speakers in this study, the Tsua alveolar and palatal click bursts are not significantly different from each other on the five acoustic dimensions. The Juǀ’hoan alveolar and palatal click bursts are significantly different from each other on all these measures.

7. Discussion

This article discussed a preliminary acoustic study comparing Tsua and Juǀ’hoan’s alveolar and palatal clicks on COG, skew, kurtosis, duration, and intensity. For Tsua, these acoustic measures appear to converge, assuming that they were distinct historically. Figure 2 suggests that the direction of the purported convergence is from the alveolar click to the palatal click. For Juǀ’hoan, these measures diverge significantly. It could be that analyzing the Tsua anterior and posterior bursts separately may yield a statistically significant result, but we are not aware of any study that provided evidence that speakers could distinguish clicks based solely on the anterior or the posterior bursts.
The marginally higher COG for palatal click bursts compared to alveolar click bursts in Tsua trend in the same direction as the results of previous studies. COG is an indication of the cavity sizes during the production of these clicks (Miller et al. 2009; see also Miller 2015 for a method estimating lingual cavity volume in clicks), although it does not indicate the cavity’s place of articulation. The similar COGs suggest similar cavity sizes for the Tsua clicks, which could be confirmed using high frame rate ultrasound (see Miller 2014). Moreover, the longer duration for alveolar click bursts versus palatal click bursts in Tsua has been reported in !Xóõ (Ladefoged and Traill 1994) and Mangetti Dune !Xung (Miller and Shah 2009). The higher intensity for Tsua alveolar click bursts compared to palatal click burst supports the intensity scale in Traill (1997) for !Xóõ.

While the click-burst acoustic properties studied here trend in the same direction as previous studies, the differences are not statistically significant in Tsua as opposed to Ju’hoan. However, the statistical results may not tell the entire story. The clicks may not be significantly different on the measures reported here but they could still be perceptually distinguishable. It could be the case that one or more of the acoustic dimensions are above the Just Noticeable Difference (JND) threshold, still helping speakers distinguish the clicks. In addition, other cues may play a more prominent role, such as the click accompaniment characteristics, Voice Onset Time (VOT), and/or aspects of the following vowel such as the formant transitions. In a study of West !Xoon, Grawunder and Naumann’s (2008) findings suggest a 3-way distinction of alveolar and palatal clicks with noisy accompaniments. Moreover, Miller et al. (2009) state: “Clear distinctive co-articulatory effects on the following vowel are seen for each click type...” in Mangetti Dune !Xung.

A potential area of future research could be to elicit speech data from many more Tsua and Ju’hoan speakers with carefully controlled segmental and syllabic structure for further analysis since there are limitations to the speech data examined here. Another area of future research could be to conduct a perceptual study to determine which cues play the most important role in differentiating the Tsua clicks.

Whether other Khoe languages with click loss display similar acoustic properties remains an open question. A large-scale study is needed to determine if the results reported here extend to other Khoe languages. What can be stated is that the Tsua alveolar and palatal click bursts converge concerning the acoustic dimensions under investigation. It may be that this is one aspect of the click loss process in Tsua. It is not apparent at which point the convergence commenced, its perceptual role, or what the phonological outcome will be for the existing alveolar and palatal clicks.

References


Beach, Douglas M. 1938. The phonetics of the Hottentot language. Cambridge: W. Heffer & Sons LTD.


Boersma, Paul & David Weenink. 2014. Praat: doing phonetics by computer, version 5.3.7. URL: http://www.praat.org/, Developed at the Institute of Phonetic Sciences, University of Amsterdam.


Tsua and Ju’hoan alveolar and palatal click consonants


Timothy K. Mathes
Audio Processing Group
Assured Information Security, Inc.
mathest@ainfosec.com

Alex de Voogt
Drew University
Madison, NJ
adevoogt@drew.edu