VOT and F_0 in Zulu Dental Clicks and Alveolar Plosives

Kasia Garbowska
International Christian University
University of Alberta

Abstract

The present study investigated the contrast in Voice Onset Time (VOT) and Fundamental Frequency (F0) between different varieties of dental clicks, and alveolar plosives. The study attempts to clarify the commonalities and dissimilarities within the two through a set of recordings with a language consultant, a native speaker of Zulu. The data findings generally held little discrepancies within current data, falling into line with Strazny's report on tonal depression in Zulu (2003), and Hanson's voiceless and voiced dichotomy in F0 at vowel onset (2009). However, there was a certain amount of deviation from Midtlyng's report on effects of Speech Rate and Place of Articulation on VOT data (2011).

Keywords: Dental clicks, alveolar plosives, Zulu, voice onset time, fundamental frequency, articulatory phonetics.

1. Introduction

This paper reports on the differences between the varying types of dental clicks and voiced and voiceless alveolar plosives used specifically in the Zulu language. A member of the Bantu Languages, specifically the Nguni linguistic family, Zulu is an official language of South Africa (Herbert 1990, p.296). Clicks are speech sounds that occur as consonants, used in many different languages in South Africa as well as a few in East Africa. (Maddieson and Ladefoged 1995, p. 246-247). The 3 main types of clicks used in the Zulu language include the lateral click, the alveolar click, and the dental click. The report will focus specifically on dental clicks, which were further broken down according to articulation type. The variations in articulation of dental clicks include Plain [|], Aspirated [|h], Nasal [n|], Depressor [g|ɦ], and Depressor Nasal [ŋ|ɦ]. The variations of the alveolar plosives that will be focused on in the paper include the Voiceless Ejective Alveolar Plosive [t'], Aspirated Voiceless Alveolar Plosive [th], Depressor Nasal [nd] (contrasted with the nasal), Ejective Alveolar Plosive [t'], and the Voiced Alveolar Plosive [d] (See tables in section 3.5). The purpose of the report is to identify the differences between dental clicks and alveolar plosives in terms of VOT and F0.

The following section summarizes background information regarding the articulatory characteristics of clicks and alveolar plosives. It will also present information on the usage of clicks in languages other than the Zulu Language. The third section will report on the data findings collected from the consultant, a native speaker of the Zulu language. The final section will discuss the findings in greater depth, summarizing the findings on the basis of commonalities and dissimilarities between the differing variations of dental clicks and voiced and voiceless alveolar plosives in relation to VOT and F0.

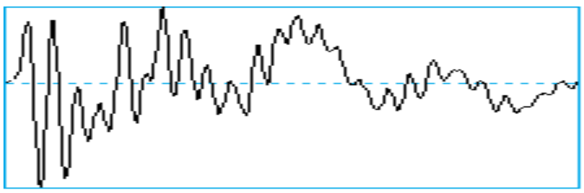
2. Literature Review

2.1 *Occurrence of Dental Clicks across Different Languages* Within South Africa's Bantu Languages it is within the Nguni subgroup that holds several major languages that clicks are more widely used. The 2 major languages Zulu and the neighbouring Xhosa in particular are known to exhibit a three-way opposition: lateral clicks, palatal clicks and dental clicks. It is estimated that approximately 15% of the Zulu and Xhosa language feature clicks. Other Bantu Languages feature either both the dental click and the palatal click, or only the palatal click (Herbert 1990, p.296). In addition to Bantu Languages, Dental Clicks are also crucial in Khoisan languages such as Sandawe or Hadza. The Cushitic language Dahalo also uses dental clicks, though minimal, with four different nasalized variations (Maddieson and

Ladefoged 1995, p. 249). The dental click has also been used para-linguistically across a variety of languages, used to convey a sense of emotion, or nuance through pitch, intonation, volume, etc. Specifically, English uses a form of the dental click in order to express annoyance, or disapproval (Laver 1994, p. 175).

2.2 *Articulatory and Acoustic Features of Alveolar Plosives* According to Nafis K, a plosive is a consonant sound that is produced through a stop in the airflow of the vocal tract (2014, p.1-2). Furthermore, there are three stages involved in the process of forming a plosive. Starting from the closing stage, two articulators make contact, closing the air passage completely and raising the soft palate. The vocal cords would vibrate in the case of a voiced plosive. Moving on to the compression stage, the flow of air is briefly halted, allowing air pressure to build up behind the closure. Lastly, during the release stage the speech organs separate abruptly, causing the built up air to escape rapidly from the oral cavity, producing a plosive, as seen below in Fig. 1.

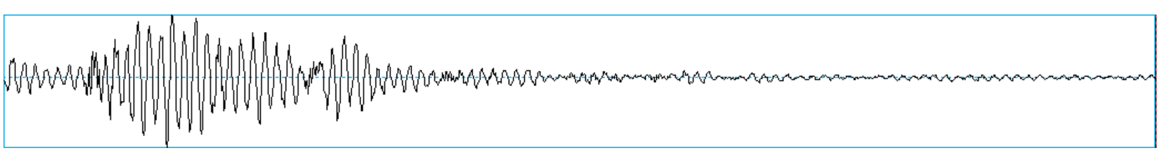
Figure 1: A wave form for a Zulu Alveolar Plosive in the word “Duck”, [ídàdà].



2.3 *Articulatory and Acoustic Features of Dental Clicks* Similar to the production of stops and affricates, in order to produce a dental click, there is a certain process that occurs in the oral cavity: a movement of air that is often referred to as the velaric air-stream mechanism (Poser, 2004). Being central consonants, movement of air occurs across the center of the tongue, rather than on the side. The process is always lingual ingressive (also known as velaric ingressive), referring to the rarefaction of an enclosed pocket of air between two points of closures. One of the closures is in front of the oral cavity, with the tip of the tongue or the lips. The other is deeper in the oral cavity, with the back of the tongue rising to form a stop at either the velar or the uvular point of articulation. The process is an essential component in the articulation of clicks, with the release of the front closure allowing air to rush into the mouth (Maddieson and Ladefoged 1995, p. 246-247).

Similar to stops, clicks hold two different acoustic components that appear as a crescendo and a decrescendo. First it is the transient, which appears as a high-amplitude, short-duration sound that takes place when the articulators come apart. The second component involves a noise that occurs as a result of a turbulent flow of air between the articulators. The transient occurs as a result of an abrupt rate of change in the vocal tract shape (Maddieson and Ladefoged 1995, p. 257-259).

Figure 2: A wave form for the Zulu Depressor Dental Click in the word “Anointing”, [úǀ̬ɔ̀wɔ̀].



As seen above in Fig.2, the two components of the dental click are present, with the noise being caused by the flow of air between the articulators being particularly evident. There is a sharp onset in the waveform, but it is followed by a noisy, sustained speech sound.

3. Data

3.1 *Stimulus Design* The materials recorded and analyzed in Praat included 3 charts of data, with each chart containing 4 phonemes of varying types of articulation so as to efficiently contrast dental clicks and alveolar plosives in terms of VOT and F0 (See Tables 1, 2 and 3). The 4 phonemes included 2 clicks and 2 plosives, each articulation having 3 words recorded and analyzed; leaving each chart with a set of 12 different words each.

3.2 *Recording* The recording of the final set of data took place on June 8th 2016 within the sound-proof room, ILC 435-A, at International Christian University. The stimuli was presented both visually and aurally in English, repeated twice in isolation. The participant's speech was recorded using the head-worn microphone, Shure WH30-XLR, connected to the solidstate digital recorder, Marantz PMD-661. The microphone was placed at approximately a 45-degree angle, 10 cm away from the corner of the speaker's mouth, so as to prevent turbulence from the airflow from interfering with the recording.

3.3 *Participant* The author collected data by recording and analyzing the voice of a female language consultant who was made available through the class LNG 397. The consultant is a bilingual speaker, who can speak and understand the language of Zulu and English at a native level. Note that the consultant has no history of a hearing or speaking impediment.

3.4 *Analysis* The analysis of the recordings was made using the Microsoft Windows Version 5.4.04 version of Praat. Dental clicks and Plosives were manually segmented and annotated, identifying the VOT of the set of words in Table 1, and the F0 for the two sets of words in Table 2 and Table 3. The average VOT in seconds and F0 in Hz for every type of articulation were also identified. The measurement of VOT was then changed into milliseconds by multiplying by 1000. All measurements were rounded up to one digit below the decimal point, the average being calculated manually.

In order to measure VOT, the initial burst was first identified, and a boundary line was designated. Next, the initial point of periodicity or voicing was found, and another boundary line was set. The time between the first burst and the start of voicing or periodicity is what is known as the VOT; written beside "Visible Part" on the bottom on the screen, appearing when the area between the two boundaries was selected.

In order to find the F0 first "Show Pitch" under the Pitch tab was selected, causing a blue line to appear on the spectrogram. Second, two boundary lines were designated to mark the vowel being analyzed. Next, along the blue line, a blue dot marked in the approximate midpoint of the vowel was selected. The F0 value was then noted down, as written on the right side of the spectrogram in Hertz.

3.5 *Results*

3.5.1 *Average VOT of Oral Clicks and Plosives* As seen below in Table 1, the first set of words contains the Plain Dental Click [|], which held a VOT average of 38.1ms. The second was the Aspirated Dental Click [|h], which had a VOT average of 139.0ms. The third and fourth included the Voiceless Ejective Alveolar Plosive [t'] with a VOT average of 45.9ms, as well as the Aspirated Voiceless Alveolar Plosive [th] which held a VOT average of 113.8ms.

Table 1: Dental Click vs Plosive

Table 1: Dental Click vs Plosive	**Meaning**	**IPA**	**Voice Onset Time [VOT] (ms)**
Plain Dental Click [ǀ]	A song	ìǀùlò	42.3
Plain Dental Click [ǀ]	A case	ìǀàlà	35.9
Plain Dental Click [ǀ]	A shoe	ìsìǀàtùlɔ̀	36.0
Aspirated Dental Click [ǀʰ]	To fascinate/explain	úɠúǀʰàzà	141.3
Aspirated Dental Click [ǀʰ]	To be bias	úɠúǀʰɛ̀mà	132.8
Aspirated Dental Click [ǀʰ]	To urinate	úgúǀʰàmà	143.3
Voiceless Ejective Alveolar Plosive [t']	A cat	ìk̬àt'ì	51.7

Voiceless Ejective Alveolar Plosive [t’]	A thing	íⁿt’ɔ̀	37.0
Voiceless Ejective Alveolar Plosive [t’]	Sunday	ísɔ̀ⁿt’ɔ̀	49.0
Aspirated Voiceless Alveolar Plosive [t^h]	Take!	tʰátʰà	109.0
Aspirated Voiceless Alveolar Plosive [t^h]	Send!	tʰúmà	114.5
Aspirated Voiceless Alveolar Plosive [t^h]	Be faithful	tʰɛ̀mbɛ́k̬à	118.0

Contrasting the VOT of the Plain Dental Click [|] with the Aspirated Dental Click [|h], a considerable difference of 100.9ms was found, with the Aspirated Dental Click holding the larger value of the two. Between the Voiceless Ejective Alveolar Plosive [t’] and the Aspirated Voiceless Alveolar Plosive [th] it is the Aspirated Plosive which holds a greater VOT time, showing a difference of 61.2ms. Comparing the VOT of the Plain Dental Click [|] with the Voiceless Ejective Alveolar Plosive [t’], the Alveolar Plosive is found to have a slightly higher value of VOT than the Plain Dental Click, with a difference of 7.8ms. Between the Aspirated Dental Click [|h] and the Aspirated Voiceless Alveolar Plosive [th], it is the Aspirated Dental Click which had a greater VOT, with a difference of 25.2ms.

3.5.2 *Average F0 of Nasal Clicks and Plosives*

Table 2 reports F0 results. The Nasal Dental Click [n|] held an average of 160.8Hz. The Depressor Nasal Dental Click [ᵑ|ʱ] had a F0 average of 154.8Hz. The last two included the Nasal [n] whose F0 sat at an average of 148.6Hz, and the Depressor Nasal [nd] which had an F0 value of 114.9Hz.

Table 2: Dental Click vs Plosive

Table 2: Dental Click vs Plosive	**Meaning**	**IPA**	**Fundamental Frequency - F^0 (Hz)**
Nasal Dental Click [^{n}|]	A letter	íⁿ|ʷàdì	152.0
Nasal Dental Click [^{n}|]	Sour porridge	íⁿ|ʷàⁿ|ʷà	168.5
Nasal Dental Click [^{n}|]	Grace	íⁿ|ʷɛ̀bà	161.8
Depressor Nasal Dental Click [ᵑ|ʱ]	A bit	íᵑ|ʱɔ̀sì	154.6
Depressor Nasal Dental Click [ᵑ|ʱ]	riches	íᵑ|ʱɛ̀bɔ̀	163.1
Depressor Nasal Dental Click [ᵑ|ʱ]	He/she is better	úᵑ|ʱɔ̀nɔ̀	146.6
Nasal [n]	(Traditional) Dance!	Sínà	142.6
Nasal [n]	I	Mìná	156.6
Nasal [n]	Seek!	Fúnà	146.5
Depressor Nasal [^{n}d]	Man	ⁿdɔ̀dà	116.2
Depressor Nasal [^{n}d]	A stick	ìⁿdùkù	112.8
Depressor Nasal [^{n}d]	The head	íkʰàⁿdà	115.6

Contrasting the F0 of the Nasal Dental Click [^{n}|] with the Depressor Nasal Dental Click [ᵑ|ʱ] shows that the Nasal Dental Click had a slightly higher F0, with a difference of 6.0Hz. Between the Nasal [n] and the Depressor Nasal [^{n}d], the Nasal had a markedly larger F0 value, with a difference of 33.7Hz. Comparing the F0 of the Nasal Dental Click [^{n}|] and the Nasal [n], the Nasal Dental Click was found to have a larger F0 value, with a difference of 12.2Hz. Between the Depressor Nasal Dental Click [ᵑ|ʱ] and the Depressor Nasal [^{n}d], it was the Depressor Nasal Dental Click which had a noticeably larger F0 value, with a difference of 39.9Hz.

3.5.3 *Average F0 of Depressors and Non-Depressors* The last results in table 3, start with the Plain Dental Click [|], with an average F0 value listed as 163.7Hz. Next, is the Depressor Dental Click [^{g}|ʱ] with an F0 average of 154.2 Hz, followed by the third set of words that contain the Voiceless Alveolar Plosive [t] holding an F0 average of 140.7Hz. The final includes the Voiced Alveolar Plosive [d], illustrating an F0 average of 111.0 Hz.

Table 3: Dental Click vs Plosive

Table 3: Dental Click vs Plosive	**Meaning**	**IPA**	**Fundamental Frequency - F^{0} [Hz]**
Plain Dental Click [|]	A song	ì|ùlò	162.9
Plain Dental Click [|]	A case	ì|àlà	158.7
Plain Dental Click [|]	A shoe	ìsì|àtùlɔ̀	169.6
Depressor Dental Click [^{g}|ʱ]	To keep	ùk̬ù|̬ʷìnà	163.7
Depressor Dental Click [^{g}|ʱ]	Tribunal	ìsì|̬àwò	160.3
Depressor Dental Click [^{g}|ʱ]	Anointing	ú|̬ɔ̀wɔ̀	138.8
Ejective Alveolar Plosive [t']	A cat	ìk̬àt'ì	137.6
Ejective Alveolar Plosive [t']	A thing	í^{n}t'ɔ̀	144.1
Ejective Alveolar Plosive [t']	Sunday	ísɔ̀^{n}tɔ̀	140.4
Voiced Alveolar Plosive [d]	Lay the table!	dɛ̀kà	123.6
Voiced Alveolar Plosive [d]	A duck	Ídàdà	102.9
Voiced Alveolar Plosive [d]	Long ago	k'ádɛ̀	106.5

In contrasting the Plain Dental Click [|] with the Depressor Dental Click [^{g}|ʱ], it was found that the Plain Dental Click had a larger F0 value, with a difference of 9.5Hz. Between the Ejective Alveolar Plosive [t'], and the Voiced Alveolar Plosive [d], it was the Ejective Alveolar Plosive which held the higher F0, with a difference of 29.7Hz. Comparing the Plain Dental Click [|] with the Ejective Alveolar Plosive [t'], it was found that the Plain Dental Click had a larger F0 value, with a difference of 23.0Hz. Between the Depressor Dental Click [^{g}|ʱ] and the Voiced Alveolar Plosive [d], it was the Depressor Dental Click which held a greater F0 value, with a difference of 43.2Hz.

4. Discussion

There are numerous factors that contribute to the VOT of a speech sound: specifically, the Place of Articulation (POA) and the rate of speech (Midtlyng 2011, p. 111). Midtlyng further states that as the POA moves farther back into the mouth, the VOT tends to increase, and as speech rate increases the VOT tends to decrease.

The Plain Dental Click contrasts greatly with the Ejective Alveolar Plosive as the source of air pressure used to produce them, the airstream mechanism, differ; the first being a velaric ingressive sound, and the latter as a glottalic egressive sound. Regarding the VOT time however, the difference is lesser at 7.80ms. The negligible difference in VOT is interesting as the plosive in this case is alveolar, the Place of Articulation being farther back in the mouth than the click; a dental. The POA is not appearing to cause as great a difference in VOT as it had in Midtlyng's data regarding isolated stimuli, where the VOT appeared jumped on average 15.57ms between values (2011, p. 110). Midtlyng's statement is further brought into question as the Aspirated Dental Click was found to hold a greater VOT as opposed to the Aspirated Voiceless Alveolar Plosive. However, at 25.2ms, the difference is not significant.

Regarding Table 1, the reason for the difference in VOT between the [|] and the Aspirated Dental Click [|h] as well as the Ejective Alveolar Plosive [t] and the Aspirated Voiceless Alveolar Plosive [th] is due to the aspiration that is applied. Aspiration in itself is a puff of air, which may accompany a voiceless stop (Cleghorn and Rugg 2011, p. 80). Whether the voiceless-ness continues momentarily before the VOT after the release of the stop, or the VOT occurs immediately right at the point of release, is an indicator or whether there is aspiration present.

According to Strazny (2003, p.223), Zulu displays a curious relationship between tone and consonants, in which an explicit set of onset consonants has a direct correlation with a considerable lowering of an immediately following high (H) or low (L) tones. Further stating that this severe lowering frequently causes a repositioning of tones, often been called "tonal depression"; with the consonants initiating the change being referred to as "depressors".

Table 2 contrasts the effect that nasalization and depressor nasals have on dental clicks and the subsequent influence on the F0 value, versus that of a nasal, and the nasalized voiced alveolar plosive. As the comparison of the Nasal Dental Click and the Depressor Nasal Dental Click showed the VOT value to be nearly identical with a difference of 6Hz, it is assumed that the natural lowering of the pitch that typically occurs in the presence of a depressor was countered by the presence of the dental click, contrarily, raising the F0 even further. The presence of the depressor nasal versus the plain nasal however shows that the depressor has a great effect on the F0, lowering the pitch of the vowel to the immediate left of the alveolar nasal, creating a difference of 33.7Hz.

Regarding the Nasal and the Nasal Dental Click, a minimal 12.2Hz difference in F0 was found. However, between the Depressor Nasal Dental Click and the depressor nasal a F0 difference of 39.9Hz was found, with the depressor once again lowering the pitch in the word containing the alveolar nasal. The presence of the dental click was found to once again counteract the depressor's effect on pitch, in fact raising the value of F0.

According to Hanson, across numerous languages when a vowel follows an obstruent, in this case the alveolar plosive, the F0 in the first few tens of milliseconds of the vowel is influenced by the voicing features of the consonant (2009, p. 425). Furthermore, Hanson stated that the F0, specifically at vowel onset, is considerably higher following voiceless obstruents, than when following the voiced counterpart.

Table 3 contrasts the dental click and the depressor dental click, exhibiting a minimal difference of 9.5Hz, exhibiting a similarity between the two dental clicks, despite the presence of the depressor in the latter. Next, the ejective alveolar plosive was found to have a F0 that was 29.7Hz greater in value than the F0 of the Voiced Alveolar Plosive. It is likely because of the voiceless and voiced dichotomy in F0 at vowel onset that Hanson stated (2009, p. 425).

In regards to the Depressor Dental Click's greater F0 value, with a difference of 43.2Hz over the Voiced Alveolar Plosive, it is likely due to the two previously mentioned factors. First, the dental click and the depressor together appear to have a raising effect on the fundamental frequency. Furthermore, as a voiced obstruent, the vowel following the Alveolar Plosive likely experienced a lowering in F0.

5. CONCLUSION

The current report compared the differences between VOT and F0 within differing varieties of dental clicks, and alveolar plosives using novel data based on the author's fieldwork. The data generally held few discrepancies with current knowledge, falling into line with Strazny and Hanson's statements. However, there was a small amount of deviation from Midtlyng's data.

Nevertheless, the deviation may be explained, as the current report was not as detailed as it may have been, with only 3 words per type of click and alveolar plosive. Further study would be required to clarify the role the Place of Articulation may hold in relation to the VOT. Additionally, an exploration into the role of rate of speech in regards to VOT would be informative as well. Regarding the Fundamental Frequency, exactly how the dental click interacts with the depressor to create an upward trend in the F0 would also be a beneficial study.

An interesting question that appeared regarding Hanson's statements in relation to Zulu is how this effect would evolve within such a language, where obstruents with the same place of articulation become distinct in more dual ways, in phonation type. Namely, the presence of [t], [th] and [t'] in the Zulu lexicon.

6. References

Hanson, Helen M. "Effects of Obstruent Consonants on Fundamental Frequency at Vowel Onset in English." In *The Journal of the Acoustical Society of America J. Acoust. Soc. Am.* 125, no. 1, 2009, 425-41.

Herbert, Robert K. "The Sociohistory of Clicks in Southern Bantu." *Anthropological Linguistics*, vol. 32, no. 3/4, 1990, 295–315.

Ladefoged, Peter, and Ian Maddieson. "Clicks." In *The Sounds of the World's Languages*. Oxford, OX, UK: Blackwell Publishers, 1996.

Laver, John. "Initiation and Phonation." In *Principles of Phonetics*. Cambridge: Cambridge University Press, 1994.

Midtlyng, Patrick J. 2011. "The Effects of Speech Rate on VOT for Initial Plosives and Click Accompaniments in Zulu". In *Selected Proceedings of the 40th Annual Conference on African Linguistics*, ed. Eyamba G. Bokamba et al., 105-118

Nafis, Kamal. "Phonetics Features of Plosive." Accessed April 20, 2014.

Poser, B. "Airstream Mechanism." *Phonetics Handouts*, 2004, 1-2.

Strazny, Philipp. "Depression in Zulu." In *The Phonological Spectrum Volume I: Segmental Structure Current Issues in Linguistic Theory*, 2003, 223-39.