ACOUSTIC AND AERODYNAMIC FEATURES OF HADZA CLICKS

Didier Demolin[1], Alain Ghio[2] & Andrew Harvey[3]

[1]LPP, UMR-7018, CNRS, Sorbonne Nouvelle, Paris
[2]Laboratoire Parole et Langage, CNRS-UMR 7309, Aix-Marseille Université
3 University of Bayreuth Faculty of Languages and Literatures

ABSTRACT

Hadza has 4 clicks in its sound inventory: bilabial, dental, alveolar and lateral [ʘ, ǀ, !, ǁ] that can be aspirated, glottal and nasal [ʘ̰, ǀˀ, ǀʰ, ᴺǀ, ᴺǀˀ, !ˀ, !ʰ, ᴺ!, ᴺ!ˀ, ǁ, ǁʰ, ᴺǁʰ, ᴺǁˀ]. The bilabial click [ʘ] has a noisy release and is defined as [grave & noisy]. Acoustically the oral clicks [ǀ, !, ǁ] are described with 2 features [grave vs. acute] and [abrupt vs. noisy]. The [grave] feature defining the alveolar click [!] has a peak in the FFT spectrum around 2 kHz. The [acute] feature defining the dental and lateral clicks [ǀ, ǁ] has a peak around 6 kHz for the dental and an emphasis in frequencies between 4 & 5 kHz for the lateral click (all peaks in rounded values). The difference between the [abrupt] [!, ǁ] also has a duration component, i.e. short burst vs short noise band.

Keywords: Hadza, Clicks, Acoustic features, Aerodynamics.

1. INTRODUCTION

Hadza an isolate endangered language spoken in Tanzania has non-pulmonic consonants, clicks and ejectives, in its phonetic / phonological inventory. [1], [2] and [3] describe Hadza with 9 clicks while [4] suggests that there are 12. Recent data show that the language has 13: [ʘ̰, ǀˀ, ǀʰ, ᴺǀ, ᴺǀˀ, !ˀ, !ʰ, ᴺ!, ᴺ!ˀ, ǁ, ǁʰ, ᴺǁʰ, ᴺǁˀ]. Thus, there are 4 types of clicks with various accompaniments [k, g, q, ʰ, ˀ, ʱ, ᴺ, ᴺ̥] that can sometimes be combined.

The aim of this paper is to describe the acoustic features of Hadza clicks, evaluate the gestures involved with a measure of time and establish their distinctive features from acoustic and aerodynamic data. A precise articulatory and acoustic understanding of the mechanisms involved in their production and perception is necessary to characterize and formalize these consonants in terms of distinctive features. Another issue that will addressed is: are there aspirated clicks in Hadza? This implies to carefully establish the characteristics of the accompaniments for the different clicks. A more general issue involved in this study if to understand the mechanisms of the production of these segments and see how this type of data assesses, or extends, the limits of our knowledge on the diversity and functioning of speech production in languages.

2. MATERIAL AND METHOD

2.1 Speakers and corpus

Words including all Hadza consonants were recorded. Each word was repeated 3 times by 5 male and 4 female native speakers. For the present study (that only analyzes data from 2 females & 2 males) the following set of data has been used to compare the production of all the speakers. 1) [ʘ] Bilabial click [mʘ̰ʷa] 'Thank God'; 2) [ǀ] Dental click (glottal, aspirated, nasal, glottal nasal) : [ǀˀu] 'To end on an errand', [ǀalʰa] 'Threshing rock', [ᴺǀ̰ata] 'Tongue', [tuᴺǀ̰ˀe] 'To pluck feathers'; 3) [!] Alveolar click (simple, nasal, aspirated, glottal): [la!o] 'To trip', [kuᴺ!̰ana] 'Kudu', [!ʰaku] 'To jump', [!ˀoje] 'Beeswax', [!ˀakxˊitɬa] 'The hard palate'; 4) [ǁ] Lateral click (simple, aspirated, glottal nasal, nasal): [ǁupʰi] 'To sleep', [ǁʰa] 'To throw', [beᴺǁ̰ˀe] 'To shine', [ᴺǁ̰a:] 'For the whole camp to visit another one'. As the bilabial click has very limited occurrences, the main focus is set on the dental, alveolar and palatal clicks.

2.2 Method

Aerodynamic and acoustic data were recorded with an EVA2 Workstation [5] allowing simultaneous recordings of several aerodynamic parameters: intraoral pressure (Po) measured in hecto-Pascal (1hPa = 1.02cm H_2O), oral airflow and if necessary nasal airflow in dm^3/s. Audio recordings, synchronized with the aerodynamic data, are recorded with a microphone integrated to the EVA2 transducers. Oral Airflow (Oaf) was collected using a flexible silicone mask set on the subjects' mouth. Intraoral pressure (Po) was recorded by inserting a small tube into the pharynx through the nasal cavity.

3. RESULTS

3.1 Introduction

Clicks have acoustically two components: an attack transient and an extinction transient. The attack transient is the explosion noise (sometimes called a burst) which is the impulse response to a change in the shape of the vocal tract. The extinction transient is the noise associated with the turbulent release. Figure 1 shows typical audio waveform shapes for the 4 types of clicks found in our Hadza data. The attack and extinction transients are indicated by arrows on the alveolar audio waveform. Acoustic measurements consist in duration for all clicks, duration of the glottal and aspirated parts following the attack transient and a FFT spectrum made in a 10 ms Hamming window around the peak of the initial transient noise.

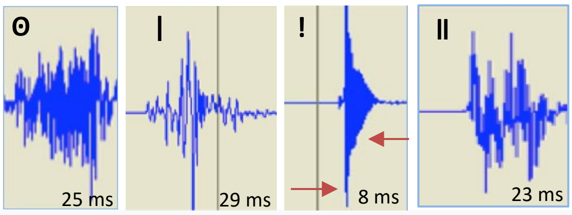

Figure 1. Audio waveform of the 4 clicks types.

3.2 Click duration

Figure 6 shows the comparison of duration, including the attack and extinction transients, between the dental, alveolar and lateral clicks made with 4 speakers. Results show that the alveolar is shorter compared to the dental and lateral clicks. The alveolar [!] is sharp compared to the dental [ǀ] that has a noisy or diffuse release and has a greater duration. This click has 2 phases: an initial rising until the peak of the release that is followed by a decreasing noise. The lateral click [ǁ] has a sharp release that lasts in intensity for an average of 23 ms before a sharp fall. This click is therefore produced as a noise band. These features are also displayed on the spectrograms of Figure 2, 3, 4 & 5.

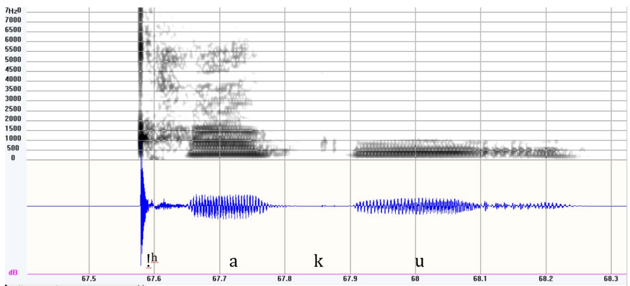

Figure 2. Audio waveform and spectrogram of the word [!ʰaku] 'to jump'

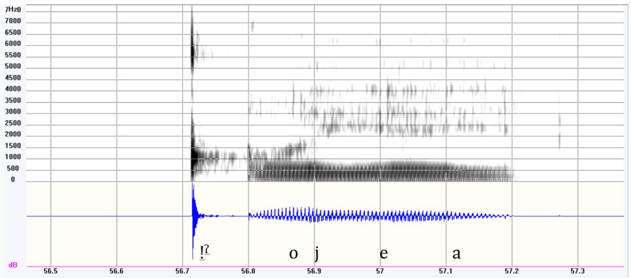

Figure 3. Audio waveform and spectrogram of the word [!ˀojea] 'beeswax'

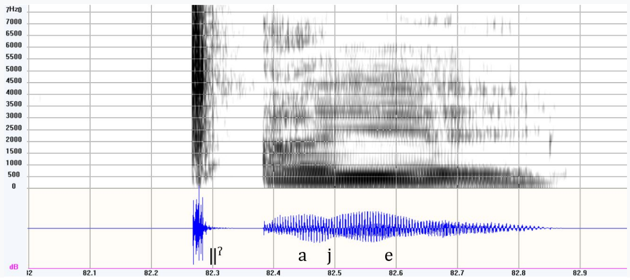

Figure 4. Audio waveform and spectrogram of the word [ǁˀaje] 'to throw'

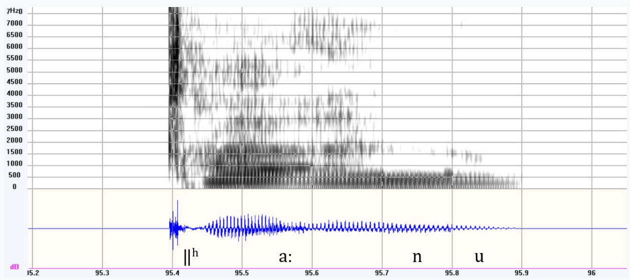

Figure 5. Audio waveform and spectrogram of the word [ǁʰanu] 'dog'

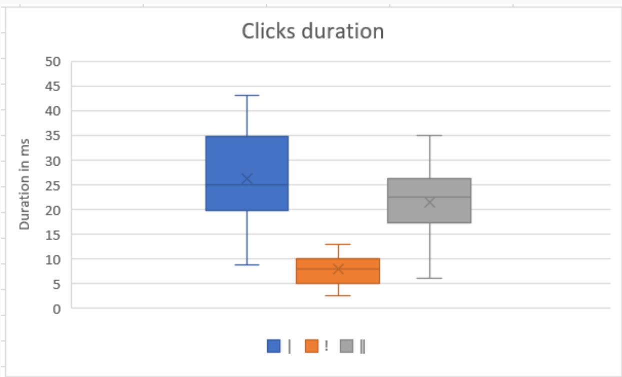

Figure 6. Mean duration of the dental, alveolar and lateral clicks measured produced by 4 Hadza speakers (n=70).

3.3 Click accompaniments

Spectrograms of Figures 2 and 5 show realizations of aspirated clicks. All clicks but the bilabial can be aspirated. The average values of the aspirated accompaniments are displayed in Figure 6. Aspiration noise following the transients has sorter duration when compared to the glottal lag accompanying all clicks. Nasal clicks start by a nasal portion (sometimes devoiced) before the transient release.

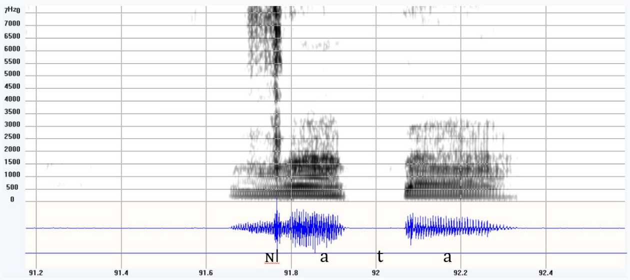

Figure 7. Audio waveform and spectrogram of the word [ᴺǀʰata] ‘tongue’

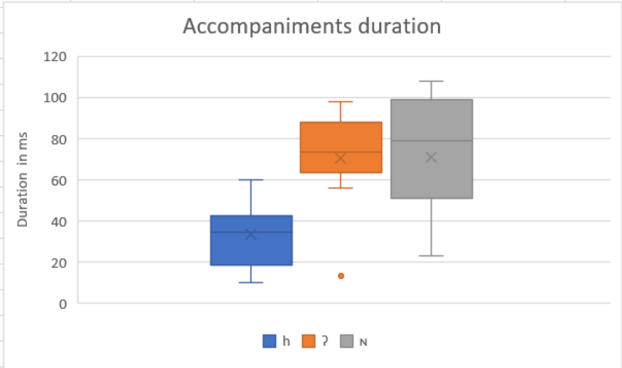

Figure 8. Duration of click accompaniments (aspirated, glottal and nasal) produced by 4 Hadza speakers (n=36).

3.4 Timing of the double closure

Figure 9 shows the audio waveform, intraoral (Po) and oral airflow (Oaf) of a dental click in the word [ǀaʔako]. There are 2 release transients that can be observed (1 & 2 in the Figure) on the waveform. The first is transducing the front release and the second the posterior (uvular) release. Aerodynamic data confirm this point. Indeed, Po curve shows that the maximum is reached after the 1st transient at the moment of the posterior release which means that there is, in this case, a short interval of time between the front and the posterior releases. Oaf become negative after the front release showing an ingressive airflow. The Oaf becomes positive after the posterior closure release.

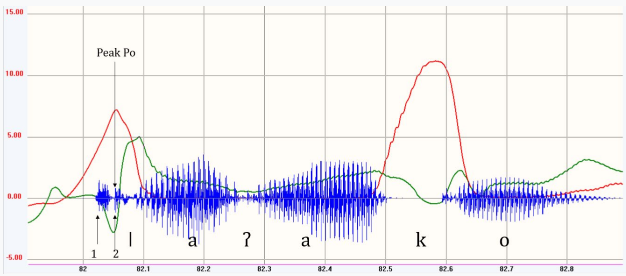

Figure 9. Audio waveform, Po in hPa (red plot) and Oaf in dcm^3/s (green plot) of the word [ǀaʔako] ‘a large tree’. Peak Po is indicated by the arrow falling to the audio waveform.

1 and 2 indicate the moment of the front and back releases of the dental click [ǀ].

3.5 spectral features

Figure 8 shows average values taken at the peak of the transient releases for all clicks. The peaks of maximum energy were measured in a 10 ms Hamming window. This shows that the dental and alveolar clicks are clearly distinguished as acute and grave. The lateral click has a frequency band of similar intensity between 4 and 5 kHz on average.

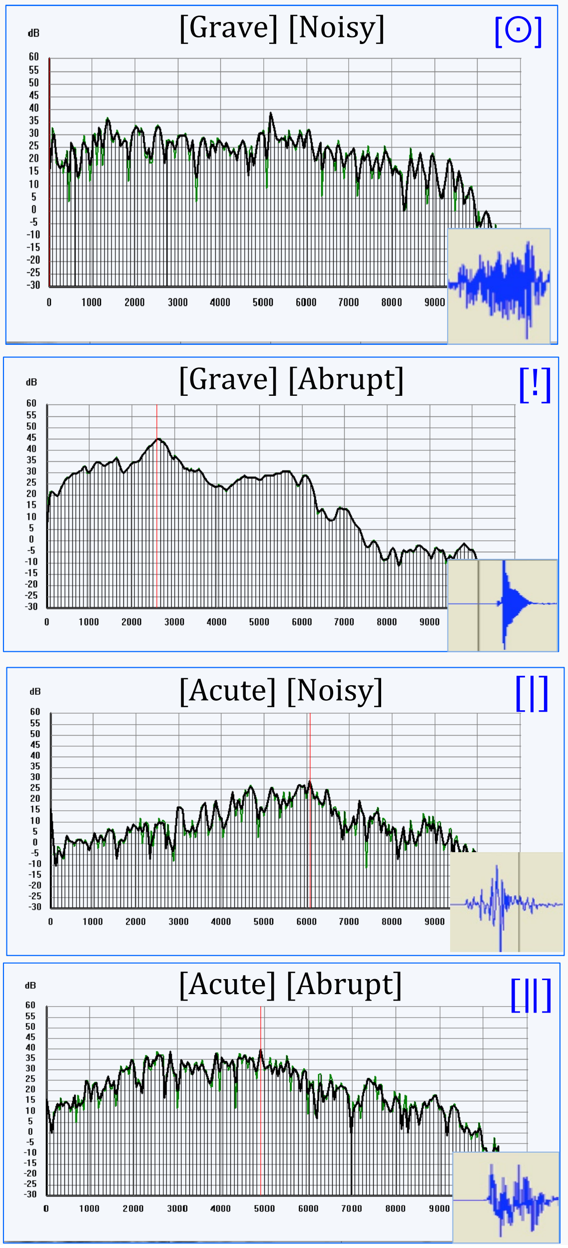

Figure 10. FFT spectra accounting the spectral characterization of the Hadza clicks.

4. DISCUSSION

The audio waveform observation of the 2 transients in clicks allow making inferences about the movement and speed of the articulators involved. The anterior articulators (apex and tongue blade) move at different speed with the clicks. There is a difference between the apical clicks [ǀ, ǃ]. The alveolar click shows a much shorter initial transient (2 ms) when compared to the dental click which takes on average 12 ms before the peak of the noise. The lateral click [ǁ] where the tongue blade extends backwards during production has a sharp release and lasts for an average of 23 ms. The release of this click which is abrupt is also different from the other clicks that have a gradual extinction noise. This movement of the tongue dorsum can vary in different click realizations. Figure 9 gives a good example of the time difference that can occur between the font and the back closure of the articulatory click components. These sequences appear randomly with the dental click in our data. In a way the release of the lateral click is also made in a sequence as the initial lateral tongue release is followed about 20 ms later by a second sharp release accounting for the complete and abrupt opening of the tongue closure (See audio waveform at Figures 4 & 5).

The role of the posterior cavity is also important to consider in click releases. The short rising resonances - the partials - [between 1 and 2 kHz] following the posterior release (Figure 2 and 3), results from the lowering and retraction of the tongue dorsum reducing the volume of the pharyngeal cavity. These short resonances are visible after the 1st transient of the glottal and aspirated clicks of Figures 2 and 3. Although this has been observed with all accompaniments, but the bilabial, this occurs more frequently with the glottal.

Feature interpretation for clicks has to integrate a time factor. All the segmental cues discussed require to account very precisely for the relative coordination of articulatory movements. In addition to these tongue movements, the adjustments of the larynx and the velum movements which regulate and control nasal flow have to be taken into consideration. Regarding the nasal clicks [ᴺǀ, ᴺǀˀ, ᴺǃ, ᴺǃˀ, ᴺǁʰ, ᴺǁˀ], there is still a substantial amount of work to be done. First it is clear that nasalization spreads over the whole click duration. This makes that these clicks fully nasal and voiced most of the time. However, there are regular occurrences of clicks starting with a voiceless or very weak nasal component. These occurrences seem to occur randomly but this will have to be checked in future data analysis.

As mentioned by [6], the distinctions between clicks are unencoded in the following context and must reside in the noise burst only. The spectral information contained in the click noise burst is delivered to the auditory system with such saliency (temporal and/or intensity) that accurate perception is ensured. This makes that acoustic features are essential in making a correct description of click features.

From an acoustic point of view, Traill [6] proposed that clicks can be described with 2 features [grave vs. acute] and [abrupt vs. noisy]. This makes that the dental [ǀ] and alveolar [ǃ] Hadza clicks are both [grave] but [ǀ] is [noisy] and [ǃ] is [abrupt]. The lateral click [ǁ] is [grave and acute]. This description is deduced from an examination of the audio waveforms and acoustic spectra taken at the clicks release. Figure 11 includes the bilabial click [ʘ] which is noisy with a grave spectrum and some high frequency emphasis due to the labio-dental component of the articulation. The lateral click [ǁ] has a short band of intense noise between 4 & 5 kHz. This is higher than the alveolar [ǃ] and lower than the dental [ǀ].

A click can also be described as a series of features or a superposition of gestures, for example, nasal, lateral and aspirated [ᵑǁ̰ʰ]. Figure 11 displays articulatory and acoustic features for Hadza clicks.

Articulatory description

labial	dental laminal	apical-alv.	lateral
ʘ	ǀ	ǃ	ǁ

Acoustic description

grave

ǃ ʘ

abrupt noisy

ǁ ǀ

acute

Figure 11. Articulatory and acoustic features for the characterization of Hadza clicks, following [6].

Finally, this contribution, coming from the partial analysis of a larger set of collected data, allows to make some observations about the relation between phonetics and phonology. The timing specifications and the time to produce some gestures are not directly related (one to one) with the phonological features. Hadza as most of the languages where clicks are found remain a challenge to give adequate answers to this question. This is why is so important to collect phonetic data in these languages.

REFERENCES

[1] Sands, B, Maddieson, I. & Ladefoged, P. 1996. The phonetic structures of Hadza. *Studies in African linguistics* (25), 2: 171-204.
[2] Sands, B. 2013. Phonetics and phonology of Hadza. In R. Vossen (ed). *The Khoesan languages*. London, Routledge: 38-42.
[3] Sands, B. 2020. Click consonants an introduction. In B. Sands (ed.) *Click consonants*. Leiden, Brill: 1-73.
[4] Miller, K. 2009. *Highlights of Hadza Fieldwork*. Ms. San Francisco.
[5] A. Ghio and B. Teston, "Evaluation of the acoustic and aerodynamic constraints of a pneumotachograph for speech and voice studies," in *International Conference on Voice Physiology and Biomechanics,* 2004: 55–58.
[6] Traill, A. 1997. Linguistic phonetic features for clicks: Articulatory, acoustic and perceptual evidence. In R.K. Herbert (ed.) *African Linguistics at the Crossroads: Papers from Kwalusemi*. Köln, Rudiger Köppe: 97-117.