Beyond the midsagittal: Toward a three-dimensional conceptualization of click consonants

Jeremy Coburn [1]

[1] Indiana University, Bloomington, IN, USA

jecoburn@iu.edu

Background: Contrary to previous claims regarding the inapplicability of ultrasonography to speech sounds with rapid articulatory transitions, such as clicks (Stone 2005:463), this study demonstrates that 3D/4D ultrasound (US) can in fact be applied to study clicks in Zulu. The coordinative sequencing of the primary phases involved in click production are illustrated in Figure 1: [1] *Shutting phase*; [2] *Closure phase (rarefaction)*; [3] *Release phase*.

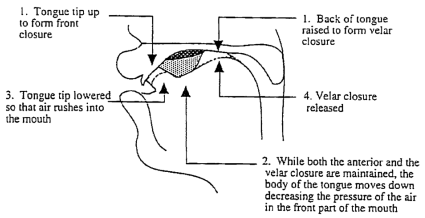

Figure 1: *Traditional midsagittal visualization of click production, taken from Ladefoged & Maddieson (1996:247)*

Although click consonants require inherently multidimensional and multidirectional tongue gestures, the illustration in Figure 1 reveals a bias toward conceptualizing clicks based upon the midsagittal plane. This bias is the result of technological constraints which, until recently, made three-dimensional imaging of lingual articulations impossible. Previous research on clicks has explored tongue shape during various phases of click articulation with a two-dimensional view along the midline of the tongue (e.g., real-time MRI, X-ray, 2D US). Information about click kinematics off of the midline have been surmised by studies which employ static and electronic palatography and linguography, providing data on tongue-palate contact. Conspicuously missing in any one of these earlier contributions is information regarding the holistic tongue shape of clicks across parasagittal, coronal, and transverse planes.

Recent advancements in ultrasound technology have made it possible to use 3D/4D US for speech science research (Lulich & Pearson 2019) and has been applied to examine other inherently multidimensional speech sounds (e.g., laterals in Charles & Lulich 2018). The present study provides novel data regarding tongue shape off the midline during click articulation.

Methods: Real-time 3D US images with synchronous audio recordings were collected from a 46-year-old female Zulu speaker from eastern South Africa. Zulu [*S.42; zul*] is a Bantu language of the Nguni group spoken in South Africa. Zulu contrasts three click types (i.e. dental [|], palato-alveolar [!] and lateral [||]) and five click accompaniments. 60 stimuli were presented, in orthography, to the participant while sitting in a sound attenuating room. The stimuli consisted of 45 target words and 15 filler words in three symmetrical vowel contexts (/i_i/, /u_u/, and /a_a/). The 45 target words contained one of 15 click type-accompaniment combinations in the three symmetrical vowel contexts.

Volumetric wedge-shaped US imaging of the tongue was collected with a Philips 3D/4D EPIQ 7G ultrasound machine and a Philips xMatrix x6-1 digital 3D/4D transducer stabilized under the jaw with a customized Articulate Instruments Ltd headset, restricting probe movement. A frame rate of 23 fps (i.e., 43.4 ms/frame) was achieved. To increase the temporal resolution, parameters were adjusted such that only the tongue body between the shadow of the mandible and the shadow cast by the hyoid bone was recorded. Secondary articulatory data was collected by creating a palate impression using hypoallergenic dental alginate and digitized using a NextEngine Desktop 3D scanner.

3D reconstructions of the tongue surface in the ultrasound frame just prior to the release of the anterior constriction were created and analyzed using a custom-built toolbox for MATLAB to compare overall tongue shape. Tongue surface reconstructions of a single click type in three sequential ultrasound frames were also created to examine various phases of click articulation (lateral click shown in Figure 2).

Results and discussion: Results show the efficacy of 3D/4D ultrasound in observing the tongue surface and contribute to a more holistic view of click production. Despite relatively lower frame rates (23 fps), this 3D/4D US methodology effectively captures the spatial and temporal complexity of clicks in a way not afforded by other methods. Three frames were captured within the time course of each click, illustrating various phases of click production (e.g. shutting (T0), rarefaction (T1), and release (T2)). Tongue shape data is simultaneously observable along the midline and lateral margins. Importantly, this speaker exhibited a high degree of anatomical and articulatory asymmetry; the midline of the probe did not correspond to the tongue midline. This fact would not have been known without observing coronal planes. 3D/4D US contributes substantially to our ability to conceptualize click production beyond the midsagittal plane.

References

Charles, S., & Lulich, S. M. (2018). Case study of Brazilian Portuguese laterals using a novel articulatory-acoustic methodology with 3D/4D ultrasound. Speech Communication, 103, 37-48

Ladefoged, P, & Maddieson, I. (1996). The Sounds of the World's Languages. Malden, MA: Blackwell Publishing.

Lulich, S. M., & Pearson Jr, W. G. (2019). Three-/Four-Dimensional Ultrasound Technology in Speech Research. *American Speech-Language-Hearing Association*, *4*(4), 733-747.

Stone, M. (2005). A guide to analysing tongue motion from ultrasound images. *Clinical linguistics & phonetics, 19*(6-7), 455-501.

Keywords: methodological research, clicks, 3D/4D ultrasound

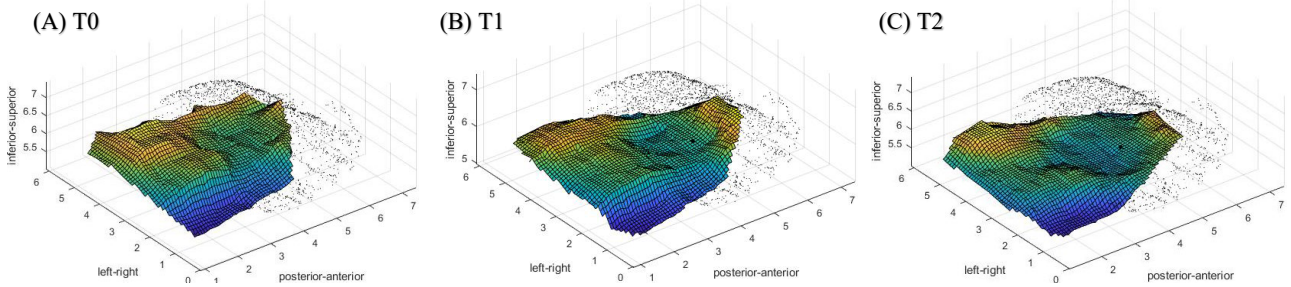

Figure 2. *Composite 3D images of ngx [ⁿǀᵍ] in /a_a/ context in three sequential frames (T0-T2) with superimposed palate.*