A Kinect 2.0 system to track and correct head-to-probe misalignment

Sam Johnston[1], **Diana Archangeli**[1,2], and **Rolando Coto**[1]

[1]*Department of Linguistics, University of Arizona*
[2]*Department of Linguistics, University of Hong Kong*

In ultrasound experimentation, a constant alignment between a subject's head and the ultrasound probe is essential to a valid analysis. This fixed head-to-probe alignment is critical to obtain accurate ultrasound images of the tongue that can be reliably compared with one another. Consequently, there has been much work to develop an effective method of securing a subject's head in relation to the ultrasound probe. Previous methods have included the HATS system (Stone and Davis, 1995), the use of a fitted helmet (McLeod and Wrench, 2008), and more recently a elastic strap (Derrick et al., 2015), all of which use a physical apparatus to manually fix the head-to-probe alignment. Two additional systems, the Palatoglossatron (Baker, 2005), and HOCUS/OptoTrak (Whalen et al., 2005) are systems which track the position of the head, instead of immobilizing the head. These each require the subject to wear additional equipment.

One limitation of the Palatoglossatron (Baker, 2005), is that it is primarily intended to correct for pitch-dimension misalignment, and does not address the dimensions of yaw and roll. HOCUS (Whalen et al., 2005) requires infrared diodes to be placed on a tiara or directly onto the head to track its possible movement and misalignment. Yet these diodes themselves are subject to possible movement during the experiment (cf. Roon et al., 2013), throwing off head tracking.

The current study utilizes the Kinect 2.0 head-tracking API (Han et al., 2013) to identify and track the the location of a head in 3D space in real time. This system allows for free head movement and also does not require any special devices to be worn, and therefore is completely non-invasive, making it particularly suitable for young children and elderly subjects. The Kinect has been integrated into a custom-designed system that will alert subject and researcher when the subject's head becomes misaligned from a stationary ultrasound probe.

The purpose of the present study was to establish the accuracy of the Kinect's head-tracking measurements. Video cameras were placed to the side, in front of, and above the subject during the experiment, capturing the angle of the head in each dimension of pitch, yaw, and roll as it moves from center. Images from the videos were taken, and the measurements of the Kinect system were verified by hand-measuring the video images.

Results indicate that the Kinect's tracking of head movement is quite similar for each of pitch, yaw, and roll. For each of these dimensions, Whalen et al. (2005) describes acceptable ranges of head movement which does not significantly alter the quality of an ultrasound image. They find that for any dimension, 5 degrees of movement is tolerable. In the present study, when the (hand-measured) head-tilt was within 5 degrees in either direction, the Kinect's measurement values diverged no more than 2 degrees from the hand-measured angle. This demonstrates that the Kinect head-tracking software can be used to set limits that will conservatively keep the subject's head within an acceptable range of movement.

References

Baker, A. (2005). Palatoglossatron 1.0. *University of Arizona Working Papers in Linguistics.*

Derrick, D., Best, C., and Fiasson, R. (2015). Non-metallic ultrasound probe holder for co-collection and co-registration with ema. pages 1–5.

Han, J., Shao, L., Xu, D., and Shotton, J. (2013). Enhanced computer vision with microsoft kinect sensor: A review. *IEEE Trans. Cybernetics*, 43(5).

McLeod, S. and Wrench, A. (2008). Protocol for restricting head movement when recording ultrasound images of speech. *Asia Pacific Journal of Speech, Language, and Hearing*, 11:23–29.

Roon, K., Jackson, E., Nam, H., Tiede, M., and Whalen, D. H. (2013). Assessment of head reference placement methods for optical head-movement correction of ultrasound imaging in speech production. *Journal of the Acoustical Society of America*, 134:4206.

Stone, M. and Davis, E. P. (1995). A head and transducer support system for making ultrasound images of tongue/jaw movement. *Journal of the Acoustical Society of America*, 98:3107–3112.

Whalen, D. H., Iskarous, K., Tiede, M., Ostry, D., Lehnert-Lehouillier, H., Vatikiotis-Bateson, E., and Hailey, D. S. (2005). The haskins optically corrected ultrasound system (hocus). *Journal of Speech, Language, and Hearing Research*, 48:543–553.