Automated tracking of hyoid movement using ultrasound

The coordinated movement of the tongue, mandible and hyoid is essential for both speech production and swallowing. While tongue-mandible coordination has been widely studied using x-ray microbeam (Edwards,1985) and electromagnetic articulography (Shaw & Chen, 2019), there have been very few studies of tongue-hyoid coordination, especially in speech production (Matsuo & Palmer, 2010). This is partly due to the lack of instrumentation to facilitate the measurement of hyoid movement. Recently, two methods of automatically tracking hyoid movement from ultrasound image sequences have been developed. They have been used to study hyoid movement in normal swallow and compared with manual tracking (Ma & Wrench, 2022). The first method, a hyoid shadow tracker, detects the position of the hyoid shadow by searching within an angular region of interest. The second uses DeepLabcut (DLC), a deep neural net training and analysis software package (Mathis et al., 2018; Nath et al., 2019), to estimate the hyoid position in ultrasound recordings of swallows. A comparison between the two tracking methods and manual labelling showed that the DLC tracker had a better agreement with the manual tracking than the shadow tracker. A new DLC network was trained to track tongue surface contours, hyoid and mandible during speech and shown to produce tongue contour estimates with equivalent accuracy to expert labelling (Wrench & Balch-Tomes, 2022a).

Extended training of the DLC neural net using both speech and swallowing data now permits estimation of tongue contour, mandible, and hyoid during bolus manipulation and swallowing. Using this net, we report on the relative timing of contractions of sectors of the tongue and geniohyoid measured during a 100ml continuous cup drinking task by 20 participants. Within each recording, multiple bolus propulsion and swallow cycles were observed. An example of a swallow sequence is displayed in Figure 1. A glossogram (Wrench & Balch-Tomes, 2022b) visualises the degree of constriction at different points along the midsagittal tongue surface normalised by the average distance measured over the whole recording and converted to grayscale. The lighter colour on the glossogram represents a contraction of the respective portion of the tongue relative to the short tendon. The hyoid-mandible distance is plotted below the glossogram. The timing relationship of the tongue during the bolus propulsion and the forward movement of the hyoid as the bolus is about to pass the oropharynx could be easily observed.

The results reported here contribute to the understanding of coordinated motor control of different tongue sections and hyoid in typical sequential swallows. The development of a reliable hyoid tracker for ultrasound offers the potential for tongue-hyoid coordination to be investigated in speech production and swallowing.

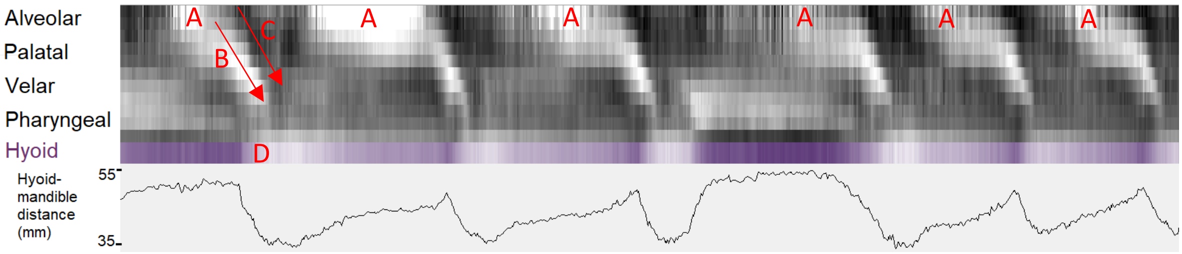

Figure 1 Grayscale glossogram showing sequence of swallows of 100ml water from a cup. A - Tongue tip lowers to accept bolus B - Tongue sectors retract in sequence to carry the bolus posteriorly. C - Sectors of the tongue extend to squeeze the bolus posteriorly. D – Hyoid advances as bolus reaches pharyngeal stage. Below is a plot of hyoid-mandible distance in mm.

References

Edwards, J. (1985). Contextual effects on lingual–mandibular coordination. *The Journal of the Acoustical Society of America*, *78*(6), 1944-1948.

Ma J.K.-Y. & Wrench A.A. (2022). Automated assessment of hyoid movement during normal swallow using ultrasound. *International Journal of Language & Communication Disorders, 57,* 615–629.

Mathis, A., Mamidanna, P., Cury, K.M., Abe, T., Murthy, V.N., Mathis, M.W. et al. (2018). DeepLabCut: markerless pose estimation of user-defined body parts with deep learning. *Nature Neuroscience, 21*(9), 1281–1289.

Matsuo, K. & Palmer, J.B. (2010). Kinematic linkage of the tongue, jaw, and hyoid during eating and speech. *Archives of Oral Biology, 55*, 325-331.

Nath, T., Mathis, A., Chen, A.C., Patel, A., Bethge, M. & Mathis, M.W. (2019). Using DeepLabCut for 3D markerless pose estimation across species and behaviors. *Nature Protocols, 14*(7), 21522176.

Shaw, J. A., & Chen, W. R. (2019). Spatially conditioned speech timing: evidence and implications. *Frontiers in Psychology*, *10*, 2726.

Wrench, A., & Balch-Tomes, J. (2022a). Beyond the Edge: Markerless Pose Estimation of Speech Articulators from Ultrasound and Camera Images Using DeepLabCut. *Sensors*, *22*(3), 1133.

Wrench, A., & Balch-Tomes, J. (2022b). pending Ultrafest 2022