Speech Initiation Measured with Simultaneous Oral Airflow and Tongue Ultrasound Registration

Pertti Palo[1,2]* and Steve Cowen[1]
[1]Queen Margaret University, Edinburgh, UK; [2]University of Strathclyde, Glasgow, UK
*pertti.palo@taurlin.org

Background

Speech initiation and pre-speech movement can be used as a window into speech movements in a context which is unconstrained by previous articulation. We are interested in how airflow fits in with previous results on speech initiation from delayed naming. An acoustic study by Rastle et al. (2005) shows that the place and manner of the first consonant in a target affects acoustic Reaction Time (RT). An articulatory study by Kawamoto et al. (2008) shows that the same effect is not present in articulatory RT of the lips and Palo et al. (2015) further show that the effect is not present in articulatory RT of the tongue. In this study, we demonstrate simultaneous Oral Airflow (OAF) and Ultrasound Tongue Imaging (UTI) and provide results from a pilot study exploring the timing relations of OAF, tongue movements, and acoustic speech.

Materials and methods

We recorded one 40-year-old native Finnish speaking participant (the first author) in a delayed naming experiment which combined simultaneous acquisitions of audio, OAF, and UTI. Audio was recorded separately in synchrony with OAF and with UTI. Overall synchrony is provided by post-hoc synchronisation of the audio tracks. OAF data was acquired using an EVA2 system (S.Q.Labs, Aix-en-Provence) and SESANE software running on a Lenovo Core-i5 notebook PC. The nasality sensor was removed from the EVA2 airflow device to allow access for the ultrasound probe. Ultrasonic and audio data were recorded using an Articulate Instruments/Telemed Echo Blaster 128 portable ultrasound scanner with a C3.5/20/128 Z-3 probe operating at 3 MHz controlled by Articulate Assistant Advanced (AAA) software (Articulate Instruments, Edinburgh). The scan depth was set to 90mm, and the field of view was reduced to approximately 70 % (88 scanlines) to give a frame rate of 78 fps. Audio data was collected using an Audio Technica AT8010 omnidirectional condenser microphone and a Focusrite Scarlett Solo2 USB interface, at 22kHz/16bit.

Procedure

The target words in the experiment were /CVCV/ Finnish phonotactically legal words: /asa/, /dasa/, /hasa/, /jasa/, /kasa/, /lasa/, /masa/, /nasa/, /pasa/, /rasa/, /sasa/, /tasa/, and /vasa/. Each trial began with the target word being displayed on a computer screen. The participant was instructed to read the word internally while remaining at rest until he heard the go-signal (50 ms long 1 kHz beep). The EVA recording was started first followed by the AAA recording, which played the randomly delayed go-signal. After the participant heard the beep, he was to produce the target word as soon as possible. No instructions were given about breathing during the experiment.

Results

Proof-of-concept is provided in Figure 1, which plots the oral airflow and pixel difference (Palo et al., 2014) of the recorded tokens. More detailed analysis will be available by the time of the conference.

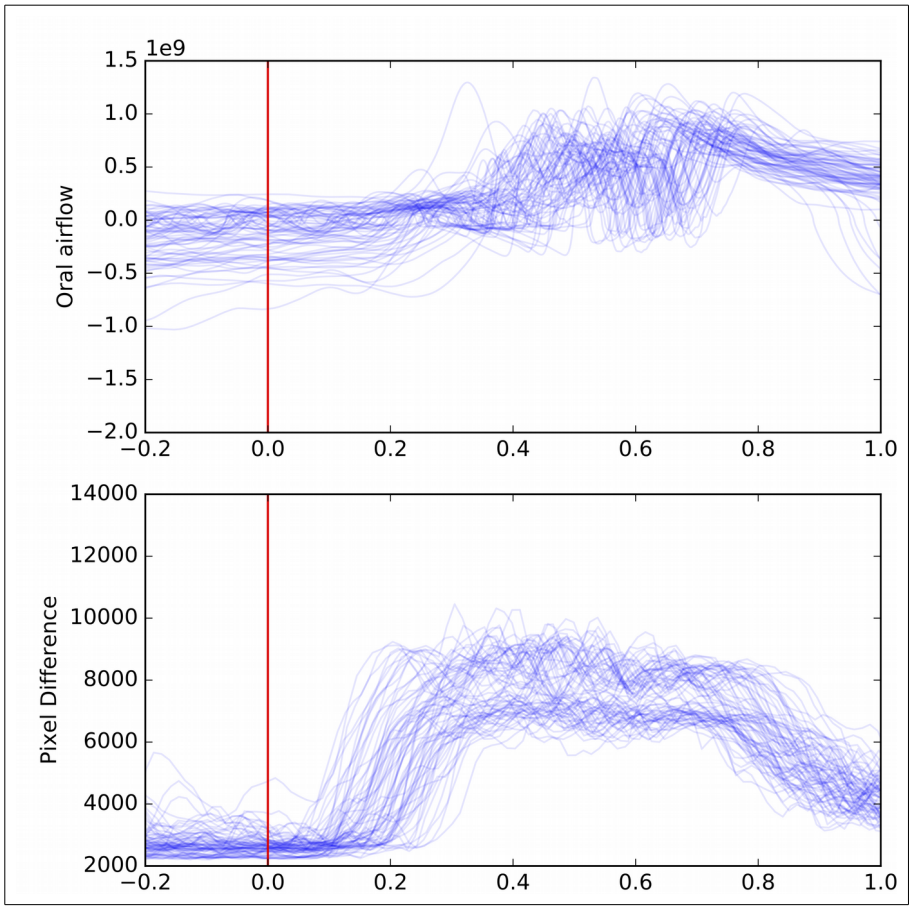

Figure 1. Oral airflow and pixel difference as function of time (s) for all of the recorded tokens synchronised on the go-signal (red, vertical line).

References

Kawamoto, A. H., Liu, Q., Mura, K., Sanchez, A. (2008). Articulatory preparation in the delayed naming task. *Journal of Memory and Language* 58(2), 347 – 365.

Palo, P., Schaeffler, S., and Scobbie, J. M. (2014). Pre-speech tongue movements recorded with ultrasound. In *Proceedings of ISSP 2014*, pages 304 – 307.

Palo, P., Schaeffler, S., and Scobbie, J. M. (2015). Effect of phonetic onset on acoustic and articulatory speech reaction times studied with tongue ultrasound. In *Proceedings of ICPhS 2015*, Glasgow, UK.

Rastle, K., Harrington, J. M., Croot, K. P., Coltheart, M. (2005). Characterizing the motor execution stage of speech production: Consonantal effects on delayed naming latency and onset duration. *Journal of Experimental Psychology: Human Perception and Performance* 31(5), 1083 – 1095.