Sonographic & Optical Linguo-Labial Articulation Recording system (SOLLAR)

Aude Noiray [a b], Jan Ries [a], Mark Tiede [b]
[a] *University of Potsdam,* [b] *Haskins Laboratories*

We present here a customized method developed jointly by scientists at LOLA (Potsdam University) and Haskins Laboratories (New Haven) for the recording of both tongue and lip motion during speech tasks in young children. The method is currently being used to investigate the development of 1) coarticulation (resistance and anticipatory coarticulation, cf. two other abstracts submitted); and 2) articulatory coordination in preschoolers compared with adults who have mature control of their speech production system.

Children are recorded with a portable ultrasound system (Sonosite Edge, 48Hz) with a small probe fixed on a custom-made probe holder and ultrasound stand. The probe holder was specifically designed to allow for natural vertical motion of the jaw but prevent motion in the lateral and horizontal translations. The set up is integrated into a child-friendly booth that facilitates integrating the production tasks into games.
Ultrasound video data are collected concurrently with synchronized audio recorded via a microphone (Shure, 48kHz,), pre-amplified before being recorded onto a desktop computer. In addition to tongue motion, a frontal video recording of the face is obtained with a camcorder (Sony HDR-CX740VE, fps: 50Hz). This video is used to track lip motion for subsequent labial measurements, and to track head and probe motion for transforming contours extracted from the ultrasound images to a head-based coordinate system. The speech signal is also recorded via the built-in camcorder microphone, and synchronization of both video signals (from the ultrasound and the camcorder) is performed through audio cross-correlation in post-processing.

Lip motion is characterized with a video shape tracking system (Lallouache 1991) previously used for examining anticipatory coarticulation in adults (Noiray et al., 2011) and children (Noiray et al., 2004; 2008). During production tasks, the lips of our young participants are painted in blue as this color maximized contrast with the skin. In post-processing these blue shapes are then tracked for calculation of lip aperture, interolabial area and upper lip protrusion.
Tongue contours derived from ultrasound are relative to the orientation of the probe with respect to the tongue surface. To correct for jaw displacement and (pitch) rotation of the head we compute two correction signals similar to the HOCUS method described in Whalen et al. (2005), but in this case derived from tracking the positions of blue reference dots in the video signal using custom Matlab procedures. The displacement of the probe relative to the centroid of dots placed on each speaker's forehead provides a vertical correction signal. The orientation of dots placed on the cheek observed within the video image through a mirror oriented at 45° giving a profile view provides a pitch rotation correction signal around the lateral axis. Application of these two signals to the extracted contours allows for their consistent comparison in a head-centric coordinate system.

Acknowledgments
This work is supported by the German DFG GZ: NO 1098/2-1.

References

Lallouache, M. T. (1991). Un poste «Visage-parole» couleur. Acquisition et traitement automatique des contours des lèvres (A «face-speech» interface. Automatic acquisition and processing of labial contours), Ph.D. ENSERG, Grenoble, France.

Noiray A Ménard L Cathiard MA Abry C and Savariaux C (2004). The development of anticipatory labial coarticulation in French: A pioneering study. In *Proceedings of Interspeech, 8th ICSLP*, 53-56.

Noiray, A., Cathiard, M. A., Me´nard, L., and Abry. C. (2008). Emergence of a vowel gesture control. Attunement of the anticipatory rounding temporal pattern in French children. in Emergence of Language Abilities, edited by S. Kern, F. Gayraud, and E. Marsico (Cambridge Scholars Publishing, Newcastle, UK), pp. 100–116.

Noiray A., Cathiard M-A., Ménard L., & Abry C. (2011). Test of the Movement Expansion Model: Anticipatory vowel lip protrusion and constriction in French and English speakers. Journal of Acoustical Society of America, 129 (1), 340-349.

Whalen, D. H., Iskarous, K., Tiede, M. K., Ostry, D. J., Lehnert-LeHouillier, H., Vatikiotis-Bateson, E., (2005). HOCUS, the Haskins Optically-Corrected Ultrasound System. Journal of Speech, Language, and Hearing Research , 48, 543-553.