A time-aligned account of lingual and laryngeal gestures using ultrasound tongue imaging (UTI) and electroglottography (EGG)

Stefano Coretta

14/01/2017

Contents

1 Purpose

The combination of techniques described in the following sections allows a synchronous mapping of lingual gestures and phonation during speech. Such methodology enables a time-aligned account of the movements of the tongue and the concomitant configuration of the glottis that characterises phonation. These techniques employ ultrasound tongue imaging (UTI) and electroglottography (EGG) for the simultaneous acquisition of articulatory data from, respectively, the oral cavity and the glottis.

2 Methodology

2.1 Ultrasound tongue imaging

Ultrasound tongue imaging (UTI) uses ultrasonography for charting the movements of the tongue into a two-dimensional image. In medical sonography, ultrasound waves (sound waves at high frequencies, ranging between 2 and 14 MHz) are emitted from a probe in a fan-like manner, and travel through organic material (such as skin and muscles). When the surface of a material with different density is hit, the ultrasound waves are partially reflected, and such "echo" is registered by the probe. These echoes can then be plotted on a two-dimensional graph, where different densities are represented by different shades (higher densities are brighter, while lower densities are darker). The graph, or ultrasound image, will show high density surfaces as very bright lines, surrounded by darker areas. By positioning the ultrasound probe in contact with the submental triangle (the surface below the chin), sagittally oriented, we can infer the cross-sectional profile of the tongue, which appears as a bright line in the resulting ultrasound image.

2.2 Electroglottography

Electroglottography is a technique that measures the size of contact between the vocal folds (the Vocal Folds Contact Area, VFCA).

2.3 Equipment setup

2.4 Post-processing: syncrhonisation and dEGG extraction

Since the signals from the ultrasound machine and the laryngograph are recorded simultaneously but separately, data from both machines need to be synchronised after acquision. Synchronisation is achieved throough the cross-correlation of the audio signals from both sources (Grimaldi et al., 2008).

References

Grimaldi, Mirko, B. Gili Fivela, Francesco Sigona, Michele Tavella, Paul Fitzpatrick, Laila Craighero, Luciano Fadiga, Giulio Sandini & Giorgio Metta. 2008. New technologies for simultaneous acquisition of speech articulatory data: 3d articulograph, ultrasound and electroglottograph. *Proceedings of LangTech* 1–5.