An EPG + UTI study of syllable onset and coda coordination and coarticulation in Italian

Cheng Chen, Chiara Celata, Irene Ricci, Chiara Bertini and Reza Falahati

Scuola Normale Superiore, Pisa, Italy

1. Introduction

This study is concerned with the methodological challenges of studying articulatory coordination of onset and coda consonants by means of an integrated system for the acquisition, real-time synchronization and analysis of acoustic, electropalatographic and ultrasound data. Electropalatographic panels (EPG) are responsible for the contact (closure/aperture) between the tongue and palate, while ultrasonographic (UTI) images provide the complementary information of the sagittal profiles of tongue synchronised with EPG data during the articulation of consonants and vowels in the speech chain. This original system makes it possible to process simultaneously the information of both linguo-palate contact and the movement of tongue for reaching its target (Spreafico et al. 2015).

The system is used to capture simultaneous data on linguo-palatal contact and tongue sagittal profiles of /s/, /l/ and /k/ adjacent to /a/ and /i/ as produced by native speakers of a Tuscan variety of Italian.

Using EPG and UTI data to investigate temporal and spatial coordination of consonant-vowel sequences is challenging to the extent that the identification of 'anchor points' for temporal measurements is not straightforward starting from information about whole tongue or tongue-palate configurations (or at least, not as straightforward as when starting from trajectories of points, as in EMA-based studies). At the same time, the two-channel experimental environment provides fine-grained spatial, in addition to temporal, information, namely, by allowing the analysis of coarticulatory activity for the selected anchor points and for the temporal lags between them.

The poster will illustrate the innovative audio-EPG-UTI synchronization system and offer some preliminary considerations about the methodological challenges related to the investigation of temporal and spatial coordination patterns in onset and coda consonants.

2. Motivation of the study

According to the articulatory model of syllable structure, the temporal and spatial coordination of articulatory gestures is conditioned by position in the syllable. The onset consonants are supposed to be more stable and to have a greater degree of constriction with respect to coda consonants (syllabic asymmetry; Krakow 1999). Moreover, the (temporal) coordination between an onset singleton consonant and the following nuclear vowel is found to be more stable than that between a nuclear vowel and the coda consonant (Browman and Goldstein 1988, 2000).

Although the stability of onset consonants is confirmed by many a study in the last ten years, recent research has revealed that the onset-nucleus coordination is also predicted by the articulatory properties of the consonant (e.g. Pastaetter & Poulier 2015); specifically, it is modulated by the degree of coarticulation resistance (Recasens and Espinosa 2009) of the consonant involved.

Such phenomena suggest that the intrinsic articulatory property of consonant might influence the temporal (and spatial) coordination between articulatory gestures. Cross-linguistic comparisons are expected to provide more evidence about the supposed interaction between coarticulatory patterns and gestural timing.

3. Description of the experiment

For this study on Italian, the corpus is composed of 12 stimuli, all disyllabic pseudo-words or very infrequent words. Each stimulus is inserted in a carrying sentence providing the same segmental context in which a bilabial consonant for all stimuli. The target consonants are /s/, /l/ and /k/; according to the DAC model (e.g. Recasens & Espinosa 2009), they have a high, intermediate and low degree of coarticulatory resistance, respectively. These consonants are analyzed both as onsets and as codas, i.e. in CV and VC contexts. The V is /a/ in one series, /i/ in another series. The stimuli with /a/ are produced twice: first in a prosodically neuter condition, then in a prosodically prominent position in which the target word bears a contrastive pitch accent.

Table 1 provides an example of carrying sentences and the list of stimuli. In the carrying sentence, the first repetition of the target stimulus corresponds to the prosodically neuter condition, while the second corresponds to the prosodically prominent condition (contrastive pitch accent). Following the hypothesis that laryngeal and supralaryngeal gestures tend to be coordinated (e.g. Ladd 2006, Muecke et al. 2012), we expect that also prosodic prominence can influence the way in which the onset-coda contrast is realized, either by enhancing or by reducing it.

Carrying-sentences	Pronuncia saba molte volte. ("He pronounces saba a lot of times.") Pronuncia seba? No, pronuncia SABA molte volte! ("Does he pronouce seba? No, he pronounces SABA a lot of times!")					
	/s/		/l/		/k/	
	CV	VC	CV	VC	CV	VC
/a/	Saba	bass	laba	bal	capa	pac
/i/	Siba	bis	liba	bill	kipa	pic

Table 1. Example of carrying sentences and list of the 12 target words in the corpus

The recordings were made in the phonetics laboratory of Scuola Normale Superiore, Pisa. Ultrasound data were captured using a MindRay device with a acquisition rate of 60 Hz, an electronic micro-convex probe (Mindray 65EC10EA 6.5 MHz) and a stabilization headset; electropalatographic data were captured via the WinEPG system by Articulate Instrument (SPI 1.0) recording palate images at 100 Hz; EPG, UTI and audio data were acquired and real-time synchronized using the Articulate Assistant Advanced (AAA) software environment and a video/audio synchronization unit. Two digital tones were produced and used to synchronize both EPG and UTI signals with the audio signal.

4. Methodological challenges

The two-channel synchronized articulatory approach allows the analysis of the temporal coordination of gestures *and* the coarticulatory patterns underpinnings gestural coordination in *one* output. For such goal to be fulfilled, it is however necessary to define a series of temporal landmarks allowing the estimation of gestures' relative distance (temporally and spatially).

Consonants and vowels are manually segmented according to the inspection of waveform and spectrogram (after exportation into Praat). In each vocalic or consonantal interval it is then possible to locate time-points for, respectively, the vocalic anchor and the reaching of maximum consonantal constriction. The vocalic anchor is the point in which the vowel reaches its target configuration (i.e., maximal predorsum lowering and tongue flattening for /a/, maximal predorsum raising for /i/). The maximum consonantal constriction is the time-point in which the articulatory target is reached (i.e. maximum constriction in the relevant lingual and palatal areas and minimal influence of V-to-C coarticulation). These two points are taken as references for the calculation of intergestural timing (or temporal distance, measured in ms) and of the coarticulatory modification of C (or spatial distance, measured in terms of changes in EPG indices, formant values and lingual profiles) as a function of V quality changes, position in the syllable and prosodic prominence.

To locate the V anchor and the maximum C constriction point, the EPG and UTI outputs for the selected acoustic intervals are first independently evaluated. E.g. for a /li/ stimulus, according to tongue profile qualitative inspection, the stable maximal constriction for /l/ is defined as the sequence of UTI frames showing apical raising and contextual dorsum flattening, before the anterodorsum fronting caused by the anticipation of the gesture for the /i/. The relevant UTI interval is labeled $\Delta t1$. Similarly, according to linguo-palatal contact patterns, the stable maximal constriction for /l/ is defined as those EPG frames in which there is maximum anterior constriction (with partial lateral contact) and before lateral obstruction and dorsum raising (also for anticipatory coarticulation). The relevant EPG interval is labeled Δt2. As a subsequent step the extension of Δt1 and Δt2 is simultaneously evaluated. The first temporal instant that falls within both Δt1 and Δt2 intervals corresponds to the maximum C constriction time-point. The V anchor time-point for /i/ is identified according to the same procedure, within the acoustic interval of the vocalic nucleus.

The temporal coordination of the consonantal and vocalic gestures in the different syllabic contexts (CV vs VC) can then be evaluated in conjunction with the spatial coarticulatory coordination for the two gestures. The study also allows the analysis of the effects of coarticulatory resistance (as evaluated from the comparison of the three consonants in the /a/ vs. /i/ context) and of prosodic prominence (as evaluated from the comparison between the prosodically neuter and the pitch accent condition) on C and V gestural organization.

Browman, C.P. & Goldstein, L. (1988). Some notes on syllable structure in Articulatory Phonology. *Phonetica* 45, 140-155.

Browman, C.P. & Goldstein, L. (2000). Competing constraints on intergestural coordination and self-organization of phonological structures. *Bulletin de la Communication Parlée*, 5, 25–34.

Krakow, R. A. (1999). Physiological organization of syllables: A review. *Journal of Phonetics* 27. 23–54.

Ladd, D. R. (2006). Segmental anchoring of pitch movements: Autosegmental association or gestural coordination? *Italian Journal of Linguistics* 18(1), pp. 19–38.

Mücke, D., Nam, H., Hermes, A. and L. Goldstein (2012). Coupling of tone and constriction gestures in pitch accents. In *Consonant Clusters and Structural Complexity*, Mouton de Gruyter, pp. 205-230.

Pastaetter M. & M. Pouplier (2105) Onset-vowel timing as a function of coarticulation resistance: Evidence from articulatory data. *Proceedings of the 18th International Congress of Phonetic Sciences*, Glasgow, 10-14 August 2015.

Recasens D. & A. Espinosa (2009) An articulatory investigation of lingual coarticulatory resistance and aggressiveness for consonants and vowels in Catalan. *Journal of the Acoustical Society of America* 125, 2288–2298.

Spreafico L., C. Celata, A. Vietti, C. Bertini & I. Ricci (2015) An EPG+UTI study of Italian /r/. *Proceedings of the 18th International Congress of Phonetic Sciences*, Glasgow, 10-14 August 2015.