DOI: 10.17469/O2103AISV000008

CHENG CHEN, CHIARA CELATA, IRENE RICCI

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

This study has two purposes. The first is that of implementing a system for the processing of multi-level speech data, which allows the synchronized acquisition and processing of electropalatographic, ultrasound tongue imaging and acoustic data. The second is that of testing the multi-level speech data system on a subset of Italian CV and VC sequences to ascertain how intersegmental coordination and coarticulation are related. We analyse onset-nucleus and nucleus-coda temporal relationships for three different consonants (/s/, /l/ and /k/) adjacent to either /a/ or /i/. Based on multi-level phonetic data on constriction location and overall tongue configuration, we show that the three consonants have different coarticulation properties and that these properties may in part explain the differences observed in the domain of temporal coordination between the consonant and the vocalic nucleus.

Key words: multi-level phonetic analysis, speech gesture coordination, coarticulation, syllable, Italian.

1. *Introduction*

This work explores the temporal coordination of articulatory gestures within the syllable in Italian, by comparing CV and VC sequences, where the consonant is, respectively, in onset and coda position. Articulatory models of syllable structure assume that the coordination between the vocalic gesture and the consonantal gesture may differ in the two cases. Based on previous literature on different languages, we expect to find differences in the temporal coordination of onset and coda consonants in Italian as well. In addition, recent literature suggests that the articulatory and coarticulatory properties of the segments play an important role in determining the details of the coordination patterns, and that not all segments or segmental sequences behave in the same way as far as gestural coordination within the syllable is concerned. Thus, an additional aim of this work is to compare consonants with different coarticulatory properties (in the sense of modifications of C articulation in varying vocalic contexts) and seek for possible relations between coarticulation and coordination patterns.

The methodology used is new. We establish an original system for the acquisition, real-time synchronization and analysis of acoustic, electropalatographic (EPG) and ultrasound tongue imaging (UTI) data, called *SynchroLing*. EPG and UTI instrumental techniques provide complementary information on, respectively, linguo-palatal contact patterns in the anterior vocal tract and midsagittal profiles of the whole tongue, including postdorsum and root. *SynchroLing* allows real-time inspection of contacts in the artificial palate and tongue midsagittal movements, coupled with acoustics. A preceding version of the multi-level system has been used for the analysis of rhotic variation in Tuscan (Spreafico, Celata, Vietti, Bertini & Ricci, 2015; Celata, Vietti & Spreafico, in press). One of the goals of this paper is that of testing the validity of this new instrumental environment for the analysis of coordination among segments and coarticulation. The experimental setting described here is challenging and beneficial at the same time. Among the challenges is the question of how we can identify landmarks (or anchor points) for temporal measurements starting from information about whole tongue or tongue-palate contact configurations. Among the benefits, a multi-level experimental environment provides fine-grained information about lingual movements, including areas of the oral cavity that are involved in the primary constriction and areas that are not, thus allowing the analysis of coarticulatory activity for the selected landmarks as well as for the temporal lags between them.

This paper is part of the first author's doctoral research project, which has a larger scope and includes the analysis of onset and coda consonant clusters and pitch accent variation. The current paper focuses on the following aspects only: the theoretical background (briefly sketched in § 2), the description of the methodology, with particular emphasis on the procedure for the identification of temporal landmarks for C and V gestures based on EPG and UTI evidence (§ 3), the coarticulatory patterns of /s/, /l/ and /k/ adjacent to /a/ and /i/ according to EPG indices and UTI profiles (§ 4.1) and the coordination patterns of /s/, /l/ and /k/ in onset vs. coda position (§ 4.2). Due to space requirements, acoustic data are not reported here, although they are an essential part of the current project.

2. *Background and hypotheses*

In articulatory models of syllable structure, the syllable is considered as a domain of articulatory timing. Different syllable structures may correspond to different characteristic patterns of phasing (Browman, Goldstein, 1988). In particular, syllable-onset positions are associated with tighter articulatory constrictions and greater stability than syllable-coda positions (since Krakow, 1989, 1999). Additionally, onset clusters are said to be globally aligned to the tautosyllabic vowel along their temporal midpoint in such away that an increase in onset complexity results in an increase of CV overlap ("c-center effect"); by contrast, coda clusters are sequentially organized (since Browman, Goldstein, 2000; De Jong, 2003; Hermes, Mücke & Grice, 2013 for Italian).

However, it has been recently shown that the articulatory and coarticulatory properties of the segments play an important role in determining the details of the coordination patterns. For instance, onset-vowel timing interacts with coarticulatory resistance, since the articulatory composition of the cluster predicts the degree of overlap between the final consonant of an onset cluster and the following vowel (e.g. Pastätter, Pouplier, 2015 on Polish). Similarly, not all coda clusters are sequentially organized with respect to the vocalic nucleus, since V shortening and an increased overlap between the vocalic nucleus and the following consonant have been observed for some complex coda clusters (e.g. Marin, Pouplier, 2010 for American English laterals; Marin, Pouplier, 2014 for Romanian rhotics).

This evidence suggests that the articulatory properties of segments (and primarily their coarticulatory resistance, e.g. Recasens, Espinosa, 2009) influence the timing of syllables and motivates a closer investigation of language-specific and consonant-specific coordination patterns.

A production experiment was run to investigate the difference between syllable onsets and codas in their temporal coordination with the syllable nucleus. Two hypotheses in particular were tested.

According to the first hypothesis, onset C gestures were expected to be more tightly coordinated to the vocalic nucleus than C gestures in syllable codas. Testing this hypothesis not only served the purpose of replicating a well-known articulatory timing effect on a different speech dataset, but – most importantly – served to establish if the procedure for the identification of temporal landmarks in *SynchroLing* (see below, § 3.4) was reliable or not. We anticipate that it was; by looking at *SynchroLing*-based gestural stability areas, it was possible to correctly identify the onset and offset of consonantal and vocalic gestures for an analysis of intergestural coordination that mirrors EMA-based coil tracking procedures.

The second hypothesis concerned the possibility that this coordination pattern varied as a function of the specific consonants involved. This hypothesis was suggested by findings reviewed above, related to the c-center effect in onset and coda clusters and the effects of coarticulatory resistance on shaping cluster-specific coordination patterns. Since the c-center effect in consonant clusters is one of the consequences of tight vs. loose coordination of syllable onsets vs. codas, respectively (since Krakow, 1999), we hypothesized that the degree of coarticulatory resistance could in principle affect also the way in which singletons are coordinated to the following or preceding vocalic nucleus. We tested three consonants, that were assumed to vary for their degree of resistance to vowel-induced coarticulatory modifications: /s/, /l/ and /k/. The former two are produced with the tongue tip as primary articulator, whereas /k/ with the tongue dorsum as primary articulator. Their degree of coarticulatory resistance was established by measuring the articulatory modifications of each consonant adjacent to /i/ as compared to adjacent to /a/. If coarticulatory resistance only affects the magnitude of the c-center effect (possibly because clusters are more susceptible than singletons to the influence of varying articulatory conditions), then the three consonants should show a similar difference between tight onset-nucleus coordination (i.e., a short latency between the consonantal and the vocalic gesture in CV sequences) and loose onset-nucleus coordination (i.e., a comparatively longer latency between the vocalic gesture and the consonantal gesture in VC sequences). If, on the contrary, consonantal properties such as coarticulatory resistance influence the onset-coda coordination pattern in singletons, the three consonants are expected to vary in the way they are timed to the vocalic nucleus as onsets or codas. In particular, spatial (coarticulatory) and temporal effects of vowel variation over adjacent consonants were expected to be positively correlated. Less resistant consonants (i.e., those that vary a lot as a function of vocalic context) were expected to be influenced by the vocalic gesture both in the sense of modifying their constriction location (spatial coarticulation) and in the sense of an increased temporal overlap between the vocalic gesture and the consonantal gesture. The effect was expected to be particularly strong for consonants in onset position. Thus, less resistant consonants in onset position were expected to show not only more variability of the constriction location in the /a/ vs. /i/ contexts, but also a stronger anticipation of the vocalic gesture. Consequently, they were also expected to show larger onset-nucleus coordination differences, compared to more resistant consonants. In the case of the consonants of the present study, /s/ was expected to be the most resistant consonant and to show the smallest onset-coda coordination difference; /k/ was expected to be the least resistant consonant and to show the largest onset-coda coordination difference; /l/ was expected to occupy an intermediate position between the two, as far as resistance to coarticulation is concerned, and therefore to show an onset-coda coordination difference smaller than /k/ and larger than /s/.

The reasons why we expect /s/ to be the most, /k/ the least and /l/ the intermediate resistant consonant are the following. Coarticulatory resistance (e.g. Recasens, Pallarès & Fontdevila, 1997; Recasens, Espinosa 2009) is known to depend on several factors, such as whether a given lingual region is involved in the constriction, jaw height, and the severity of the manner of articulation requirements. With respect to jaw height, /s/ having a higher jaw is known to be more resistant to coarticulation than both /l/ and /k/ (Recasens, 2012). Note that /l/ in Italian is always clear, i.e. with no involvement of the tongue dorsum in articulation. Frication with turbulence is also the most demanding manner of articulation, which reinforces the view that /s/ should show the comparatively highest degree of coarticulatory resistance. However, /l/ as a lateral continuant also has more demanding requirements than stops. In addition, the places of articulation of dorsal /k/ and /g/ are known to vary to a very large extent with vowel fronting (since Öhman, 1966; see also Fowler, Brancazio, 2000).

3. *Methodology*

3.1 Materials

For the current study, the corpus was composed of 12 disyllabic pseudo-words. They are shown in Table 1. The target consonants were /s/, /l/ and /k/. These consonants were included in non-word or very low-frequency word stimuli either as onsets (e.g. /saba/) or as codas (e.g. /bas/). The vowel was /a/ in one series, /i/ in another series.

Each stimulus was included in the carrier sentence *Pronuncia … molte volte* ("He pronounces … a lot of times"). In this sentence, the stimuli are produced with broad focus. The stimuli with /a/ were additionally produced in a prosodically prominent position in which the target word bore a corrective focus. In this case, the carrying sentence was structured as to favouring the emergence of a contrastive pitch accent on the experimental stimulus (e.g. *Pronuncia seba? No, pronuncia SABA molte volte!* "Does he pronounce seba? No, he pronounces SABA a lot of times!"). However, in this paper we deal with broad focus stimuli only. The total number of stimuli elicited was therefore 228 (3 consonants x 2 vowels x 2 syllable structures x 19 repetitions; see below, § 3.2, for details about repetitions). Due to mispronunciations (e.g., VC stimuli produced with a final schwa) or recording errors, the total number of stimuli analysed in the current paper was 216.

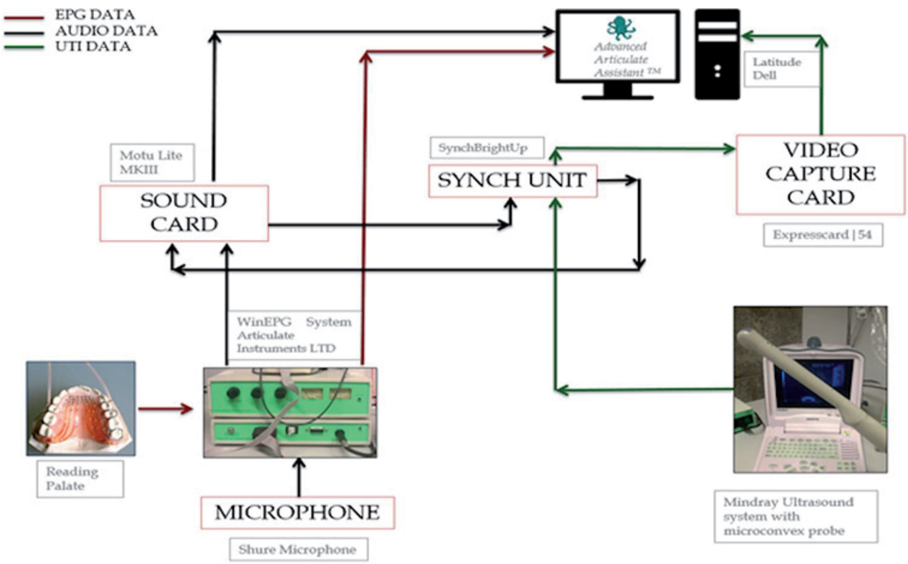
Table 1 - *Experimental stimuli used for the present experiment. For each stimulus, the orthographic and the phonetic transcriptions are given*

	/s/		/l/		/k/	
	CV	VC	CV	VC	CV	VC
/a/	saba[ˈsaba]	bass [bas]	laba [ˈlaba]	bal [bal]	capa [ˈkapa]	pac [pak]
/i/	siba [ˈsiba]	bis [bis]	liba [ˈliba]	bill [bil]	kipa [ˈkipa]	pic [pik]

3.2 *SynchroLing*: the audio-EPG-UTI synch system and the experimental procedure

This sub-section provides a detailed description of how *SynchroLing* works. The innovative aspect of the system lies in the real-time, automatic synchronisation of the three channels, which allows obtaining multi-level phonetic information simultaneously.

Figure 1 - *The instrumental setting of SynchroLing*



The system combines three parallel channels – one for the audio signal, one for the electropalatographic data (EPG) and one for the ultrasonographic data (UTI) (Figure 1). The EPG, UTI and audio signals were synchronized using a synchronization unit

(*SynchBright*) controlled by the Articulate Assistant Advanced software (AAA; version 2.16.14).

The acquisition platform was based on hardware and software components provided by Articulate Instruments Ltd[1]. The data were collected in the sound-proof studio in the linguistics laboratory of Scuola Normale Superiore. UTI data were collected at 60 Hz via a Mindray ultrasound machine with a microconvex probe (Mindray 65EC10EA 6.5 MHz). EPG data were collected at 100 Hz via the WinEPG™ (SPI 1.0) system.

Four subjects were recorded for this study. Each subject was recorded separately, according to the following procedure.

In the first place, the participant was asked to wear her/his personal artificial palate, and to practice with the word stimuli contained in the prompt sentences, by reading them aloud a sufficient number of times. The familiarization phase also served as an adaptation phase to reduce salivation and improve the overall comfort of the subjects in wearing the artificial palate. After 10-15 minutes, the ultrasonic transducer was fixed beneath the mandible of the participant with the help of a stabilizing helmet. The position of the probe was supposed to be orthogonal to the tongue surface so that the ultrasonic wave emitted from the probe was able to catch the entire lingual configuration during the movement of lingual gestures. Approximately half an hour after the wearing of the palate, the recording session began.

During the recording session, the speaker read the prompt sentences one by one, as they appeared in the AAA window on the computer screen in front of her/him. Five repetitions of the prompt list were recorded, except in one case, when one of the participants repeated the prompt list only four times. The order of the prompts was different across participants, but identical across repetitions for the same participant. The total duration of each recording session was never longer than 30 minutes.

The speech material was segmented and annotated according to one tier, i.e. a segmental tier, using the Praat software (version 6.0.29).

3.3 From gestural stability areas to segmental landmarks

3.3.1 Definition of segmental landmarks

For the analysis of onset-coda coordination difference, four articulatory landmarks were defined: Ctarget, Vonset and Vtarget for CV syllables; Vtarget, Conset and Ctarget for VC sequences in CVC syllables. We defined:

- Ctarget as the timepoint at which the consonantal gesture reaches its maximum constriction in the relevant lingual and palatal areas and bears minimal influences of coarticulation from the adjacent vowels;
- Vtarget as the timepoint at which the vocalic gesture reaches its target configuration and bears minimal influences of coarticulation from the adjacent consonant;
- For CV syllables, Vonset as a timepoint within the acoustic interval of the preceding consonant at which the nuclear vowel gesture begins;

[1] We acknowledge the contribution and technical support of Alan Wrench (Edinburgh) as well as the collaboration of Chiara Bertini (SNS Pisa) during the implementation procedures.

– For VC syllables, Conset as a timepoint within the acoustic interval of the preceding vowel at which the constriction gesture for the coda consonant begins.

To define the relative position of the segmental landmarks defined above, gestural stability areas were first detected on the two articulatory channels (EPG and UTI) separately. Then the two sources of information were combined to identify one reference gestural stability area and the landmarks for coordination measurements. The procedure is described below.

3.3.2 Gestural stability area: UTI evidence

The analysis of ultrasound tongue images was based on the idea that the target lingual configuration of a consonantal gesture differs maximally from the target lingual configuration of the adjacent vocalic gesture. The analysis of ultrasound tongue images was divided into three steps:

(i) Selection of a reference spline for the characterization of the vocalic gesture. The spline that happened to be the closest to the acoustic midpoint of the vowel was selected as the reference.

(ii) Identification of relevant fan radii for tongue vertical displacement measurement. Depending on the different articulatory properties shown by the lingual sagittal profile, two radii – the 7th (called "first front") and the 14th (called "first middle") from the right – were identified to maximally capture the articulatory differences between vocalic and consonantal gestures in terms of tongue vertical displacement in a relevant area of the oral cavity. The first front radius was seen to be responsible for the difference between the /a/ gesture and the consonantal gestures of /s/ and /l/, as well as for the difference between the /i/ gesture and the consonantal gesture of /k/. The first middle radius was seen to be responsible for the difference between the /i/ gesture and the consonantal gestures of /s/ and /l/, as well as for the distinction between /a/ and /k/;

(iii) Identification of a gestural stability area for the consonant involved and of the relevant Ctarget landmark. Once the reference spline for the vocalic gesture and the fan radius for capturing the C-V difference were established, the distance between the tongue position in the reference spline and the tongue position in all the splines included in the acoustic interval of the consonant along the selected fan radius was calculated (in millimetres). The temporal interval during which this distance reached its maximum was defined as the UTI gestural stability area of the consonant. So the UTI gestural stability area corresponded to the interval during which the tongue was maximally different (in terms of vertical displacement along the selected fan radius) from the vowel.

3.3.3 Gestural stability area: EPG evidence

A traditional procedure based on the contact index method (Fontdevila, Pallarès & Recasens, 1994) was elaborated for the identification of vocalic and consonantal gestures from EPG data. Since EPG allows the detection of the linguo-palatal contact during speech production, it is possible to get different kinds of information for various regions on the artificial palate by referring to different EPG indices, implemented in the AAA software.

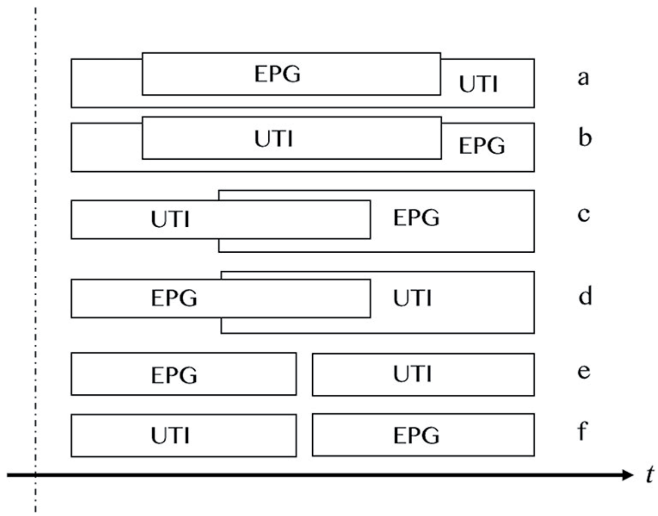
Two of these indices were used to detect the EPG gestural stability area. Specifically, the CAa index (contact anteriority in the anterior palate) was calculated on the three most anterior rows of each EPG frame included in the acoustic intervals of each /s/ and /l/ token. CAa served as an indicator of the anteriority of the linguopalatal contact of /s/ and /l/: the higher the CAa index, the more apical contact on the most anterior region of the palate. Consequently, the temporal intervals characterized by the maximum CAa values within the acoustic intervals of /s/ and /l/ were identified as the gestural stability areas of /s/ and /l/. For /k/, the Qp index (percentage of activated electrodes in the four back rows of the palate) was used: the higher the Qp value, the more contacted the posterior palate for the realization of the consonant. Thus the temporal intervals characterized by the maximum Qp values within the acoustic intervals of /k/ were identified as the gestural stability areas of /k/.

3.3.4 Integration of the two information sources: UTI + EPG gestural stability areas and calculation of gestural coordination latencies

We finally put the two gestural stability areas together, the intersection of them being identified as the stability area for the consonantal gesture, because both lingual profiles and palatal contact maximally distinguished such gesture from the adjacent vocalic gestures.

Six different conditions of temporal overlap of UTI and EPG stability areas were found in the data. They are schematically represented in Figure 2. The shadowed regions represent the final gestural stability areas (a-d). In only two items, the UTI and EPG stability areas were found not to overlap (e-f). In those cases, the EPG gestural stability area was taken as the gestural stability area.

Figure 2 - *Six idealised patterns of EPG-UTI gestural overlap, and resulting gestural stability areas*

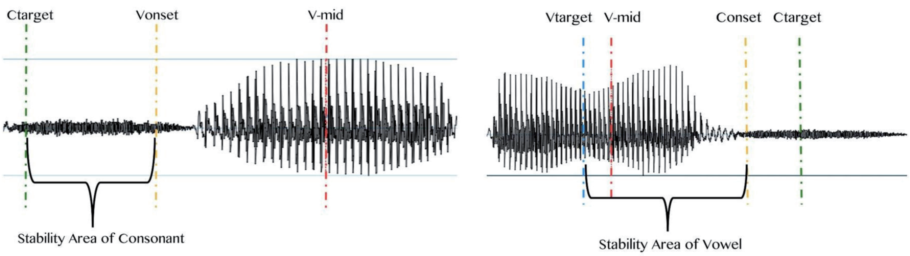


In CV sequences, the time point corresponding to the beginning of the gestural stability area was identified as Ctarget, whereas its offset was identified as Vonset (because we assume that when the position of the tongue starts to move from the consonantal target, the gesture for the following vowel begins). As anticipated in § 3.3.2, Vtarget corresponded to the acoustic midpoint of the vowel (Figure 3, left).

As to the VC sequence, the time point corresponding to the beginning of the gestural stability area was identified as Ctarget. A further step was therefore needed to identify Vtarget and Conset in VC sequences. In these cases, the UTI spline of Ctarget previously obtained was treated as the reference for the identification of the stability area for the vocalic gesture. The C-V distance along the same fan radius was calculated within the acoustic interval of the vowel. Similar to the third step described in § 3.4.1, the distance values plotted as a function of the temporal development gave a plateau representing the gestural stability area of the vowel involved. Then the same criteria were applied: the time point corresponding to the first spline of the gestural stability area was identified as Vtarget, the last one as Conset (Figure 3, right).

Once all the relevant gestural landmarks were identified for both CV and VC sequences, the gestural coordination latencies were calculated. The latency between Ctarget and Vonset was taken as a cue of CV coordination in CV syllables, whereas the latency between Vtarget and Conset was taken as a cue of VC coordination in CVC syllables.

Figure 3 - *Selection of temporal landmarks and identification of the gestural stability area in CV (left) and VC (right) sequences*



3.4 Analysis

The analysis of temporal coordination was based on the quantitative evaluation of gestural coordination latency variations, as illustrated above; the dependent variable was the milliseconds.

To evaluate if the three consonants differed for their coarticulatory properties, qualitative inspection of UTI profiles and quantitative evaluation of EPG indices variation over the acoustically defined consonantal interval were also run. As for EPG indices, CAa (contact anteriority in the anterior palate) for /s/ and /l/ and Qp (percentage of activated electrodes in the posterior palate) for /k/ were the dependent variables.

For both spatial analyses (temporal coordination and coarticulatory properties), the factors under evaluation were Syllable (CV vs. VC), Consonant (/k/ vs. /l/ vs. /s/) and Vowel (/a/ vs. /i/). The statistical analyses were univariate analyses of variance and the non-parametric Mann-Whitney test for distribution differences in independent samples. Non-parametric statistics was run whenever the requirements for ANOVA were not met (i.e., asymmetry and kurtosis higher than |1| and/or unequal variances, as attested by the Levene's test). The statistics were computed using SPSS (22.0.0).

4. *Results*

The results are presented in the following order: first, we present the results of the coarticulation analysis (lingual profiles and EPG indices); second, we present the results of the temporal coordination analysis.

4.1 Spatial coordination: coarticulatory patterns

4.1.1 Lingual profiles

Figures 4 to 6 display the average lingual profile for the three different consonants of the study (average of all lingual profiles included in the acoustic interval of the consonant of each stimulus) when adjacent to /a/ and to /i/, on a subject-by-subject comparison. In each figure, tongue root is to the left and tongue tip to the right of the images. The average tongue profile is represented by the darker line, whereas the upper and lower lighter lines identify the variance (standard deviation of the curves).

Figure 4 - *Comparison between average tongue profiles (with standard deviations) of /s/ adjacent to /a/ and /i/ in CV (upper half) and VC syllables (lower half) for the four speakers of the study. Upper graph: subject's palate in violet; /si/ in dark green. Lower graph: subject's palate in light blue; /is/ in black*

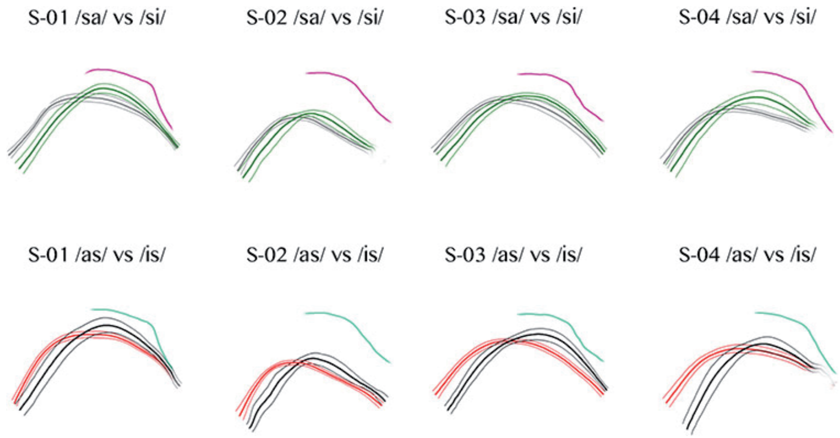


Figure 4 displays average tongue profiles during the production of /sa/ and /si/ (upper part) and /as/ and /is/ (lower part) across speakers. For all of them, the midsagittal profile is higher and more advanced in /si/ as opposed to /sa/; the difference is visible in the predorsal and dorsal areas, whereas in three out of four subjects the lingual profiles do not differ in the region of the anterior tongue (blade and tip). Only S04 appears to neatly differentiate the production of the sibilant in /sa/ vs. /si/ as far as the anterior tongue is concerned, with a higher position in /si/ than in /sa/. As far as coda /s/ is concerned, overall tongue advancement in the /i/ context is visible in all subjects, whereas anterior tongue advancement is present in one or maybe two out of four subjects (S03 and possibly also S02). Note also that for some of the speakers, the tongue tip is raised to a very limited extent, in both /a/ and /i/ contexts. These results therefore show that, although /s/ is produced with the lateral sides of the tongue raised towards the alveolar ridge, and the medial part of the tongue and the tongue tip slightly lowered to allow the air to escape from the front, there are nevertheless significant differences in the midsagittal profile as far as the dorsal and pre-dorsal areas are concerned. Additionally, there may be differences also in the case of the anterior tongue, both when /s/ is in syllable onset (one subject) and when it is in coda (two subjects). In one subject (S03), vowel-induced variation is larger in coda than in onset /s/.

Figure 5 - *Comparison between average tongue profiles (with standard deviations) of /l/ adjacent to /a/ and /i/ in CV (upper half) and VC syllables (lower half) for the four speakers of the study. Upper graph: subject's palate in violet; /li/ in dark green. Lower graph: subject's palate in light blue; /il/ in black*

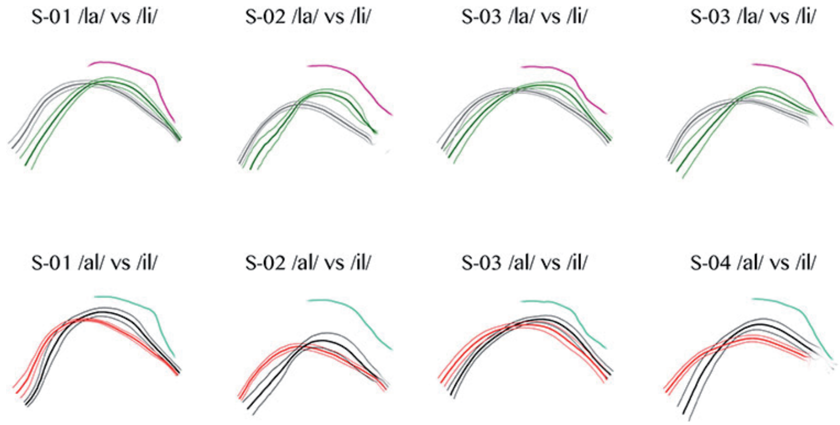
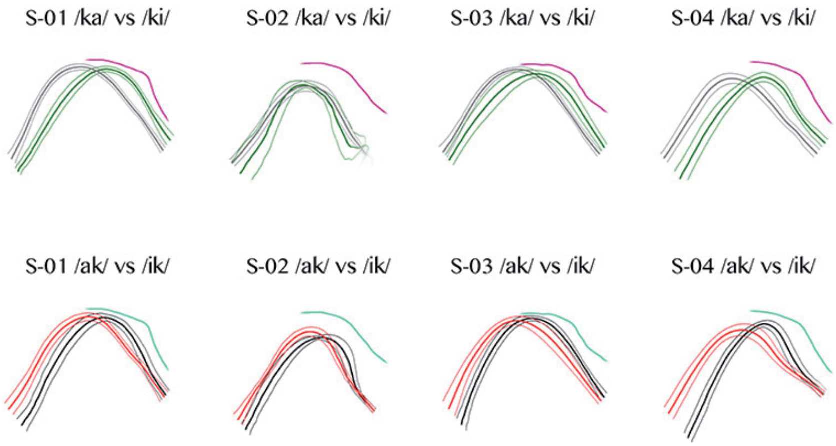


Figure 5 displays average tongue profiles during the production of /la/ and /li/ (upper part) and /al/ and /il/ (lower part) across speakers. As in the case of /s/, the tongue tip gesture is sometimes hardly visible. However, for all of the speakers, the onset consonant is produced with the tongue in a higher and more advanced position before /i/ than before /a/. However, the tongue tip (which is revealing of the constriction location for this consonant) is significantly higher in only two of the subjects (S02 and S04). In coda, tongue tip differences are visible in only one of the subjects (i.e., S04), thus suggesting that there is slightly less coarticulatory variation in coda than in onset position.

Figure 6 - *Comparison between average tongue profiles (with standard deviations) of /k/ adjacent to /a/ and /i/ in CV (upper half) and VC syllables (lower half) for the four speakers of the study. Upper graph: subject's palate in violet; /ki/ in dark green. Lower graph: subject's palate in light blue; /ik/ in black*



Finally, Figure 6 displays average tongue profiles during the production of /ka/ and /ki/ (upper part) and /ak/ and /ik/ (lower part) across speakers. As far as the onset consonants are concerned, three subjects (i.e., S01, S03 and S04) displayed neatly different tongue configurations, with overall tongue advancement before /i/, whereas S02 displayed undifferentiated tongue profiles in the two phonetic contexts. The pattern for coda /k/ is similar to that for onset /k/. Thus we can conclude that in three out of four subjects, the dorsal consonant displayed a more advanced lingual configuration when adjacent to /i/ than /a/, independently of its syllabic status.

4.1.2 EPG indices

The CAa index was expected to be higher when the consonant is adjacent to an /i/ than when it is adjacent to an /a/, because there is more fronting in the production of a high vowel than a low vowel. Table 2 shows the average CAa values of /s/ as a function of vocalic context and position in the syllable (onset vs. coda /s/). The data did not confirm the expectation regarding higher CAa values in /i/ than in /a/ sequences, since the Vowel factor did not differentiate among groups (CV items, Mann-Whitney test: $U = 201$, $z = 1.854$, $p > .05$; VC items: $U = 155.500$, $z = -.206$, $p > .05$). Thus the gesture for /s/ as measured by the CAa index turned out to be insensitive to variations in the vocalic context. However, there was a significant effect of Syllable, with coda /s/ exhibiting higher CAa values in coda than in onset position ($U = 817.500$, $z = 2.156$, $p < .05$). Thus the linguopalatal contact was overall significantly more fronted for coda than for onset /s/.

Table 2 - *Average CAa values for /s/ in CV and VC sequences as a function of vowel quality (/a/ vs. /i/)*

Syllable Type	*Vowel*	*N*	*average CAa*	*st. dev.*
CV	/a/	17	.918	.063
	/i/	18	.949	.032
	all	35	.934	.051
VC	/a/	18	.958	.024
	/i/	18	.953	.039
	all	36	.955	.032

Table 3 shows the average CAa values of /l/ as a function of vocalic context and position in the syllable (onset vs. coda /l/). The data confirmed the expectation regarding higher CAa values in /i/ than in /a/ sequences for onset /l/ ($U = 221$, $z = 2.657$, $p < .05$). By contrast, CAa did not change significantly across vocalic contexts in the case of coda /l/ ($U = 144$, $z = -.297$, $p > .05$). Thus, differently from /s/, the anteriority of the linguopalatal contact for /l/ was sensitive to variations in the vocalic context with more spatial overlap with the /i/ gesture when the consonant was in onset position. However, similar to /s/, there was also a significant effect of Syllable and the linguopalatal contact for /l/ was significantly more fronted in coda than in onset position ($U = 770.500$, $z = 2.107$, $p < .05$).

Table 3 - *Average CAa values for /l/ in CV and VC sequences as a function of vowel quality (/a/ vs. /i/)*

Syllable Type	*Vowel*	*N*	*average CAa*	*st. dev.*
CV	/a/	18	.951	.036
	/i/	16	.984	.017
	all	34	.967	.033
VC	/a/	17	.985	.025
	/i/	18	.982	.023
	all	35	.983	.024

The Qp index was expected to be higher when /k/ was adjacent to an /i/ than when it was adjacent to an /a/, because there is more dorsal activity in the production of a high vowel than a low vowel. Table 4 shows the average Qp values as a function of vocalic context and position in the syllable (onset vs. coda /k/). The data showed that, according to the expectations, Qp values were significantly higher close to /i/ than close to /a/ ($F(1,66) = 208.985$, $p < .001$). The interaction Syllable*Vowel was not significant ($F(3, 66) = 1.997$, $p > .050$), thus suggesting that the difference was equally present in CV and VC sequences.

Table 4 - *Average Qp values for /k/ in CV and VC sequences as a function of vowel quality (/a/ vs. /i/)*

Syllable Type	*Vowel*	*N*	*average Qp*	*st. dev.*
CV	/a/	18	.105	.036
	/i/	16	.283	.054
	all	34	.189	.101
VC	/a/	16	.105	.035
	/i/	17	.252	.050
	all	33	.181	.086
Total	/a/	34	.105	.036
	/i/	33	.267	.054

4.1.3 Summary of coarticulation results

The cumulative evidence of UTI lingual profiles and EPG contact indices suggested that there were varying patterns of coarticulatory activity and varying degrees of coarticulatory resistance across vowel contexts. However, we found variable coarticulation across syllable contexts and (limited to UTI data) across subjects as well. This provided additional information about the dynamics of spatial overlap of /s/, /l/ and /k/in different /a/ and /i/ contexts, proving that the way in which the articulatory gestures of an onset consonants vary to anticipate the lingual properties of the following vowel may be different from the way in which the same consonant, when in coda, accommodates to the lingual properties of the vowel, which precedes.

The sibilant was expected to show a comparatively higher degree of coarticulatory resistance, i.e., to vary the least as a function of vocalic variations. The analysis of linguopalatal contact in the anterior palate during /s/ production confirmed the hypothesis, inasmuch as no variation was found in the /a/ vs. /i/ contexts. However, the midsagittal lingual configuration was slightly different in the two vocalic contexts, according to the UTI analysis. Thus the tongue accommodated to the position required for the vowel (by reaching an overall more anterior position in the /i/ context and, for some of the speakers, a higher blade and tip), at the same time warranting the achievement of the target constriction with central groove in the anterior palate. There was, however, syllable-induced variation, with coda /s/ more fronted (both on average and for some specific speakers in particular) than onset /s/.

The lateral consonant was expected to vary more than /s/ as a function of vocalic variation, thus exhibiting less coarticulatory resistance than the sibilant. The expectation was met, inasmuch as the linguopalatal contact was found to vary in a statistically significant way, at least when the consonant was in onset position. However, the lateral was similar to the sibilant in two aspects. First, both consonants showed vowel-induced variation in the UTI midsagittal view, with an overall more anterior tongue configuration and – for a subset of the subjects – a higher blade and tip in the /i/ context. This variation affected the anteriority of the linguopalatal contact in /l/ only, consistently with the observation that /l/ is produced with a central apical contact. Second, both consonants exhibited syllable-induced variation, being more fronted in coda than in onset position.

Finally, the dorsal consonant showed a consistently variable lingual configuration and a statistically significant change in anteriority values as a function of changing vocalic contexts. The tongue during the production of /k/ was overall more advanced when adjacent to an /i/ than to an /a/. The fact that the vocalic nucleus precedes or follows the consonant did not impact on its coarticulatory modification pattern, which was strong in both syllabic contexts.

We could therefore conclude that, on the basis of additive evidence from midsagittal tongue imaging and linguopalatal contact analysis, a hierarchy can be confirmed for the three consonants of the study, with /s/ exhibiting the highest, /k/ the lowest and /l/ an intermediate degree of coarticulatory resistance.

4.2 Temporal coordination of C and V gestures

In this section, we verified if there is variation in the consonant-vowel temporal coordination patterns according to the different syllabic status of the consonant (onset vs. coda). In addition, we verified whether the coordination patterns are affected by the articulatory properties of the consonants involved, with specific reference to the degrees of coarticulatory resistance outlined above.

Table 5 shows the gestural coordination latencies across stimuli and subjects.

For the onset-nucleus coordination (CV syllables), the mean value of the Ctarget-Vonset interval across all valid tokens was 38 msec. For the nucleus-coda coordination (VC sequences in CVC syllables), the mean value of the Vtarget-Conset interval across all valid tokens was 105 msec. The temporal interval occurring between the consonantal gesture and the vocalic gesture in onset-nucleus sequences was therefore much shorter than the temporal interval between the vocalic gesture and the consonantal gesture in nucleus-coda sequences (negative difference: 67 msec), which suggested an opposition between tight coordination (quasi-simultaneous activation) and loose coordination (sequential activation) in CV vs. VC sequences. In an ANOVA with C-V latency as dependent variable and Syllable as between subject factor, the difference between CV and VC was statistically significant ($F(1, 205) = 523.869$, $p < .001$).

Table 5 - *Gestural coordination latencies in CV and VC sequences across subjects, vocalic contexts and prosodic conditions*

Syllable Type	*N*	*mean (ms)*	*st. dev.*
CV	103	38	20
VC	102	105	20.4

Table 6 additionally shows the gestural coordination latencies as a function of consonantal (/k/, /l/, /s/) and vocalic variations (/a/, /i/).

The mean values confirmed that the sharp distinction between CV and VC coordination patterns was present in all three consonants of the study, with a negative difference between the two conditions of 55 msec for /k/, 67 msec for /s/and 77 msec for /l/.

The interaction between Syllable and Consonant turned out to be statistically significant ($F(5, 201) = 5.906$, $p < .05$), thus confirming that the coordination pattern was different across consonants. In particular, post-hoc Tukey HSD test revealed that /l/ was significantly different from both /s/ ($p < .010$) and /k/ ($p < .050$), whereas /s/ and /k/ were not significantly different. Average latency values for VC sequences were similar across consonants, whereas for CV sequences there was a much lower latency value for /l/ (27 msec on average) than for the two other consonants (45 msec for /k/ and /42 msec for /s/).

The difference between CV and VC coordination was equally present in /a/ (70 msec) and /i/ stimuli (65 msec), as confirmed by the non-significant Syllable by Vowel interaction ($F(3, 201) = 0.713$, $p > .05$).

We could therefore conclude that the properties of the consonantal gesture influenced the gestural coordination patterns, whereas vowel quality did not.

Table 6 - *Gestural coordination latencies in CV and VC sequences as a function of consonantal and vocalic variations*

		CV			*VC*		
C	*V*	*mean (ms)*	*st. dev.*	*N*	*mean (ms)*	*st. dev.*	*N*
k	a	48	14	18	109	16	16
	i	41	10	16	91	13	17
	Total	45	13	34	100	17	33
l	a	29	13	17	104	22	16
	i	24	13	16	104	27	17
	Total	27	13	33	104	24	33
s	a	34	19	17	108	14	18
	i	51	29	18	109	16	16
	Total	42	26	35	109	15	34
Total	a	37	17	52	107	17	50
	i	39	22	50	101	21	50
	Total	38	20	102	104	19	100

5. *Discussion*

This study had two purposes. The first was that of implementing a methodologically challenging experimental setting, designed to advance our knowledge of intersegmental coordination and coarticulation and the relationship between the two; the second was that of verifying whether intersegmental coordination is influenced by the articulatory properties of the segments in a selection of stimuli produced by Italian speakers.

Concerning the first purpose, the experimental recordings were successfully run in *SynchroLing* and the speech materials were easily and accurately analysed within the AAA software environment, which allowed us to manage complex data sources in a unified environment. Therefore, *SynchroLing* proved itself a fruitful instrument for the processing of multi-level speech data. It invites further explorations in the domain of speech instruments synchronization and development.

The data provided by *SynchroLing* were used to characterize the dynamics of coarticulation and coordination between adjacent segments. This enriched window over speech movements was exploited here in two different ways. On one hand, it was used to determine the temporal dynamics of the consecutive constriction and opening gestures in sequences of consonants and vowels. This allowed us to identify onset and offset gestures from multiple sources of information, which cover together ample areas of the articulatory organs (as opposed to e.g. EMA-based tracing of individual coil movement). On the other hand, the multi-level analysis provided us with information on spatial adjustments of consonant and vowel lingual gestures from a variety of perspectives, e.g. by showing the effects of variable constriction points for those lingual regions that are not involved in the constriction themselves.

This pilot study was very limited in scope and additional work will be needed to improve methodological aspects and develop more sophisticated techniques and measures to map UTI-based tongue profile information onto EPG-based palate surface information (and viceversa). This will provide advantages both for articulatory phonetics (e.g. to quantify cross-instrument reliability of instrument-specific measures) and for phonetic theory (e.g. to determine the accommodation patterns of the tongue back to movements in the tongue front etc.) (see also Recasens, Rodriguez 2016; Spreafico *et al.*, 2016 for similar considerations).

A production experiment was run to investigate the difference between syllable onsets and codas in their temporal coordination with the syllable nucleus. The hypothesis that we wanted to test with Italian CVCV and CVC word stimuli was based on well-known effects of syllable structure on gestural coordination (see above, § 2). In particular, we tested whether onset C gestures are more tightly coordinated to the vocalic nucleus than C gestures in syllable codas. The hypothesis was confirmed in the present study, by showing that the temporal interval between Ctarget and Vonset in CV sequences is much smaller than the temporal interval between Vtarget and Conset in VC sequences. We use this evidence to suggest that the procedure for the identification of temporal landmarks in *SynchroLing* (see above, § 3.4) was reliable and that, by looking at EPG and UTI stability areas, it was possible to correctly identify the onset and offset of consonantal and vocalic gestures in the mono-and disyllabic experimental stimuli used here.

An additional hypothesis concerned the possibility that this coordination pattern varied as a function of the specific consonants involved. This hypothesis was suggested by the finding that, in some languages, different consonants may have different temporal behaviours within the syllable and the onset-coda opposition can show up differently according to the articulatory properties of the segments involved. As reviewed above (see § 2), the most frequently investigated coordination pattern in this respect is the so-called c-center effect for consonant clusters. The c-center effect is a consequence of tight vs. loose coordination of syllable onsets vs. codas and its magnitude and regularity have been shown to be influenced by segmental properties such as the degree of resistance to coarticulation with neighbouring segments.

The analysis presented here was limited to singleton onsets and codas; as already mentioned, clusters will be dealt with in a prosecution of the study. We hypothesized that different consonants may be timed to the vocalic nucleus as onsets or codas in different ways. In particular, less resistant consonants were expected to show a larger onset-coda coordination difference than more resistant consonants. To test this hypothesis, /s/, /l/ and /k/ were compared; /s/ was expected to be the most resistant, /k/ the least resistant and /l/ to occupy an intermediate position between them. Consequently, /s/ was expected to show a smaller onset-coda coordination difference than /l/, and /l/ was expected to show a smaller onset-coda coordination difference than /k/.

We discuss the results of the coarticulation analysis first. The multi-level approach adopted in this study provided us with information on variation in both constriction location and uncontacted portions of the tongue (midsagittal profile). Constriction location was predominantly evaluated on the basis of EPG index values. In this respect, /s/ was found to maintain the same linguopalatal constriction location before /a/ and /i/, in both onset and coda positions; /l/ showed fronting before /i/ in onset position only; /k/ changed in both onset and coda positions. EPG evidence appeared therefore to support our hypothesis concerning coarticulation degrees across consonants. The target constriction for /s/ notoriously requires fine control over the articulators in order to narrow the air channel and sustain turbulence during frication, which can explain the absence of vowel-induced variation. However, the same EPG indices also revealed that /s/ and /l/ were overall more fronted in coda than in onset. This effect of syllable position on constriction location did not generalize to /k/. According to Recasens (2004: 451-452), in Catalan /VC_1#C_2V/ stimuli /s/ and /l/ are more fronted when they are in coda position (C_1) than when they are in onset (C_2), provided that the adjacent consonant is articulatorily less constrained at tongue front. By contrast, if the adjacent consonant is more constrained at tongue front (e.g. /ʃ/), then the pattern is reversed (i.e., /s/ and /l/ have higher fronting values in onset than in coda). Since our stimuli include intervocalic /s/ and /l/, and vowels have a less constricted lingual configuration than the two apical consonants, we can conclude that our results are consistent with what has been observed for Catalan and lend further support to the generalizations stemming from an account of coarticulation patterns in terms of segments coarticulatory resistance.

Besides constriction location, tongue dorsum and postdorsum as revealed by UTI images were found to vary more for /l/ than for /s/. For both consonants, tongue tip position changed as a function of vowel quality only in a minority of the cases, thus indirectly confirming the results of the EPG analysis. However, there were nevertheless significant differences in the midsagittal profile as far as the dorsal and pre-dorsal areas were concerned. This clearly revealed that, even for strongly coarticulation-resistant consonants, such as /s/ (and slightly less resistant consonants, such as /l/), areas of the tongue that are not directly involved in constriction realization may accommodate to the articulatory requirements of the adjacent vowel, without substantially modifying the position and the amount of linguo-palatal contacts. In the UTI analysis, /s/ and /l/ were therefore similar with respect to coarticulation (whereas EPG data differentiated more clearly between the two). However, while /s/ showed (in one of the subjects) more vowel-induced variation in coda than in onset, /l/ showed (in two of the subjects) the opposite pattern, i.e. less vowel-induced variation in coda than in onset. Whether and how this finding is important for the overall picture of /s/ and /l/ coarticulatory behaviour sketched so far should be ascertained in future analyses, using quantitative measures of tongue profile variations and possibly comparing the behaviour of a larger number of informants. The tip and blade of the tongue were also found to consistently and significantly vary before /a/ as compared to /i/ for the dorsal consonant /k/. This finding suggested that variation in constriction location additionally implies a change in the overall tongue configuration. Thus UTI-based analysis apparently highlights a basic distinction between dorsal /k/, whose place of articulation changes a lot as a function of a modification of the vocalic context, and the two anterior consonants, that still change, but not necessarily for constriction location, and to a smaller extent.

In sum, the multi-level approach adopted here was found to increase the amount of information on vowel-induced modifications in lingual consonant production, thus showing up as a potentially relevant way of improving our understanding of coarticulation patterns. There are major and minor differences among consonants. Major differences (e.g. the difference between dorsal and apical consonants) are highlighted by both EPG and UTI. Minor differences (such as the extent to which the overall configuration of the tongue accommodates to varying constriction locations) may emerge differently, according to the specific source of articulatory data from which the information is drawn. More sophisticated ways of establishing correspondences and overlapping functions between the two instrumental outputs (as well as between the articulatory output and the acoustic data) have to be developed, to fully exploit the advantages of the multi-level analysis of speech movements.

We additionally tested the hypothesis that the onset-coda coordination difference varied as a function of the specific consonants involved. We found that the three consonants did vary in the way they differentiate onset-nucleus from nucleus-coda coordination, but not in the expected direction, which was a larger latency difference for less resistant consonants and a smaller latency difference for more resistant consonants. As a matter of fact, our findings met the expectations to the extent that the latency difference was bigger for /l/ than for /s/, but the finding that the latency difference for /k/ was even smaller than for /s/ did not fit our initial hypothesis. In a comparison between /s/ and /l/, the lateral showed a very short CV latency in onset position compared to /s/. This finding can be interpreted that, consistently with our initial hypothesis, not only the articulation of the following vowel influences the constriction location of the lateral more than it influences the constriction location of the sibilant, but also its temporal anticipation is stronger in the case of an onset /l/ than of an onset /s/. However, if this was the case, onset /k/ should have shown an even stronger anticipation of the vowel and a consequently shorter CV latency than /l/, which was not the case.

We think that there are at least two possible explanations for these findings. The first is that the onset-coda coordination effect is actually *not* affected by the coarticulatory resistance of singleton consonants. Resistance to coarticulation has been found to influence the c-center effect in clusters and maybe singletons are not as prone as clusters to suffer the influence of varying articulatory conditions. The second possibility is that the onset-coda coordination effect *is* affected by the coarticulatory resistance of the singletons, but only for a subset of them. In particular, one could hypothesize that it is only within the range of apical consonants that the degree of articulatory constraint determine the temporal effects of segments coordination; dorsal consonants could be hypothesized to be influenced by adjacent vowels only in terms of variable tongue configuration and changing constriction location. This difference might be due to the fact that dorsal consonants are produced by the movement of a less flexible and therefore slower part of the tongue, which might obscure the effects of tight C-V coordination in syllable onsets. The finding that /k/ showed the smallest difference between onset C-V and coda V-C coordination latencies (55 msec, as opposed to 67 msec for /s/ and /77/ msec for /l/) might bring support to this view, since it shows that the vocalic gesture is not strongly anticipated during the production of onset /k/. However, to be fully ascertained this hypothesis should be tested by comparing two dorsal (or non-apical) consonants with different degrees of coarticulatory resistance, to verify if the same hierarchy found for /l/ and /s/ applies to consonants produced more in the back in the oral cavity. Additional experiments would therefore be needed.

Bibliography

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

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

Celata, C., Vietti, A. & Spreafico, L. (in press). An articulatory account of rhotic variation in Tuscan Italian: synchronized UTI and EPG data. In Gibson, M., Gil, J. (Eds.), *Romance Phonetics and Phonology*. Oxford: Oxford University Press.

De Jong, K.J. (2003). Temporal constraints and characterizing syllable structure. In Local, J., Ogden, R. & Temple, R. (Eds.), *Papers in laboratory phonology*, 6. Cambridge: Cambridge University Press, 253-268.

Fontdevila, J., Pallarès, M.D. & Recasens, D. (1994). The contact index method of electropalatographic data reduction. In *Journal of Phonetics*, 22, 141-154.

Fowler, C., Brancazio, L. (2000). Coarticulation resistance of American English consonants and its effects on transconsonantal vowel-to-vowel coarticulation. In *Language and Speech*, 43(1), 1-41.

Hermes A., Mücke, D. & Grice, M. (2013). Gestural coordination of Italian word-initial clusters: the case of 'impure s'. In *Phonology*, 30(1), 1-25.

Krakow, R.A. (1989). The Articulatory Organization of Syllables: A Kinematic Analysis of Labial and Velar Gestures. PhD Dissertation, Yale University.

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

Marin, S., Pouplier, M. (2010). Temporal organisation of complex onsets and codas in American English: Testing the prediction of a gestural coupling model. In *Journal of Motor Control*, 14(3), 380-407.

Marin, S., Pouplier, M. (2014). Articulatory synergies in the temporal organization of liquid clusters in Romanian. In *Journal of Phonetics*, 42, 24-36.

Öhman, S.E.G. (1966). Coarticulation in VCV utterances. In *Journal of the Acoustical Society of America*, 39, 151-168.

Pastätter, M., Pouplier, M. (2015). Onset-vowel timing as a function of coarticulation resistance: Evidence from Articulatory Data. In The Scottish Consortium for ICPhS (Ed.), *Proceedings of the 18th International Congress of Phonetic Sciences*, Glasgow, UK, the University of Glasgow. Paper number 783.

Recasens, D. (2004). The effect of syllable position on consonant reduction (evidence from Catalan consonant clusters). In *Journal of Phonetics*, 32, 435-453.

Recasens, D. (2012). A study of jaw coarticulatory resistance and aggressiveness for Catalan consonants and vowels. In *Journal of the Acoustical Society of America*, 132(1), 412-420.

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

Recasens, D., Pallarès, M.D. & Fontdevila, J. (1997). A model of lingual coarticulation based on articulatory constraints. In *Journal of the Acoustical Society of America*, 102, 544-561.

Recasens, D., Rodríguez, C. (2016). A study of coarticulatory resistance and aggressiveness for front lingual consonants and vowels using ultrasound. In *Journal of Phonetics*, 59, 58-75.

Spreafico, L., Celata, C., Vietti, A., Bertini, C. & Ricci, I. (2015). An EPG+UTI study of Italian /r/. In The Scottish Consortium for ICPhS (Ed.), *Proceedings of the 18th International Congress of Phonetic Sciences*. Glasgow, UK, the University of Glasgow. Paper number 775.