International Conference
ICT for Language Learning

Investigating Speech Gesture Coordination in the Production of L2 Consonant Clusters: the Added Value of Electromagnetic Articulography

Sonia d'Apolito[1], Barbara Gili Fivela[1]

Abstract

The study describes how non-native consonant clusters are realized by Italian learners of French L2. In particular it aims at describing articulatorily the coordination across successive consonant gesture in order to understand which strategies are at play when learners produce non-native clusters. The electromagnetic articulography (AG500), a technology that allows a precise tracking of position and movement of tongue, lips and jaw during speech, is employed to observe L2 productions. The attention is focused on the production of sibilant clusters since they are phonotactically marked in Italian, while in French are frequent so much so that place assimilations occur. Sibilant clusters were studied across words into a carrier phrase at different speech rate (normal/fast) as a faster rate can facilitate coproduction. In French, results show that vowel insertion is realized only at normal speech rate, while at faster speech rate place assimilations are realized instead. Italian learners' of French-L2main strategy is the vowel insertion between the two sibilants in order to repair the non-native clusters, independently of speech rate. On the acoustic level, learner's productions show a dichotomic result (vowel segment between consonants/no segment). On the articulatory level, learners' productions reveal a greater variability of articulatory strategies, due to their attempts to produce non-native clusters. However, speech gesture tracking data suggest that the inserted vowel results from the failure in reaching a felicitous coordination between gestures. Indeed, vowel insertion seems to be due to a transitional passage between the two consonants. These results highlight how it is important not only to learn to produce correctly a segment but also to concatenate it with other segments. The electromagnetic articulography, capturing tongue/lip movement and position, is crucial for investigating these features of L2 speakers' production.

1. Introduction

This study focuses on articulatory strategies that Italian leaners of French as L2 realize during the production of non-native consonant clusters. The theoretical account to which we refer is the Articulatory Phonology [1, 2, 3] which describes speech in terms of articulatory gestures which occur in space and in time. Their coordination and overlapping can account for a variety of phonetic and phonological consequences of what is well known as coarticulation [3]. In order to observe the coordination and the overlapping between neighbouring consonant gestures of a cluster as well as which articulatory strategies are at play in L2, a detailed articulatory investigation of L2 productions has been performed using electromagnetic articulography (AG500). This is a high-precise technology which allows to measure position and movements of sensor coils placed on tongue, lips and jaw during speech, obtaining tracking data that are time-aligned with audio recordings. Such information are of great importance in phonetics and speech pathology, as they are useful to understand the way sounds are produced and how the articulators are coordinated with each other. Here, our focus is the production of French sibilant sequences across word boundary, as they are very frequent in French so much so that place assimilations occur [4], but they are also phonotactically marked according to Italian phonetics and phonology [5]. Therefore, Italian leaners have troubles in producing such sequences and may quite easily insert a vowel segment between the two sibilants as general articulatory repair strategy [6]. Articulatorily, this may correspond to a full vowel insertion (i.e. there is an articulatory target) or to a gestural mistiming (i.e. no articulatory target is detected) because learners fail to coordinate the consonants by sufficiently overlapping them [7,8]. Moreover, speech rate has been taken into account as an important factor for coarticulation since a faster speech rate can facilitate the coproduction between successive segments [9].

2. Goal and hypotheses

This articulatory study (AG500) will allow us to observe in-depth the articulatory strategies that Italian learners realize during the production of non-native marked sequences and to understand if the vowel

[1] University of Salento, Centro di Ricerche Interdisciplinare sul Linguaggio (CRIL), Italy

insertion corresponds to a full vowel insertion, in order to reorganize the sequence VC#CV as VCV#CV according to the influence of mother tongue's characteristics, or rather a gestural mistiming takes place, due to a low degree of overlapping between consonant gestures. Our hypothesis is that the vowel insertion is a general strategy to repair marked sequences due to strong influence of mother tongue. However, from articulatory point of view a gestural mistiming is also expected due to learners' difficulties in coordinating correctly the consonant gestures. Concerning the influence of speech rate, it is expected a stronger coarticulation at faster rate and a reduction of vowel insertion.

3. Method

Target heterosyllabic sibilant clusters /sʃ - ʃs - sʒ - ʒs - zʃ - zʒ/ were studied in a /a_i/ surrounding vowel context, proposed within a carrier sentence in which stimuli were inserted within a phonological phrase (e.g. "Il dit sa**ge sy**rien rapidement"; *He said Syrian wise quickly*). Three Italian advanced learners of French L2 (PI1, PI2, PI3) and two French native speakers (PF4, PF5) read 7 times the randomized phases in the corpus, both at normal and at faster speech rate (84 phrases per speaker= 6 clusters x 2 speech rates x 7 repetitions). Acoustic and articulatory data were collected at CRIL laboratory (University of Salento). Articulatory data were collected by placing coils on the tongue midsagittal plane (4), on upper and lower lip (2), upper and lower incisors (2) and behind the ears (2) for normalization. The audio signal was recorded simultaneously onto a Edirol sound card at 44.1kHz. Data were analyzed auditorily, acoustically and articulatorily by means of PRAAT [11] and MATLAB scripts [12].

4. Analyses

4.1 Auditory and acoustic analysis

An auditory analysis was performed by both authors who verified independently the realization of the sequences (vowel insertion, assimilation, deletion). In the acoustic signal, segment boundaries were labeled in the $V_1C_1\#C_2V_2$ including possible inserted vowels (V0) and/or pauses (P). Acoustic measurements were: i) each segment duration; ii) sequence duration; iii) formants (F1, F2) in /a/, /i/ and V0, if present, calculated at the vowel steady-state.

4.2 Articulatory analysis

As for the articulatory labeling, closing and opening gestures were labeled on the following trajectories (see Fig.1):

- Tongue Tip (TT): vertical (z-axis) and horizontal (x-axis) movements;
- Tongue Dorsum (TD): vertical (z-axis) and horizontal (x-axis) movements for the V-to-V articulation ([a]-to-[i]);
- Lower Lip (LL): horizontal (x-axis) movement for lip protrusion.

The following measurements were then calculated:

- Duration (ms) and displacement (mm) of closing gesture for each fricative;
- Absolute timing (ms) between the target of the second and that of the first fricative (C2-to-C1);
- Relative phasing (ms): the ratio between C2-to-C1 interval and the duration of [a]-[i] articulation.

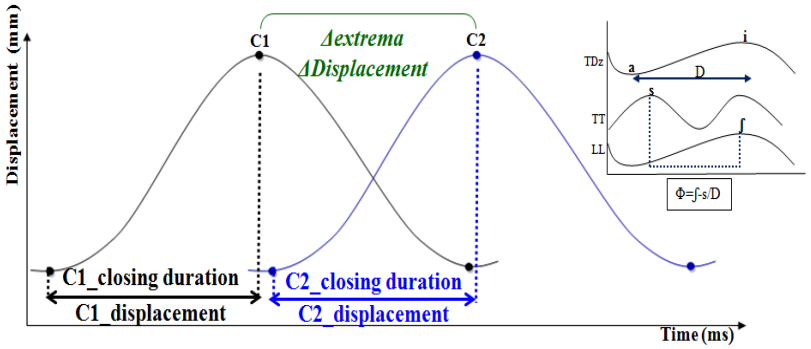


Figure 1: Schema of articulatory measurements

Non-parametric Kruskal-Wallis as well as Mann-Whitney tests were performed using SPSS, comparing data for each subject separately ($p<0.05$). For space limits, only the graphical representation of significant differences in the realization of the clusters are reported and discussed (for details on statistical analyses see [12]).

5. Results

At normal speech rate, all speakers insert a vowel within both clusters (alveolar-postalveolar; postalveolar-alveolar) – see Fig. 2. As the graphs show, few cases are found in which the Italian learner PI3 and the French speaker PF5 do not realize vowel insertion, 16,7% and 12,5% respectively (PI3: 5 cases out 30 for the alveolar-postalveolar cluster; PF5: 2 cases out 16 for the postalveolar-alveolar cluster).

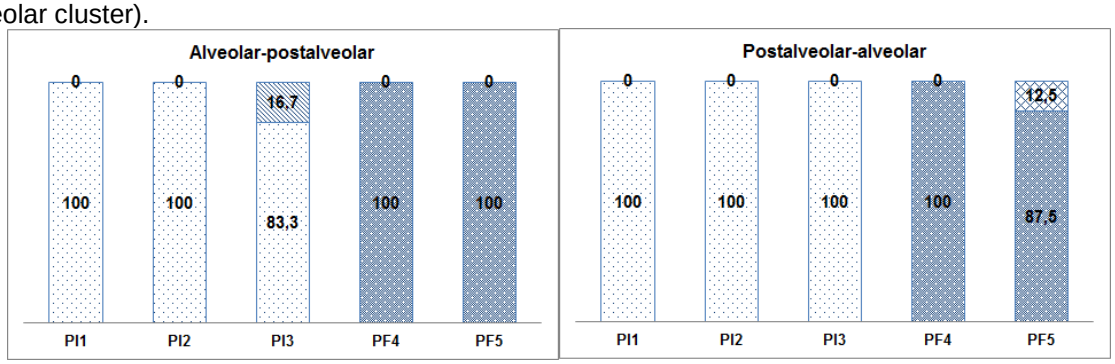


Figure 2: Vowel insertion frequencies at normal speech rate for all speakers.

In fast speech, as the graph in Fig. 3 clearly shows, the Italian learner PI1 and both French native speakers do no insert a vowel at all, while the Italian learners PI2 and PI3 keep inserting a vowel above all for the postalveolar-alveolar cluster.

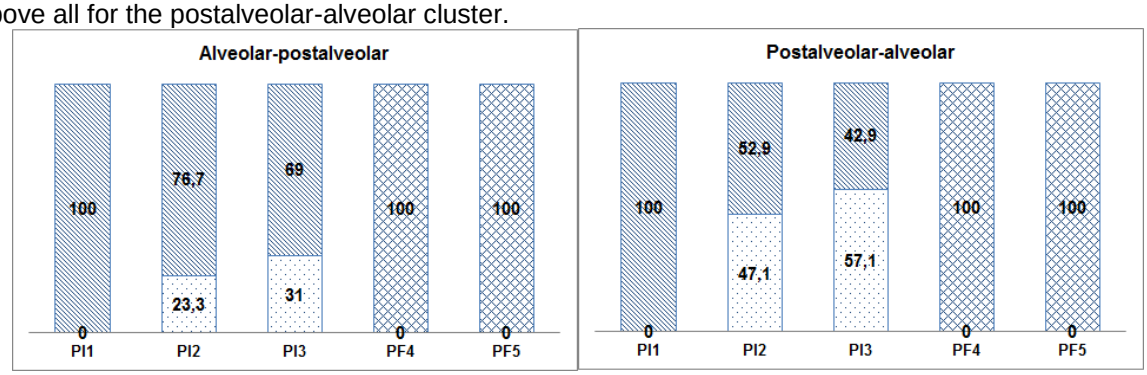


Figure 3: Vowel insertion frequencies at fast speech rate for all speakers.

As for articulatory results, we focused our attention on tongue dorsum movement trajectory in order to identify a possible target corresponding to vowel insertion. In both speech rate productions, the targets for [a] and [i] gestures were easily found while no specific target for vowel insertion was identified. In particular, on tongue dorsum trajectory a dip was visible as an opening space sufficient for an acoustic output. This dip coincided with an opening gesture as no linear interpolation between the two consonant targets was found. Indeed, on the tongue tip trajectory the consonant gesture for fricatives as well as an opening gesture were identified. The schema in fig. 4 represents the typical articulatory pattern found in both clusters at normal speech rate (alveolar-postalveolar; postalveolar-alveolar). As the schema shows an opening gesture is realized between the two consonant targets.

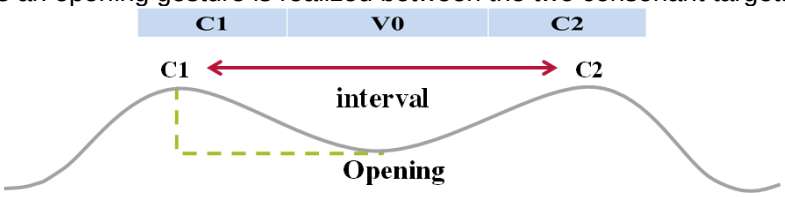
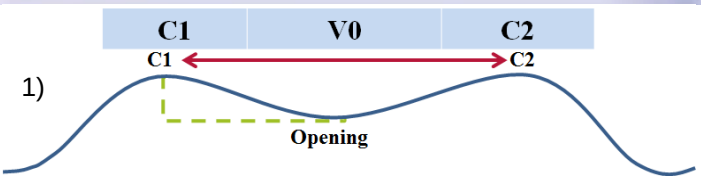


Figure 4: Schema for acoustic output (above) and articulatory trajectory movement (below).

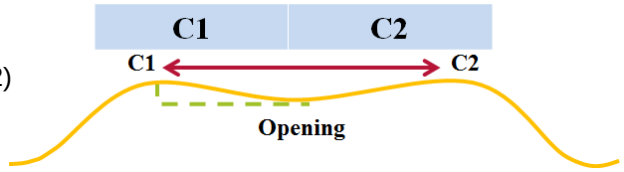
At faster speech rate, learners' productions reveal a greater variability of articulatory strategies. The Italian learner PI1 and the two native speakers produce the clusters with a high degree of gesture overlapping realizing different coarticulatory process such as place assimilation or deletion of the first consonant. In any case, on tongue dorsum trajectory only the passage from [a] to [i] was visible. On the contrary, the Italian learners PI2 and PI3 produce the sequences with a lower degree of coarticulation realizing three articulatory patterns:

- an opening gesture with an acoustic output, as represented in 1)



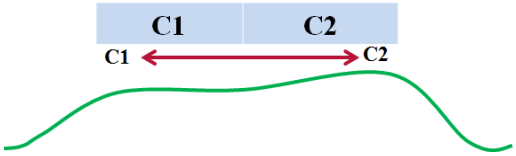
1)

- an opening gesture without an acoustic output between the two fricatives, as represented in 2)



2)

- no opening gesture and no acoustic output between the two fricatives, as represented in 3)



3)

The three articulatory patterns are realized within the alveolar-postalveolar sequence, while for the postalveolar-alveolar sequence only the first two patterns appear. As the graphs in Fig. 5 show, in the above mentioned cases the first pattern has the longer interval duration and it differs significantly from the second pattern (except for PI2 for the alveolar-postalveolar sequence) and from the third one (asterisks indicate significant differences).

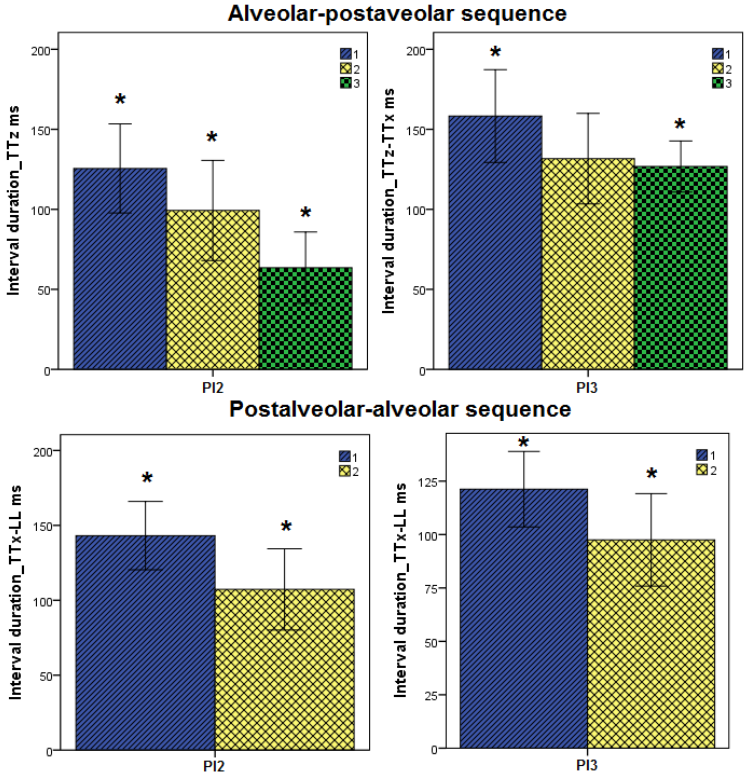


Figure 5: Absolute timing for the alveolar-postalveolar (above) and postalveolar-alveolar sequence (below) for PI2 (on the left) and PI3 (on the right).

6. Discussion and conclusion

In this study, both acoustic and articulatory (AG500) data have been collected to observe: 1) the articulatory strategies that Italian learners of French-L2 realize during the production of non-native marked consonantal sequences; 2) if the (frequent) vowel insertion corresponds to a full vowel or rather to a gestural mistiming; and 3) an expected stronger coproduction in faster speech rate.

In line with our hypotheses, results clearly reveal the influence of mother tongue's phonetics and phonology, as the typical strategy is the vowel insertion between two fricatives in order to repair a marked structure. Acoustic results show a dichotomic result (vowel segment between consonants/no segment), while articulatory results show a greater variability. At normal rate, the articulatory pattern is

characterized by an opening gesture between the two consonant targets with an acoustic output. At faster speech rate the coproduction between segments is stronger. In particular, PI1 and the two French native speakers do not insert any vowel, realizing place assimilation and/or deletion of a consonant. On the contrary PI2 and PI3 keep inserting a vowel realizing three articulatory patterns: 1) an opening gesture with an acoustic output; 2) an opening gesture without an acoustic output; and 3) no opening gesture and no acoustic output.

Variation in gestural patterns suggest that the inserted vowel may result from a gestural mistiming because learners fail in reaching a felicitous coordination between gestures (in line with [7]), taking also into account that no specific target is found for the inserted vowel on the tongue dorsum trajectory. Therefore, articulatory data capture fine tongue movements, showing the strategies followed by learners in their attempts to produce non-native clusters. On the contrary, no vowel insertion is observed in native speaker productions. Thus the opening gesture, with and without an acoustic output seems to be due to a transitional passage from one consonant to the other in the clusters [13], at least at normal speech rate.

To conclude, these results show that the electromagnetic articulography, capturing tongue/lip movement and position, is crucial for investigating in-depth features of L2 speakers' production and how it is important to take into account the coordination between segments. Moreover, it can also be used as learning and/or corrective pronunciation training because articulatory target can be presented in real-time [14].

References

[1] Browman, C.P., Goldstein, L., Towards an Articulatory Phonology, Phonology Yearbook, 3, 1986, 219-252.

[2] Browman, C.P., Goldstein, L., Tiers in Articulatory Phonology with some implications for casual speech, in Haskins Laboratories Status Report on Speech research SR-92, 1987, 1-30.

[3] Browman, C.P., Goldstein, L., Articulatory gestures as phonological unit, Phonology, 6, 1989, 201-251.

[4] Niebuhr, O., Lancia, L, Meunier, C., On place assimilation in French sibilant sequences, in Proceedings of the VII International Seminar on Speech Production, Strasbourg France, December 8-12, 2008, 221-224.

[5] Eckman, R., F., Typological markedness and second language phonology, in Zampini & Hansen (Eds.), CUP, 2008.

[6] Davidson, L., Phonology, Phonetics, or frequency: Influences on the production of non-native sequences, Journal of Phonetics, 34, 2006, 104-137.

[7] Davidson, L., Stone, M., Epenthesis versus gestural mistiming in consonant cluster production: An ultrasound study, in Proceeding of the West Coast Conference on Formal Linguistics, MA: Cascadilla Press, 2003, 165-178,

[8] Davidson, L., Addressing phonological questions with ultrasound, Clinical Linguistics & Phonetics, 19(6/7), 2005, 619-633.

[9] Byrd, D. & Tan, C., C., Saying consonant clusters quickly, Journal of Phonetics, 24, 1996, 263-282.

[10] Boersma, P., Weenink, D., PRAAT: Doing phonetics by computer, version 5.2, 2010.

[11] Sigona, F., Stella, A., Grimaldi, M., Gili Fivela, B., MAYDAY:A Software for Multimodal Articulatory Data Analysis, Atti del X Convegno AISV, 2015, 173-184.

[12] d'Apolito, S., Gili Fivela, B., "Targetless schwa" in francese L2, in Atti del XII Convegno AISV, 27-29 gennaio, Università degli Studi di Salerno, 2016 – under revision.

[13] Rialland, A., Schwa et syllables en français, in Studies in compensatory lengthening, Dordrecht: Foris Publications, 1986, 187-226.

[14] Suemitsu, A., Dang, J., Ito, T., Tiede M., A real-time articulatory visual feedback approach with target presentation for second language pronunciation learning, JASA 138(4), 2015, EL382-EL387.