Place assimilation and articulatory strategies: The case of sibilant sequences in French as L1 and L2

Sonia d'Apolito, Barbara Gili Fivela

Università del Salento – CRIL (Centro di Ricerche Interdisciplinare del Linguaggio), Lecce, Italy

sonia.dapolito@unisalento.it; barbara.gili@unisalento.it

Abstract

This study focuses on how French heterosyllabic sibilant clusters are produced by one French native speaker and by three Italian learners of French-L2. In French, these clusters are frequent and are reported to show place assimilation; in Italian, on the contrary, they are very rare and speakers are expected to repair the phonotactically marked sequences by epenthesizing a schwa. In the current study, acoustic and articulatory (AG500) data have been collected in order to observe the production of clusters in French L1 and L2, under the influence of two factors - that is speech rate and presence of prosodic boundaries - that may affect the realization of assimilation or repairs by acting on coarticulation.

Results reveal that place assimilation occurs only at faster speech rate and it favors the realization of the postalveolar over the alveolar fricative; as expected, the presence of prosodic boundaries interferes with this process. Place assimilation is realized by the French speaker and one Italian speaker, although the latter produces it in fewer cases and using a different articulatory strategy.

Index Terms: French L1 and L2, place assimilation, speech rate, prosodic structure, AG500

10.21437/Interspeech.2013-162

1. Introduction

Coarticulatory and assimilatory processes are considered as phonetic and phonological events, respectively, due to the influence that segments in the speech flow may play on each other, both within and across word boundaries. According to Articulatory Phonology [1,2,3,4], which describes speech in terms of articulatory gestures and captures both phonetic and phonological properties, a variety of phonological and phonetic consequences of coarticulation, such as assimilation, may be accounted for by gestural overlapping [2]. In particular, the nature of the overlapping of gestures may have different consequences on the articulatory and acoustic output: 1) if two gestures are on different articulatory tiers (e.g., lips and tongue tip closures), both gestures achieve their target, but a gesture may obscure the other, due to *gestural hiding*, and the latter cannot be perceived [2]; 2) if two gestures are on the same articulatory tier (e.g., tongue tip closures), the dynamical parameters of both gestures may be "*blended*" and their motions differ from that of the individual gestures [2].

Two relevant factors that are known to affect coarticulatory and assimilatory processes are prosodic structure, as the presence of a prosodic boundary can interfere with coarticulation [5], and speech rate, as a fast speech rate can facilitate coarticulation, because of increasing coproduction between successive segments [6].

In this paper, we focus on coarticulatory and assimilatory processes in sibilant sequences across word boundaries. Such contexts are very common in some languages, such as French, and offer interesting data on coarticulatory and assimilatory processes. For instance, a systematic study on the production of alveolar-postalveolar (AP) and postalveolar-alveolar (PA) sequences by French speakers at a normal speech rate [7] reported that sibilants may undergo place assimilation, although French was claimed not to show such a process [8]. In particular, according to [7], place assimilation is not restricted to a specific direction (it was regressive in AP and progressive in PA sequences) and may even show up without simultaneous schwa deletion. In other languages, such as Italian, consonantal clusters - including sibilant clusters - are less frequent and, thus, fewer coarticulatory processes may be observed across consonants. In Italian, consonant sequences across word boundaries appear only in a few cases, mainly where prepositions and loanwords occur or in contexts in which word final vowel truncation takes place [9,10]. Thus, consonantal sequences are phonotactically marked for Italians who find it difficult to produce them and tend to insert epenthetic vowels, as a general repair strategy [11,12].

An interesting research question then concerns what happens to such sequences when they are produced in an L2 in which they are frequent. Indeed, even though the literature shows that marked structures are difficult to learn [13], advanced L2 learners may have developed specific strategies to cope with phenomena that are marked and infrequent in their L1 but frequent and unmarked in the L2. Thus, in this study we will focus on how three Italian learners of French and one French native speaker realize coarticulatory and assimilatory processes in French sibilant sequences across word boundaries. Moreover, given their relevance, we will take into account the influence of higher level prosodic boundaries and differences in speech rate.

2. Goals of the study and hypotheses

The goal of the present study is twofold. First, we will observe how Italian advanced learners of French realize in French-L2 sibilant sequences that are phonotactically marked in Italian and how their production differs from that of the native French speaker, in particular in relation to place assimilation. Our hypothesis is that Italian learners tend to insert schwa vowels within consonant clusters and, in the case of coarticulatory processes, they may show different articulatory strategies in comparison to the French speaker. Second, we will observe the influence of speech rate and prosodic structure. Our hypothesis is that place assimilation in French-L2, if any, is realized at fast speech rate and when no strong boundary is present, i.e. in contexts in which coarticulation is favored.

3. Method

Target heterosyllabic sibilant clusters, /sʃ/,/ʃs/,/sʒ/,/ʒs/,/zʃ/,/zʒ/, are studied in an /a_i/ surrounding vowel context, proposed within a carrier sentence. Stimuli were inserted at a phonological phrase boundary (*weak-boundary condition*: e.g. "Il dit sa**ge sy**rien rapidement" *He said Syrian wise person quickly*) and at an intonational phrase boundary (*strong-boundary condition*: e.g. "D'abord il a dit sa**ge, sy**rien il l'a dit après" *First he said wise person, he said Syrian afterwards*). All the carrier sentences were judged as naturally sounding by the French native speakers who were involved in preliminary test or in the experimental session. Three Italian advanced learners of French (SI1, SI2, SI3) and a native French speaker (SF4) read the corpus 7 times, once at normal and once at a faster speech rate, while both acoustic and articulatory data were recorded. Articulatory data (AG500) were collected at the CRIL lab, in Lecce, by placing coils on the tongue midsagittal plane (4), on upper and lower lips (2), upper and lower incisors (2) and behind the ears (2). The audio signal was simultaneously recorded (Edirol sound card, 44.1 kHz). Data were analyzed auditorily, acoustically and articulatorily by means of PRAAT and MATLAB.

3.1 Auditory check

An auditory analysis was performed by both authors, who independently verified whether the sibilant sequences underwent place assimilation and expected prosodic boundaries were realized (in case of disagreement, the auditory test was repeated after a period of time). The auditory classification was used to perform acoustic and articulatory measures, by separating alveolar, postalveolar and place assimilated tokens.

3.2 Acoustic labeling and measurements

Acoustic labeling was performed at segment boundaries in the $V_1C_1\#C_2V_2$ sequence, even when the sequence included a schwa and/or a pause at the prosodic boundary. It was followed by a detailed acoustic analysis, performed by means of spectral, amplitudinal and durational measurements. Several spectral measures of the frication noise (range 1.5kHz-15kHz) were calculated, however here we will focus on the Centre of Gravity – CoG [14] – a spectral moment which is particularly useful to distinguish the place of articulation [5] - for a discussion of the other measures see [19,20]. We also computed the duration and normalized duration of all segments, i.e. the ratio between the segment duration and the duration of the two words including the target context. Finally, speech rate was quantified as the ratio between the number of syllables realized in the sentence and sentence duration.

3.3 Articulatory labeling and measurements

Articulatory data were labeled on the following trajectories, as for gesture onset and offset and velocity peaks for both fricatives and adjacent vowels:

- Vertical (*z*-axis) and horizontal (*x*-axis) movement of tongue tip (TT) for both alveolar and postalveolar fricatives, since they are coronal, and tongue dorsum (TD) for the V-to-V articulation ([a]-to-[i]);
- Horizontal (*x*-axis) movement of lower lip (LL) for lip protrusion mainly due to the postalveolar fricative (though alveolars may also have a protrusion gesture [21 in 22; 23]).

The following measurements were then calculated:

1) Duration (ms) and displacement (mm) of "closing" gesture for each fricative (Fig.1: *C1/C2 closing_duration; C1/C2_displacement*) [5];
2) Absolute timing (ms) and displacement (mm) between the target of the second fricative and that of the first fricative (Fig.1: Δextrema; Δdisplacement) [5];
3) Relative phasing (ms): the ratio between the fricative target-to-fricative target interval and the duration of [a]-[i] articulation, [18] (Fig.1, right box).

Statistical analyses, conducted in SPSS, were based on non-parametrical tests (Kruskal-Wallis and Mann-Whitney), run separately for each subject. Here, only results relative to closing gesture duration and displacement on TT and LL trajectories are discussed (for the others, see [19, 20]).

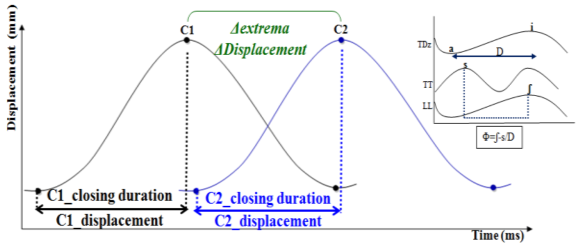

Figure 1: *Schema of articulatory measurements*

4. Results

4.1 Auditory check results

At normal speech rate, all speakers inserted a schwa and/or a pause in the strong-boundary condition. At faster speech rate, two out of three Italians, SI2 and SI3, inserted a schwa in both prosodic conditions, while SI1 and SF4 inserted a schwa in few cases and only in presence of a strong prosodic boundary. Moreover, at faster rate some cases of both progressive and regressive place assimilation were observed and they correspond to the realization of a postalveolar segment. In particular, the French native (SF4) realized place assimilation within both AP and PA sequences and in both prosodic conditions (though almost always in the weak-boundary condition); the Italian (SI1) realized place assimilation only within the PA sequence (progressive assimilation), and in the weak-boundary condition (Tab. 1).

Consequently, measurement results on assimilation, discussed in the following sections, only relate to SI1 and SF4 productions.

Spk	Postalveolar-alveolar		Alveolar-postalveolar	
	Weak-B	Strong-B	Weak-B	Strong-B
SF4	14 of 14	6 of 14	26 of 28	8 of 28
SI1	9 of 14	0 of 14	0 of 28	0 of 28
SI2	2 of 14	1 of 14	0 of 28	0 of 28
SI3	3 of 14	0 of 14	0 of 28	0 of 28

Table 1: *Place assimilations realized by speakers within both sibilant sequences in weak-boundary (Weak-B) and strong-boundary (Strong-B) condition at fast rate.*

4.2 Acoustic results

In both boundary conditions, speakers effectively varied their speech rates, as the ratio between the number of syllables realized in the sentence and sentence duration was significantly smaller at normal than at fast speech rate [*weak-bnd*: SI1 Z=-6,952 p=,000; SF4 Z=-7,176 p=,000; *strong-bnd*: SI1 Z=-7,535 p=,000; SF4 Z=-7,975 p=,000].

At normal rate, for both speakers and boundary conditions mean CoG was significantly lower for postalveolars than for alveolars [*weak-bnd*: SI1 $\chi^2(1,82)$= 61,555 p=,000; SF4: $\chi^2(1,109)$=66,285 p=,000; *strong-bnd*: SI1 $\chi^2(1,82)$=60,759 p=,000; SF4: $\chi^2(1,109)$=78,941 p=,000] – Fig. 2.

At fast speech rate, as mentioned above, both speakers realized place assimilations, showing a shift from the alveolar to the postalveolar fricative. When comparing CoG mean values for alveolar, postalveolars and assimilated segments, results showed that, in both boundary conditions, alveolars correspond to higher mean CoG values than postalveolars and assimilated segments, while there is no difference between assimilated segments and postalveolars (Fig. 3) [*weak-bnd*: SI1 $\chi^2(2,70)$=38,638 p=,000; *strong-bnd*: SI1 $\chi^2(2,80)$= 53,108 p=,000; SF4 $\chi^2(2,70)$=38,288 p=,000]; the Mann-Whitney test showed a significant difference between alveolar and postalveolar [*weak-bnd*: SI1 Z=-5,876 p=,000; *strong-bnd*: SI1 $\chi^2(2,80)$=53,108 p=,000; SF4: Z=-6,010 p=,000] and between alveolar and place assimilated segments [*weak-bnd*: SI1 Z=-4,163 p=,000; *strong-bnd*: SF4 Z=-3,259 p=,000].

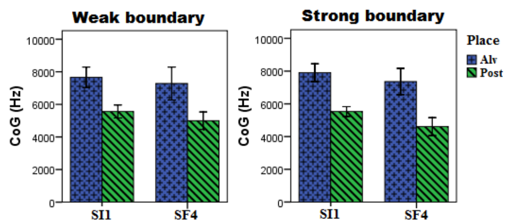

Figure 2: CoG mean values (and STD) for alveolar and postalveolar fricatives, realized by both speakers in weak (left) and strong boundary condition (right) at normal speech rate.

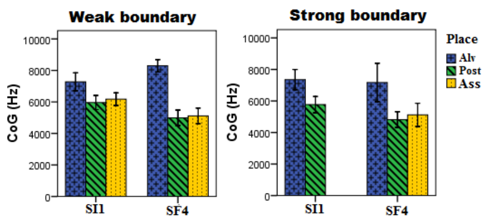

Figure 3: *CoG mean values (and STD) for alveolar, postalveolar and place assimilated fricatives, realized by both speakers in weak (left) and strong boundary condition (right) at fast speech rate.*

4.3 Articulatory results

Articulatory results are discussed in relation to diagrams, constructed from statistical analyses, which make clear how speakers realized sibilant sequences. In the diagrams, the closing gesture duration and displacement are represented for each fricative gesture by the length and height of the box, respectively (intra-gesture changes are not represented). All gestures are aligned according to the absolute timing within the same articulator (dotted black line) and to the absolute timing between different articulators (green line).

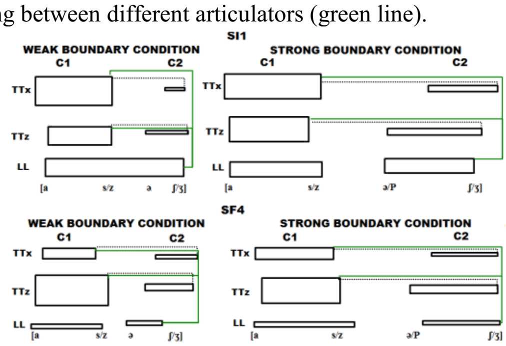

Figure 4: *Diagram of articulatory results for AP sequence realized at normal speech rate by SI1 (upper panels) and by SF4 (lower panels), in weak (left) and strong boundary condition (right).*

At normal speech rate, in both boundary conditions and in no-assimilation cases, the closing gestures of the two fricatives within both AP and PA sequences was visible on TT (*x* and *z*-axes), due to the insertion of a schwa (and pause in the case of a strong boundary). On LL, in the weak-boundary condition one closing gesture for SI1 (whose peak is aligned towards the end of the postalveolar) and two gestures for SF4 were found; in the strong-boundary condition two LL closing gestures were identified for both speakers (Fig.4 for AP sequence).

On the contrary, at fast speech rate, a stronger coarticulation occurred between the two fricatives which favours place assimilations. Generally, it was possible to identify only one closing except when an opening gesture occurred between fricatives due to the presence of a strong prosodic boundary.

As for the PA sequence, both SI1 and SF4 realized place assimilation, by means of different articulatory strategies (Fig. 5-6). In the SI1 productions, one closing gesture is visible on TT, and one protrusion gesture is visible on LL. In no-assimilation cases, the TT gesture reaches its target for the alveolar fricative and the same seems to be true also for the assimilation cases. Indeed, when no assimilation occurred, the alveolar closing gesture on TT (x) was slightly longer (TTx: 207ms *vs* 182ms; TTz: 225ms *vs* 175ms) and higher in displacement (TTx: 12,80 *vs* 6,97; TTz: 7,22mm *vs* 4,35mm) in comparison to that found in assimilated cases; however, only the difference for displacement on TTz was significant [$\chi^2(1,12)$=4,521 p=,033]. As for the protrusion gesture on LL, it starts earlier in assimilation cases than in no-assimilation ones, and is lengthened by means of a plateau (dotted blue box) that lasts even after the alveolar closing gesture. Thus, in SI1 assimilation cases, the lengthening of the protrusion gesture on LL through a plateau seems to hide the alveolar gesture on TT (Fig. 5).

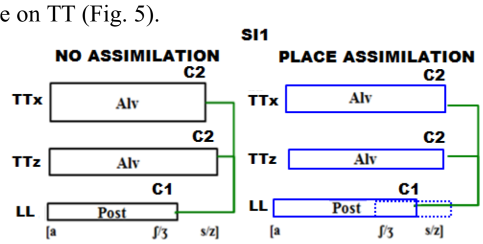

Figure 5: *Diagram of articulatory results for PA sequence in weak-boundary condition: no place assimilations (left) and place assimilations (right).*

The French native speaker systematically realized place assimilation in the weak-boundary condition, while in the strong-boundary condition she realized only some cases of place assimilation and in the other she inserted a schwa. When place assimilations occurred in PA clusters, and in both prosodic conditions, only one closing gesture was visible on TT (Fig. 6). On TTz the target seems to be due to a blending between the two fricative gestures. Indeed, statistics performed on the strong-boundary condition showed that the closing gesture duration on TTz was longer [$\chi^2(1,12)$=7,180 p=,007; (186,69ms *vs* 119,42ms)] and its displacement was greater than the gesture for C_1 [$\chi^2(1,12)$=8,077 p=,004; (11,24mm *vs* 6,98mm)] when assimilation took place. However, the duration of the closing gesture on TTz of assimilated cases was shorter, though not significantly different, than the sum of C_1 and C_2 fricative closing gestures (180,36ms *vs* 220,49ms). On TTx only one gesture is visible, although its duration appears to be possibly influenced also by the following anterior vowel [i]. On LL a protrusion gesture ending in a plateau is found (Fig. 6, dotted blue box).

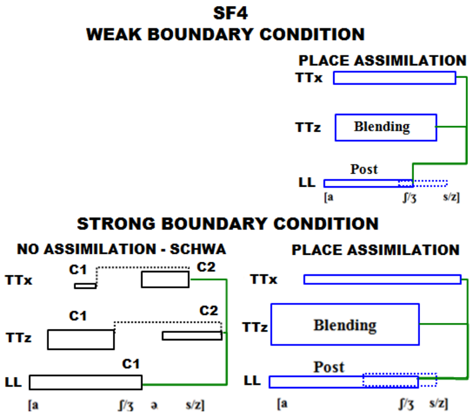

Figure 6: *Diagram of articulatory results for PA sequence in weak (upper panel) and strong boundary condition (lower panels); no place assimilations (left) and place assimilations (right).*

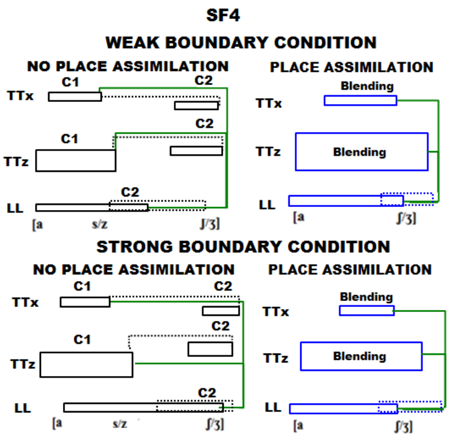

Figure 7: *Diagram of articulatory results for the AP sequence in weak (upper panels) and strong boundary condition (lower panels); no place assimilations (left) and place assimilations (right).*

As for the AP sequence, only SF4 realized place assimilations that are regressive in both prosodic conditions. Her place assimilations corresponded to a single closing gesture on TT and LL (Fig. 7). In particular, the closing gesture on TTx and TTz seems to be due to a blending between the two fricative gestures. Indeed, in both prosodic conditions, the closing gesture on TTz was significantly longer and wider than that due to C_1 if no-assimilation took place (length: [*weak-bnd*: $\chi^2(1,28)=4,159$ p=,041 (90ms vs 172ms); *strong-bnd*: $\chi^2(1,15)=9,843$ p=,002 (88,91 vs 185,71)]; height: [*weak-bnd*: $\chi^2(1,28)=5,327$ p=,021 (2,22 vs 10,82); *strong-bnd*: $\chi^2(1,15)=5,898$ p=,015 (3,89 vs 9,91)]. The closing gesture on TTx shows a very similar trend though only its duration in the strong-boundary condition was significantly greater in assimilated cases than for C_1 in non-assimilated cases [*strong-bnd* $\chi^2(1,15)=6,876$ p=,009 (45,65ms vs 80,64ms)]. However, the duration of closing gesture on both TTx and TTz was significantly shorter than the duration corresponding to the sum of C_1 and C_2 gestures [*weak-bnd*: TTx $\chi^2(1,28)=5,379$ p=,020 (85,52ms *vs* 235,39ms); TTz: $\chi^2(1,28)=5,396$ p=,020 (171,67ms *vs* 262,50ms); *strong-bnd*: TTx: $\chi^2(1,15)=8,014$ p=,005 (93,94ms *vs* 242,34ms); TTz: $\chi^2(1,15)=8,029$ p=,005 (193,38ms *vs* 256,28ms)]. Only in some cases did the LL gesture end with a plateau (dotted black/blue boxes).

5. Discussion and conclusions

Both acoustic and articulatory (AG500) data have been collected in order to observe: 1) how French sibilant sequences are realized by three Italian advanced learners of French-L2 and by one French native speaker, in particular as for place assimilations; and 2) how speech rate and prosodic structure influence the coarticulatory process.

In line with our hypotheses, the productions of the Italian learners were affected by mother tongue phonetics and phonology. Two out of three speakers inserted a schwa even at faster speech rate to repair the phonotactically marked clusters. Only one Italian learner realized place assimilations, although in fewer cases in comparison to the French speaker, and only within the postalveolar-alveolar sequence and in the weak-boundary condition. On the other hand, the French speaker realized assimilations in both sibilant sequences and prosodic conditions (though mainly in the weak-boundary condition). She inserted a schwa at faster speech rate only in presence of a prosodically strong-boundary and, contrary to findings in [7], she never realized assimilations in case of schwa insertion.

Auditory and acoustic results show that place assimilations were characterized by a (spectral) change from an alveolar to a postalveolar fricative. However, even if from a perceptual and acoustic point of view the outputs of the assimilatory processes were the same for both speakers, different articulatory strategies seemed to be at play. The French speaker realized a blending between the two fricatives, through a complex strategy that requires modification of the timing and displacement of each single gesture and, thus, of intergestural timing; furthermore, she realized labial gesture prolongation only in some cases, but not all. On the contrary, the Italian learner realized one TT gesture both in assimilation and no-assimilation cases, and obtained a 'place assimilation effect' by hiding the alveolar gesture through the prolongation of the postalveolar lip protrusion. This strategy requires the modification of the duration of the gesture of the lower lip only (and probably of the upper lip too). More data are needed to establish whether the strategies adopted by the French and by the Italian speaker are speaker dependent or somehow affected by the L1. However, the latter option seems plausible, as the gestural hiding appears to be simpler and more conservative than the blending produced by the French speaker. In this respect, the hiding strategy could be ideal for an L2-learner trying to approach an unfamiliar L2 cluster.

Finally, results relative to the influence of speech rate and prosodic structure confirmed our expectations. For the native French speaker and one Italian speaker, a fast speech rate facilitated coarticulation and thus the realization of place assimilation, while the presence of a prosodically strong-boundary interfered with the process, since place assimilation occurred especially in the weak-boundary condition.

6. Acknowledgements

We would like to thank engineer Francesco Sigona for his help in EMMA data processing and for the PRAAT and Matlab scripts.

7. References

[1] Browman, C., P., Goldstein, L., "*Towards an articulatory phoonlogy*", Phonology Yearbook, 3, 219-252, 1986.

[2] Browman, C., P., Goldstein, L., "*Articulatory gestures as phonological units*", Phonology, 6, 201-251, 1989.

[3] Browman, C., P., "*Dynamics and Articulatory Phonology*", Minds as motion, 175-193, 1995.

[4] Browman, C., P., Goldstein, L., "*The gestural phonology model*" Speech production: motor control, Brain research and Fluency disorders, 57-71, 2007.

[5] Byrd, D., Kaun, A., Narayanan, S., Saltzman, E., "*Phrasal signatures in articulation*", Papers in Laboratory Phonology V, CUP, 70-87, 2000.

[6] Byrd, D. and Tan, C. C., "*Saying consonant clusters quickly*", Journal of Phonetics, vol. 4, 263-282, 1996.

[7] Niebuhr O., Lancia, L., Meunier, C., "*On place assimilation in French sibilant sequences"* Proc. of the VII ISSP, 221-224, Strasbourg, France, 2008.

[8] Walker, D. C., "*On a phonological innovation in French*", Ed. Cambridge University Press, vol. 12, 72-77, 1982.

[9] Farnetani, E., Busà, M. G., "*Italian clusters in continuous speech*" Proc. of the 3rd ICSLP, vol. 1, 359-362, Yokohama, Japan, 2004.

[10] Muliačić, Z., Fonologia della lingua italiana, ed. Il Mulino, Bologna, 1973.

[11] Davidson, L., "*Phonology, phonetics or frequency: influences on the production of non-native sequences*" Jour. of Phon., vol. 34,104-137, 2006.

[12] Oh, E., "*Coarticulation in non-native speakers of English and French: an acoustic study*", Jour. of Phon., vol. 36, 361-384, 2008.

[13] Eckman R., F., *Typological markedness and second language phonology*, Hansen-Edwards J., G., Zampini, m., L., (editors), J. Benjamins Publishing, 2008.

[14] Jongman, A., Wayland, R., Wong, S., "*Acoustic characteristics of English fricatives*", JASA, 108(3), 1252-1263, 2000.

[15] Evers, V., Reetz, H., Lahiri, A., "*Crosslinguistic acoustic categorization of sibilants independent of phonological status*", Journal of Phonetics, 26, 345-370, 1998.

[16] Jesus, L., M., T., Shadle, C., H., "*A parametric study of the spectral characteristics of European Portuguese fricatives*", Journal of Phonetics, 30, 437-464, 2002.

[17] Maniwa, K., Jongman, A., Wade, T., "*Acoustic characteristics of clearly spoken English fricatives*", JASA, 125(6), 3962-3973, 2009.

[18] Tiede, M., Shattuck-Hufnagel, S., Johnson, B., Ghosh, S., Mattheis, M., Zandipour, M., Perkell, J., "*Gestural phasing in /kt/ sequences contrasting within and cross word contexts*", ICPhS XVI, Saarbrucken, 2007.

[19] d'Apolito S., (2012),La coarticolazione: Studio acustico, cinematico e percettivo di sequenze di sibilanti della lingua francese nelle produzioni di studenti italofoni, Phd thesis.

[20] d'Apolito S., Sigona, F., Gili Fivela, B., (in prep) "Acoustic and articulatory study on French sibilant sequences", Proceedings of 9th ASIV Congress, Università Cà Foscari, Venice, 21-23 January, 2013.

[21] Perkell, J., (1986), "*Coarticulation strategies: preliminary implications of a detailed analysis of lower lip protrusion movements*", Speech Communication, 5, 47-86.

[22] B. Kühnert & F. Nolan, (1999) *The origin of coarticulation,* in W.J. Hardcastle and N. Hewlett (eds), Coarticulation: Theory, Data and Techniques in Speech Production. Cambridge: CUP. pp 7-30.

[23] Engwall, O., (2000), "*Dynamical aspects of coarticulation in Swedish fricatives – a combined EMA & EPG study*" TMH Quarterly Status and Progress Report, 4, KTH, Stockholm, 49–73.