On Place Assimilation in French Sibilant Sequences

Oliver Niebuhr, Leonardo Lancia, Christine Meunier

Laboratoire Parole & Langage, CNRS UMR 6057, Université de Provence, Aix-en-Provence, France
[oliver.niebuhr],[leonardo.lancia],[christine.meunier]@lpl-aix.fr

Abstract

A corpus of systematically constructed sentences read by 4 female speakers revealed the existence of place assimilation in sequences of French alveolar and postalveolar sibilants. The assimilation manifests itself gradually in time and frequency measurements and is directed towards 'postalveolar'. Thus, it can be regressive and progressive, depending on the order of the place features in the sibilant sequences.

1 Introduction

It is well known for British and American English that sequences of alveolar and postalveolar sibilants across word boundaries show systematic phonetic patterns that may be interpreted as a regressive, gradual process of assimilation of place of articulation which is directed towards 'postalveolar'. Hence, it affects the /s/ in /sʃ/ sequences, cf. [1,2,3,4]. In contrast to English, place assimilation in general is *"thought to be non-existent in French"* (cf. [5,6]). Apart from such explicit statements, this view is mirrored in the fact that descriptions of French assimilation patterns solely deal with voice assimilation. An exception is [7], who briefly noted that /z/ at the end of *"quinze"* in *"quinze juin"* ('June 15th') can become a [ʒ]. That is, it takes over regressively the place of articulation of the following word-initial /ʒ/. This and analogues examples are taken up and discussed by [5,8] in more detail, and they conclude – contrary to the general view sketched above – that place assimilation can actually occur in French sibilant sequences, but just regressively and in connection with simultaneous Schwa deletion. So, the assimilatory change from /z/ to /ʒ/ in *"quinze juin"* should only take place, if the basic phonological representation of *"quinze"* shows a final /ə/, whose deletion then creates the sequence of adjacent sibilants. Therefore, the assimilation is said to be restricted to certain regional variants of French.

This conclusion might still be too narrow. Expressions like *"je suis"* (/ʒə sɥi/, 'I am') and *"je sais"* (/ʒə sɛ/, 'I know') may be realized as [ʃːɥi] or [ʃːɛ], respectively. On the one hand, the place assimilations involved in these examples are accompanied by Schwa deletion, as claimed by [5,8]. Contrary to their claims, however, the process that changes /ʒs/ into [ʃː] is progressive, not regressive. Moreover, the assimilations within *"je suis" and "je sais"* are very widespread within the francophone community. It might be argued that these two examples are exceptions, since they are fixed expressions of function words, for which there are separate (non-productive) rules. Moreover, from a synchronic point of view, these fixed expressions might even be new words so that the alternation between [ʒə sɥi], [ʒə sɛ] on the one and [ʃːɥi], [ʃːɛ] on the other hand is no longer a matter of phonological processes at all.

Irrespective of these arguments, it has never been systematically investigated so far, if, to what extent, and in which way place assimilation occurs within French sibilant sequences. The present study aims at shedding first light on these questions. They will be addressed by means of time and frequency measurements. The great diversity in the temporal dynamics of spectral changes that can take place in gradual assimilations (cf. [1]) as well as the correlation between different types of spectral changes in English sibilant sequence and the overall durations of these sequences (cf. [9]) suggest consideration of both domains of the signal, which has rarely been done in previous (acoustic) studies.

2 Method

The recorded speech corpus is based on 72 sentences, which may be divided into 3 subsets. The primary subset contains all the eight possible sibilant sequences across word boundaries that result from the cross-combination of the features 'alveolar', 'postalveolar' and 'voiced', 'voiceless', i.e. (a) /sʃ/, (b) /ʃs/, (c) /zʒ/, (d) /ʒz/, (e) /sʒ/, (f) /ʒs/, (g) /ʃz/, and (h) /zʃ/. They are placed in the symmetrical vowel contexts /i/___/i/, /a/___/a/, and /u/___/u/. In total, this yields 24 test sentences. The secondary subset of test sentences only includes the 4 completely voiced or voiceless sibilant sequences (a)-(d), but combines them with all 6 asymmetrical vowel contexts, which leads to another 24 sentences. The remaining 24 sentences form a complementary subset, in which each of the 4 individual sibilants /s/, /ʃ/, /z/, and /ʒ/ is paired across word boundaries with a labial consonant (C) in the two possible orders __C and C__. The labial consonants do not interfere with the voicing of the adjacent sibilant in terms of voice assimilation in any of the sequences. As in the primary subset, the 8 sequences of the complementary subset were combined with the three symmetrical vowel contexts. The individual sibilants of the complementtary subset will be used as reference sounds.

The sentence list was produced subsequently in 4 different randomizations by 4 female subjects. They were 20-50 years old and had no regional accent. So, none of the subjects regularly produces a Schwa at the end of the investigated target words. Prior to the individual recordings, the speakers were instructed to read the sentences displayed on the screen as naturally as possible with a constant loudness suitable for a normal conversation. The subjects were recorded in an anechoic chamber at the *Laboratoire Parole et Language*. The first of the 4 repetitions of the sentence list was preceded by 10 dummy sentences, which were similar to the ones of the sentence list and served to familiarize the subjects with their task. Overall, each recording session took 30-40 minutes. A short interview after the session showed that none of the subjects guessed the actual aim of the recording.

The sentences were labelled, among others, with regard to the following aspects:

- (a) vowel onset in the first target syllable, vowel offset in the second one,
- (b) onset and offset of the sibilant sequence,
- (c) a spectral bipartition of the sibilant sequence that is also mirrored in perception.

The latter label is optional. In the following, it is referred to as the 'B' (=boundary) label. It is a general marker of discontinuitiy, i.e. it is set at an abrupt transition between two constant sibilant sections as well as at the boundary between a constant and a (following or preceding) dynamic section. So, sibilant sequences without a 'B' label are either marked by a constant sibilant sound quality or by a continuous transition between two different sibilant sound qualities (cf. [1]). All labelling was done in *praat*. In some cases, a pause was realized between the two sibilants of a sequence. It is very likely that these interrupted articulations are different from the ones in which the phonological sibilant sequences are produced continuously. Therefore, the corresponding sentences have been excluded from further analyses.

The acoustic analyses comprised durational as well as spectral measurements. The former were determined on the basis of the labels and include:

- the vowel durations in the first and the second syllable of the target sequences, using labels (a), (b),
- the total duration of the sibilant sequence, based on the (b) labels,
- the durations of the two sibilant sections and the corresponding ratios, by means of labels (b), (c).

The latter values are suitable to measure the (degree of) assimilation in the temporal domain. However, since not every sibilant sequence requires a 'B' label annotation, and since the sections separated by 'B' need not have clear postalveolar and alveolar friction qualities, the durational measurements on the basis of the 'B' label must be complemented by a measure in the frequency domain. In the present study, this is the centre of gravity (CoG). For each sibilant or sibilant sequence, two CoG values were determined (in Hz) for frequencies between 1.5-15kHz: the mean CoG and the CoG range. The measurements were done automatically. Finally, in addition to the acoustic measurements, the frequencies of the words that contain the target sequences were determined. They were estimated by means of a corpus of movie subtitles with about 52 million words, cf. [10].

Two analogues mixed-model, linear-regression analyses were performed. They differed in the predicted variable, which was either the duration ratio of the first and second sibilant sections or the mean CoG value of the complete sibilant unit (i.e. single sibilants as well). The two analyses included the following independent variables: (a) order of place of articulation within the sibilant sequences, voicing features within the sibilant sequences, (b) vowel context (i.e. symmetric vs. asymmetric), frequency of each word (recalculated as a logarithmic value), duration of the two vowels surrounding the sibilant sequences (as an indicator of the highlighting of the corresponding word), and (c) speaker as well as sentence. While the latter two variables (c) were considered as random effects, the remaining ones constituted fixed effects that represented either control variables (b) or experimental variables (a).

3 Results

The results of the statistical analyses for the two different predictor variables, i.e. duration ratios and mean CoGs, are summarized in Table 1. Moreover, Figure 1 shows the mean durations of the alveolar and postalveolar sibilant sections in the sequences, while Figure 2 presents the averages of the mean CoGs and of the CoG ranges for the single sibilants as well as for the sibilant sequences.

The statistical analyses of the phonetic patterns revealed a significant effect of the order of the places of articulation in the sibilant sequences on the duration ratios (Tab.1). That is, if there is a discontinuity in the sibilant sequences, the postalveolar section is longer than the alveolar one (Fig.1). There were no differences in the length of alveolar and postalveolar single sibilants. However, both were on average clearly shorter (around 40-50%) than the sequences.

In addition to the boundary shift in favour of 'postalveolar', the mean CoGs are more similar to the postalveolar than to the alveolar references (Fig.2). Regression models comparing the sibilant sequences with the single sequences found that this shift of the CoG values towards 'postalveolar' is even stronger in alveolar-postalveolar than in post-alveolar-alveolar sequences (Tab.1)

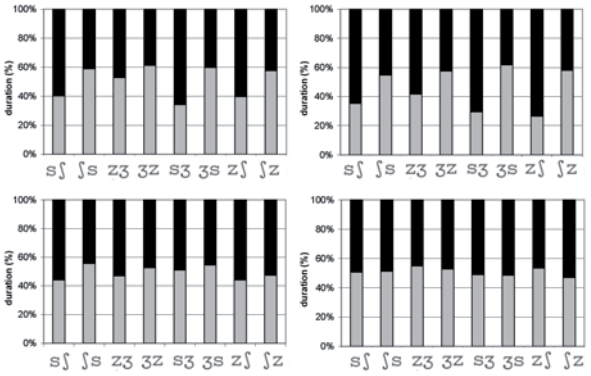

Figure 1. *Relative mean durations of the first (grey) and the second (black) sibilant sections in the sibilant sequences for the 4 speakers ABO, MTE, CDL, and IVI (clockwise); n=36, 12.*

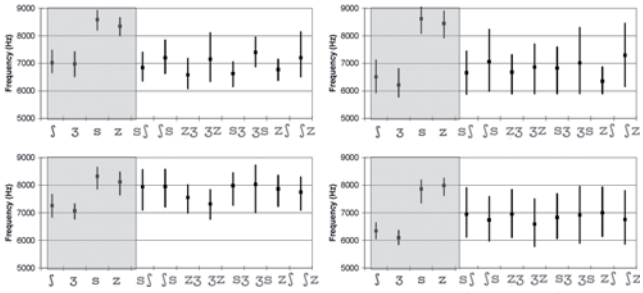

Figure 2. *Averages of the mean CoGs (dots) and CoG ranges (bars) of the single reference sibilants (left) and the sibilant sequences (right) for the 4 subjects ABO, MTE, CDL, and IVI (clockwise); n=24.*

Table 1. *Coeffecients of the mixed-model, linear-regression analyses for the duration ratios and the mean CoGs. For each analysis, the fixed effects as well as their regression coefficients, standard errors, and T values are given; 'vd'/'vl' refer to 'voiced'/'voiceless'. Grey lines indicate significant contributions to the model.*

	analysis on duration ratios			analysis on mean CoGs		
comp. var. levels	coefficient	std. error	T value	coefficient	std. error	T value
vowel (symmetry)	-0.02329	0.06489	-0.36	105.06	64.28	1.63
vowel (1st syllable)	2.50752	1.36649	1.84	1456.88	944.21	1.54
vowel (2nd syllable)	3.68464	1.26588	2.91	1691.33	1000.81	1.69
frequ. (1st word)	-0.00147	0.00668	-0.22	3.05	6.89	0.44
frequ. (2nd word)	0.00985	0.00636	1.55	-5.12	6.49	-0.79
order place artic.	0.35630	0.05429	6.56	135.81	53.53	2.54
voice(vlvl-vdvd)	0.18182	0.07364	2.47	-260.85	71.99	-3.62
voice(vlvl-vlvd)	0.06713	0.10678	0.63	-143.58	105.58	-1.36
voice(vlvl-vdvl)	0.08249	0.09739	0.85	-73.16	96.26	-0.76

4 Discussion

In combination, the above findings represent for the first time systematic empirical evidence for the presence of assimilation of place of articulation in French, based on sibilant sequences. This assimilation process manifests itself gradually in the time and

the frequency domain. It is directed towards the post-alveolar feature, like the English process. However, different from what we know about the latter, the French assimilation is not restricted to a specific direction. Thus, it is regressive in alveolar-postalveolar and progressive postalveolar-alveolar sequences, although the latter seems to be somewhat weaker than the former. For both directions, a similar spectrum of phonetic patterns was found that covers all the assimilation types differentiated by [1], cf. also [2,4,9]. This also includes (spectrally) complete assimilations. Finally, contrary to previous claims ([5,7,8]), the assimilation shows up without simultaneous Schwa deletion and beyond the frequent (and hence possibly fixed) expressions *"je sais"* and *"je suis"*.

However, in view of the Figures 1 and 2, it has to be considered that this bidirectional, feature-determined process was only found clearly for 2 of the 4 subjects, viz. ABO and MTE. It was only marginally present for CDL; and IVI shows no phonetic patterns at all that point to an assimilation process. On the other hand, pronounced speaker-specific differences that can concern degree and overall pattern of the assimilation are known to be a characteristic of assimilatory processes across languages, cf. [2]. Yet, more subjects should be recorded in follow-up studies to examine how pervasive the process actually is.

Furthermore, the statistical analyses revealed that the vowel durations in the first and the second target syllables increased the duration ratios and the mean CoGs values of the sibilant sequences (Tab.1). The vowel durations were to measure to what extent the corresponding word is highlighted by the speaker. It may be assumed that longer vowels indicate more pronounced highlighting. Hence, the vowel-duration measure may be seen in analogy to F0 heights or F0 ranges measured in pitch-accent languages like English and German. So, in view of the French assimilation towards 'postalveolar', a positive correlation between vowel durations and mean CoGs could indicate that the (degree of) assimilation is reduced for stronger highlighted words. This is in line with findings from other languages, in which accentuation counteracts assimilation.

Finally, two other factors have significantly influenced the measurements. Phonologically voiced sequences were marked by longer duration ratios and lower mean CoGs than phonologically voiceless ones. While the duration changes may be explained by less articulatory effort in voiced than in voiceless sounds (shorter and less tight strictures), the lower mean CoGs are likely due to the vocal fold vibration itself. It reduces the airflow velocity and hence the friction energy in higher frequencies, while the harmonics of the glottal excitation increase the energy of the lower part (particularly in the lower formants) of the spectrum. The fact that the second sibilant was shorter, if the corresponding second target word was more frequent points to a positive correlation between articulatory reduction and word frequency, which also goes well with previous findings. However, it may be assumed that none of these additional influences on the time and frequency measurements created the observed assimilation process.

References

[1] T. Holst and F. Nolan, The influence of syntactic structure on [s] to [ʃ] assimilation. In B. Connell, A. Arvaniti (eds.), *Papers in Laboratory Phonology IV: Phonology and Phonetic Evidence*. CUP, Cambridge, pp. 315-333, 1995.

[2] E. Zsiga, An acoustic and electropalatographic study of lexical and postlexical palatalization in American English. In B. Connell, A. Arvaniti (eds.), *Papers in Laboratory Phonology IV: Phonology and Phonetic Evidence*. CUP, Cambridge, pp. 282-302, 1995.

[3] J. C. Catford, *Fundamental problems in phonetics.* Indiana University Press, Bloomington, 1977.

[4] V. Zue and S. Shattuck-Hufnagel, Palatalization of /s/ in American English: when is a /ʃ/ not a /ʃ/? *JASA* 67:S27, 1980.

[5] Z. Fagyal, D. Kibbee, and F. Jenkins, *French. A linguistic introduction*. CUP, Cambridge, 2006.

[6] F. Ramus, Outstanding questions about phonological processing in dyslexia. *Dyslexia* 7:197-216, 2001.

[7] F. Carton, *Introduction à la phonetique du francais.* Bordas, Paris, 1974.

[8] F. Gadet, *Le francais populaire*. PUF, Paris, 1992.

[9] C. Browman, Assimilation as gestural overlap: comments on Holst and Nolan. In B. Connell, A. Arvaniti (eds.), *Papers in Laboratory Phonology IV: Phonology and Phonetic Evidence*. CUP, Cambridge, pp. 334-342, 1995.

[10] B. New, M. Brysbaert, J. Veronis, C. Pallier, The use of film subtitles to estimate word frequencies. *Applied Psycholinguistics* 28: 661-677, 2007.