



HAL
open science

On place assimilation in sibilant sequences – comparing French and English

Oliver Niebuhr, Meghan Clayards, Christine Meunier, Leonardo Lancia

► To cite this version:

Oliver Niebuhr, Meghan Clayards, Christine Meunier, Leonardo Lancia. On place assimilation in sibilant sequences – comparing French and English. *Journal of Phonetics*, 2011, 39 (3), pp.429-451. 10.1016/j.wocn.2011.04.003 . hal-01491742

HAL Id: hal-01491742

<https://hal.science/hal-01491742>

Submitted on 5 Jan 2022

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

On place assimilation within sibilant sequences – comparing French and English

Oliver Niebuhr¹, Meghan Clayards², Christine Meunier¹, Leonardo Lancia¹

¹ Laboratoire Parole et Langage, Université de Provence, Aix-en-Provence, France

² Department of Psychology, University of York, York, UK

Abstract

Two parallel acoustic analyses were performed for French and English sibilant sequences, based on comparably structured read-speech corpora. They comprised all sequences of voiced and voiceless alveolar and postalveolar sibilants that can occur across word boundaries in the two languages, as well as the individual alveolar and postalveolar sibilants, combined with preceding or following labial consonants across word boundaries. The individual sibilants provide references in order to determine type and degree of place assimilation in the sequences. Based on duration ratios and centre-of-gravity values that were determined for each individual sibilant and sibilant sequence, we found clear evidence for place assimilation not only for English, but also for French. In both languages the assimilation manifested itself gradually in the time as well as in the frequency domain. However, while in the English assimilation occurred strictly regressive and primarily towards postalveolar, the French assimilation was solely towards postalveolar, but in both directions regressive and progressive. Apart from these basic differences, the degree of assimilation varied in French and English due to the individual subjects, the vowel contexts, the frequencies of the target words, and the voicing features of the sibilants in ways that match well with previous findings.

1. Introduction

1.1 Aim of the study in the light of assimilation and co-articulation

While assimilation and co-articulation both extend the phonetic properties of speech segments, the demarcation line between them runs in a traditional perspective parallel to the one of phonology and phonetics. Assimilation is anchored in the phonology and is hence a cognitive phenomenon, in which a phonemic feature is substituted by one of a (typically) adjacent segment (cf. Chomsky and Halle 1968; McCarthy 1988; Spencer 1996). In this way, assimilation affects the articulation in a categorical way. That is, if the feature substitution takes place (which depends, e.g., on the speaking style, the presence of an accent, or on the word class and frequency; cf. Recasens 1993; Kohler 2001) it leaves no phonetic trace of the original feature. On the other hand, co-articulation refers to a phonetic phenomenon that results from the biomechanical properties of the articulatory apparatus, particularly with regard to its restrictions in the temporal execution of movements and their co-ordinations. Thus, in contrast to assimilation, co-articulation is obligatory and leads to gradual variation within phonological categories. While the kind of variation is determined by the neighboring segments, its degree changes, for example, in connection with the speaking rate (cf. Amerman et al. 1970; Daniloff and Hammarberg 1973; Gay 1978; Barry 1985; Wood 1996).

Assimilation and co-articulation are still treated as separate phenomena in the literature in some more recent studies (cf. Recasens 1993; Recasens and Pallarès 2001). However, a growing number of investigations cross their conceptual differentiation. For example, production experiments of Whalen (1990) as well as of Wood (1991, 1996) provide evidence that phenomena subsumed under co-articulation can also be pre-planned and hence rooted at a cognitive level of the speech production process. Moreover, several studies showed that assimilation does not only create categorical, but also gradual phonetic variation (cf. Wright and Kerswill 1989; Nolan 1992; Hardcastle 1994), while simultaneously coarticulatory effects may also be characterized as categorical (cf. Ambrazaitis and John 2005). The data of Ellis and Hardcastle (2002) as well as of Heuvel et al. (1996) even suggest that identical segmental and prosodic contexts can create categorical and gradual assimilations, depending on the strategies of the individual speakers. Analogously, the findings of Manuel and Krakow (1984)

and of Hoole et al. (1993) contradict the view of co-articulation as a biomechanical automatism by revealing speaker and language-specific patterns. In consequence, there are several discussions in the literature on how assimilation and co-articulation can be defined adequately (e.g., Barry and Hawkins 1992), while some authors have started to use the two terms synonymously (cf. Ohala 1993; Clark and Yallop 1990).

It is not the primary aim of the present study to look for empirical evidence against or in favour of a separate treatment of assimilation and co-articulation, although our findings may contribute to this discussion. Instead, we aim at describing the acoustic interactions of alveolar and postalveolar sibilants across word boundaries with regard to the place of articulation on the basis of segmentally-controlled read-speech material. We will compare these interactions for French and English, assuming that the two languages show quantitative and qualitative differences. At the same time, we assume for both languages with regard to previous studies that the interactions have the potential to cross phonemic boundaries. For example, in a sequence like /sʃ/ the /s/ could be realized as a [ʃ] which is spectrally indistinguishable from an allophone of /ʃ/ in a comparable context. On this basis, we prefer to use the term ‘assimilation’ to refer to the investigated effects.

1.2 Research background

English and French are both known to have regressive voice assimilation across word boundaries (cf. Roach 1983; van Dommelen 1985; Snoeren et al. 2008). However, while English additionally shows assimilation of place of articulation, which affects plosives and fricatives as well as nasals (cf. Roach 1983), this type of assimilation “*is thought to be non-existent in French*” (Fagyal et al. 2006:49; cf. also Ramus 2001 and Gow 2003). Apart from such explicit statements, this view is mirrored in the fact that descriptions of French assimilation patterns just deal with voice assimilation (cf. Armstrong 1932; Batt 1960; Malmberg 1969; Price 1991).

An exception is Carton (1974), who briefly noted that /z/ at the end of “*quinze*” in “*quinze juin*” (‘June 15th’) can become a [ʒ], i.e. it takes over regressively the place of articulation of the following word-initial /ʒ/. Gadet (1992) as well as Fagyal et al. (2006) take up and discuss this and other examples in more detail and conclude – contrary to the above statement – 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 the example “*quinze juin*” should only take place, if the 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 like the Northern Metropolitan French. However, this conclusion might still be too narrow. Expressions like “*je suis*” ([ʒə sɥi], ‘I am’) and “*je sais*” ([ʒə sɛ], ‘I know’) may become [ʃ:ɥi] or [ʃ:ɛ], respectively. On the one hand, it is in line with Gadet and Fagyal et al. that the place assimilation involved in these two examples is accompanied by Schwa deletion. But on the other hand, it starts from /ʒs/ and is hence not regressive, but progressive. Moreover, the assimilations within “*je suis*” and “*je sais*” are not dialectally restricted, but very widespread within the francophone community and even entered into informal orthographical representations. 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 (cf. Kohler 1990). Moreover, from a synchronic point of view these fixed expressions might even be regarded as new words so that the alternations between [ʒə sɥi] and [ʒə sɛ] on the one and [ʃ:ɥi] and [ʃ:ɛ] on the other hand are 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 will shed some light on these questions.

The situation is different for English. Holst and Nolan (1995) have investigated /sʃ/ sequences across word boundaries in British English. They found not an assimilation effect which is gradual, as the ones of other consonant sequences across languages (cf. 1.1). At the one end of the continuum, there were examples of non-assimilations represented by two individual spectrally stable sibilant sections. At the opposite end, there were sibilant sections, which were spectrally indistinguishable from /ʃ/ in a comparable environment, and which may thus be regarded as examples of complete regressive assimilations. Additionally, Holst and Nolan report different intermediate forms. They comprise successive spectral changes from more [s]-like to more [ʃ]-like sibilant qualities (i.e. ‘contour segments’ in the sense of Hayes 1992) and constant frictions with spectra in between the ones of [s] and [ʃ]. Parallel to this continuum of different spectral patterns, Holst and Nolan found a duration continuum, in which the non-assimilated sibilant sequences had the longest and the complete assimilated ones the shortest overall durations. The spectrally intermediate forms were also marked by intermediate durations (cf. Browman 1995). The EPG-based investigation of Zsiga (1995) revealed similar gradual regressive assimilation patterns within the /sʃ/ sequences of American English that fit in with previous studies by Catford (1977) as well as by Zue and Shattuck-Hufnagel (1980). Moreover, the data of Zsiga (1995) points to a considerable inter-speaker variation, not only for the realization of the /sʃ/ sequences, but also for the production of single /s/ and /ʃ/ sounds in similar environments.

French and English differ in several aspects. For example, Delattre and Olsen (1969) found on the basis of written-text analyses that French mainly consists of open syllables (i.e. CV or V, 76%), whereas closed syllables are predominant in English (e.g., CVC, CVCC, 60%). Similar values are reported by Goldman et al. (1996) for larger corpora. In consequence, consonant sequences across syllable and word boundaries are much more likely to occur in English than in French. Hence the two languages provide a different potential for assimilations within such sequences. In addition, the syllable as a (cognitive) unit of speech seems to play a much more important rule in French than in English (cf. Cutler et al. 1986). Furthermore, French and English tend to differ in speech rhythm. French is less stress-timed than English (cf. Roach 1983), which means that there is less temporal compression and hence less need for articulatory simplifications between two perceptually salient syllables. Furthermore, the study of Dart (1991) shows that French /s/ is more often produced with a laminar articulation in the dental-alveolar region, whereas English /s/ typically articulated in an apical way in the alveolar-postalveolar region. So, place assimilations within sibilant sequences would involve different articulators and hence require different spatial-temporal movements in French and English. Finally, French speakers are known to use a higher speaking rate than English ones (in terms of syllables per second, cf. Dellwo et al. 2006)

1.3 Research assumptions

Differences like the ones summarized above are likely or even known to influence the interaction of adjacent sounds (cf. 1.1). Thus, if French shows place assimilation within sibilant sequences, it is suggestive that it will be qualitatively and quantitatively different from the English one. Specifically, based on the notes on French sibilant assimilation in 1.2, we start from the following assumptions:

- (F1) In principle, assimilation of place of articulation does occur within French sibilant sequences.
- (F2) This assimilation is a much more pervasive phenomenon than postulated in the literature. That is, it takes place without accompanying Schwa deletion (i.e. for speakers without regional accent) and irrespective of the phonological voicing features of the sibilants. Moreover, it is a productive phenomenon and can be found beyond fixed expressions like “*je suis*” and “*je sais*”.
- (F3) The assimilation can be progressive as well as regressive, depending on whether the postalveolar sibilant is the first or the second element in the sequence. In both cases, the assimilation is gradual, ranging from non-assimilations over – temporally and spectrally – intermediate forms to complete assimilations.

As for the English sibilant sequences, we assume that

- (E1) the present study replicates the spectrum of – exclusively regressive – assimilation patterns in the time and frequency domains that showed up in previous studies, particularly in the one of Holst and Nolan (1995).
- (E2) overall, the regressive assimilation patterns are asymmetrical. That is, alveolar sibilants can be assimilated by following postalveolar ones, whereas there is no comparable assimilation towards alveolar in postalveolar-alveolar sequences. The latter process is unlikely, since, cross-linguistically, assimilations lead away from (apical-) alveolar articulations. Considering that such articulations may be viewed as short, but complex movements that are saddled up on the global (vowel-related) tongue-body gestures, the consistent assimilation in favour of non-alveolar sounds gave rise to the claim that assimilatory processes aim at reducing effort in speech production, while maintaining salient acoustic-perceptual properties (cf. Kohler 1990; Jun 2004).

Finally, with regard to the assimilation patterns that were observed within consonant sequences across languages (cf. 1.1), we assume that the degree of assimilation in the French and English sibilant sequences is affected by variables like (a) the speaker, (b) the frequency of the two words involved, and (c) the salience of the corresponding syllables.

The investigation is based on acoustic measurements in the time *and* frequency domains within comparably structured corpora of read, segmentally-controlled sentences. The great diversity in the temporal dynamics of spectral changes that can occur in gradual assimilations as well as the correlation between different types of spectral changes in English sibilant sequence and the overall durations of these sequences suggest to consider both domains of the signal, which has rarely been done in previous (acoustic) studies on place assimilation. Further details of the method are summarized in the following sections.

2. Method

2.1 Sentence material

The speech corpus for the investigation of the French sibilant sequences is based on 72 sentences that result from three different 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 inserted into the three symmetrical vowel contexts /i/___/i/, /a/___/a/, and /u/___/u/. In combination, this yields 24 test sentences. The three non-nasalized vowels are articulatory and acoustically very distinct (cf. Blache and Meunier 2004 for characteristic formant values of the French vowels) in order to determine

whether the vowel contexts affect the degree of assimilation (cf. Hughes and Halle 1956; Nartey 1982; Soli 1981). For the same reason, the secondary subset of test sentences combines the four phonologically voiced or voiceless sibilant sequences (a)-(d) with all six asymmetrical vowel contexts, which leads to another 24 sentences.

The remaining 24 sentences form a complementary subset, in which each of the four individual sibilants /s/, /ʃ/, /z/, and /ʒ/ is paired across word boundaries with a labial consonant (C) like /p/ or /v/ in the two possible orders __C and C__. The labial consonants do not involve the tongue (tip) as active articulator. So, articulatory interferences with the sibilants are minimized. As in the primary subset, the eight sequences of the complementary subset were combined with the three symmetrical vowel contexts. The individual sibilants of the complementary subset function as reference sounds. That is, their acoustic measurements are used as baseline values that provide a parametric frame. Relating the sibilant measurements from the primary and secondary subsets to this frame allows for conclusions about the direction and degree of assimilation in the sibilant sequences. Particularly with regard to duration, individual sibilants were preferred over sibilant geminates (e.g., /ss/, /ʒʒ/), since their values are more informative references. Moreover, including sequences with labial consonants should help to make the aim of the study less obvious for the subjects.

Overall, the structure of the English sentence list is comparable to the French one. However, since English does not have /ʒ/ at word edges, the primary subset only comprises the four sibilant sequences (a), (b), (g), (h). For the same reason, the secondary subset contains just the two voiceless sequences (a)-(b). Thus, the two subsets are reduced to 12 sentences each. The complementary subset again combines the individual sibilants with labial consonants in the two possible orders __C and C__ across word boundaries. Due to the lack of /ʒ/ at word edges, it consists of 18 sentences. So, in total, the English list comprises 42 sentences. Furthermore, the vowel contexts the sibilant sequences were placed in are similar to the French ones. However, the number of vowels had to be extended in order to find suitable English word pairs. Three pairs of vowel phonemes are used instead of three single phonemes. That is, in some cases /i/, /u/, and /a/ are substituted by /ɪ/, /ʊ/, or /ʌ/ respectively. Although there are differences in the second formant frequency, the vowel phonemes in each pair differ mainly in the frequency of the first formant (cf. Wells 1962). Since the spectra of sibilants vary primarily according to the F2 of the surrounding vowels (e.g., Soli 1981), the different vowels in each pair should not have substantial influences on the acoustic manifestations and hence on the spectral measurements of the sibilants. Moreover, in view of Jongman (1998) it may be assumed that the quantity differences within each pair do not significantly affect the sibilant durations.

Finally, in both the French and the English sentences, the word pairs which contain the target sequences of vowels and consonants consist of one to three syllables and are placed towards the end of the sentences. Moreover, the word boundary between the sibilants is mainly located within a syntactic noun or verb phrase. The lists of French and English sentences are given in the Appendices I and II.

2.2 Recording procedure

The recording procedure was parallel for French and English. That is, for each of the two languages the sentence list was produced in four different randomizations by four female subjects. They were 20-56 years old and known to speak without a noticeable regional accent. In the case of French, this also means that none of the subjects regularly produces a Schwa at the end of the investigated target words. The restriction to female speakers was to exclude

gender-specific effects on the durational and spectral measurements (cf. Gordon et al. 2002). Prior to the individual recordings, the speakers were instructed to read the sentences displayed on a screen in front of them as naturally as possible with a constant loudness suitable for a normal conversation. Moreover, each sentence had to be said twice with a different speaking style: first, with a slow and careful pronunciation, as in a talk in front of an audience; then with a fast, casual style, as in an informal conversation with a good friend. We expected that the direct repetition and the resulting contrast would help the speakers to produce the two styles in consistent and clear ways.

The French subjects produced the sentences at the Laboratoire Parole et Language of the Université de Provence. The English ones were recorded at the Department of Psychology of the University of York. The recordings were made digitally in anechoic chambers at a sampling rate of 44.1 kHz and with a 16 bit amplitude resolution. The first of the four repetitions of the sentence list was preceded by 10 dummy sentences, which resembled the ones on the list, and which were to familiarize the subjects with their task. The subjects were given a short break after each repetition of the sentence list. Overall, the recording sessions took about 40 minutes. Since the French list comprises 72 sentences, produced in four rounds with two different speaking styles, each of the four subjects elicited 576 sentences, which results in a total number of 2,304 sentences. The English corpus comprises 1,344 sentences. A short interview after each session showed that none of the 4 French and English speakers guessed the actual reason of the recording. Most of the subjects speculated that the study's aim was phonetic effects of speaking-rate differences; others assumed that it was concerned with speech-melody patterns or with consonant-vowel interactions in general. The interviews were also recorded to get some spontaneous speech (around 3 minutes) from each subject.

2.3 Labelling and measurements

The French and English sentences were labelled with regard to the following aspects:

- (a) beginning and end of the sentence,
- (b) vowel onset in the first target syllable, vowel offset in the second one,
- (c) onset and offset of the sibilant sequence,
- (d) onset and offset of the voiceless section within the sibilant sequences (only for sequences with heterogeneous voice features),
- (e) onsets and offsets of pauses within the sentence,
- (f) a spectral bipartition of the sibilant sequence that is also mirrored in perception.

Like (e), (f) is an optional label. In the following, (f) is referred to as the 'B' (=boundary) label. It is a general marker of discontinuity, i.e. it is placed 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. Consequently, sibilant sequences without a 'B' label are either marked by a constant sibilant sound quality or (which was rather rare) by a continuous transition between two different sibilant sound qualities (cf. assimilation types A-B vs. C-D of Holst and Nolan 1995). Figure 1 presents an example of a sibilant sequence with and without the 'B' label. It underlines that the 'B' label does not necessarily separate two clear alveolar and postalveolar sibilants, but that it may be regarded as a phonetic mirror of a succession of two different phonological place features (i.e. alveolar-postalveolar or postalveolar-alveolar) that form two sections, which are primarily shaped by more postalveolar or more alveolar articulations. Therefore, bipartite sequences are referred to as consisting of postalveolar and alveolar *sections*, not of postalveolar and alveolar *sibilants*.

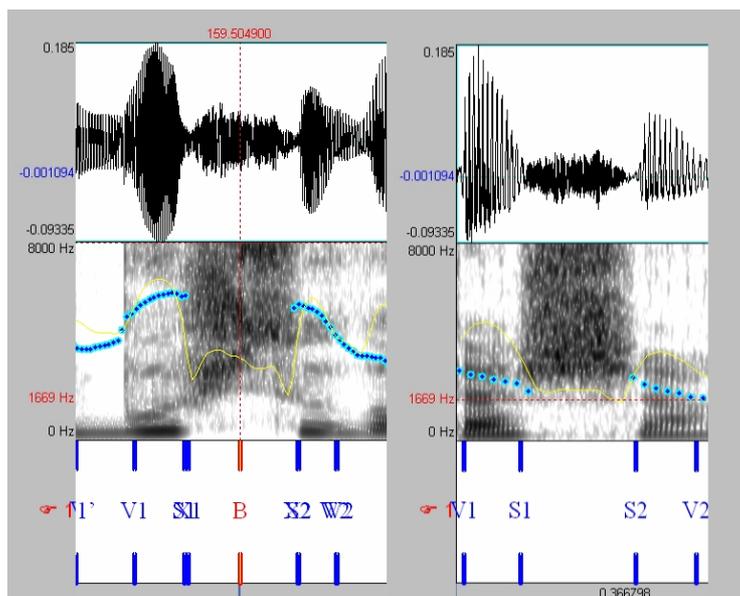


Figure 1. Left: the acoustic and perceptual bipartition of the sibilant section (/ʃs/) in the utterance “*Tu te cache si mal*” is indicated by a ‘B’ label at the boundary between the two parts (French speaker MTE). Right: the phonological sibilant sequence /sʃ/ in the utterance “*Cette une classe chargée*” is realized as a spectrally constant section. Thus, no ‘B’ label is set (French speaker ABO). From the top, oscillogram, sonagram, and label tier are shown. The dotted line within the sonagram represents the F0 contour.

All labelling was done (within a point tier) in *praat*. In some cases, a pause (cf. e) 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 the further analyses of the sibilant assimilation patterns, and only the labels (a) and (e) were set.

The acoustic analyses of the data comprised durational as well as spectral measurements. As regards the former, the following values were determined by means of the labelling:

- the overall sentence duration without all pauses, based on the labels (a) and (e),
- the vowel durations in the first and second syllable of the sibilant sequence, using the (b) and (c) labels,
- the overall duration of the sibilant sequence, based on the (c) labels,
- the proportion of voicelessness in the overall sibilant sequence for those sequences with heterogeneous phonological voice features, using the (c) and (d) labels,
- the relative durations and the corresponding duration ratio of the two sibilant sections within the sequence, based on the labels (c) and (f).

The ‘B’-label-based duration ratios are suitable to measure the degree of assimilation in the temporal domain. They are complemented by a measure in the frequency domain: the centre of gravity (CoG). The CoG represents the average value of the frequencies in the spectrum, weighted by their amplitudes. In previous studies on a number of languages, the CoG values or similar measures turned out to be good representatives of the acoustic and perceptual differences between alveolar and postalveolar sibilants (cf. Hughes and Halle 1956; Heinz and Strevens 1961; Jassem 1968; Nartey 1982; Hoole et al. 1993; Dart 1998; Jongman et al. 2000; Gordon et al. 2002; Goodacre and Nakajima 2005; Toda 2007). The CoGs of postalveolar sibilants are found between 2-4kHz, while the alveolar ones are substantially higher, typically between 4-8kHz, and they show more variation due to the speaker or the phonetic context. In the present study, two CoG values were determined for each sibilant or sibilant sequence: the

mean CoG and the CoG range, i.e. the difference between the smallest and the largest value. The CoG measurements are to some extent related to the ‘B’-label-based duration measurements. For example, sibilant sequences with longer postalveolar sections may also show lower mean CoG values, and the spectral discontinuity indicated by the presence of a ‘B’ label may be reflected in a larger CoG range. In principle, however, the two types of measurements are independent from each other.

The CoG measurements were based on spectral analyses, using a 30ms (Hamming) window (512 FFT points). It was shifted through the sibilant or the sibilant sequence in steps of 7ms. Hence, number of spectral slices varied according to the overall duration of the sibilant section. For each slice, the CoG was calculated (in Hz) within a frequency range between 1.5 and 15 kHz. This band pass covers the section of the spectrum which was found to be decisive for the distinction between alveolar and postalveolar sibilants (cf. also above), while it simultaneously excludes irrelevant components like the fundamental frequency. In a following step, the mean CoG and the CoG range were determined, considering the values of all slices of the corresponding sibilant section. All analyses and measurements were done automatically, using a MATLAB script.

The vowel durations of each of the target syllables were measured in order to estimate the salience of the corresponding syllable or word with regard to the findings of, for example, Fry (1958) and Klatt (1976). Finally, in addition to the acoustic measurements, the frequencies of the words that contain the sibilants of the sequence were determined. The French word frequencies were estimated on the basis of a corpus of movie subtitles which comprises about 52 million words (cf. New et al. 2007). The English estimations were based on the British National Corpus (BNC, cf. Leech et al. 2001), which includes around 18 million words.

2.4 Statistical analyses

The primary aim of this study is the investigation of sibilant assimilation patterns in French and English. Therefore, it concentrates on the sentences of the fast, casual condition. Three analogues mixed-model, linear-regression analyses were performed on the sibilant sequences for each language to examine each of the dependant variables separately, i.e. duration ratio, mean CoG, and CoG range. All three analyses included three groups of independent variables: (a) order of place of articulation within the sibilant sequence, phonological voicing features within the sibilant sequence, (b) following or preceding vowel context, frequency of each word (recalculated as a relative, logarithmic value), duration of the two vowels surrounding the sibilant sequence, and (c) speaker as well as sentence. While the two variables in the latter group (c) were considered as random effects, the remaining ones were fixed effects that represented either control variables (group b) or experimental variables (group a). The vowel-context variable (group b) accounts for the findings of Hoole et al. (1993), who showed that French is mainly characterized by carryover co-articulation between sibilants and adjacent vowels, while the opposite is true for English. Therefore, the analysis of the vowel context, constituted by the contrasts /u/ vs. /a/ and /u/ vs. /i/, is restricted to the following vowels in French and to the preceding ones in English. In connection with the CoG measures, four further mixed-model, linear-regression analyses were performed for each language to compare the sibilant sequences with the single alveolar and postalveolar sibilants.

In all analyses variables with coefficients more than two standard deviations from zero ($T > 2$) are regarded as reliable. The results of the analyses are summarized for each language in 3.2 and 3.3. The preceding section 3.1 briefly presents some general characteristics of the sentence and sibilant realizations within the French and English corpora.

3. Results

3.1 General findings

In addition to the main analyses, a further series of t tests were performed to address general aspects of the data, which go beyond the assumptions (F1)-(F3) and (E1)-(E2) that guided this study. For example, one series was to compare the sentence durations (without pauses) in the two speaking-style conditions for each speaker in the two languages. The results of the t tests show consistently that, overall, the fast, casual sentences were highly significantly shorter than the slow, careful ones. For most of the speakers, this shortening amounts to average values of about 30-35%. Only the French subject CDL deviates in this respect from all other speakers. The means for her slow, careful sentences are already slightly smaller than the values for the fast, casual sentences of all other speakers; and her fast, casual sentences are even 30% shorter.

Complementary, a spontaneous-speech section of around 10 seconds was selected by chance from the interview of each speaker. After subtracting all pause durations, its speaking rate was determined in phonological syllables per second and compared with the corresponding mean values of the read sentences from the two speaking-style conditions. For most of the speakers the speaking rate of the spontaneous section did not differ substantially from the fast, casual condition. Again, the French subject CDL is the only exception. She had had considerably more syllables per second in the fast, casual sentences than in the spontaneous-speech section. Furthermore, from a more general point of view, the syllables-per-second values of the French subjects were slightly higher than the ones of the English subjects, which goes well with the findings of Dellwo et al. (2006).

Further t tests based on the overall durations of the sibilant sections within the French and English fast, casual sentences revealed that the single sibilants were significantly shorter than the sibilant sequences. The sequences were on average around twice as long (i.e. 80-120%) as the single sibilants. The latter had mean durations of 60-90ms, whereas the sibilant sequences were mostly marked by values between 140ms and 180ms. However, a closer look at the individual data yielded two exceptions: CDL (French) and CMO (English). In their cases, the sibilant sequences are on average just about 50% longer than the single sibilants. In the case of CDL, this is due to extremely short mean sibilant-sequence durations of 80-90ms. In another test series, it was consistently found across the speakers of both languages that voiceless sibilants or sibilant sequences were significantly longer than voiced ones and that single, alveolar sibilants were significantly shorter than single, postalveolar ones.

Finally, contrary to the slow, careful sentences, the fast and casual ones form an almost complete dataset. That is, the number of analyzable sentences with no pauses within the sibilant sections was 1,092 (or 94.8%) in the case of French and 663 (or 98.7%) in the case of English. In the slow, careful sentences, the subjects paused considerably more often within the sibilant sequences, particularly in French.

3.2 Analyses of the French sibilants

3.2.1 Duration ratios within the sequences

The results of the mixed-model, linear-regression analysis given in Table 1 show a strong main effect of the place of articulation within the sibilant sequences. The duration ratio was significantly higher for postalveolar-alveolar than for alveolar-postalveolar sequences (45.1% vs. 76.6%, cf. Tab.1). That is, the postalveolar section was always the longer one. Figure 2 illustrates the relative mean durations of the eight sibilant sequences for each of the four

subjects. It may be seen that the described effect of the order of the place of articulation is not consistent across the subjects. It primarily shows up for ABO and MTE. To a lesser extent it is also present for CDL. However, the effect cannot be observed at all for IVI. In her case, the duration ratio is closely around 50% for all eight sibilant sequences.

An additional descriptive analysis of the mean absolute durations within the sibilant sequences revealed further speaker-specific patterns. In the case of MTE, for example, the differences in the duration ratios are primarily established by a shortening of the alveolar section in the postalveolar-alveolar sequences, while the mean duration of the first sibilant remains similar in both orderings, the postalveolar-alveolar as well as the alveolar-postalveolar one. In addition to the shift in the duration ratio, this also results in longer overall durations for alveolar-postalveolar than for postalveolar-alveolar sequences. The latter also holds for IVI. However, the longer overall durations of the alveolar-postalveolar sequences involve both sibilants to a comparable extent. Therefore, this difference to the postalveolar-alveolar sequences is not reflected in the duration ratios. The shifts in the duration ratios that were found for CDL are mainly based on a strategy that is diametrically opposed to the one of MTE. That is, compared with the alveolar-postalveolar sequences, she lengthens the initial (postalveolar) section in the postalveolar-alveolar ones, while the duration of the subsequent (alveolar) section remains similar. Hence, the postalveolar-alveolar sequences have longer overall durations than the alveolar-postalveolar ones. Finally, ABO is the only subject whose differences in the duration ratios are caused by complementary changes of the durations of both sibilants within the sequences, which leave the overall durations of all eight sequences approximately constant.

Table 1. Results of the mixed-model, linear-regression analyses, based on the duration-ratio measurements (predictor variable). From left to right, the names of the (levels of the) fixed-effects variables as well as the corresponding regression coefficients, standard errors, and T values are given. As for the voice features, ‘vd’ and ‘vl’ refer to ‘voiced’ or ‘voiceless’. Asterisks indicate significant outcomes.

	comp. var. levels	coefficient	std. error	T value
group(a)	order place of artic.	0.35	0.05	6.48*
	voice(vlvl-vdvd)	0.15	0.08	2.03*
	voice(vlvl-vlvd)	0.01	0.10	0.13
	voice(vlvl-vdvl)	0.07	0.09	0.82
	follow. vowel (/i/-/a/)	0.04	0.07	0.59
group (b)	follow. vowel (/i/-/u/)	0.19	0.07	2.6*
	vowel dur. (initial)	388.97	161.12	2.41*
	vowel dur. (final)	357.55	127.14	2.81*
	frequ. (1 st word)	0.01	0.01	0.61
	frequ. (2 nd word)	0.01	0.01	1.00

The duration ratio was also significantly affected by the phonological voicing patterns within the sibilant sequences. The ratio was higher, i.e. the first sibilant was longer, in completely voiced than in completely voiceless sequences. Sibilant sequences with mixed voicing features (i.e. voiceless-voiced or voiced-voiceless) yielded intermediate average ratios, which were not significantly different from the ones of the completely voiced and voiceless sequences.

As regards the control variables, no significant effect was found for the frequencies of the words involved. However, there is a clear effect of the vowel context, i.e. the duration was

lower for following /u/ than for /a/ and /i/. Moreover, the duration ratio increased significantly with longer durations of the preceding and following vowels.

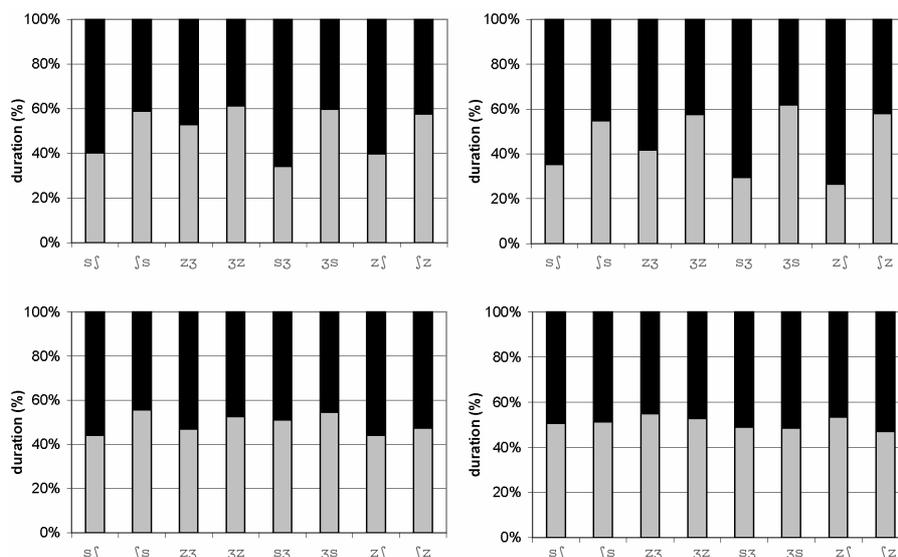


Figure 2. Relative mean durations of the first (grey) and the second sibilant (black) in the eight different sibilant sequences of the French fast, casual sentences for the four speakers ABO (top, left), MTE (top, right), CDL (bottom, left), and IVI (bottom, right).

3.2.2 Centre-of-gravity values within the sequences

As regards the experimental variables (group a), the results of the mean CoGs are similar to the ones of the duration ratios. That is, there are significant effects for both the order of the place of articulation as well as the voicing pattern (cf. Tab. 2 and Fig. 3). Phonologically voiced sequences yielded lower mean CoGs than voiceless sequences and sequences with heterogeneous voicing features. Significantly lower mean CoGs were also found for the alveolar-postalveolar sequences compared with the postalveolar-alveolar ones. At the same time, these two sequences differ significantly in CoG means and ranges from both the single postalveolar and the single alveolar reference sibilants. Overall, however, the mean CoG values of the sibilant sequences were more similar to the single postalveolar than to the single alveolar reference sounds (cf. T values in Tab.3 and Fig. 3), while the CoG ranges of the sequences are equally larger than the ones of the references.

To further clarify this pattern, Figure 4 shows the mean CoG (horizontal axis) and the CoG range (vertical axis) values for the individual tokens of the sibilant sequences and reference sibilants. The reference sibilants form two clear groups of tokens, differing in mean CoG. The distributions of sibilant sequences are more variable in both their means and their ranges; and the two orderings, alveolar-postalveolar and postalveolar-alveolar, overlap considerably. For both orderings, some tokens are produced with a large range and intermediate mean CoG, which points to transitioning CoGs across the sequences; other tokens are produced with restricted range and mean CoG similar to the single postalveolar sibilants, particularly in the case of ABO. There are few or no sibilant sequences similar to the alveolar referents in both mean CoG and CoG range.

None of the group-(b) control variables (vowel context and word frequency) had an influence on the mean CoG values. This holds for the single reference sibilants as well as for the sibilant sequences. However, both single sibilant and sibilant sequences have in common that

the CoG ranges were significantly affected by the following vowel quality. That is, the ranges are larger before /u/ than before /a/ and /i/ (cf. Tab. 2-3).

Table 2. Results of the mixed-model, linear-regression analyses for the sibilant sequences, based on the predictor variables mean CoG (left) and on the CoG range (right). From left to right, the names of the (levels of the) fixed-effects variables as well as the corresponding regression coefficients, standard errors, and T values are given. As for the voice features, ‘vd’ and ‘vl’ refer to ‘voiced’ or ‘voiceless’. Asterisks indicate significant outcomes.

		mean CoG			CoG range		
comp. var. levels		coefficient	std. error	T value	coefficient	std. error	T value
group(a)	order place of artic.	138.85	53.71	2.58*	301.39	89.21	3.38*
	voice(vl vl-vd vd)	240.93	72.86	3.31*	153.06	121.03	1.26
	voice(vl vl-vl vd)	189.03	86.56	2.18*	119.28	143.79	0.83
	voice(vl vl-vd vl)	241.80	93.73	2.58*	67.28	155.64	0.43
group(b)	follow. vowel (/i/-/a/)	-104.31	69.06	-1.51	289.94	114.62	2.53*
	follow. vowel (/i/-/u/)	-115.71	67.72	-1.71	235.94	112.46	2.10*
	vowel dur. (initial)	1.690.13	1.068.96	1.58			
	vowel dur. (final)	1.447.75	946.23	1.53			
	frequ. (1 st word)	-0.58	7.18	-0.08			
	frequ. (2 nd word)	-1.55	6.73	-0.23			

Table 3. Excerpt of the results of the mixed-model, linear-regression analyses for the comparison between sibilant sequences and single sibilants, starting from the alveolar (top) or the postalveolar (bottom) reference, with the predictor variables mean CoG (left) and CoG range (right). From left to right, the names of the (levels of the) fixed-effects variables as well as the corresponding regression coefficients, standard errors, and T values are given. As for the voice features, ‘vd’ and ‘vl’ refer to ‘voiced’ or ‘voiceless’; ‘av’ and ‘pav’ refer to alveolar and postalveolar.

		mean CoG			CoG range		
comp. var. levels		coefficient	std. error	T value	coefficient	std. error	T value
	place (av-av pav)	-1323.62	82.52	-16.04*	523.98	100.93	5.19*
	place (av-pav av)	-1194	82.05	-14.63*	822.17	100.34	8.19*
	place (av-pav)	-1606.65	90.27	-17.87*	-16.99	110.29	-0.15
	voice(vl vl-vd vd)	229.55	64.68	3.54*	126.39	79.09	1.60
	follow. vowel (/i/-/a/)				183.96	82.12	2.24*
	follow. vowel (/i/-/u/)				176.26	79.39	2.22*
	place (pav-av pav)	283.03	82.58	3.43*	540.97	101.06	5.35*
	place (pav-pav av)	412.65	81.76	5.05*	839.17	99.99	8.39*
	place (pav-av)	1606.65	90.27	17.8*	16.99	110.29	0.15
	voice(vl vl-vd vd)	229.55	64.68	3.55*	126.39	79.09	1.64
	follow. vowel (/i/-/a/)				183.96	82.12	2.24*
	follow. vowel (/i/-/u/)				176.26	79.39	2.22*

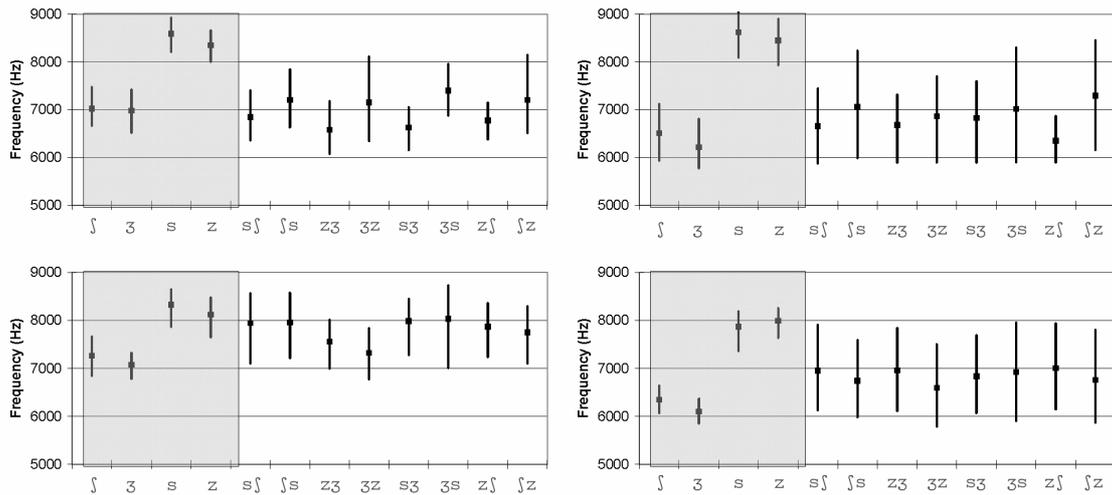


Figure 3. Average of the mean CoG (dots) and the CoG range (vertical bars) of the single reference sibilants (grey, left) and the sibilant sequences (right) for the four French subjects ABO (top, left), MTE (top, right), CDL (bottom, left), and IVI (bottom, right).

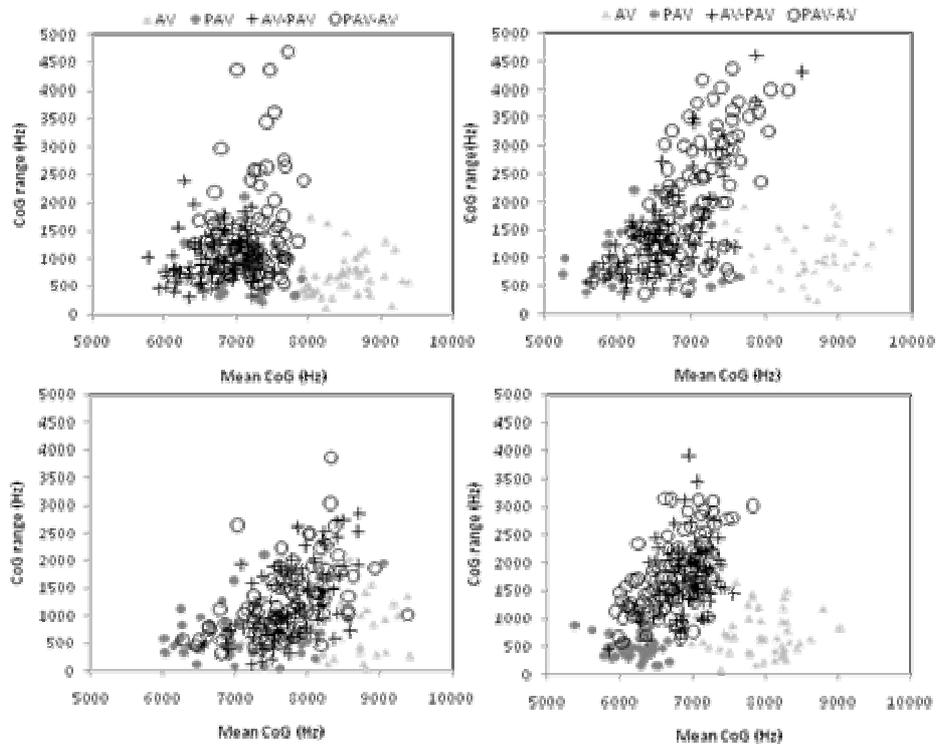


Figure 4: Distribution of individual tokens for the reference sibilants (postalveolar: grey circles, alveolar: grey triangles) and sibilant sequences (alveolar-postalveolar: crosses, postalveolar-alveolar: open circles) for each of the four French subjects ABO (top, left), MTE (top, right), CDL (bottom, left), and IVI (bottom, right).

3.3 Analyses of the English sibilants

3.3.1 Duration ratios within the sequences

The mixed-model, linear-regression analysis yielded only one significant trend, which concerns the order of the place of articulation. The alveolar-postalveolar sequences tended to show smaller duration ratios than the postalveolar-alveolar ones (cf. Tab.4). This shift in the duration ratios, which is comparable with the French one, though less pronounced, may also

be observed in Figure 5. Moreover, Figure 5 shows that the second sibilant section was consistently longer than the first one, irrespective of the corresponding place-of-articulation feature. In general, the English temporal patterns are marked by less speaker-specific differences than the French ones.

Table 4. Results of the mixed-model, linear-regression analyses, based on the duration-ratio measurements (predictor variable). From left to right, the names of the (levels of the) fixed-effects variables as well as the corresponding regression coefficients, standard errors, and T values are given. As for the voice features, ‘vd’ and ‘vl’ refer to ‘voiced’ or ‘voiceless’. Asterisks indicate significant outcomes.

	comp. var. levels	coefficient	std. error	T value
group(a)	order place of artic.	0.26	0.13	1.99
	voice(vl vl-vl vd)	0.25	0.16	1.61
	voice(vl vl-vd vl)	0.12	0.40	0.30
group(b)	preced. vowel (/a/-/u/)	0.09	0.07	1.26
	preced. vowel (/i/-/u/)	0.09	0.10	0.92
	frequ. (1 st word)	0.05	0.03	1.75
	frequ. (2 nd word)	0.01	0.01	0.56

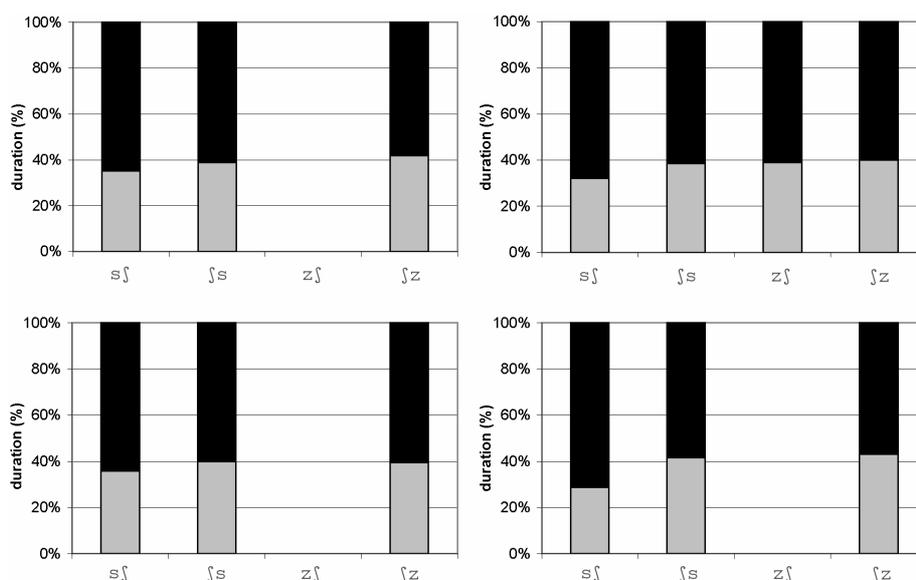


Figure 5. Relative mean durations of the first (grey) and the second sibilant (black) in the four different sibilant sequences of the English fast, casual sentences for the four speakers CMO (top, left), FAY (top, right), DGO (bottom, left), and FDU (bottom, right).

3.2.2 Centre-of-gravity within the sequences

Like in French, order of place of articulation had an effect on the mean CoG and the CoG range. The mean CoG was lower for the alveolar-postalveolar than for the postalveolar-alveolar sequences. The CoG range was significantly larger for postalveolar-alveolar sequences (cf. Tab. 5). Both can clearly be seen in Figure 6. Moreover, the voiceless sequences had lower mean CoGs and smaller CoG ranges than the mixed-voice sequences, particularly the voiceless-voiced ones. Like in French, the spectral patterns of the English sibilant sequences were also influenced by the group-(b) control variables (cf. Tab.5). For example, mean CoGs were significantly different following the front or open vowel qualities /i, ɪ/ and /a, ʌ/ compared to following the close, back, rounded vowels /u, ʊ/. However, different from French, they were not lower, but higher for /u, ʊ/. A high frequency of

occurrence of the second word in the target-word pairs also led to higher mean CoGs and larger CoG ranges in the sibilant sequences.

As shown in the Figures 6-7, the mean CoGs of the postalveolar-alveolar sequences were found to differ from both the alveolar and the postalveolar references (cf. Tab.6), while the alveolar-postalveolar sequences differ *only* from the alveolar, and *not* from the postalveolar references. This represents a major difference from the French results. Moreover, the majority of the mean CoGs of the postalveolar-alveolar sequences are located almost exactly in between the two reference sibilants. As for the CoG ranges, the English results parallel the French ones in that the postalveolar-alveolar sequences show significantly larger CoG ranges than the reference sibilants, which points to successive CoG transitions across the sequences. The alveolar-postalveolar CoG ranges, however, are only marginally larger than the postalveolar reference, and do not differ at all from the alveolar references (cf. Tab.6). This is again an important deviation to the spectral patterns in French. In addition, the English single alveolar sibilants show larger CoG ranges than the postalveolar ones, cf. Tab.6). Finally, it must be noted – different from French, but analogues to the English duration ratios – that the spectral patterns of the individual sibilant sequences are very similar between the four subjects.

Table 5. Results of the mixed-model, linear-regression analyses for the sibilant sequences, based on the predictor variables mean CoG (left) and on the CoG range (right). From left to right, the names of the (levels of the) fixed-effects variables as well as the corresponding regression coefficients, standard errors, and T values are given. As for the voice features, ‘vd’ and ‘vl’ refer to ‘voiced’ or ‘voiceless’. Asterisks indicate significant outcomes.

		mean CoG			CoG range		
comp. var. levels		coefficient	std. error	T value	coefficient	std. error	T value
group(a)	order place of artic.	843.2	51.5	16.4*	1187.54	62.23	19.08*
	voice(vl vl-vlvd)	259	127.2	2*	417.9	156.43	2.67*
	voice(vl vl-vdvl)	-78.3	68.2	-1.1	-143.85	82.5	-1.74
group(b)	preced. vowel (/a/-/u/)	-297.5	51.7	-5.8*	16.96	62.58	0.27
	preced. vowel (/i/-/u/)	-217.1	61	-3.6*	5.61	74.15	0.08
	frequ. (1 st word)	16.1	17.5	0.9	-10.46	21.27	-0.49
	frequ. (2 nd word)	29.6	13	2.3*	40.09	16.12	2.49*

Table 6. Excerpt of the results of the mixed-model, linear-regression analyses for the comparison between sibilant sequences and single sibilants, starting from the alveolar (top) or the postalveolar (bottom) reference, with the predictor variables mean CoG (left) and CoG range (right). From left to right, the names of the (levels of the) fixed-effects variables as well as the corresponding regression coefficients, standard errors, and T values are given. As for the voice features, ‘vd’ and ‘vl’ refer to ‘voiced’ or ‘voiceless’; ‘av’ and ‘pav’ refer to alveolar and postalveolar.

		mean CoG			CoG range		
comp. var. levels		coefficient	std. error	T value	coefficient	std. error	T value
place (av-av pav)		-2107.3	113.6	-18.5*	-90.4	76.9	-1.18
place (av-pav av)		-1284.2	109.9	-11.7*	1091.9	74.6	14.64*
place (av-pav)		-2044.3	119.6	-17.1*	-247.7	80.9	-3.06*
frequ. (2 nd word)		-13	12.1	-1.1	19.6	8.5	2.31*
place (pav-av pav)		-62.9	112.6	-0.6	157.3	76.2	2.07*
place (pav-pav av)		760.1	109.5	6.9*	1339.6	74.1	18.07*
place (pav-av)		2044.3	119.6	17.1*	247.7	80.9	3.06*
frequ. (2 nd word)		-13	12.1	-1.1	19.6	8.5	2.31*

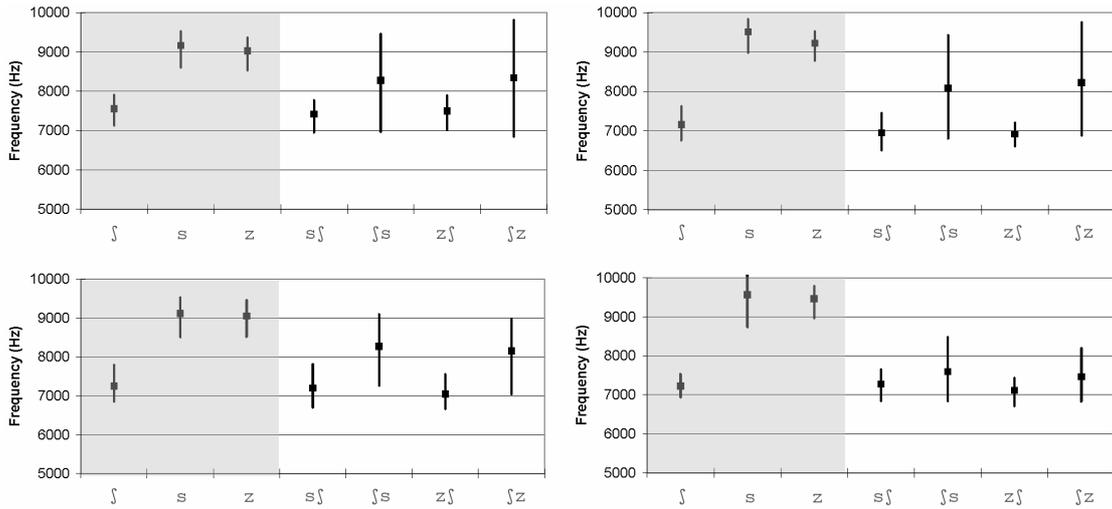


Figure 6. Average of the mean CoG (dots) and the CoG range (vertical bars) of the single reference sibilants (grey, left) and the sibilant sequences (right) for the four English subjects CMO (top, left), FAY (top, right), DGO (bottom, left), and FDU (bottom, right).

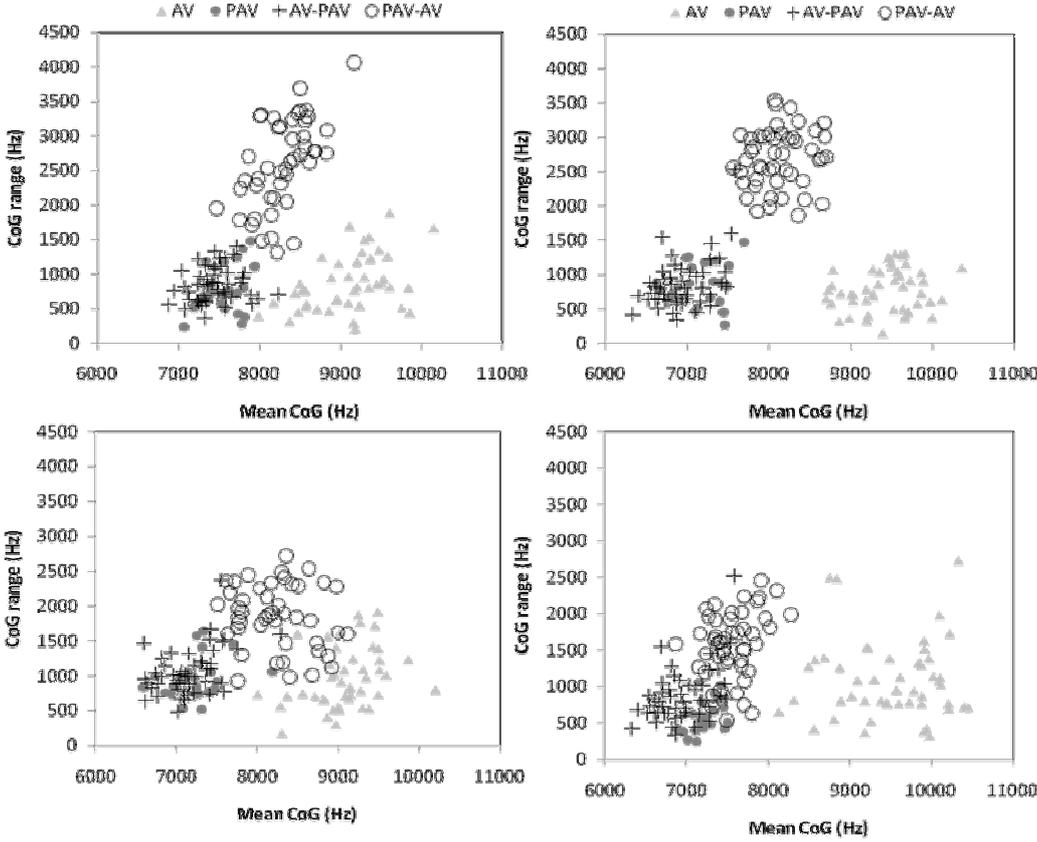


Figure 7: Distribution of individual tokens for the reference sibilants (postalveolar: grey circles, alveolar: grey triangles) and sibilant sequences (alveolar-postalveolar: crosses, postalveolar-alveolar: open circles) for each of the four English subjects CMO (top, left), FAY (top, right), DGO (bottom, left), and FDU (bottom, right).

4. Discussion

4.1 Basic interpretation of the French results

Overall, ABO and MTE showed clear systematic shifts in the relative durations of the two sibilant sections in favour of postalveolar regardless of sibilant order. So, postalveolar sections were consistently longer than alveolar ones. Moreover, the sibilant sequences of ABO and MTE are characterized by mean CoGs that are more similar to the postalveolar than to the alveolar references. Also the average CoG ranges overlap almost exclusively with the ones of the postalveolar reference sibilants. At the same time, all sibilant sequences were significantly longer than the reference sibilants. This holistic pattern may be interpreted as an assimilation phenomenon towards postalveolar. It can be progressive as well as regressive, depending on whether the sibilant sequences are postalveolar-alveolar or alveolar-postalveolar ones. However, for alveolar-postalveolar sequences the duration ratios deviate stronger from 50%. Additionally, the mean CoGs of these sequences turned out to be significantly lower than the ones of postalveolar-alveolar sequences. These two findings indicate that the regressive place assimilation is stronger (i.e. more often complete) than the progressive one.

For IVI, there was no assimilation of place of articulation in the time domain. In the cases the sibilant sequences were realized as discontinuous, bipartite spectral patterns (so that a ‘B’ was set), the two sibilant sections had always about the same durations. However, this does not rule out that these sibilant sections are spectrally shifted towards postalveolar (cf. 2.3). And, as clearly illustrated in Figure 5, this was true for IVI, independent from the order of the sibilants in the sequences. On the whole, the spectral assimilation effects of IVI are similar to the ones of MTE, though less pronounced.

The results of CDL are contradictory. On the one hand, the shifts in the duration ratios are qualitatively comparable with the ones of ABO and MTE. That is, the durations within the sequences change in favour of longer postalveolar sibilant sections. On the other hand, the CoG values, i.e. means and average ranges, are more similar to the alveolar than to the postalveolar references. This overall pattern may be explained with regard to the speaking rate. It was noted in 3.1 that compared with the other subjects the sentence durations of CDL are extraordinarily short, which, in turn, is caused by an extremely high speaking rate, accompanied by a high degree of reduction. For example, the sibilant sequences of CDL are on average only marginally longer than the single reference sibilants of the other speakers. Furthermore, it must be considered that ‘B’ labels were set in most of the sibilant sequences of CDL, which indicates the presence of two spectrally and perceptually different sections (cf. 2.3). The latter requires complex articulatory movements, which should be to some extent incompatible with the extremely high speaking rate of CDL. Thus, it is likely that CDL was not able to reach the intended articulatory targets within the sibilant sequences of the fast, casual sentences and that this led to a general upward shift of the CoG values. This assumption is supported by two aspects. First, a closer perceptual investigation of the CDL’s sibilant sequences showed that they sound softer, more “windy”, and less distinct than the ones of the other speakers. Overall, this suggests less close strictures between the active and passive articulators. Secondly, preliminary acoustic and perceptual analyses of CDL’s sibilant sequences in the slow, careful sentences yielded CoG patterns that are more similar to the ones of the other subjects shown in Figure 2.

However, irrespective of whether the contradictory findings for CDL (particularly the CoG-related ones) are or are not due to artefacts, the present study provides for the first time systematic empirical evidence for the presence of assimilation of place of articulation in French, based on sibilant sequences. Thus, in general the present findings are in line with the

assumption (F1) made in 1.3. Furthermore, as claimed in (F2) the assimilation showed up without simultaneous Schwa deletion and beyond the frequent (and hence possibly) fixed expressions “*je sais*” and “*je suis*”. So, considering the descriptions in the literature, assimilation of place of articulation in French sibilant sequences may be regarded as a productive, widespread phenomenon that takes place across subjects (from different regions), and combinations of words, and that can be progressive as well as regressive, as assumed in (F3). However, there are also substantial speaker-specific differences along all dimensions of the assimilation process. For example, the speakers differ in the extent they show assimilation in the time and frequency domains (IVI vs. ABO, MTE), in the extent they show regressive and progressive assimilation (ABO vs. IVI), or in the assimilatory strategy, which can be a strengthening of the postalveolar and/or the weakening of the alveolar element (MTE vs. CDL).

4.2 Basic interpretation of the English results

By and large, the phonetic patterns of the individual sibilants or sibilant sequences are similar across the four English subjects. The alveolar-postalveolar sequences are marked by longer postalveolar sections, smaller CoG ranges, and lower mean CoGs than the postalveolar-alveolar ones, while, at the same time, both kinds of sequences are longer than the single sibilants. Moreover, the CoG ranges of the alveolar-postalveolar sequences are very similar or even statistically comparable with the ones of the single reference sibilants, and the mean CoGs of the alveolar-postalveolar sequences differ only from the alveolar, but not from the postalveolar references. Thus the alveolar-postalveolar sequences are almost indistinguishable from the postalveolar references, except that they are consistently longer. These findings may be regarded as clear manifestations of regressive place assimilation. Thus, the present findings confirm the assumption (E1) and are hence also in line with previous studies (cf. 1.2, 1.3).

At the same time, there is no clear evidence that this regressive assimilation process is symmetrical. That is, the assimilation seems to be oriented towards postalveolar and is hence restricted to alveolar-postalveolar sequences, as assumed by (E2). However, it must be noted that there were a few exceptions to this general finding. The majority of them occurred for the speakers CMO and DGO. Two examples of the latter speaker are given in Figure 8. It can clearly be seen that the /ʃs/ sequence in “Spanish sea” has just a very short postalveolar section, whereas the /ʃs/ sequence in “Rosebush season” is spectrally entirely comparable with the adjacent alveolar sibilants (and with the alveolar references in the same vowel contexts, cf. the individual tokens in Fig.7). This is true acoustically, as well as perceptually. It is possible that examples like these are due to slips of the tongue. On the other hand, it must be considered the speakers were instructed to repeat sentences with mispronunciations. So, the fact that the speakers did not repeat sentences with /ʃs/ realizations like the ones in Figure 8 suggests that they judged their own productions as acceptable. In view of this, the following temporal aspect of the results is also striking. Despite the significant boundary shift towards ‘postalveolar’ in the alveolar-postalveolar sequences, the postalveolar-alveolar sequences are also dominated by a (clearly and consistently) longer alveolar section. Consequently, the exceptional sequences, which were mainly found for CMO and DGO, and for which there are no counterparts in French, raise doubts about the general validity of the claim that regressive place assimilation in English sibilant sequences is an asymmetrical process, which is solely directed towards postalveolar and hence restricted to alveolar-postalveolar sequences. Follow-up studies should deal with this issue in more detail.

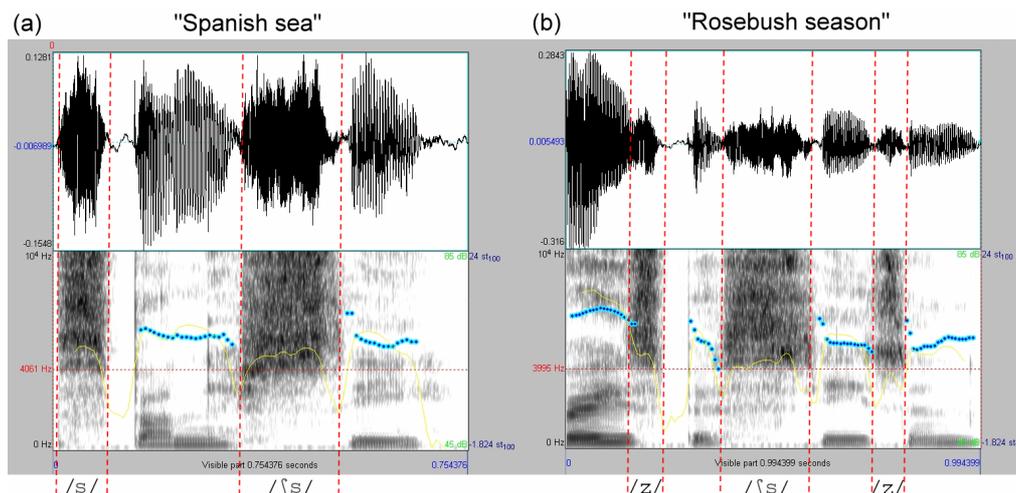


Figure 8. Examples of the two target-word pairs “*Spanish sea*” (a, left) and “*Rosebush season*” (b, right), produced by DGO. The /ʃs/ sequences in both pairs are overall spectrally very similar to the adjacent single alveolar sibilants (/s,z/). Only in (a), there is a discontinuity in the sequence that may be regarded as reflecting a short initial postalveolar section.

4.3 General discussion

The present study demonstrated by means of sibilant sequences that assimilation of place of articulation does not only exist in English, but also in French. It was further revealed that the assimilation processes in the two languages show fundamental differences. The differences are best described with reference to the two dimensions that characterize every assimilation process: its direction and its target quality. On this basis, the French sibilant assimilation is clearly determined by the target quality, since all assimilations are in favour the postalveolar sibilant. Simultaneously, direction does not play a major role in French. However, it is also not completely irrelevant, since we found evidence that the regressive assimilation was stronger than the progressive one. By contrast, in the case of English, the direction is the central characteristic of the assimilation process. It is strictly regressive. In addition, this regressive assimilation is dominated by the postalveolar target quality. But, in view of the examples of assimilations towards alveolar (cf. Fig.8), the target quality is not as important and restricting as in French. So, from a simplified point of view, the assimilations of place of articulations within the sibilant sequences of French and English may be described as *primarily quality-guided* or *primarily direction-guided*, respectively. This contrastive finding represents the major new insight of the present study, together with the fact that the assimilations in both languages not only manifested themselves in the frequency domain, but also in the time domain. Thus, such a two-dimensional perspective should also be taken in follow-up studies on assimilation in general.

Moreover, the processes in both languages are gradual rather than categorical. So, although there are instances of spectrally complete assimilations (type D) in French (progressive and particularly regressive ones towards postalveolar) as well as in English (regressive ones towards postalveolar and, rarely, towards alveolar), the great diversity of spectral patterns in both languages cover also all the other types of patterns (A-C) that are differentiated in Holst and Nolan (1995) and that are schematically illustrated in Figure 9 for alveolar-postalveolar sequences. On the other hand, there were considerably less instances of spectrally complete assimilations in French than in English; and for English the assimilation process was also found to be more consistent across sibilant sequences and speakers.

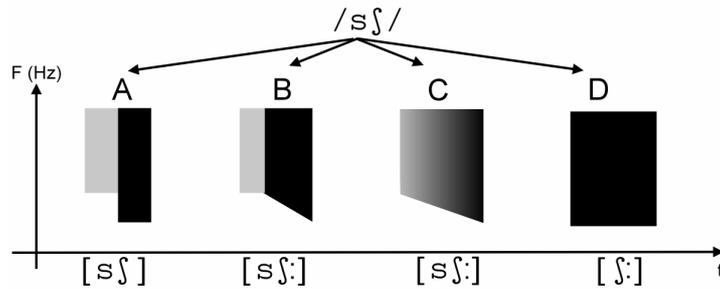


Figure 9. Schematic illustrations of spectral patterns found for alveolar-postalveolar sequences in French and English, based on the types differentiated by Holst and Nolan (1995). Type A and B require a ‘B’ label, type C and D not.

However the data of four speakers are not sufficient for general comparative conclusions about the consistency, the frequency, and the completeness of the assimilation processes in French and English. For example, a recent EPG study with a male French subject by Fiasson (2008) came up with the same assimilation pattern that was revealed in the present study, i.e. a consistent and strong (or even complete) progressive and regressive assimilation towards postalveolar in sibilant sequences. So, in connection with the descriptions from the literature (cf. 1.2), place assimilation in French sibilant sequences may be a more consistent and pervasive process than the present findings suggest. At the same time, for English the studies of Holst and Nolan (1995), Zsiga (1995), or Ellis and Hardcastle (2002) revealed substantial speaker-specific differences in place assimilation processes, which are largely absent in the present study. The present study also yielded clearly more cases of (spectrally complete) type-D assimilations (cf. Fig.9) in alveolar-postalveolar sequences than the one of Holst and Nolan (1995), i.e. more than 80% vs. less than 60%. Thus, basically further studies with more subjects are needed in order to make more detailed quantitative comparisons between the assimilation processes in both languages. However, with regard to the phonotactic, cognitive, and rhythmic differences between French and English summarized in 1.2, it is reasonable to assume that in general English place assimilation processes are more consistent and homogeneous across subjects and segmental sequences than French ones.

In connection with quantitative comparisons, it should also be noted that the CoG values that were measured for the alveolar and postalveolar reference sibilants in French and English are considerably higher than the ones reported for comparable speech sounds in previous studies (cf. 2.3). However, this is primarily due to the large frequency range used in the present study, which included frequencies up to 15kHz, whereas the frequency ranges in other studies were already cut off at much lower frequencies. That is, the relatively high CoG values do not reflect any extraordinary articulatory, acoustic, and perceptual phenomena, but are mere methodological artefacts. The significantly smaller CoG ranges that characterized the postalveolar sibilants compared with the alveolar ones in French and English indicate that the alveolar ones were acoustically more variable. This is again in line with findings from other cross-linguistic studies (cf. Nartey 1982; Gordon et al. 2002). In general, the CoG measurements of the present study underline the statement of Ladefoged (1971:57) that sibilants have “*a comparatively large amount of energy at high frequencies*”.

Furthermore, the place assimilation patterns found for French and English are not artefacts of the remaining experimental and control variables, i.e. voicing, adjacent vowel context, and word frequency. Yet, these variables had significant influences on the temporal and spectral measurements. For example, it was found for English that a higher frequency of the second word in the target-word pairs is accompanied by higher mean CoGs and larger CoG ranges. Comparisons between the sibilant sequences of the three target-word pairs with the highest

and the lowest second-word frequencies suggest that these significant effects are based on a shortening of the sibilants in the words with higher frequencies. Such a shortening goes well with the positive correlation between word frequency and articulatory reduction that was noted, for example, by Ernestus et al. (2006). Shorter speech sounds with more reduced articulations are marked by less close strictures and more pronounced co-articulations, which further explain the higher mean CoGs and the larger CoG ranges. Co-articulation is also likely to be the reason for the significant effects of the back, close, rounded vowel qualities /u, ʊ/ on the spectral measurements in French and English. A temporal overlap between the rounding of the (following or preceding) vowel and the sibilant sequence explains the larger CoG ranges (French) as well as the lower mean CoGs (English), compared with the other, unrounded vowels. It is also well known from other studies (on other languages) that rounding is a feature which typically spreads over adjacent segments (e.g., Bell-Berti and Harris 1982; Perkell and Matthies 1992). At the same time, the lack of significant differences between /i, ɪ/ and /a, ʌ/ shows that the open-close dimension of the vowel space had no effect on the sibilant-sequence realization. Finally, the lower mean CoGs that were found for single sibilants and sibilant sequences containing voicing are likely to have a twofold origin: First, the vibration of the vocal folds slows down the supraglottal airflow, which in turn reduces the acoustic energy of the friction, particularly for higher frequencies. Secondly, the vibration adds formants that increase the acoustic energy in the lower region of the frequency spectrum (cf. Jongman et al. 2000). Moreover, it is known and supported by the present findings (cf. 3.1) that (phonologically) voiced segments are typically shorter than (phonologically) voiceless ones (while simultaneously the preceding vowel is longer, e.g., Klatt 1976; Flege 1992). This explains the greater duration ratios that were found for sibilant sequences with a voiced segment in second position in French, and a similar tendency in English.

As regards voicing, the present study also determined the duration of the voiceless interval within the sibilant sequences that have heterogeneous phonological voice features. These measurements may be used to get a basic idea of the voice assimilation that was produced in French and English, in addition to the place assimilation processes of primary interest. The corresponding findings are summarized in Figure 10. Figure 10 shows for French that voiced-voiceless sequences were largely phonetically voiceless, whereas voiceless-voiced sequences were produced almost completely voiced. This pattern, which was consistently found for all four speakers, fits in well with the regressive voice assimilation that is known to occur in French (cf. Malmberg 1969; Carton 1974; Fagyal et al 2006). Thus, French sibilant sequences can undergo two simultaneous phonological processes, which concern place and voice, and which can even go into different directions. That is, progressive place assimilation can coincide with regressive voice assimilation (cf. “*je sais*” and “*je suis*”, 1.2). As for the English sibilant sequences, different voicing patterns were found for CMO and the remaining three subjects FAY, DGO, and FDU. While for CMO the voiced-voiceless sequences were marked by clearly less voicing than the voiceless-voiced ones, which as in French points to regressive voice assimilation, the other three subjects produced both kinds of sibilant sequences almost completely voiceless. Together, this again matches perfectly with descriptions from the literature. For instance, Roach (1983:14) wrote that “*only regressive assimilation is found across word boundaries, and then only of one type [...]. If C^f [i.e. the word-final consonant] is a lenis (i.e. “voiced”) consonant and Cⁱ [i.e. the word-initial consonant] is fortis (“voiceless”) we often find that the lenis consonant has no voicing; this is not a very noticeable case of assimilation, since [...] initial and final lenis consonants usually have little or no voicing anyway*”. However, there are no indications that the voice assimilation patterns of CMO interact with her place assimilation patterns. The same holds for all French subjects.

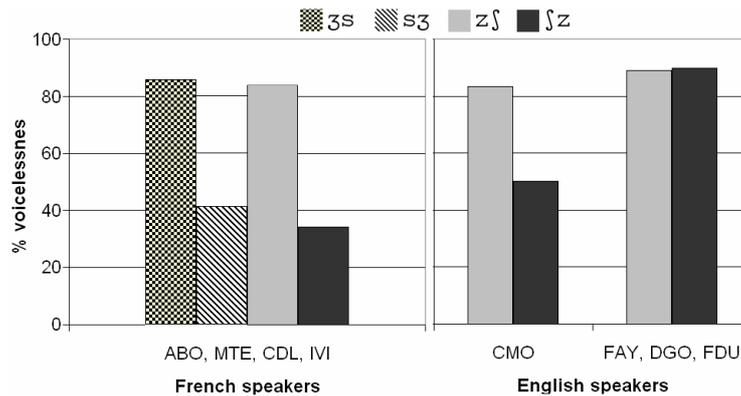


Figure 10. Mean percentages of voicelessness in the French (left) and the English (right) sibilant sequences with heterogeneous voice features across. The French percentages refer to all four subjects, while for English the values of CMO are shown separately from the ones of the remaining three subjects FAY, DGO, and FDU.

The place assimilation continua of French and English, with non-assimilated and fully assimilated sibilant sequences at its extremes, are incompatible with the view of assimilation as a phonological process which takes place at a cognitive level prior to the actual speech production and in which a feature of a certain segment is replaced by the feature of another (usually adjacent) segment (cf. 1.1). Such a process requires categorically different phonetic patterns. Several approaches have been made to fill this gap between empiricism and theory. For example, it was proposed by Holst and Nolan (1995:330) that in general “*an apparent continuum of assimilation involves two separate phenomena*”. While the non-assimilated and completely assimilated extremes are manifestations of the actual, categorical phonological process, the spectrum of intermediate forms is caused by mere “*mechanical effects*” (Holst and Nolan 1995:330) that occur during the articulatory implementation of the unchanged sequence of phonological features and that vary gradually according to phonetic factors. This proposal is not supported by the results of the present study. In particular, two aspects pose a problem for the proposal of Holst and Nolan.

First, their proposal implies that if there are differences between the assimilation processes of two languages, they will show up solely for the extremes of the observed phonetic continua. However, the present study found systematic, language-specific differences between the sibilant sequences of French and English that were not completely assimilated. They may be observed most clearly for the ‘B’-label based duration ratios (cf. Fig.2 and Fig.5). Such differences would not be expected, if these intermediate forms actually have a mere mechanical origin, at least not as long as phonetic factors like phrasing or speaking rate are largely comparable between the two compared languages, as in the present study. Rather, the systematic shifts in the duration ratios of the non-complete assimilations suggest a pre-planned assimilatory basis. Moreover, the shifts in the duration ratios that were found within French as well as between French and English are also difficult to explain by different degrees of temporal overlap of two articulatory gestures (one for each phonological feature), as it is done within the framework of articulatory phonology (cf. Browman and Goldstein 1992; Browman 1995). In order to model the temporal patterns of the sibilant sequences of French and English within this framework, it is necessary to include changes in the temporal extensions of the involved gestures, in addition to changes in their temporal overlap. Secondly, it is also problematic for the proposal of Holst and Nolan (1995) that spectrally constant type-D assimilation patterns (cf. Fig.9) can have friction qualities which are in between the ones of the two alveolar and postalveolar reference sibilants. According to Holst and Nolan, they should all be indistinguishable from the postalveolar reference. This was

actually true in many, but – crucially – not in all cases. Different from the approach of Holst and Nolan, the latter spectral variation can basically be explained within the framework of articulatory phonology by assuming two overlapping gestures which differ in the range of their movements. In the case of spectrally complete assimilation, one of them is reduced to zero. However, we agree with Ellis and Hardcastle (2002:394), who characterize this as an “*unsatisfying ad hoc account*”.

Although the approach of Ellis and Hardcastle (2002) is different from the one of Holst and Nolan (1995), it also starts from the assumption that assimilation is essentially categorical at the phonetic level. In general, this assumption is supported by results of Ellis and Hardcastle (2002), which suggest that the gradual variation found for the investigated assimilatory process may be a methodological artefact, based on the lack of consideration for inter-speaker differences. While some speakers did not assimilate at all or consistently produced spectrally complete assimilations, other speakers’ productions were characterized by intermediate forms. So, the individual speakers showed categorically different behaviours, but overall the assimilation appears to be gradual. On the one hand, this goes well with the present findings. Particularly the gradual overall picture of French may partly be ascribed to inter-speaker differences (cf. 4.1). On the other hand, however, there are also considerable intra-speaker variations and/or contradictory time and frequency patterns for individual subjects. Hence, on the basis of the present results it is not entirely possible to shift the categorical patterns that should characterize assimilation according to the traditional theoretical point of view from the inter-subject level to the level of individual subjects.

In summary, it may be concluded that none of the above phonological frameworks is able to model and to explain the complete spectrum of observed time and frequency patterns in French and English in a straightforward way. Hence, further improvements are necessary. For this, it seems promising to go beyond the level of the traditional segment and the rigid distinction between phonetics and phonology (cf. Flemming 1997), as it is done within the articulatory phonology as well as in the more recent exemplar-based approaches (cf. Hawkins 2003). Furthermore, more attention should be paid to the perception of assimilation or to the interrelations between production and perception. For instance, not every assimilation that is gradual in articulatory and acoustic terms is actually perceived as gradual. Holst and Nolan (1995) note that many spectrally incomplete assimilations of the types B and C (cf. Fig.9) appeared to be complete in perception, even for trained listeners. According to perceptual analyses of the present data by the main author (ON), the same holds for many sibilant sequences with ‘B’-labels of the present study, at least if they are played in larger contexts like the corresponding target-word pairs. On the other hand, particularly in view of Nolan (1992) and similar studies, it seems unlikely that a perceptual perspective will yield solely categorical assimilations. Theories of speech perception have already acknowledged that assimilation is essentially a gradual phenomenon. The same must hold for theories from the other side of the speech chain. Within this perspective, further investigations on how assimilation is perceived, particularly from a comparative, cross-linguistic point of view, may also contribute indirectly to improve phonological theories and production models.

An improvement may also be achieved by regarding assimilation as a meaningful phenomenon in a general communicative sense, not (just) as a means of reducing effort by articulatory re-organizations. That is, by the presence or absence of an assimilation the speaker indicates a certain speaking style, which in turn signals a certain relationship to the dialogue partner. The fact that assimilation is optional and hence “*affects instantiations of specific phonemes in defined situations*” rather than “*all instantiations of all phonemes*” (Wood 1996:160) is a prerequisite to be used as a carrier of meaning in speech. The

possibility that assimilation conveys communicative meaning, and its potential to yield phonetic patterns that are (almost) indistinguishable from other patterns with different phonological starting points – potentially requiring special perceptual processing – are aspects which may be suitable to keep assimilation separate from co-articulation.

References

Ambrazaitis, G.I and T. John (2005). On the allophonic behaviour of German /x/ vs /k/ - an EPG investigation. *Working papers of the Institute of Phonetics and digital speech processing Kiel 35*, 1-14.

Amerman, J.D., R. Daniloff, K. Moll (1970). Lip and jaw coarticulation for the phoneme /XXX/. *Journal of Speech and Hearing Research 13*, 147-161.

Armstrong, L. E. (1932). *The Phonetics of French: a practical handbook*. London: G. Bell and Sons

Barry, M. (1985). A palatographic study of connected speech processes. *Cambridge Papers in Phonetics and Experimental Linguistics 4*, 1-16.

Barry, W. and S. Hawkins (1992). Comments on “On types of coarticulation”. In D.R. Ladd, G.J. Docherty (Eds.) *Papers in Laboratory Phonology 2* (pp. 138–145). Cambridge: CUP.

Batt, J.C. (1960). *French pronunciation and diction*. New York: St. Martin’s.

Bell-Berti, F. and K. Harris (1982). Temporal patterns of coarticulation: lip rounding. *JASA 71*, 449-454.

Blache, P. and C. Meunier (2004). Language as a complex system: the case of phonetic variability. *Proc. Congreso de Lingüística General, Saint-Jacques de Compostelle, Spain, XXX-XXX*.

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

Browman, C. and L. Goldstein (1992). Articulatory phonology: an overview. *Phonetica 49*, 155-180.

Carton, F. (1974). *Introduction à la phonétique du français*. Paris: Bordas.

Catford, J.C. (1977). *Fundamental problems in phonetics*. Bloomington: Indiana University Press.

Chomsky, N. and M. Halle (1968). *The sound pattern of English*. New York: Harper and Row.

Clark, J. and C. Yallop (1990). *An introduction to phonetics and phonology*. Oxford: Blackwell.

- Cutler, A., J. Mehler, D. Norris, and J. Segui (1986). The syllable's differing role in the segmentation of French and English. *Journal of Memory and Language* 25, 385-400.
- Daniloff, R.G. and R.E. Hammarberg (1973). On defining coarticulation. *Journal of Phonetics* 1, 239-248.
- Dart, S. (1991). *Articulatory and acoustic properties of apical and laminal articulations*. PhD dissertation, UCLA, USA.
- Dart, S. (1998). Comparing French and English coronal consonant articulation. *Journal of Phonetics* 26, 71-94.
- Delattre, P. and C. Olsen (1969). Syllabic features and phonic impression in English, German, French, and Spanish. *Lingua* 22, 160-175.
- Dellwo, V., E. Ferragne, and F. Pellegrino (2006). The perception of intended speech rate in English, French, and German by French speakers. *Proc. of the 3rd speech prosody conference, Dresden, Germany*, XXX-XXX.
- Dommelen, W. van (1985). Further experiments in the perception of French plosives. *Phonetica* 42, 100-115.
- Ellis, L. and W.J. Hardcastle (2002). Categorical and gradient properties of assimilation in alveolar to velar sequences: evidence from EPG and EMA data. *Journal of Phonetics* 30, 373-396.
- Ernestus, M., M. Lahey, F. Verhees, and R.H. Baayen (2006). Lexical frequency and voice assimilation. *JASA* 120, 1040-1051.
- Fagyal, Z., D. Kibbee, and F. Jenkins (2006). *French. A linguistic introduction*. Cambridge: CUP.
- Fiasson, R. (2008). XXX. MA thesis, Université de Provence, France.
- Flege, J.E. (1992). Speech learning in a second language. In C.A Ferguson, L. Menn, and C. Stoel-Gammon (Eds.), *Phonological Development: Models, Research, Implications* (pp. 565-604). York: York Press.
- Flemming, E. (1997). Phonetic details in phonology: evidence from assimilation and coarticulation. In K. Suzuki, D. Elzinga (eds.) *Southwest Workshop on Optimality Theory: Features in OT* (pp. XXX-XXX), Coyote Papers.
- Fry, D. B. (1958). Experiments in the perception of stress. *Language and Speech* 1, 126-152.
- Gadet, F. (1992). *Le français populaire*. Paris : PUF.
- Gay, T. (1978). Effect of speaking rate on formant movements. *JASA* 63, 223-230.
- Goldman, J.P., A. Content, and U.H. Frauenfelder (1996). Comparison des structures syllabiques en français et en anglais. *Proc. Of the 21st Journées d'Etude sur la Parole, Avignon, France*, 119-122.

- Goodacre, J. and Y. Nakajima (2005). The perception of fricative peaks and noise bands. *Journal of Physiological Anthropology* XXX, 151-154.
- Gordon, M., P. Barthmair, and K. Sands (2002). A cross-linguistic study of voiceless fricatives. XXX.
- Gow, D.W. (2003). How representations help define computational problems: commentary on Grossberg, Gaskell and Greenberg. *Journal of Phonetics* 31, 487-493.
- Hardcastle, W.J. (1994). Assimilation of alveolar stops and nasals in connected speech. In J. Windsor (ed.), *Studies in general and English phonetics in honour of professor J.D. O'Connor* (pp. 49-67). London: Routledge.
- Hawkins, S. (2003). Roles and representations of systematic fine phonetic detail in speech understanding. *Journal of Phonetics* 31, 373-405.
- Hayes, B. (1992). Commentary on F. Nolan, "The descriptive role of segments: evidence from assimilation". In D.R. Ladd, G.J. Docherty (Eds.) *Papers in Laboratory Phonology 2* (pp. 280-286). Cambridge: CUP.
- Heinz, J.M. and K.N. Stevens (1961). On the properties of voiceless fricative consonants. *JASA* 33, 589-596.
- Heuvel, H. van den, B. Cranen, and T. Rietveld (1996). Speaker variability in the coarticulation of /a,i,u/. *Speech Communication* 18, 113-130.
- Holst, T. and F. Nolan. (1995). 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* (pp. 315-333). Cambridge: CUP.
- Hoole, P., N. Nguyen-Trong, and W.J Hardcastle (1993). A comparative investigation of coarticulation in fricatives: electropalatographic, electromagnetic, and acoustic data. *Language and Speech* 36, 235-260.
- Hughes, G.W. and M. Halle (1956). Spectral properties of fricative consonants. *JASA* 28, 303-310.
- Jassem, W. (1968). Acoustic description of voiceless fricatives in terms of spectral parameters. In W. Jassem (ed.), *Speech Analysis and Synthesis* (pp. 189-206). Warsaw: PWN.
- Jongman, A. (1998). Effects of vowel length and syllable structure on segment duration in Dutch. *Journal of Phonetics* 26, 207-222.
- Jongman, A., R. Wayland, and S. Wong (2000). Acoustic characteristics of English fricatives. *JASA* 108, 1252-1263.
- Jun, J. (2004). Place assimilation. In B. Hayes, R. Kirchner, D. Steriade (eds.), *The phonetic basis of markedness* (pp. 377-407). Cambridge: CUP.

- Klatt, D.H. (1976). Linguistic uses of segmental duration in English: acoustic and perceptual evidence. *JASA* 59, 1208-1221.
- Kohler, K.J. (1990). Segmental reduction in connected speech in German: phonological facts and phonetic explanations. In A. Hardcastle, A. Marchal (eds.), *Speech Production and Speech Modelling* (pp. 69-92). Dordrecht: Kluwer.
- Kohler, K.J. (2001). Articulatory dynamics of vowels and consonants in speech communication. *Journal of the International Phonetic Association* 31, 1-16.
- Ladefoged, P. (1971). *Preliminaries to linguistic phonetics*. Chicago: University of Chicago Press.
- Leech, G., P. Rayson, and A. Wilson (2001). *Word Frequencies in Written and Spoken English: based on the British National Corpus*. London: Longman.
- Malmberg, B. (1969). *Phonétique française*. Lund: LiberLäromedel.
- Manuel, S.Y. and R.A. Krakow (1984). Universal and language particular aspects of vowel-to-vowel coarticulation. *Haskins Status Reports on Speech Research* 78.
- McCarthy, J.J. (1988). Feature geometry and dependency. *Phonetica* 43, 84-108.
- Nartey, J. (1982). *On fricative phones and phonemes*. PhD dissertation, UCLA, USA.
- New, B., M. Brysbaert, J. Veronis, C. Pallier (2007). The use of film subtitles to estimate word frequencies. *Applied Psycholinguistics* 28, 661-677.
- Nolan, F. (1992). The descriptive role of segments: evidence from assimilation. In D.R. Ladd, G.J. Docherty (Eds.) *Papers in Laboratory Phonology 2* (pp. 261–280). Cambridge: CUP.
- Ohala, J.J. (1993). Coarticulation and Phonology. *Language and Speech* 36, 155-170.
- Perkell, J. and M. Matthies (1992). Temporal measures of anticipatory labial coarticulation for the vowel /u/: within- and cross-subject variability. *JASA* 91, 2911-2925.
- Price, G. (1991). *An Introduction to French Pronunciation*. Oxford: Blackwell.
- Ramus, F. (2001). Outstanding questions about phonological processing in dyslexia. *Dyslexia* 7, 197-216.
- Recasens, D. (1991). An electropalatographic and acoustic study of consonant-to-vowel coarticulation. *Journal of Phonetics* 19, 177-192.
- Recasens, D. and M.D. Pallarès (2001). Coarticulation, blending and assimilation in Catalan consonant clusters. *Journal of Phonetics* 29, 273-301.
- Roach, P. (1983). *English phonetics and phonology. A practical course*. Cambridge: CUP.

Snoeren, N.D., J. Segui, and P.A. Hallé (2008). Perceptual processing of partially and fully assimilated words in French. *Journal of Experimental Psychology: Human Perception and Performance* 34, 193-204.

Soli, S. (1981). Second formants in fricatives: acoustic consequences of fricative-vowel coarticulation. *JASA* 70, 976-984.

Spencer, A. (1996). *Phonology: theory and description*. Oxford: Blackwell.

Toda, M. (2007). Speaker normalization of fricative noise: considerations on language-specific contrast. *Proc. of the 16th ICPHS, Saarbrücken, Germany*, 825-828.

Wells, J.C. (1962). *A study of the formants of the pure vowels of British English*. MA thesis, University of XXX, XXX.

Whalen, D.H. (1990). Coarticulation is largely planned. *Journal of Phonetics* 18, 3-35.

Wood, S.A.J. (1991). X-ray data on the temporal co-ordination of speech gestures. *Journal of Phonetics* 19, 281-292.

Wood, S.A.J. (1996). Assimilation or coarticulation? Evidence from the temporal co-ordination of tongue gestures for the palatalization of Bulgarian alveolar stops. *Journal of Phonetics* 24, 139-164.

Wright, S. and P. Kerswill (1989). Electropalatography in the study of connected speech processes. *Clinical Linguistics and Phonetics* 3, 49-57.

Zsiga, E. (1995). 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* (pp. 282-302). Cambridge: CUP.

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

Appendix 1

The French sentence list, subdivided into the three subsets

The primary subset

Toi tu triches si j'veux

C'est la Miss chignon

Tu te couches sous l'drap

C'est une rousse chouchou

Elle remâche sa viande

C'est une classe chargée

C'est un vestige Zimbabwéenne

C'est une devise gitane

C'est un rouge zoulou

J'ai vendu douze journaux

C'est un mariage zaïrois

C'est une phrase japonaise

Ils ont trouvé la biche zigouillée

Ils annoncent une crise chinoise

C'est un prestige si convoité

Si c'est à Nice j'y vais

J'ai rencontré de farouches zoulous

Il y a douze chouchous

Il bouge souvent

C'est une douce journée

Cette fumée cache Zagreb

C'est une danoise charmante

Il part en voyage samedi

Il ne chasse jamais

The secondary subset

Il se niche sous l'toit

C'est la Miss chouchou

Tu te couches si j'veux

C'est la housse chinoise

Tu te caches sous l'lit

C'est un passe chouré

Une grosse mouche ça pique

C'est une trousse chargée

J'étais sur la corniche samedi

C'est un fils charmant

Tu te caches si mal

C'est une tasse chinoise

C'est un vrai prodige zoulou

Nous passons une exquise journée

Ce peau-rouge zigouille ses ennemis

Il y avait douze gitanes

J'étais invité à un mariage zoulou

C'est une bourgeoise joufflue

C'est un peau-rouge zaïrois

C'est une andalouse jalouse

iJ'ai vu les vestiges zaïrois

Je ne le précise jamais

On le considère comme un sage zigoto

C'est un vase gitan

The completemtary subset

Rejoins-moi sur la place face au port
Tu te caches facilement
C'est une chinoise particulière
Ce message va à la mer
C'est un service financier
C'est une belle pouliche finnoise
Il a une devise vitale
C'est un vestige biblique
Il a fait un couscous pour toi
Il achète des cartouches pour son arme
Il regarde sa pelouse foutue
Il voit rouge pour tout

Elle tape sa soeur
C'est une gaffe charmante
C'est un pape zairois
Il ne la baffe jamais
Il se rebiffe si souvent
C'est une équipe chinoise
C'est une initiative Zimbabwéenne
Si elle s'active, j'y vais
Il bouffe sous les ponts
C'est un coupe-choux-fleur
C'est un groupe zoulou
C'est une soupe journalière

Appendix 2

The English sentence list, subdivided into the three subsets

The primary subset

They crossed the Spanish sea
He will steer this ship
He wanted to push Susan
He entered the house shoeless
The dog needs a brush sometimes
It's great how the glass shines

They have shorn these sheep
For the cake you may use sugar
He broke Anna's shower
She has a childish zeal
They want to ambush Zulus
In the football match, France tried to crush Zaire

The secondary subset

He buys fish soup
The girl has nice shoes
In May, it's rosebush season
The cold wind made the mouse shiver
hearts was the flush suit
He is a first-class shoemaker
The diplomat found Bush sarcastic
The gardener fixed the loose shovel
She likes the British south
Her friends miss Charlotte
The spy tried to rush secretly
This is a brass shield

The complementary subset

The students went to the blue-grass party
The driver received a harsh fine
This is Barbara's mouse
Do not try to kiss me
This is a fish filet
He can touch his feet
The park keeper found the goose food
The negotiations made Bush moody
The contestant was careful not to choose foolishly

I've had enough sardines
The bay was known for tough sharks
They planned to halve Zagreb
The team arrived at the crime scene
We have visited a steam ship
I bought a cheap zipper
We have to cut the bloom soon
During work, he has to wear fireproof shoes
The camera has a childproof zoom