QMC Working Papers in Speech and Language Sciences

Edited by James M. Scobbie
Volume 1 (1996) 13-25

Noël Nguyen, Fiona Gibbon, and William J. Hardcastle: Articulatory and perceptual aspects of fricative-stop coarticulation: a pilot study

1 Introduction

This paper is concerned with the influence of a preceding fricative on the production and perception of a stop consonant. In a well-known series of experiments (Mann & Repp, 1981; Repp & Mann, 1981, 1982), Mann and Repp showed that, when preceded by a fricative, an ambiguous stop acoustically halfway between /t/ and /k/ is identified differently depending on the place of articulation of the fricative. Specifically, listeners tend to identify stops more frequently as velars following /s/ than following /ʃ/. This effect was shown to occur regardless of the presence/absence of a syllable boundary between the two consonants. It also appeared to decrease gradually in magnitude as a silence of increasing duration was inserted after the fricative, although it was still significant with silent gaps as long as 375 ms. As the listeners' responses were affected by both the specific acoustic structure of the fricative and its perceived category, it was suggested that the perceptual mechanism responsible for this effect was partly continuous and partly categorical, i.e. operated both before and after phonetic categorisation.

These findings were accounted for in the framework of the motor theory of speech perception (see Liberman, 1996, for a recent synthesis). In this view, context-dependent variations in the perceived place of articulation of a stop demonstrate that some “tacit knowledge” of articulatory dynamics is involved in processing speech sounds. It was assumed that fricatives have an influence on how a following stop is produced, and, more particularly, that the place of articulation of a velar stop is shifted forward in the context of /s/. As the perceptual effect observed goes, as it were, in the opposite direction (more velar responses following /s/ than following /ʃ/), this tends to indicate that compensatory adjustments are made in the identification of a stop depending on the adjacent fricative. In other words, a perceptual mechanism of some kind was assumed to be employed to factor out coarticulatory interactions between the fricative and the stop.

For this explanation to be valid, one has to show that /s/ does induce a shift in the place of articulation of a following velar. To our knowledge however, there is still little data available on fricative-stop coarticulation. In an acoustic investigation involving eight speakers of American English, Repp and Mann (1982) measured the formant onset frequencies following the stop release in utterances such as [stɑ], [skɑ], [ʃtɑ], [ʃkɑ], but their results proved to be difficult to interpret. The F3 onset frequency was higher following /s/ than following /ʃ/, as predicted, but the fricative had no clear effect on the onset frequency of F2, and the pattern observed for F4 was at variance with M&R's expectations. The few acoustic studies conducted earlier along the same lines (Bailey & Summerfield, 1980; Malécot & Chermak, 1966; Schwartz, 1967; Uldall, 1964) were mainly concerned with the rôle of fricative transitions as cues to the place of articulation of the following stop. Patterns of coarticulation in fricative-stop sequences do not seem to have been studied extensively in a direct manner, by gathering articulatory data (although this issue is partially addressed in Byrd, 1996, and in Borden & Gay, 1979). This is somewhat surprising, given the high frequency of consonant sequences such as [st], [sk] and [ʃt] in English. In addition, the “compensation-for-coarticulation” perceptual effect turns out to have aroused a great deal of attention in several recent psycholinguistic studies, as it is considered to shed some light on lexical effects in speech perception (Elman & McClelland, 1988; McQueen, unpublished; Nguyen, in prep.; Norris, 1993; Shillcock et al., 1993). For these reasons, we believe that a more extensive study of fricative-stop coarticulation at the articulatory, acoustic, and perceptual level, is called for.

In this paper, we present the results of two pilot experiments carried out in the framework of a larger research project on this topic.

2 Perceptual data

The first experiment aimed to replicate Mann and Repp's main finding in the perceptual domain, i.e. that stop consonants acoustically midway between [t] and [k] are perceived more frequently as velars in the context of [s] than in the context of [ʃ]. The main difference with Mann and Repp's experiments was that we used /i/ and /ɑ/ as final vowels, instead of /u/ and /ɑ/. The vowel /u/ was dismissed because it brings another articulatory parameter into play, namely lip rounding, in addition to place of articulation and more general aspects of tongue dynamics. As it is well-known, anticipatory coarticulation of lip rounding can extend over quite a long time interval (see Perkell & Matthies, 1992, inter alia). Thus, in sequences such as [stu], [sku], [ʃtu] and [ʃku], lip rounding is likely to start at the beginning of the fricative, and to affect the spectral characteristics of both the fricative and the stop burst, making it difficult to isolate the specific influence of fricative-stop coarticulation on the listeners' responses (note that anticipatory labial coarticulation was not taken into consideration by M&R, in that the spectral characteristics of their synthetic fricatives were the same regardless of whether the final vowel was /ɑ/ or /u/).

The vowel /i/ was used in a first attempt to elucidate the exact nature of fricative-dependent variations in stop identification. As indicated above, Mann and Repp espoused a motoric view of speech perception, and claimed that such variations are due to the listener's compensating for an assimilatory phenomenon which takes place at the articulatory level. However, the observed perceptual effects may also be ascribed to some general property of the auditory system itself, regardless of any reference to the articulatory movements, in a view reminiscent of the auditory theories of speech perception (e.g. Diehl & Kluender, 1989). Effects of auditory contrast, for example, may be assumed to occur in fricative-stop-vowel sequences (Mann & Repp, 1981). More precisely, the high-frequency noise (typically above 4000 Hz, see Hughes & Halle, 1956) associated with /s/ may induce a lowering in the perceived frequency of the stop release burst and of the third formant at the vowel onset, therefore causing the stop to be perceived as more compact (i.e. closer to a velar) than would be the case in the context of /ʃ/. In this account, one might expect the influence of /s/ on stop identification to be reduced in the context of /i/, as the onset frequencies of F3 (typically around 2800-3000 Hz for both velar and alveolar stops) and F4 (ranging from 3400 Hz for a velar to 3900 Hz for an alveolar in a male speaker, see Stevens & Blumstein, 1978) approaches that of the lowest main spectral peak for /s/. Auditory contrast effects between /s/ and a following stop may be further reduced because of the frequency of the burst for velars being shifted upwards, in the context of /i/. As has been established by the acoustic theory of speech production (e.g. Stevens, 1972), the resonance affiliated to the front cavity at the release of a velar is continuous with F3 in the vicinity of a front vowel, whereas it is continuous with F2 in the vicinity of a back vowel.

2.1 Stimuli

Each stimulus was composed of two monosyllables. The first monosyllable was either the word “plus” or the word “dash”, thus ending either with /s/ or /ʃ/. The second monosyllable was a CVC synthetic sequence with the initial consonant taking place on a 7-step /d/-/g/ continuum[1], and the vowel being either /ɑ/ or /i/.

These stimuli were constructed using the following method. First, a male native speaker of Southern British English was asked to produce four single CVC words, “dark”, “guard”, “deed”, and “geese”, and four two-word sequences, “plus guard”, “dash dark”, “plus geese”, and “dash deed”. The utterances were recorded on a DAT tape, low-pass filtered (cut-off frequency: 9800 Hz), and digitised at a sample rate of 10 kHz with a 16-bit resolution on a SiliconGraphics workstation.

The two versions of “plus” and “dash” were then manually extracted from the two-word sequences using the signal editor Waves+. Our purpose in recording “plus” spoken in combination with both “guard” and “geese” was twofold. First, we wanted to maximise the listener's bias towards perceiving a /g/ following /s/. As the fricative was produced prior to /g/, the fricative offset was expected to contain cues pointing to a velar place of articulation for the stop. Note that trading relationships between acoustic cues are not distinguished from genuine context effects (Repp, 1982) in the present work. Second, we thought it necessary to take into account the potential long-range coarticulatory interactions between the vowel of “plus” and that in the second monosyllable (either /ɑ/ or /i/). It was assumed that the articulatory and acoustic characteristics of /ʌ/ could be slightly different in the context of /i/ as opposed to /ɑ/. Using two different versions of “plus” spoken in these two vocalic environments allowed us to maintain, to a certain extent, the acoustic coherence (Hawkins, 1995) of our stimuli. Finally, our procedure ensured that “plus” would show an appropriate prosodic pattern. The word “dash” was recorded in combination with both “dark” and “deed” for the same reason.

We used the four single words[2] “dark”, “guard”, “deed”, and “geese” as models to synthesise four CVC stimuli with the same phonetic structure (/dɑk/, /gɑd/, /did/, /gis/), using the Klatt formant synthesiser (Klatt, 1980). In a first step, the formant frequencies, F0 contour, and RMS amplitude of the acoustic signal were computed automatically for each natural utterance, by means of a set of ESPS scripts. From these values, a Klatt parameter file was generated using an ESPS/Klatt interface specifically developed for this experiment. The Klatt parameter values were then modified until the similarity between the synthesised signal and the natural one was judged to be satisfactory. The synthesiser output sampling rate was set to 10 kHz and the parameter values were updated every 5 ms. Although rather long and laborious, this copy-synthesis method had the advantage of making the synthetic stimuli acoustically coherent with the natural utterances “plus” and “dash”, i.e. of making them sound as if they originated from the same speaker.

---

[1] This consonant was a voiceless unaspirated stop, phonetically transcribed as [t] or [k]. The phonological labels /d/ and /g/ will however be used throughout this section, so as to bring the identity of the carrier words (“dark”, “guard”, “deed” and “geese”) to the attention of the reader.

[2] These words were originally selected within the framework of a larger study on the combined effects of the phonetic context and of the lexicon on speech perception (Nguyen, in preparation).

In synthesising the stimuli, we also drew to quite a large extent on the procedure employed by Stevens and Blumstein (1978), in their work on the acoustic cues to place of articulation in stop consonants. In their initial part (stop burst + vowel onset) at least, our stimuli were very similar to those of Stevens and Blumstein. From each of the four initial CVC sequences, a 7-step continuum was generated, with the initial consonant ranging from /d/ (stimulus 1) to /g/ (stimulus 7). For each continuum, the parameters manipulated were the onset frequencies of F2, F3 (for both /ɑ/ and /i/) and F4 (for /i/), the duration of the transition for F1 (shorter at the alveolar endpoint), and the duration of the interval between the burst onset and the beginning of the vowel (shorter at the alveolar endpoint). Table I gives the parameter values for each stimulus on the continuum and each vowel. The duration of the formant transition had a fixed value of 40 ms for F2 and the upper formants.

A great deal of attention was also paid to the spectral characteristics of the stop burst. As in Stevens and Blumstein (1978), the burst main spectral peaks were considered to be continuous with formants at the onset of voicing in the following vowel. For alveolars, the noise burst was produced by exciting a resonator continuous with F4 at the onset of /ɑ/, and with F4/F5 at the onset of /i/. For velars, the burst main peak was continuous with either F2 (/ɑ/) or F3 (/i/). For intermediate stimuli on each /d/-/g/ continuum, the noise burst had two spectral peaks, whose relative amplitudes were interpolated linearly (on a dB scale) between the amplitude values associated with the two end-points. In all cases, noise bursts were generated using the parallel branch of the synthesiser. The overall amplitude of the burst was adjusted by trial and error so as to be constant for all stimuli. The duration of the burst was set to 10 ms. Amplitude values (in dB) for each stimulus are listed in Table I.

stim.	F1 transition duration	F2 onset freq.	F3 onset freq.	burst onset	AF	A2F	A4F
1	30 ms	1700 Hz	2800 Hz	-10 ms	38 dB	0 dB	80 dB
2	30 ms	1690 Hz	2683 Hz	-10 ms	48 dB	13 dB	70 dB
3	35 ms	1680 Hz	2567 Hz	-15 ms	58 dB	26 dB	60 dB
4	35 ms	1670 Hz	2450 Hz	-15 ms	67 dB	40 dB	50 dB
5	35 ms	1660 Hz	2333 Hz	-15 ms	61 dB	54 dB	40 dB
6	40 ms	1650 Hz	2217 Hz	-20 ms	47 dB	67 dB	30 dB
7	40 ms	1640 Hz	2100 Hz	-20 ms	35 dB	80 dB	20 dB

Table I: Klatt parameter values for each of the seven stimuli on the /dɑ/-/gɑ/ continua. AF: overall amplitude of frication. A2F: Amplitude of noise-excited parallel second formant. A4F: Amplitude of noise-excited parallel fourth formant. See text for details.

Stim	F1 transition duration	F2 onset frequency	F3 onset frequency	F4 onset frequency	burst onset	AF	A3F	A4F	A5F
1	20 ms	2000 Hz	2800 Hz	3900 Hz	-10 ms	34 dB	0 dB	80 dB	80 dB
2	20 ms	2067 Hz	2833 Hz	3817 Hz	-10 ms	44 dB	13 dB	70 dB	70 dB
3	25 ms	2133 Hz	2867 Hz	3733 Hz	-15 ms	54 dB	26 dB	60 dB	60 dB
4	25 ms	2200 Hz	2900 Hz	3650 Hz	-15 ms	64 dB	40 dB	50 dB	50 dB
5	25 ms	2267 Hz	2933 Hz	3567 Hz	-15 ms	63 dB	54 dB	40 dB	40 dB
6	30 ms	2333 Hz	2967 Hz	3483 Hz	-20 ms	49 dB	67 dB	30 dB	30 dB
7	30 ms	2400 Hz	3000 Hz	3400 Hz	-20 ms	36 dB	80 dB	20 dB	20 dB

Table I (cont.): Klatt parameter values for each of the seven stimuli on the /di/-/gi/ continua.

The formant target frequencies for each vowel are indicated in Table II. Note that vowel /i/ was diphthongised.

	/dɑd/	/dɑk/	/did/	/gis/
F1	750 Hz	750 Hz	300-220 Hz	300-200 Hz
F2	1280-1100 Hz	1280-1100 Hz	2200-2000 Hz	2200-2300 Hz
F3	2400 Hz	2400-2250 Hz	3000-2800 Hz	3000-3200 Hz
F4	3700-3350 Hz	3700-3300 Hz	3600 Hz	3600 Hz
F5	4500 Hz	4500 Hz	4500 Hz	4500 Hz
duration	385 ms	285 ms	385 ms	235 ms

Table II: formant target frequencies and duration for /ɑ/ and /i/ in each context.

The frequencies of the fricative main spectral peaks and the fricative duration are listed in Table III for the two versions of /s/ and /ʃ/. Spectral peak frequencies were determined on the basis of a wide-band spectrogram supplemented with a DFT computed over an 80-ms Hamming window centred at the fricative midpoint.

fricative	context	duration	peak 1	peak 2
/s/	/plʌs gɑd/	133 ms	3883 Hz	4776 Hz
/s/	/plʌs gis/	143 ms	3913 Hz	4817 Hz
/ʃ/	/dæʃ dɑk/	116 ms	2866 Hz	3780 Hz
/ʃ/	/dæʃ did/	137 ms	3010 Hz	3996 Hz

Table III: duration and frequencies of the main spectral peaks for /s/ and /ʃ/ in each context.

2.2 Subjects

Eight adult subjects took part in the experiment. All were native speakers of English, with no known hearing disorder, nor any prior experience of phonetics. Each participant was paid 4 pounds. The experiment was carried out in the department of Linguistics of the University of Cambridge.

2.3 Procedure

Disyllabic sequences were formed by combining “plus” as spoken in the context of “guard” with each of the CɑC synthetic stimuli, and “plus” as spoken in the context of “geese” with each of the CiC synthetic stimuli. Likewise, “dash” spoken in the context of “dark” was combined with each CɑC stimulus, and “dash” spoken in the context of “deed” with each CiC stimulus. The duration of the interval between the fricative and the onset of the following burst was set to 70 ms.

Subjects were asked to identify the first sound of each CVC sequence as /d/ or /g/, by clicking on a button in a response panel displayed on a computer screen. Each stimulus was presented 8 times. The test itself was preceded by a trial run in which the subject was presented with the alveolar and velar endpoints played twice each for each continuum, and was given the correct response after having heard the stimulus.

There was a beep between each stimulus and the following one. The inter-stimuli interval was 3 sec. The total number of stimuli was 448. The experiment was run on a Silicon Graphics workstation used to play the stimuli and to collect the responses.

2.4 Results

Average percentages of “d” responses are displayed in Figure 1 as a function of the position of the synthetic stimuli on a /d/-/g/ continuum. Separate identification curves were drawn for each of the two preceding fricatives and each of the two following vowels. Each percentage was computed from a total of 128 responses given by 8 subjects.

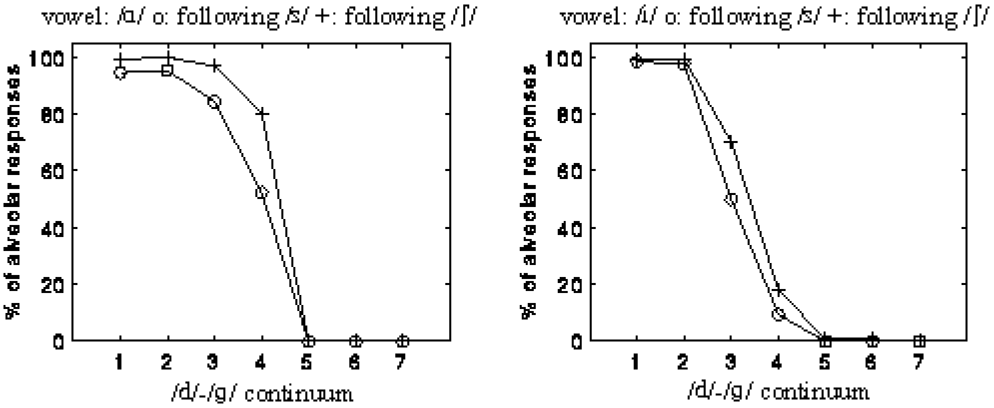

Figure 1: Percentages of alveolar responses as a function of the position of the synthetic stops on a /d/-/g/ continuum. Position 1: alveolar endpoint. Position 7: velar endpoint.
Left panel: preceding /ɑ/. Right panel: preceding /i/.

As Figure 1 shows, the endpoint stimuli were clearly perceived as belonging to different phonemic categories. Stimuli 1 and 2 on each continuum were categorised as alveolars in the great majority of cases, regardless of the phonetic context. Inversely, stimuli 5, 6 and 7 were systematically categorised as velars. In addition, the categorical boundary between /d/ and /g/ fell approximately in the middle of each continuum, as was expected, although some differences were observed depending on the following vowel (the most ambiguous stimulus being the fourth one in the context of /ɑ/, and the third one in the context of /i/). Note also that the identification functions had a rather sharp slope in the vicinity of the /d/-/g/ boundary, i.e. the subjects' responses switched rather abruptly from “d” to “g” across each continuum.

Figure 1 also reveals that stops were categorised differently depending on the preceding fricative, as they were more frequently identified as alveolars following /ʃ/ than following /s/. The difference in the percentages of alveolar responses following /ʃ/ and /s/ was significant for Stimulus 4 in the context of /ɑ/ (diff.: 28.1%, *F*(1.7) = 13.186, $p < .01$), and for Stimulus 3 in the context of /i/ (diff.: 20.3%, *F*(1,7) = 13.598, $p < .01$).

Thus, the context effect found by Mann and Repp was basically replicated. In the present experiment however, the effect turned out to be relatively small in magnitude and confined to the most ambiguous stimuli. This may be partly due to the fact that we used a /d/-/g/ continuum with only seven steps, which means that the distance between two adjacent stimuli in the acoustic space was rather large. As a result, the subjects responded in a strongly categorical manner, by tending to associate the same stimulus with the same phonetic label. This might have left less room for the fricative to exert an influence on the perception of the stop. The presence of a stop burst could also have contributed to a reduction in that influence, by making the stop still a little less ambiguous. However, fricative-dependent variations in the identification of stops proved to be quite systematic, in that they occurred for all the subjects tested.

Finally, the magnitude of the context effect did not show any significant variations depending on the following vowel ($F(1,7) = .612$, NS).

3 Articulatory data

In this second experiment, an attempt was made to characterise the influence of [s] and [ʃ] on the articulatory movements involved in the production of [t] and [k]. Articulatory data were gathered for one speaker using electropalatography.

3.1 Method

The monosyllables “da” ([tɑ]), “ga” ([kɑ]), “di” ([ti]), “gi” ([ki]), “sta” ([stɑ]), “shta” ([ʃtɑ]), “ska” ([skɑ]), “shka” ([ʃkɑ]), “sti” ([sti]), “shti” ([ʃti]), “ski” ([ski]) and “shki” ([ʃki]) were spoken 5 times each in a random order by a male, phonetically trained, native speaker of Southern British English. Syllable-initial stops were phonologically voiced but phonetically voiceless unaspirated[3]. Stops preceded by a fricative were phonologically voiceless and produced as voiceless unaspirated owing to the phonology of English. Thus, all the stops here were phonetically voiceless unaspirated, as in the first experiment and in M&R’s work. The above utterances were identical to the CV and FCV syllables used in Repp and Mann (1982)[4]. Note that our corpus did not include any VFCV sequence and, therefore, that the potential effects of the presence/absence of a syllable boundary between the fricative and the following stop on fricative-stop coarticulation were not dealt with in the present work.

EPG data were recorded at a sampling rate of 100 Hz, using the Reading EPG3 system (Hardcastle et al., 1989). In this system, dynamic variations in the configuration of the linguo-palatal contacts are monitored by means an artificial palate with 62 electrodes arranged in eight rows and eight columns. The acoustic signal was simultaneously digitised at a sampling rate of 10 kHz with a 12-bit resolution. The EPG and acoustic data were then transferred on to a Unix workstation, and analysed using a set of Matlab programs (Nguyen, this volume).

For each item, EPG patterns were extracted at the following locations:

- The onset of stop closure, defined as the first EPG frame showing a full closure in any of the three frontmost rows for alveolars, or any of the three backmost rows for velars.
- The release of stop closure, also determined on the basis of the EPG trace. The pattern retained was the one immediately preceding that release.

These two articulatory events were identified automatically. In several instances of [skɑ] however, no full velar closure could be located in the EPG trace during the stop (this point is discussed further below). In such cases, the onset of stop closure was considered to coincide with a maximum in the number of tongue-palate contacts in the most posterior row, and the stop release was taken as the first EPG frame showing a decrease in the number of contacts in the same row. Note that these annotation criteria were probably rather conservative, in that the first annotation point could in some cases be situated posterior to the actual onset of closure, and the second annotation point be prior to the actual stop release.

---

[3] Although some of them might have been prevoiced, we will use the phonetic labels [t] and [k] in all cases for the sake of simplicity.

[4] The material we used in Experiment 1 was different for being designed for a larger study which also aimed to investigate lexical effects in speech perception, as already indicated.

In addition, for each fricative we identified:

- The onset of the fricative, on the basis of the acoustic signal.
- The offset of the fricative, arbitrarily taken as the EPG frame located 20 ms before the onset of stop closure.

The EPG patterns thus identified were analysed as follows. First, a centre-of-gravity index (referred to as COG hereafter) was computed to determine the location of the main concentration of electrodes along the front-back axis, for each pattern. The index used in the present work had a value between 1 (centre of gravity coinciding with the backmost row) and 8 (centre of gravity coinciding with the frontmost row). In addition, an “average” EPG pattern was generated for each item and each of the annotation points defined above, by computing the percentage of contacts for each electrode over all repetitions.

3.2 Results

In this section we examine *a)* anticipatory coarticulation patterns in fricative-stop sequences, i.e. variations in the configuration of the tongue-palate contacts for the fricatives depending on the following stop, and *b)* carry-over coarticulatory influences of fricatives on stops.

3.2.1 Influence of stops on the preceding fricative

Average EPG centre-of-gravity values at the onset and offset of the fricative are listed in Table IV for each fricative-stop-vowel utterance.

	[stɑ]	[skɑ]	[sti]	[ski]	[ʃtɑ]	[ʃkɑ]	[ʃti]	[ʃki]
onset	4.00	3.96	3.71	4.10	3.42	3.39	3.38	3.55
offset	5.11	4.02	4.91	4.42	4.26	3.38	4.36	3.79

Table IV: mean centre of gravity at the onset and offset of the fricative for each FCV utterance. Values averaged over 5 repetitions.

Table IV shows that stops had little influence on the EPG centre of gravity at the onset of the fricative. There was a slight backward shift in the COG for both fricatives, in the context of [kɑ] as opposed to [tɑ]. Inversely, the COG moved forward when the fricative was followed by [ki] rather than [ti]. On the whole however, the observed differences were quite small (lower than 0.4).

As could be expected, variations in the COG depending on the place of articulation of the stop were greater at the fricative offset. For /s/ as well as for /ʃ/, the centre of gravity was located further back in the context of [k] than in that of [t]. This difference was greater when the following vowel was [ɑ] rather than [i].

Average EPG patterns computed at the onset and offset of each fricative for each FCV utterance are shown in Figure 2. In each pattern, the percentage of tongue-palate contacts over five repetitions for a given electrode is represented as a shade of grey (black: 100%, white: 0%). The frontmost row (which has 6 electrodes only) is on top. The fricative-stop coarticulatory patterns visible in this figure are consistent with the results of the centre-of-gravity analysis.

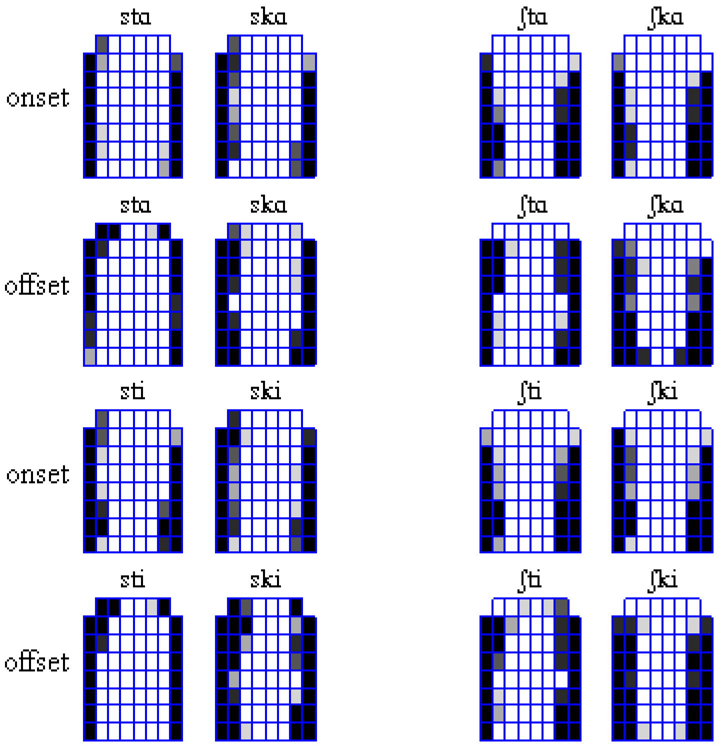

Figure 2: Average EPG patterns at the onset and offset of [s] and [ʃ] for each FCV utterance, computed over 5 repetitions. See text for details.

3.2.2 Influence of fricatives on the following stop

COG values at the onset of closure and just before the release are listed in Table V for single stops and stop-fricative sequences.

	[tɑ]	[stɑ]	[ʃtɑ]	[ti]	[sti]	[ʃti]	[kɑ]	[skɑ]	[ʃkɑ]	[ki]	[ski]	[ʃki]
onset	4.18	5.32	5.27	4.78	5.35	4.96	2.75	3.68	2.96	2.81	4.13	3.52
rel.	5.21	5.24	5.24	4.92	5.07	4.88	2.16	3.03	2.52	2.86	3.21	2.87

Table V: mean centre of gravity at the onset of stop closure and just before the release for each CV and each FCV utterance. Values averaged over 5 repetitions.

At the onset of closure, systematic variations were observed for each stop depending on the adjacent segments. Stops following a fricative had a more anterior COG than initial stops. This forward shift was larger in the context of [s] (+.99 on average) than in that of [ʃ] (+.55 on average). It was also greater for velar stops (+.97 on average) than for alveolar stops (+.74), and adjacent to [ɑ] (+.84) as opposed to [i] (+.69).

Just before the release however, the COG showed little variation for the alveolar stop, although it appeared to be slightly further back prior to [i] than prior to [ɑ]. The trends observed for the velar stop were in general the same as at the onset of closure, i.e. the COG was further forward following a fricative than in initial position (average difference: +.4), even more so when the fricative was [s] (+.61) rather than [ʃ] (+.18).

Average EPG patterns illustrating the configuration of the tongue-palate contacts at the onset of stop closure and just before the release are shown in Figure 3 for each CV and each FCV utterance.

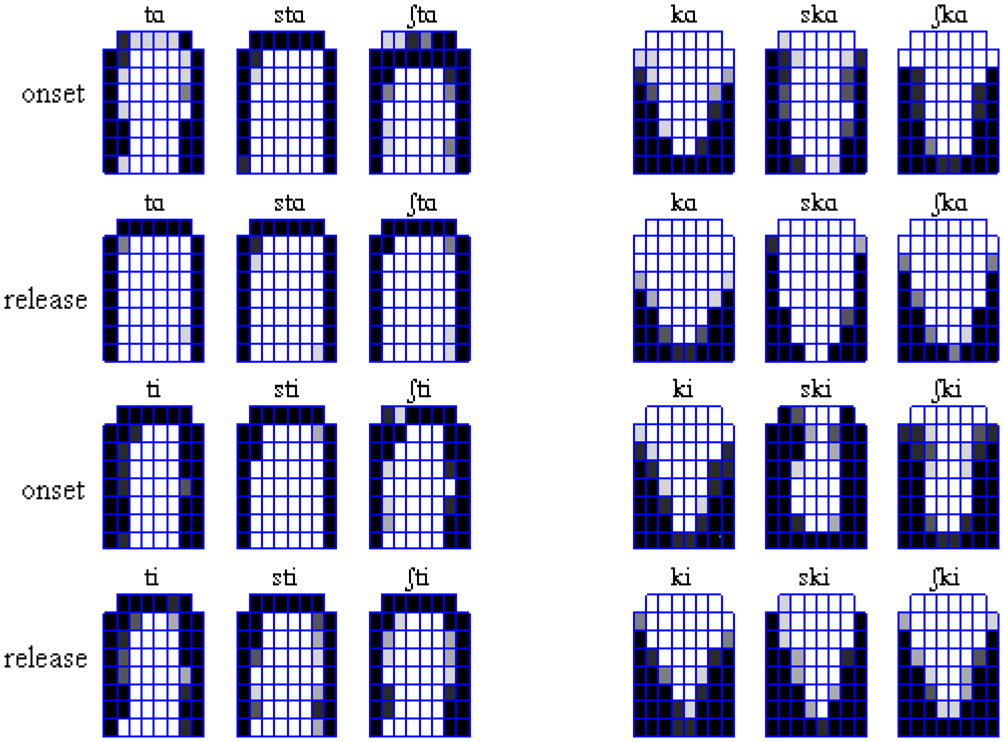

Figure 3: Average EPG patterns at the onset of stop closure and just before the release for each CV and each FCV utterance, computed over 5 repetitions.

These patterns basically confirm the trends outlined above, with one noticeable exception. By comparison with the CV sequence [kɑ], the centre of gravity for the velar stop does appear to move forward in [skɑ], in accord with our COG analysis. Note in particular the increased side contact in the alveolar region following the fricative. However, Figure 3 also reveals that in no instance did the EPG patterns show a complete velar closure, whether one looks at the stop onset or just before the stop release, as the two centremost electrodes in Row 8 (the most posterior one) always remained deactivated. This tends to suggest that the velar closure remained incomplete throughout, and that the stop was actually produced as a fricative in this context[5]. Alternatively, it can be hypothesised that the velar closure was achieved further back in the vocal tract, i.e. posterior to the back edge of the artificial palate.

In the latter case, one could speculate that (at least for this particular subject) alveolar fricatives have a dissimilatory influence on a following velar, rather than an assimilatory one, as has been assumed until now in the literature. In other words, the place of articulation for [k] would shift backward following [s], as compared to syllable-initial [k]s. However, it is clear that EPG does not provide us with enough information about tongue configuration to investigate this hypothesis. More articulatory data will have to be gathered to deal with this issue in a more extensive manner.

---

[5] Unfortunately, the acoustic signal recorded with the EPG data did not offer clear-cut clues on this point, because of its poor quality (owing to the low signal-to-noise ratio, the actual resolution of this signal was estimated to be 8 bits in preliminary tests).

4 Conclusion

Two pilot experiments on fricative/stop coarticulation have been presented in this paper. The first experiment was conducted with the aim of replicating the effect that fricatives have been reported to have on the perception of a following stop. For that purpose, subjects had to identify synthetic stops ranging on a /d/-/g/ continuum following /s/ and /ʃ/. In agreement with all the studies previously carried out on this topic, we found that a stop acoustically halfway between /d/ and /g/ was more frequently identified as a velar in the context of /s/ than in the context of /ʃ/, although this effect was restricted to the most ambiguous stimuli.

In our experiment, fricative-stop sequences were followed either by /ɑ/ or /i/. Our goal in manipulating the vowel was to determine whether the context effect results from a “compensation-for-coarticulation” mechanism (as assumed by motor theories of speech perception), or from some kind of auditory contrast between the fricative and the stop. Specifically, if it is to be accounted for in terms of auditory contrast, we hypothesised that this effect should be reduced in the environment of /i/. However, our results showed that the magnitude of the context effect did not significantly vary as a function of the following vowel. It is difficult to draw any definite conclusion from these preliminary results, and new tests involving a larger number of subjects are clearly needed here.

Experiment 2 was aimed at studying anticipatory and carry-over coarticulation in fricative-stop sequences at the articulatory level, using electropalatography. Our results can be summarised as follows. Stops were found to have an influence on the tongue-palate contact pattern at the offset of the preceding fricative. More precisely, the EPG centre of gravity for the fricative was located further back in the context of [k] than in the context of [t]. Inversely, fricatives had an effect on the EPG pattern at the onset of closure for the following stop (for both [t] and [k]) and at the stop release (for [k]). In general, the EPG centre of gravity appeared to be further forward following [s] than following [ʃ].

However, the average EPG patterns offered a more complex picture of fricative-stop coarticulation. Specifically, we found no complete velar closure for velar stops in [ska] sequences. This may suggest that [k] was actually produced as a fricative or, alternatively, that the place of articulation of the stop was shifted backward in the context of [s], when syllable-initial [k]s are taken as reference. This latter hypothesis, if it was to be confirmed, would fit in nicely with a series of recent findings on the dynamics of articulatory movements in the production of velar stops (e.g. Moshammer, Hoole & Kühnert, 1995; Hardcastle, Vaxelaire, Gibbon, Hoole & Nguyen, 1996; Kühnert & Hoole, in preparation). Single velar stops are known to be produced with the tongue body following a forward elliptical trajectory during the occlusion (see e.g. Perkell, 1969). In some contexts however, this elliptical movement appears to be inhibited. For example, Hardcastle et al. (1996) found that the “looping” trajectory of the tongue body for /k/ was abruptly halted during the production of /kl/ clusters. The authors suggested that the tip/blade raising for the following /l/ is incompatible with the characteristic velar trajectory. Kühnert and Hoole (in preparation) found that articulatory “loops” associated with velar stops can also be affected by jaw height, i.e. be reduced in magnitude or even totally stopped in the context of another consonant characterised by a high position of the jaw. This is precisely the case for strident fricatives such as /s/ and /ʃ/ whose spectral characteristics partly depend on the airflow impinging against the lower incisors (see e.g. Stevens, 1989). Thus, apparent backward shifts in the place of articulation for [k] following [s] may be simply due to the blocking of the velar articulatory loop in that environment. We are planning to conduct a more extensive study on this topic, using EMA (electromagnetic articulography, Perkell et al., 1993) to monitor the movements of the tongue body in the midsagittal plane.

Acknowledgements

Work supported by a grant (#8210-043017) from the Swiss National Foundation for Science. The natural utterances used in Experiment 1 were recorded by N. Lovric. We are grateful to Sarah Hawkins, Ruth Kearns and Barbara Kühnert for very useful discussions. Thanks are also due to Jonathan Rodgers for his participation as subject in the EPG study, and to Geoffrey Potter for his help with the EPG3 system.

References

Bailey, P.J., and Summerfield, Q. (1980). “Information in speech: observations on the perception of [s]-stop clusters,” *Journal of Experimental Psychology: Human Perception & Performance* **6**, 536-563.

Byrd, D. (1996). “Influences on articulatory timing in consonant sequences,” *J. Phonetics* **24**, 209-244.

Borden, G.J., and Gay, T. (1979). “Temporal Aspects of articulatory movements of /s/-stop clusters,” *Phonetica* **36**, 21-31.

Diehl, R.L., and Kluender, K.R. (1989). “On the objects of speech perception,” *Ecological Psychology* **1**, 121-144.

Elman, J.L., and McClelland, J.L. (1988). “Cognitive penetration of the mechanisms of perception: compensation for coarticulation of lexically restored phonemes,” *Journal of Memory and Language* **27**, 143-165.

Hardcastle, W.J., Jones, W., Knight, C., Trudgeon, A., & Calder, G. (1989). “New developments in electropalatography: a state-of-the-art report,” *Clinical Linguistics & Phonetics* **3**, 1-38.

Hardcastle, W.J., Vaxelaire, B., Gibbon, F., Hoole, P., & Nguyen, N. (in press). “Tongue kinematics in /kl/ clusters and singleton /k/: A combined EMA/EPG study,” *Proceedings of the Australian Speech Science and Technology Conference*, Adelaide, Dec. 96.

Hawkins, S. (1995). “Arguments for a nonsegmental view of speech perception,” *Proceedings of the XIIIth International Congress of Phonetic Sciences*, Stockholm, 13-19 August 1995, vol. 3, 18-25.

Hughes, G.W., & Halle, M. (1956). “Spectral properties of fricative consonants,” *J. Acoust. Soc. Am.* **28**, 303-310.

Klatt, D.H. (1980). “Software for a cascade/parallel formant synthesizer,” *J. Acoust. Soc. Am.* **67**, 971-995.

Kühnert, B., & Hoole, P. “On loops: additional data” (in preparation).

Liberman, A.M. (1996). *Speech: A Special Code* (MIT Press, Cambridge, Mass.).

Malécot, A., & Chermak, A. (1966). “Place cues for /ptk/ in lower cutoff frequency shifts of contiguous /s/,” *Language and Speech* **9**, 162-169.

Mann, V.A., and Repp, B.H. (1981). “Influence of preceding fricative on stop consonant perception,” *J. Acoust. Soc. Am.* **69**, 548-558.

McQueen, J.M. “Lexical or non-lexical effects in compensation for coarticulation?”, unpublished manuscript.

Mooshammer, C., Hoole, P., & Kühnert, B. (1995). “On loops,” *J. Phonetics* **23**, 3-21.

Nguyen, N. (1996). “A Matlab package for the analysis of EMA/EPG data,” *Queen Margaret College Working Papers in Speech and Language Sciences* **1**, 4-12.

Nguyen, N. “Phonetic context effects and lexical effects in speech perception” (in preparation).

Norris, D. (1993). “Bottom-up connectionist models of 'interaction',” *in Cognitive Models of Speech Processing: the Second Sperlonga Meeting*, edited by G. Altmann and R. Shillcock (Lawrence Erlbaum, Hillsdale, N.J.), pp. 211-234.

Perkell, J.S. (1969). *Physiology of Speech Production. Results and Implication of a Quantitative Cineradiographic Study* (MIT Press, Cambridge, Mass.).

Perkell, J.S., & Matthies, M.L. (1992). “Temporal measures of anticipatory labial coarticulation for the vowel /u/: within- and cross-subject variability,” *J. Acoust. Soc. Am.* **91**, 2911-2925.

Perkell, J.S., Cohen, M.H., Svirsky, M.A., Matthies, M.L., Garabieta, I., & Jackson, M.T.T. (1992). “Electromagnetic midsagittal articulometer (EMMA) systems for transducing speech articulatory movements,” *J. Acoust. Soc. Am.* **92**, 3087-3096.

Repp, B. (1982). “Phonetic trading relations and context effects: new experimental evidence for a speech mode of perception,” *Psychological Bulletin* **92**, 81-110.

Repp, B.H., and Mann, V.A. (1981). “Perceptual assessment of fricative-stop coarticulation,” *J. Acoust. Soc. Am.* **69**, 1154-1163.

Repp, B.H., and Mann, V.A. (1982). “Fricative-stop coarticulation: Acoustic and perceptual evidence,” *J. Acoust. Soc. Am.* **71**, 1562-1567.

Schwartz, M.F. (1987). “Transitions in American English /s/ as cues to the identity of adjacent stop consonants,” *J. Acoust. Soc. Am.* **42**, 897-898.

Shillcock, R., Levy, J., Lindsey, G., Cairns, P., & Chater, N. (1993). “Connectionist modelling of phonological space,” *Computational Phonology, Edinburgh Working Papers in Cognitive Science*, edited by T.M. Ellison and J.M. Scobbie, vol. 8, pp. 179-195

Stevens, K.N. (1972). “The quantal nature of speech: evidence from articulatory-acoustic data,” *in Human Communication: A Unified View*, edited by E.E. David and P.B. Denes (McGraw-Hill, New York), pp. 51-66.

Stevens, K.N. (1989). “On the quantal nature of speech,” *J. Phonetics* **17**, 3-45.

Stevens, K.N., and Blumstein, S.E. (1978). “Invariant Cues for Place of Articulation in Stop Consonants,” *J. Acoust. Soc. Am.* **64**, 1358-1368.

Uldall, E. (1964). “Transitions in fricative noise,” *Language and Speech* **7**, 13-14.