informa
healthcare

The quantification of place of articulation assimilation in electropalatographic data using the similarity index (SI)

KARITA M. GUZIK[1] & JONATHAN HARRINGTON[2]

[1]*Institute of Phonetics and Digital Speech Processing, University of Kiel, Germany, and* [2]*Institute for Phonetics and Speech Processing (IPS), University of Munich, Germany*

Abstract
The present study is concerned with an electropalatographic (EPG) and acoustic analysis of coarticulatory influences in fricative consonant clusters that span a word boundary. We aimed to test whether, as has been found for other languages and consonant-classes, final consonants are more prone to coarticulatory influences than initial ones, and also whether there is any evidence for a relationship between a consonant's susceptibility to coarticulation and the extent to which it exerts a coarticulatory influence on flanking consonants. We developed an algorithm, the similarity index (SI), which quantifies coarticulation in EPG data by measuring the extent of deviation of consonants in heterorganic clusters from their homorganic counterparts. This algorithm was applied to EPG data recorded from three native speakers of Polish producing word-pairs in a carrier sentence such that word-final and word-initial fricatives occurred across a word-boundary for all 16 possible combinations of the four fricatives [s, ʃ, ɕ, x]. The same data were analysed acoustically using Bark-scaled cepstral coefficients. The results showed that word-final consonants were more susceptible to coarticulatory influences than were initial ones; and there was some evidence that the alveolopalatal fricative [ɕ] was most resistant to coarticulation, and exerted the greatest coarticulatory influences.

Keywords: *Coarticulation, electropalatography, Polish, fricatives, clusters.*

Introduction

This paper is concerned with an analysis of the relative coarticulatory influence of consonants in fricative clusters in Polish. The starting point for the analysis is the extensive phonetic and phonological evidence which shows that domain final consonants are inherently weaker in many languages than domain initial consonants (Hock, 1991, 1992; Martinet, 1955; Vennemann, 1993). There are, for example, several cases of diachronic final consonant loss or lenition, as in the development of Italian "applicare" from Latin "adplicare". Synchronically, in German, English and many other languages, final consonants are far more likely to assimilate to the place of articulation of a following consonant than the other way round (Guy, 1980; Kohler, 1976). Compatibly Byrd (1996) showed in an electropalatographic study of $C_1\#C_2$ consonants that spanned a word boundary that final consonants generally showed more variability and spatial reduction and that they were often substantially overlapped by a following initial consonant.

The precise reasons for this difference in the direction of assimilation remain unclear. Ohala (1990) and Ohala and Kawasaki (1984) propose an auditory explanation: since syllable-initial consonants tend to be auditorily more salient than syllable-final consonants, they can mask perceptually the final-consonant place of articulation. There is some experimental evidence in support of this view. After splicing syllable-final V-Stop sequences onto a syllable-initial Stop-V sequence to create a VC_1C_2V sequence with different places of articulations, listeners were much more likely to hear a single consonant whose place of articulation was predominantly that of C_2 if the closure duration was shortened (Fujimura, Macchi, & Streeter, 1978; Streeter & Nigro, 1979).

As various experiments by Recasens have shown, the extent to which a final consonant is lenited or overlapped also depends on the phonetic attributes

of the consonant itself (Recasens, 2002, 2004; Recasens & Pallarès, 1999, 2001; Recasens, Pallarès & Fontdevila, 1997). Recasens developed the notion of the degree of articulatory constraint (DAC) to express the idea that inherently physiological reasons account for different types of assimilation patterns. According to Recasens (2002, 2004), consonants like the trill [r] that are produced with a high degree of articulatory precision also exert a strong coarticulatory effect on neighbouring segments and are themselves resistant to coarticulatory influences. While the idea of the DAC is itself interesting, this relationship between resistance and dominance has been demonstrated predominantly for Catalan; moreover, Recasens (2004) found that fricatives patterned differently from other consonants in not showing an asymmetrical dominance between initial and final position.

In Polish, initial consonants are diachronically more stable and exert more influence on preceding final consonants than the other way round (e.g., voicing assimilation in the sound change from early to modern Polish, /babka/ > /bapka/, "grandmother"), (see Długosz-Kurczabowa & Dubisz, 2001; Mazur 1993). Synchronically Ostaszewska & Tambor (2000) observed assimilations which were more likely to be anticipatory than perservatory, as in *zszyć*: /sʃɪtɕ/ > /ʃːɪtɕ/, "to stitch down". There is also evidence for anticipatory assimilation at word boundaries, e.g., *nad ciałem*: /nad#ɕauɛm/ > /natɕ#ɕauɛm/, "over the skin" (Morciniec, 1990). But none of these studies have been conducted with articulatory methods.

The motivation for the present study was threefold. Firstly, we wanted to explore whether, as for many other languages, Polish final consonants were weaker than initial ones; secondly, we wanted to carry out this analysis for sequences of fricatives because the behaviour of fricative-clusters has not been studied in this way before. Moreover, in contrast to the studies of Catalan by Recasens, Polish has several fricative places of articulation. Finally, we wished to explore further whether there was any evidence for a relationship between coarticulatory resistance and dominance.

Polish has 29 consonant phonemes (Biedrzycki, 1974) including four voiceless lingual fricatives: an alveolar [s], produced with a contact between the tongue tip and roof of the mouth; a post-alveolar [ʃ] which is produced with a constriction of the tip or blade of the tongue against the hard palate (although based on acoustic data only Hamann, 2004, suggests it might be retroflex); an alveolopalatal [ɕ] in which the constriction is formed between the blade of the tongue and the hard palate; and a velar fricative [x] which has its primary constriction in the velar region. Fricatives in Polish can occur both in syllable-onset and syllable-final position; consonant clusters are especially prevalent at word boundaries (Rochoń, 2000).

Method

Speakers, materials, pre-processing

Electropalatographic and acoustic data were simultaneously recorded from three native Polish speakers (two male, one female) from the northern area of Poland. The speakers' age ranged between 29 and 35 years and all of them had previous experience of using EPG palates. The recording was carried out with the Reading EPG 3+ System, consisting of an artificial palate with 62 electrodes. All EPG recordings were carried out at the Zentrum für allgemeine Sprachwissenschaft, Berlin. The EPG data were digitized with a frame rate of 10 ms and the synchronized acoustic signal was digitized at 10 kHz. The speakers, who had no knowledge of the aims of the experiment, were seated in a sound-treated room and provided with a list of the sentences. They were instructed to use a normal speech rate and a normal level of speaking loudness while reading the sentences. If a word was produced with errors, such as hesitations or false starts, the speaker was asked to repeat the whole sentence.

The speech material consisted of two word sequences such that a sequence of two fricatives spanned the word boundary. These sequences contained every possible combination of a two token sequence of the fricatives [s, ʃ, ɕ, x] including the homorganic ones. All word sequences were embedded in the carrier sentence "Nie powiedziała *word sequence 1* tylko *word sequence 2*", ("She didn't say…but…"). The sentences were randomized and repeated 10 times, so every speaker produced 160 tokens. The words with the relevant fricatives are shown in Table I.

The speech samples were labelled acoustically from waveform and spectrographic displays. The labelling consisted of marking the acoustic boundaries of the relevant fricatives as well as the boundaries of the entire word. All acoustic and physiological analyses were carried out using EMU (Cassidy & Harrington, 2001) and the R programming language. Examples of spectrograms and palatograms for the four fricatives in the homorganic context are shown in Figure 1.

A spectral analysis of the fricatives was carried out in order to establish whether the same pattern of results emerged from both EPG and from acoustic data. dB-spectra were calculated for all homorganic

Table I. The fricative types and the words from which they were extracted.

	/s/	/ɕ/	/ʃ/	/x/
/s/	nos sarni	głos siostry	włos szary	nos chomika
/ɕ/	weś sanki	weś siostre	weś szafe	weś hak
/ʃ/	pytasz sąsiada	pytasz siostre	kosz szary	kosz chaty
/x/	dach sali	dach siostry	dach szopki	dach chaty

Note: Within the orthography the relevant fricatives are underlined.
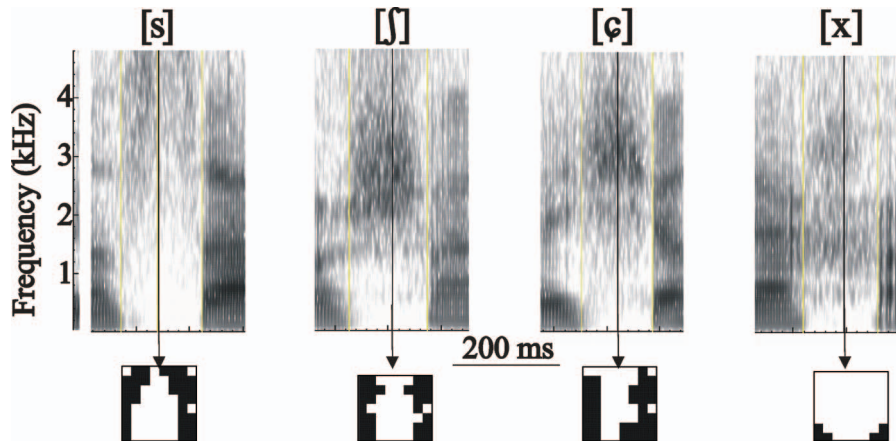
Figure 1. Spectrograms (0–5 kHz) range of four homorganic Polish fricatives produced by the same speaker showing palatograms near the temporal midpoint of the homorganic sequence in each case.

and heterorganic fricatives. From these, the dB-energy values were averaged in the first 19 critical (Bark) bands up to 5000 Hz using the formula given in Traunmüller (1990). These were then transformed into 19 Bark-scaled cepstral coefficients using the discrete cosine transformation following the methodology in Watson and Harrington (1999). For the homorganic fricatives, the Bark-scaled cepstral coefficients were calculated from dB-spectra obtained with a 25.6 ms Hamming window centred at their temporal midpoint. For the heterorganic fricatives, two sets of Bark-scaled cepstral coefficients were obtained per heterorganic cluster with 25.6 ms Hamming windows centred at the two SI-targets as defined in below.

Quantification of coarticulation

The magnitude and extent of coarticulation was quantified using an EPG-measure, the similarity index targets (SI-targets) and an acoustic measure, the Euclidean distance ratio (E_{RATIO}). These two algorithms are described in below.

Similarity index (SI). The extent of coarticulation was quantified electropalatographically as a parameter, the similarity index (SI), which is defined in more detail below, by measuring the deviation of a target consonant in a heterorganic sequence from the corresponding consonant in a homorganic sequence. This procedure is analogous to the way in which Moon and Lindblom (1994) quantified the extent of vowel undershoot by measuring the deviation of a vowel in context from the same vowel produced in isolation.

In general, the SI is based on the idea that the more the context consonant encroaches on the target consonant, the greater the deviation of the target consonant from the corresponding homorganic sequence. For a sequence of heterorganic consonants [AB], the similarity index provides a measure for every EPG frame between the onset and the offset of [AB] of the extent to which the sequence is similar to homorganic [A:] or similar to homorganic [B:].

The SI varies between 1 and 0: when the SI has a value of 1, then a palatogram is maximally similar to [A:]; and when the SI has a value of 0, a palatogram is maximally similar to [B:].

This is illustrated in further detail in Figure 2 for the heterorganic fricative sequence [ɕs]: the closer the SI trajectory is to 1, the more a palatogram in [ɕs] approximates homorganic [ɕ:] (i.e. a sequence of [ɕ#ɕ] spanning a word boundary) and the closer the trajectory is to 0, the closer a palatogram is to [s:] (i.e., [s#s]). Consequently, the SI-index falls throughout this heterorganic sequence.

The actual calculation of the SI index is accomplished with a training and a testing stage. The training stage consists of two parts. Firstly, the palatograms at the temporal midpoint of all of the tokens of homorganic [s:] and of all the tokens of homorganic [ɕ:] are extracted and these are separately averaged. The averaging is binary: if a particular cell is contacted in more than 50% of the tokens, then the cell is given a value of 1 (black), otherwise it is 0 (white). Therefore, two averaged palatograms are obtained, one for homorganic [s:] and one for homorganic [ɕ:] and these are shown in the left and middle panels of Figure 3 respectively. Secondly, a difference palatogram (DP) is obtained by subtracting from each other the values in each respective cell of the two averaged palatograms in the two homorganic sequences.

When this subtraction is done, there can only be one of three outcomes for each cell in the DP:

- If the cells in averaged [s:] and [ɕ:] are both "on" or both "off", then the result of the subtraction is 0. Thus, since the cells are "on" in row 8 and column 1 of both [s:] and [ɕ:] (Figure 3), the DP in the same cell registers a value of 0; and the DP also registers a value of 0 in row 8 column 4, where the corresponding cells for [s:] and [ɕ:] are both "off".
- If a cell is "on" in [s:] but "off" in [ɕ:], the result of the subtraction is 1 (shown by the hatched cells in the DP in Figure 3).
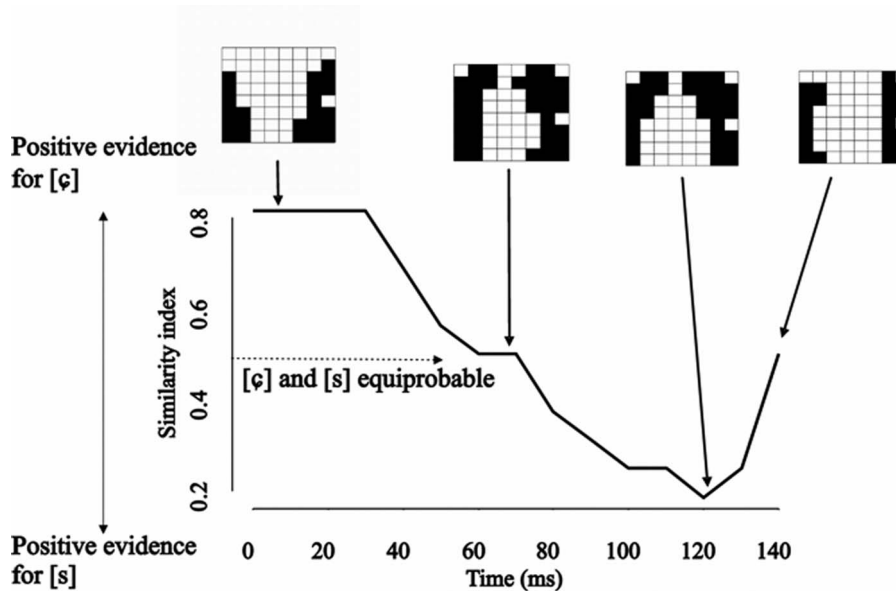
Figure 2. The SI-metric applied to the palatograms from the start to the end of a [ɕs] token. The MAX-SI is the point at which there is maximum evidence for [ɕ].

average palatogram for [s#s]   average palatogram for [ɕ#ɕ]   difference palatogram for [s]-[ɕ]
▨ a contact for [s] and not for [ɕ]
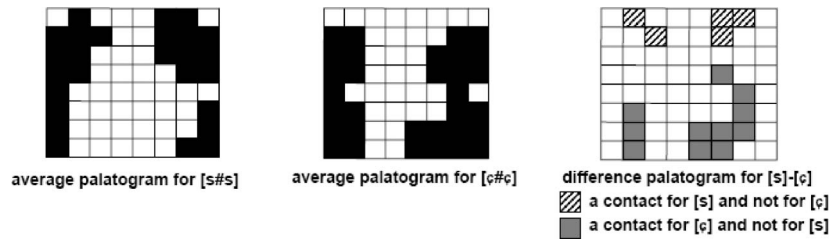▦ a contact for [ɕ] and not for [s]

Figure 3. Training stage: Averaged palatograms for homorganic [s:] and homorganic [ɕ:] and the corresponding difference palatogram for [s] – [ɕ] (right). An averaged palatogram shows the cells that are contacted more than 50% of the time at the temporal midpoint.

- If a cell is "on" in [ɕ:] but "off" in [s:], the result of the subtraction is −1 (shown by the grey cells in the DP in Figure 3).

The DP registers, therefore, the differences in the contact patterns for homorganic [s:] and homorganic [ɕ:] showing both where there is a contact for [s:] but not for [ɕ:] (the hatched cells) and where there is a contact for [ɕ:] but not for [s:] (the grey cells).

In the testing stage, each palatogram in the heterorganic sequence is assigned a score which is given by the number of hatched cells in the DP that are "off" in the palatogram to be tested, plus the number of grey cells that in the DP that are "on". All other cells are disregarded. The SI is a normalized equivalent of this raw score and lies between 0 and 1: it is obtained by dividing the raw score by the total number of hatched and grey cells in the DP. When the hatched cells in the DP are found to be "in contact" in a test palatogram, no count is made. Similarly, when the grey cells in the DP are "not in contact" in the test palatogram, no count is made. As a result, the test palatogram will be assigned a raw score of 0. After normalization, the SI of the test palatogram will be 0 (maximally like [s:]). Conversely, a count is made when all the hatched cells in the DP are "not in contact" in the test palatogram. When all the grey cells in the DP are "in contact", a count is also made. As a result, the test palatogram will be assigned a raw score of 16. After normalization, the SI will come to 1 (maximally like [ɕ:]). SI values between these extremes represent intermediate articulatory positions between [s:] and [ɕ:].

An example of the SI calculation is shown for the palatogram taken from near the beginning of a [ɕ#s] sequence in Figure 4. In this case, all five hatched cells are "off" and eight of the grey cells are "on". Thus this palatogram's raw score is $5 + 8 = 13$ from which the SI is given by $13/16 = .8125$, 16 being the sum of the hatched and grey cells. This score is much closer to 1 than to 0, commensurate with this palatogram's close similarity to homorganic [ɕ:] (middle panel of Figure 3).

SI-targets. The maximum and minimum in the SI-trajectory for a consonant sequence provide the information about coarticulatory influences.

Recall that for a consonant sequence [AB], the SI-index typically falls: a high SI-value corresponds to a time point that tends to occur early in [AB] when the vocal tract is predominantly given over to [A]; and a low SI-value corresponds to a time point that mostly

occurs late in [AB] when the vocal tract is predominantly influenced by [B]. The value of the SI-maximum can therefore give some information about the extent to which [B] influences [A]; and the value of the SI-minimum information about the extent to which [A] influences [B]. Consider as an example Figure 5 which shows two trajectories, one for each of two tokens of [sç]. It is clear that the SI-trajectory starts at a much higher value (.62) for the trajectory on the left compared with the one on the right (.28). The interpretation of these data is, then, that the extent of influence of word-initial [ç] on the preceding word-final [s] is a good deal greater in Figure 5b than in Figure 5a and this is borne out by the palatograms extracted at the beginning of the these tokens: in the leftmost palatogram in Figure 5a there is much clearer evidence for an alveolar pattern (as shown by the contact in the front row and fewer contacts in the back two rows) than for the leftmost palatogram of Figure 5b; indeed, in Figure 5b, the alveolar has almost been completely assimilated to a following [ç] and this is reflected in a low SI-index throughout the [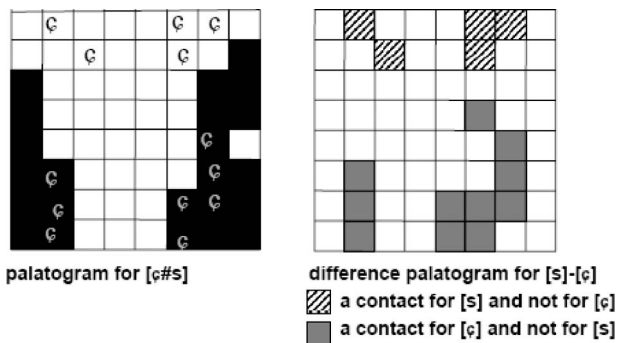sç] sequence (i.e., the vocal tract is almost entirely given over to [ç] for the duration of the entire cluster). As far as the coarticulation influence *exerted* by word-final [s] on word-initial [ç] is concerned, both trajectories reach the same minimum (.045) leading to the interpretation that the influence of word-final [s] on word-initial [ç] is about the same for both, and also minimal (given that the SI-minimum is so close to 0). This interpretation is consistent with the rightmost palatograms extracted at these SI-minima: these show negligible differences between Figures 5a and 5b and no real evidence of any palatographic patterning due to [s].

Both maxima and minima together are referred to as SI-targets after subtracting the SI-minimum (.045 in Figure 5) from 1 (resulting in .955). The purpose of this subtraction is to allow us to make the following generalization across both final and initial consonants: SI-targets close to 1 are associated with minimal coarticulatory perturbation and SI-targets close to 0 with maximum coarticulatory perturbation. Since for the present example in Figure 5, the SI-targets for word-initial [ç] $(1 - .045 = .955)$ are much higher than for word-final [s] (.62 and .28), we would conclude that the word-initial [ç] in [sç] is less prone to coarticulatory perturbation than is the word-final consonant [s].

Euclidean distance ratio. Our aim in the acoustic analysis was to determine the extent to which the cepstra at the SI-targets deviated from those in the corresponding homorganic fricatives. This was done by measuring in the Bark-scaled cepstral space the log. Euclidean distance ratio from the heterorganic fricatives to their corresponding homorganic fricatives. The Euclidean distance measurements were made in a three-dimensional space formed by the first three cepstral coefficients (excluding the 0th coefficient, which is proportional to the Bark-scaled spectral mean). This space was chosen because it provided a very effective separation between the four fricative types (see, for example, Figure 6 for the distribution of the fricatives on two of these coefficients).

Figure 4. Testing stage: a palatogram taken from a production of [çs] (left) and the difference palatogram derived from the differences between homorganic [s:] and homorganic [ç:] (right). The superimposed "ç" symbols on the palatogram on the left mark any grey cell that was contacted (black symbol) and any hatched cell that was not contacted (white symbol) in the difference palatogram.
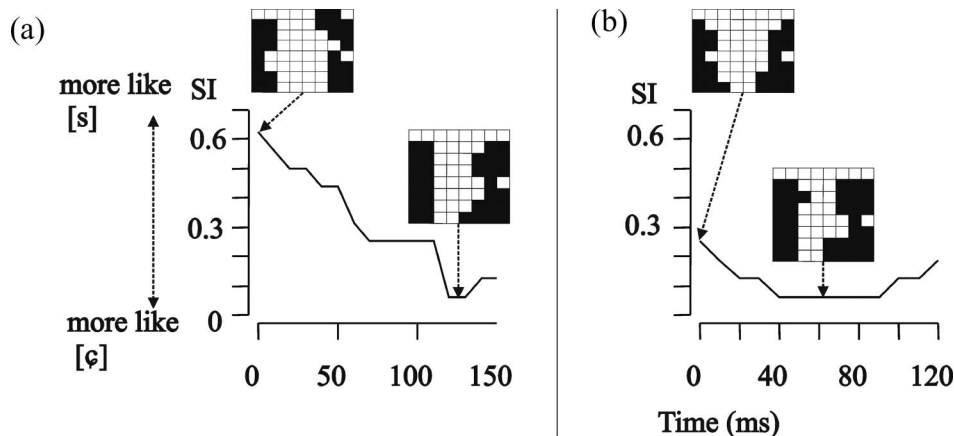

Figure 5. Similarity index trajectories calculated over two separate tokens of [sç]. Higher SI-values correspond to a closer palatographic approximation to [s:], lower SI values to [ç:]. The palatograms were extracted at the time points marked by arrows, i.e., at the SI-targets (SI-maximum and SI-minimum) in each case.
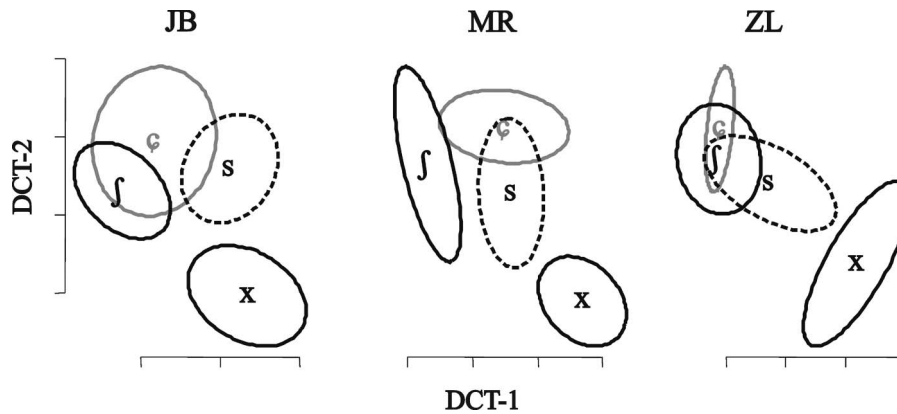
Figure 6. 95% confidence ellipses for the four homorganic fricatives on the first two Bark-scaled cepstral coefficients at the temporal midpoint separately for the three speakers.

The log. Euclidean distance ratio is intended to provide a measure of the relative proximity of a fricative [A] in a heterorganic cluster [AB] to the corresponding two homorganic fricatives [A:] and [B:]. This is illustrated in Figure 7 for a two-dimensional Bark-scaled cepstral space for [s] in a heterorganic cluster [sʃ] (we made corresponding calculations in three dimensions but the principle is the same). Two Euclidean distances are measured: E_1 which is the distance of [s] in [sʃ] to the centroid (mean) of [s:]; and E_2 which is the distance of the same [s] in [sʃ] to the centroid of [ʃ:]. The log. Euclidean distance ratio, E_{RATIO}, is then defined as:

$$E_{RATIO} = \log(E_1/E_2) = \log(E_1) - \log(E_2) \quad (1)$$

If E_{RATIO} is 0 (i.e., $E_1/E_2 = 1$), then [s] in [sʃ] is equidistant between the centroids of [s:] and [ʃ:] . If E_{RATIO} is negative (i.e., E_1 is small relative to E_2), then [s] in [sʃ] is acoustically closer to the centroid of [s:]; and if E_{RATIO} is positive (i.e., E_2 is small relative to E_1), then [s] in [sʃ] is acoustically closer to [ʃ:] than to [s:]. Therefore, the relative influence of [ʃ] on [s] in [sʃ] can be inferred from E_{RATIO}: if there is no influence, [s] in [sʃ] should be acoustically close to [s:] and E_{RATIO} has a large negative value; and if the influence is maximal (i.e., [s] assimilates to [ʃ]), then [s] in [sʃ] will be acoustically much closer to [ʃ:] and E_{RATIO} will have a large positive value (see Harrington, in press, for the application of the same metric to vowels). E_{RATIO} was calculated in this way for each member of the heterorganic cluster to their respective homorganic fricatives.

Predictions

The authors' predictions are as follows. If final consonants are weaker than initial consonants, then their SI-targets should be lower and E_{RATIO} is expected to be positive and large. Secondly, if there is a relationship between the extent to which a fricative is *susceptible* to coarticulatory influences and the extent to which it *exerts* an influence on other
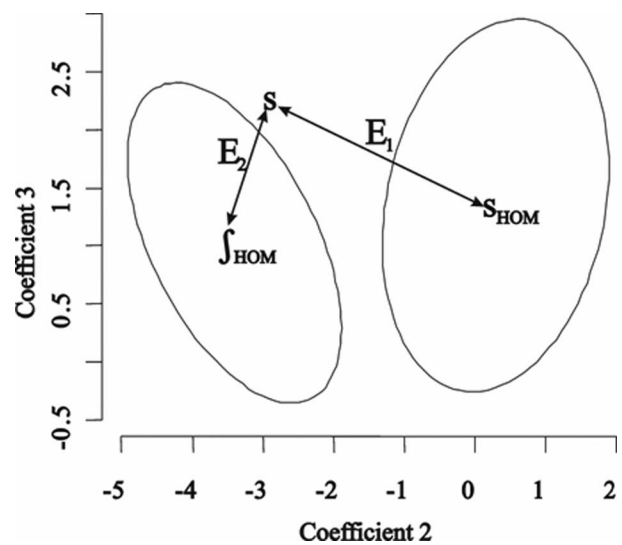

Figure 7. "s" is the position at the SI-target of a final [s] token produced by JB in the context of heterorganic [sʃ] in the space of the Bark-scaled 1st and 2nd cepstral coefficients. "sHOM" and "ʃHOM" are the centroids (mean positions) in this space of homorganic [s:] and homorganic [ʃ:] produced by the same speaker showing 95% confidence interval ellipses. E_1 and E_2 are the Euclidean distances from this final [s] token to "sHOM" and "ʃHOM" respectively. The log. Euclidean distance ratio for "s" is given by $\log(E_1) - \log(E_2) = 0.86$ for this token.

fricatives, then, any fricative's SI-target and E_{RATIO} values are not expected to change very much in the context of another fricative which is characterized by a low SI-target and a large, positive E_{RATIO} (conversely, a fricative's SI-target and E_{RATIO} values should be substantially affected next to another fricative with a high SI-target and negative E_{RATIO}).

Results

The results can be divided into two parts: In the first part we aim to establish whether final fricatives were more prone to coarticulatory influences than initial ones, using both the SI measure and the separate acoustic measure based on the Bark-scaled cepstral coefficients discussed above. Within this section, we

also seek to establish whether there were any differences between the four fricative types in the extent to which they were susceptible to coarticulatory perturbations. In the second part we seek to determine how far the fricative types differed in the extent to which they exerted coarticulatory influences on their neighbouring fricative in initial or final position. In the final part of this paper, we consider whether there was a relationship between a fricative's susceptibility to coarticulatory perturbation and the extent to which it exerted a coarticulatory influence on other fricatives.

Position and coarticulatory perturbation

EPG analysis. The magnitude of the SI-targets in final and initial position were compared since, as argued above, lower SI-targets are indicative of greater coarticulatory influences due to the context consonant. A summary of the mean SI-target by Consonant and Position in Table II supports the view that final consonants are more prone to coarticulatory influences than are initial consonants. With the exception of [ʃ] for speaker MR, for all speakers and for all four fricative types, the SI-target was lower in final than in initial position.

A two-factor ANOVA was carried out separately for each speaker with independent variables Consonant ([s, ʃ, ç, x]) and Position (initial, final) and dependent-variable the arc-sine transformed SI-target. Speaker JB showed a significant effect for Position ($F = 71.2$, $df = 1$, 230, $p < .001$) and for Consonant ($F = 8.4$, $df = 3$, 230, $p < .001$) and a significant interaction between Position and Consonant ($F = 9.3$, $df = 3$, 230, $p < .001$). The interaction is expected in view of the results in Table II showing that [x] for JB showed virtually no position-dependent difference on the SI-target. The significant effect for Consonant comes about because the SI-target means were higher for the dorsals [ç, x] than for the lingual consonants [s, ʃ].

The pattern of results was very similar for ZL. He showed a significant effect for Position ($F = 19.6$, $df = 1$, 218, $p < .001$) and a significant effect for Consonant ($F = 5.2$, $df = 3$, 218, $p < .01$) and also an interaction between Position and Consonant ($F = 3.4$, $df = 3$, 218, $p < .05$). As for JB, the interaction comes about because there were no position-dependent differences for [x] (see also Table II). Averaged across position, the dorsals had higher SI-targets which once again suggests that the dorsals [ç, x] were in general somewhat less prone to coarticulatory influences than were the lingual consonants [s, ʃ].

In contrast to the other two speakers, speaker MR showed no significant effect for Position, although as already indicated, Table II shows higher average SI-targets for [s, ç, x] in initial compared with final position. MR showed a significant effect for Consonant ($F = 25.5$, $df = 3$, 236, $p < .001$) and a significant interaction between Position and Consonant ($F = 7.2$, $df = 3$, 236, $p < .001$). This significant interaction comes about because, as shown by post-hoc t-tests carried out separately on the fricative categories, none of [s, ç, x] showed any position-dependent significant differences. Contrary to all other results presented in this section, MR's [ʃ] had a higher SI-target in final than in initial position ($t = 3.69$, $df = 70.0$, $p < .001$).

Acoustic analysis. Table III shows the average values of E_{RATIO} taken at the SI-target by Position (initial, final) and Consonant. As discussed above, when E_{RATIO} is large and negative, then the interpretation is that a fricative in a heterorganic cluster is scarcely susceptible acoustically to coarticulatory perturbations due to the flanking consonant; and when E_{RATIO} is positive and large, then the coarticulatory perturbation due to the flanking consonant was interpreted to be very high. For example, Table III shows for JB that final [s] in heterorganic [s#F] ($F = [ʃ, ç, x]$) had a value of .089 (this value, being close to 0, suggests that final [s] was acoustically about midway between homorganic [s:] and [F#F]) whereas the corresponding value for initial [s] in heterorganic [F#s] was -1.213: taken together these two results show that, compatibly with JB's EPG data below, final [s] was perturbed acoustically more by the flanking consonant than was initial [s].

The main way in which there is consistency between the acoustic and EPG data is in the effect

Table II. Mean SI-target values by consonant and position for the three speakers.

	s	ʃ	ç	x
JB				
final	.470	.559	.668	.751
initial	.791	.790	.858	.753
ZL				
final	.722	.799	.829	.928
initial	.919	.893	.906	.937
MR				
final	.625	.958	.774	.777
initial	.761	.834	.784	.823

Table III. Average E_{RATIO} values separately for the three speakers at the SI-target for heterorganic [s, ʃ, ç, x] in final (row 1) and in initial position (row 2).

	s	ʃ	ç	x
JB				
final	.089	$-.205$	$-.164$	$-.150$
initial	-1.213	$-.860$	$-.176$	$-.974$
ZL				
final	$-.888$	$-.461$	$-.455$	$-.606$
initial	-1.158	-1.292	$-.700$	-1.190
MR				
final	-1.012	$-.507$	$-.799$	-1.349
initial	$-.769$	$-.837$	$-.622$	-1.399

of Position. As Table III shows, the average E_{RATIO} values were all smaller (i.e., larger negative values) for initial than final position except for MR's [s] and [ɕ] fricatives. Consistently with the EPG data, these results can be interpreted to show that final fricatives were more prone to coarticulatory perturbations than were initial fricatives. The results of a two-factor ANOVA with independent variables Consonant ([s, ʃ, ɕ, x]) and Position (initial, final) and dependent-variable E_{RATIO} is consistent with this view. There were significant effects for Position for JB (F = 52.6, df = 1, 230, p < .001), for ZL (F = 22.23, df = 1, 218, p < .001), and for MR (F = 14.4, df = 1, 278, p < .001), the latter having shown no position-dependent effect from the EPG data.

On the other hand, there is somewhat less consistency between the EPG and acoustic data as far as differences between [s, ʃ, ɕ, x] are concerned. Two issues stand out:

- [x] showed no effect for Position for JB and ZL from the EPG data, but it clearly did in the acoustic data as post-hoc t-tests showed (for JB: t = 6.7, df = 35.0, p < .001; for ZL: t = 5.0, df = 49.2, p < .001). This discrepancy could well come about because the EPG palate does not extend far back enough to include the full range of lingual-palatal contacts for this fricative.
- The EPG data showed that the dorsals [ɕ, x] were less prone to coarticulatory influences for JB and ZL than were [s, ʃ]. Leaving aside [x], which may be problematic for the reason stated above, if [ɕ] had been acoustically less prone to coarticulatory perturbation, then the E_{RATIO} values should all be lower for [ɕ] than for [s] or [ʃ] and this, as Table III shows, was not the case. On the other hand, Table III does show that [ɕ] had the smallest change on the E_{RATIO} parameter between final and initial position of all the fricatives. Indeed, whereas post-hoc t-tests showed significant position-dependent change in [s, ʃ, x] for JB and in [ʃ, x] for ZL, there were no significant position-dependent changes for [ɕ]. So it can be argued that [ɕ] is less likely to show a difference in its susceptibility to coarticulatory influences in final compared with initial position.

Finally, we note two correspondences between the EPG results below and these acoustic data as far as speaker-differences are concerned. Firstly, the SI-target values for ZL were generally higher than for JB which suggests that the coarticulatory influences on fricatives were less for this speaker. This interpretation is consistent with the acoustic data which show that ZL has, in general, lower E_{RATIO} values than JB (so the extent to which a fricative was deflected away from its target acoustically was less for ZL than for JB). Secondly, of the three speakers, MR had the least consistent position-dependent results for the EPG data and this was also the case as far as the acoustic data are concerned. However, it must be emphasized that the correspondences between the EPG and acoustic data for MR were weak at best: whereas for the EPG data MR showed no significant differences due to position in [s, ɕ, x] and a significant effect in the "wrong" direction for [ʃ] (i.e., final [ʃ] was more stable than initial [ʃ]), in the acoustic data, there were no significant differences for [s, ʃ, ɕ] and a significant position-dependent effect for [x] (t = 5.5, df = 53.3, p < .001).

Position and coarticulatory dominance

We have so far considered how susceptible the separate fricative categories were to coarticulatory influences in final and initial position; but we have not yet said anything about the extent to which they exerted a coarticulatory influence on other fricatives. In general, if a context-fricative F_A exerts more coarticulatory influence than does another context-fricative F_B on a flanking target-fricative F_T, then the target-fricative F_T should show a lower SI-target in [F_T #F_A] than in [F_T #F_B]; and (similarly) an initial target-fricative F_T should show a lower SI-target in [F_A# F_T] than in [F_B# F_T], if F_A exerts more coarticulatory influence than F_B (as argued above, the lower the SI-target, the more a consonant is prone to coarticulatory perturbations due to the context-consonant).

Table IV shows the average SI-target values for F_T separately as a function of the four different context-fricatives [s, ʃ, ɕ, x]. (Thus row 1, column 1 is the average SI-target pooled across F_T = initial [#ʃ, #ɕ, #x] when the context-fricative was a preceding final [s#]; row 2 column 1 shows the average SI-target pooled across F_T = final [ʃ#, ɕ#, x#] when the context-fricative was a following initial [#s]). For both JB and ZL, it is clear that the SI-target values were lower for final target fricatives than for initial target fricatives (compare row 2 with row 1 in Table IV). This means that initial context-fricatives exerted a greater coarticulatory influence on final-target fricatives than final context-fricatives did on initial target-fricatives. This result is expected and compatible with the results above. MR showed a somewhat different pattern of results: for MR, SI-targets were lower for final target fricatives compared with initial target fricatives only in the context of [ɕ].

The results of a two-factor ANOVA with dependent variable the arc-sine transformed SI-target of the target-fricative and independent variables Position and Context-fricative showed the following results. For JB, there was a significant effect for Position (F = 82.8, df = 1, 230, p < .001) and for Context-fricative (F = 22.8, df = 3, 230, p < .001) and a significant interaction between Position and Context-fricative (F = 10.3, df = 3, 230, p < .001). The interaction comes about because context-[x], in

Table IV. SI-target values separately for the three speakers averaged across three of the four fricative categories.

Context-F	s	ʃ	ç	x
JB				
initial target-F	.851	.870	.659	.816
final target-F	.650	.524	.490	.791
ZL				
initial target-F	.940	.944	.836	.925
final target-F	.909	.762	.741	.866
MR				
initial target-F	.788	.731	.863	.851
final target-F	.880	.747	.693	.864

Notes: Initial target-F: these are average SI-targets pooled across three of initial [#s, #ʃ, #ç, #x] in the context of a preceding final [s#] (column 1), a preceding final [ʃ#] (column 2), a preceding final [ç#] (column 3) and a preceding final [x#] (column 4) for all heterorganic clusters only: thus the entry of .851 in row 1 column 1 for JB is the SI-target in the initial target fricative averaged across [s#ʃ], [s#ç], [s#x]. Final target-F: these are the average SI-targets pooled across three of the four final [s#, ʃ#, ç#, x#] in the context of a following initial [#s] (column 1), a following initial [#ʃ] (column 2), a following initial [#ç] (column 3), and a following initial [#x] (column 4) for heterorganic clusters only: thus the entry of .650 in row 2 column 1 for JB is the SI-target in the final target fricative averaged across [ʃ#s], [ç#s], [x#s].

contrast to the other three fricatives, showed hardly any position-dependent effect (see Table IV).

For speaker ZL, there was a significant effect for Position (F = 20.0, df = 1, 218, p < .001) and for Context-fricative (F = 7.7, df = 3, 128, p < .001) and no interaction between the two independent variables. For speaker MR, there was no significant effect for Position, a significant effect for Context-fricative (F = 8.9, df = 3, 278, p < .001), and a significant interaction between Position and Context-fricative (F = 9.4, df = 3, 278, p < .001). Post-hoc tests showed no significant effects for position for [ʃ] nor for [x]. Contrary to the pattern of results presented in this section, SI-targets were significantly higher for MR's final target-fricatives than for his initial target-fricatives in the context of [s] (t = 3.42, df = 78.0, p < .001); but compatibly with the data from JB and ZL, MR's SI-targets were significantly lower for final target-fricatives than for initial target-fricatives in the context of [ç] (t = −3.43, df = 47.9, p < .01).

Based on these statistical results and the pattern of averages shown in Table IV, we can make the following conclusions about how the four fricatives were differentiated as far as their coarticulatory influence on other fricatives is concerned:

- [ç] exerted the greatest coarticulatory influence of all four fricative types in both positions for JB and ZL; for MR this applied only when the target-fricative was in final position (i.e., when context-[ç] is initial).
- [x] exerted very little coarticulatory influence in either position.

- For speakers JB and ZL only, initial [ʃ] exerted a strong coarticulatory influence on final target fricatives.
- For speakers JB and ZL only, both final [s] and final [ʃ] exerted very little coarticulatory influence on initial target fricatives.

Discussion

The study has shown that the word-final fricatives in Polish examined in this paper were more unstable and more prone to coarticulatory influences than those in domain-initial position. This is consistent with findings for English (e.g., Byrd, 1996; Hardcastle & Roach, 1979), German (Kühnert, 1996) and Catalan (Recasens & Pallarès, 2001).

There is also evidence from the EPG data that the alveolopalatal fricative [ç] both resists coarticulation and exerts the greatest coarticulatory influence. This is compatible with the findings of Recasens, Fontdevila, Pallarès, & Solanas (1993) who have suggested that a high degree of coupling between the primary articulator and other tongue regions for palatals and alveolopalatals accounts for their reduced susceptibility to coarticulatory influences. On the other hand, these results from electropalatography were not consistent with our acoustic findings which showed no evidence that [ç] was acoustically less prone to coarticulatory influences than the lingual consonants [s, ʃ]. There are four possible reasons for this discrepancy. One is the well-known finding that articulatory-acoustic relationships are non-linear (Stevens, 1989), and so we need not expect differences between these consonants in their EPG patterns to result in corresponding acoustic differences. Another is that the data in the present study was band-limited acoustically to 5000 Hz due to the 10,000 Hz sampling rate of the EPG system; it may well be that a different pattern of results in the acoustic data would emerge if the frequency range beyond 5000 Hz could be included, possibly allowing a greater acoustic differentiation of [ç] from the other fricatives. A third reason is that [ç] tends to occupy a somewhat central position in the Bark-scaled cepstral space which means that it is acoustically quite close to all three fricative categories (in particular to [ʃ]). But this is not so electropalatographically: there is no analogous sense in which [ç] falls centrally in a parameterized EPG-space between the fricatives [s, ʃ, x]. Consequently, [ç] may well be physiologically distinct but acoustically quite close to some of the other fricatives and this may account for the different patterns of results in the EPG and acoustic data. Finally, the discrepancy between the acoustic and EPG data may arise because [ç] may be produced with lip-rounding which would have an acoustic effect that could not show up in our physiological measures; however, we do not know of any prior studies that have shown that [ç] is inherently lip-rounded in the same way that

lip-rounding can occur in the production of [ʃ] even before unrounded vowels in Southern British English (Gimson & Cruttenden, 1994, p. 173).

The alveolar fricative [s] was found to be especially unstable in final position and this seems to corroborate other findings that final alveolars are prone to reduction and assimilation. Based on an analysis of stops, Byrd (1992) suggests that alveolars might be unstable for acoustic reasons: since the acoustic cues to final alveolars are already weak, weakening these cues further is unlikely to entail a great deal of loss of information for the listener (see also Kohler, 1992). However, this argument cannot be readily carried over to fricatives since the acoustic cues to [s] are quite salient (indeed [s] is according to Stevens, 1972, 1989, a quantal articulation that is therefore acoustically stable). It seems, therefore, that the explanation for the final-[s] instability in the present study may have to be couched in articulatory terms. However, the nature of this articulatory explanation is unclear. Although alveolar instability could be explained in terms of its relatively simple articulation in the case of stops (Barry, 1992; Hardcastle & Roach, 1979), this argument does not apply to [s] which requires a very precise articulatory positioning. For this reason, Recasens & Pallarès (2001) assign [s] a high DAC value in their degree-of-articulatory constraints model, which means, compatibly with their findings, that [s] should not be especially susceptible to coarticulatory influences. On the other hand, we found that [s] showed the greatest variability of all fricatives especially in final position (see also Holst & Nolan, 1995, and Nolan, Holst, & Kühnert, 1996, for [s] assimilation in the context of [ʃ]). This discrepancy is likely to come about because of the different nature of the flanking consonants in the two studies: thus, Recasens & Pallarès (2001) investigated [s] in the context of different manners of articulation (oral stops and nasals), whereas in our study, the context consonants were always fricatives.

Finally, we believe that the SI-index is a useful way of quantifying coarticulation electropalatographically in heterorganic consonant clusters and we have shown that the results from the SI-index are at least partially corroborated by an acoustic analysis of the same data. From a clinical point of view, the SI-index might provide a useful addition to electropalatographic studies of phonological disorders (e.g., Gibbon, 1990) and of cleft palate speech which have often shown that alveolar and velar places of articulation are not sufficiently distinguished (e.g., for English: Gibbon & Crampin, 2001; Gibbon & Hardcastle, 1989; Hardcastle, Morgan Barry, & Nunn., 1989; Hardcastle, Gibbon, & Nicolaidis, 1991; Palethorpe, Croot, Molloy, Gibbon, & Harrington, 2003; for Cantonese: Whitehill, Stokes, Hardcastle, & Gibbon, 1995; Whitehill, Stokes, & Man, 1996; for Japanese: Yamashita Michi, Imai, Suzuki, & Yoshida, 1992). The targets from which the SI-index is calculated in the training stage could be the therapist's production of alveolar and velar consonants rather than homorganic stops across word boundaries produced by the same subject, as in this study. An SI-index that remained level at around a value of .5 in a child's attempted production of the medial cluster in e.g., ''biscuit'' would suggest little or no differentiation across the cluster, whereas a steeply falling SI-index as a function of time would be indicative of a close approximation to the therapist's [s] and [k] targets. Moreover, SI-trajectories, despite their quite intricate calculation, are quite easy to interpret: a line that falls between 1 and 0 means that [sk] is distinctively produced, whereas a straight line close to .5 means it is not. For this reason, if the SI-trajectory were displayed online during an articulation, it might provide additional useful data to the child for interpreting the success in producing a heterorganic sequence, beyond the rich data provided by a sequence of palatograms.

Acknowledgements

References

Barry, M. (1992). Palatalization, assimilation and gestural weakening in connected speech. *Speech Communication*, *11*, 393 – 400.

Byrd, D. M. (1992). Perception of assimilation in consonant clusters. *Phonetica*, *49*, 1 – 24.

Byrd, D. (1996). Influences on articulatory timing in consonant sequences. *Journal of Phonetics*, *24*, 209 – 244.

Biedrzycki, L. (1974). *Abriss der polnischen Phonetik*. Warszawa: Wiedza Powszechna.

Cassidy, S., & Harrington, J. (2001). Multi-level annotation in the Emu speech database management system. *Speech Communication*, *33*, 61 – 77.

Długosz-Kurczabowa, K., & Dubisz, S. (2001). *Gramatyka Historyczna Języka Polskiego*. Warszawa: Polskie Wydawnictwo Naukowe.

Fujimura, O., Macchi, M., & Streeter, L. (1978). Perception of stop consonants with conflicting transitional cues: A cross-linguistic study. *Language and Speech*, *21*, 337 – 346.

Gibbon, F. (1990). Lingual activity in two speech-disordered children's attempts to produce velar and alveolar stop consonants: evidence from electropalatographic (EPG) data. *British Journal of Disorders of Communication*, *25*, 329 – 340.

Gibbon, F., & Crampin, L. (2001). An electropalatographic investigation of middorsum palatal stops in an adult with repaired cleft palate. *Cleft Palate Craniofacial Journal*, *38*, 96 – 105.

Gibbon, F., & Hardcastle, W. (1989). Deviant articulation in a cleft palate child following late repair of the hard palate: a description and remediation procedure using electropalatography. *Clinical Linguistics and Phonetics*, *3*, 93 – 110.

Gimson, A., & Cruttenden, A. (1994). *Gimson's pronunciation of English*, fifth edition. London: Edward Arnold.

Guy, G. (1980). Variation in the group and the individual: The case of final stop deletion. In W. Labov (Ed.), *Locating language in time and space* (pp. 1 – 36). New York: Academic Press.

Hamann, S. (2004). Retroflex fricatives in Slavic languages. *Journal of the International Phonetic Association, 34,* 53–67.

Hardcastle, W., & Roach, P. (1979). An instrumental investigation of coarticulation in stop sequences. In H. Hollien & P. Hollien (Eds.), *Current issues in the phonetic sciences* (pp. 531–540). Amsterdam: John Benjamins.

Hardcastle, W., Morgan Barry, R., & Nunn, M. (1989). Instrumental articulatory phonetics in assessment and remediation: case studies with the electropalatograph. In J. Stengelhofen (Ed.), *Cleft palate: The nature and remediation of communication problems* (pp. 136–164). Edinburgh: Churchill Livingstone.

Hardcastle, W. J., Gibbon, F., & Nicolaidis, K. (1991). EPG data reduction techniques and their implications for studies of lingual coarticulation. *Journal of Phonetics, 19*(3/4), 251–266.

Harrington, J. (2006). An acoustic analysis of 'happy-tensing' in the Queen's Christmas broadcasts. *Journal of Phonetics, 34,* 439–445.

Hock, H. (1991). Initial strengthening. In W. U. Dressler (Ed.), *Phonologica 1988* (pp. 101–110). Cambridge: Cambridge University Press.

Hock, H. (1992). Causation in language change. In W. Bright (Ed.), *Oxford international encyclopedia of linguistics*, Volume 1 (pp. 228–231). Oxford: Oxford University Press.

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

Kohler, K. (1976). The instability of word-final alveolar plosives in German: An Electropalatographic investigation. *Phonetica, 33,* 1–30.

Kohler, K. J. (1992). Gestural reorganization in connected speech: A functional viewpoint on articulatory phonology. *Phonetica, 49,* 205–211.

Kühnert, B. (1996). Die alveolare-velar Assimilation bei Sprechern de Deutschen und des Englischen—kinematische und perzeptive Grundlagen. *Forschungsberichte des Institut für Phonetik und Sprachliche Kommunikation der Universität München, 34,* 175–392.

Martinet, A. (1955). *Economie des Changements Phonétiques.* Bern: Francke.

Mazur. J. (1993). Geschichte der polnischen Sprache, *EuropäischeHochschulschriften: Reihe 16, Slavische Sprachen und Literaturen; 44,* Frankfurt am Main: Peter Lang.

Moon, S.-J., & Lindblom, B. (1994). Interaction between duration, context and speaking-style in English stressed vowels. *Journal of the Acoustical Society of America, 96,* 40–55.

Morciniec, N. (1990). *Das Lautsystem des Deutschen und des Polnischen.* Heidelberg: Groos.

Nolan, F., Holst, T., & Kühnert, B. (1996). Modelling [s] to [ʃ] accommodation in English. *Journal of Phonetics, 24,* 113–137.

Ohala, J. J. (1990). The phonetics and phonology of aspects of assimilation. In J. Kingston & M. E. Beckman (Eds.), *Papers in Laboratory Phonology I* (pp. 258–275). Cambridge: Cambridge University Press.

Ohala, J. J., & Kawasaki, H. (1984). Prosodic phonology and phonetics. *Phonology Yearbook, 1,* 113–127.

Ostaszewska, D., & Tambor, J. (2000). *Fonetyka i Fonologia Współczesnego Języka Polskiego.* Warszawa: Polskie Wydawnictwo Naukowe.

Palethorpe, S., Croot, K., Molloy, N., Gibbon, F., & Harrington, J. (2003). *An electropalatographic investigation of sibilant production by children with repaired cleft palate.* Paper presented at the 6th International Seminar on Speech Production, Sydney, December.

Recasens, D. (2002). An EMA study of VCV coarticulation direction. *Journal of the Acoustical Society of America, 111,* 2828–2841.

Recasens, D. (2004). The effect of syllable position on consonant reduction: Evidence from Catalan consonant clusters. *Journal of Phonetics, 32,* 435–453.

Recasens, D., Fontdevila, J., Pallarès, M. D., & Solanas, A. (1993). An electropalatographic study of stop consonant clusters. *Speech Communication, 12,* 335–356.

Recasens, D., Pallarès, M., & Fontdevila, J. (1997). A model of lingual coarticulation based on articulatory constraints. *Journal of the Acoustical Society of America, 102,* 544–561.

Recasens, D., & Pallarès, M. (1999). A study of /R/ and /r/ in the light of the DAC coarticulation model. *Journal of Phonetics, 27,* 143–169.

Recasens, D., & Pallarès, M. (2001). Coarticulation, assimilation and blending in Catalan consonant clusters. *Journal of Phonetics, 29,* 273–301.

Rochoń M. (2000). *Optimality in complexity: The case of Polish consonant clusters.* Berlin: Akademie Verlag.

Stevens, K. N. (1972). The Quantal nature of speech; Evidence from articulatory-acoustic data. In P. B. Denes & E. E. David Jr. (Eds.), *Human communication: A unified view* (pp. 51–66). New York: McGraw-Hill.

Stevens, K. N. (1989). On the quantal nature of speech. *Journal of Phonetics, 17,* 3–45.

Streeter, N., & Nigro, G. (1979). The role of medial consonant transitions in word perception. *Journal of the Acoustical Society of America, 65,* 1533–1541.

Traunmüller, H. (1990). Analytical expressions for the tonotopic sensory scale. *Journal of the Acoustical Society of America, 88,* 97–100.

Vennemann, T. (1993). *Language changes as language improvement. Historical linguistics: Problems and perspectives.* London: Longman.

Watson, C. I., & Harrington, J. (1999). Acoustic evidence for dynamic formant trajectories in Australian English Vowels. *Journal of the Acoustical Society of America, 106,* 458–468.

Whitehill, T., Stokes, S., Hardcastle, W. J., & Gibbon, F. (1995). Electropalatographic and perceptual analysis of the speech of Cantonese children with cleft palate. *European Journal of Disorders of Communication, 30,* 193–202.

Whitehill, T. L., Stokes, S. F., & Man, Y. H. Y. (1996). Electropalatography treatment in an adult with late repair of cleft palate. *Cleft Palate Craniofacial Journal, 33,* 160–168.

Yamashita, Y., Michi, K., Imai, S., Suzuki, N., & Yoshida, H. (1992). Electropalatographic investigation of abnormal lingual-palatal contact patterns in cleft palate patients. *Clinical Linguistics and Phonetics, 6,* 201–217.