Journal of Phonetics (2000) **28**, 69–102
doi:10.006/jpho.2000.0109
Available online at http://www.idealibrary.com on IDEAL®

Phonetic alignment constraints: consonant overlap and palatalization in English and Russian

Elizabeth C. Zsiga*

Department of Linguistics, Georgetown University, Washington, DC 20057, U.S.A.

Received 20th Januarry 1999, and in revised form 10 March 2000

The possibility of differences in phonetic alignment of phonological segments was examined by contrasting consonant sequences at word boundaries in Russian and English. In Experiment 1, two acoustic measures of consonant overlap, percent released and duration ratio, are computed for stop sequences at word boundaries, and results for English and Russian are compared. While significant variation due to subject, phrase, and cluster type was found, English consistently showed significantly greater overlap than Russian. A formal account of both the variation and the consistent difference is offered, incorporating the idea of phonetic alignment constraints within Byrd's (1996*b*) "phase window" framework. Experiment 2 examines the relationship between overlap and palatalization (or lack thereof) at word boundaries, with an acoustic study of /s + j/ sequences in both English and Russian. The claim of Zsiga (1995) that the apparent change from /s/ to /ʃ/ in phrases such as "press your point" can be attributed to overlap between the /s/ and /j/ gestures is tested and partially supported. In addition to alignment constraints, however, additional phonetic constraints must be taken into account. It is concluded that phonetic constraints must differ from phonological both in containing quantitative information and in being evaluated through weighting rather than strict dominance.

1. Introduction

This paper investigates consonant overlap at word boundaries in English and Russian, with three aims in view. First, the investigation aims to extend our cross-linguistic knowledge of articulatory coordination. While consonant overlap in English has been extensively studied (e.g., Catford, 1977; Hardcastle & Roach, 1977; Hardcastle, 1985; Barry, 1985, 1991; Browman & Goldstein, 1986, 1989, 1990; Nolan, 1992; Zsiga, 1994, 1995; Byrd, 1994, 1996*a*; Byrd & Tan, 1996), little experimental work has been conducted on Russian. (An exception is Barry, 1992; see below.) In the study presented here, two acoustic measures of consonant overlap, percent released and duration ratio, are

* E-mail: zsigae@georgetown.edu

0095–4470/00/010069 + 34 $35.00/0

computed for stop sequences at word boundaries, and results for English and Russian are compared. The second goal of this paper is to investigate the relationship between overlap and palatalization (or lack thereof) at word boundaries. Zsiga (1995) claimed that the apparent change from /s/ to /ʃ/ in phrases such as "press your point" could be attributed to overlap between the /s/ and /j/ gestures. That claim is further examined here, with an acoustic study of /s + j/ sequences in both English and Russian. Finally, the third goal of the study is to integrate these findings into an overall theory of phonology and phonetics. Byrd's (1996*b*) "phase window" framework is supported, and the idea of establishing "phonetic alignment constraints" for the two languages is explored.

Timing of consonants at word boundaries in English has been the subject of a number of studies. Previous studies agree in finding that there is both substantial articulatory overlap between stop consonant gestures at word boundaries in English, and that there can be variability, based on a number of factors, in just how much overlap there is. Catford (1977), for example, found that hetero-organic stop sequences were overlapped for between 29 and 45% of their duration. Zsiga (1984) found acoustic evidence that the second consonant of a C1 # C2 sequence could begin moving toward its target even before closure for the first consonant had been reached. Larger effects were seen when C2 was /k/ as opposed to /p/. Hardcastle & Roach (1977), Barry (1991) and Byrd (1994) each used electropalatographic (epg) data to investigate overlap in coronal–dorsal and dorsal–coronal sequences at word boundaries in English. These sources all found substantial overlap in both cases, and also found that dorsals overlap a preceding consonant to a greater extent than do coronals. Barry (1991), for example, found that in a /g # d/ sequence, the dorsal was overlapped by the coronal for 53% of its duration, and in a /d # g/ sequence, the coronal was overlapped for 87% of its duration. Byrd (1996*a*) went on to compare these hetero-syllabic sequences with onset clusters, and found that hetero-syllabic sequences were both more overlapped and more variable than the onset sequences were. This is in agreement with the findings of Hardcastle (1985), who collected epg data on /kl/ sequences. Hardcastle also found that variability in overlap increased as boundary strength increased.

The present paper aims to replicate and extend these findings on overlap at word boundaries with an acoustic study. Articulatory contact is not directly measured here; however, an acoustic study offers the possibility of examining all places of articulation, including labials. Here, all possible two-consonant combinations of labials, coronals, and dorsals are compared. Two measures of overlap are considered: "percent released", the percentage of clusters in which the release of the first consonant is audible, and "duration ratio", a comparison of the duration of the cluster with the durations of the two consonants (in corresponding word-initial and word-final positions) occurring singly between vowels.

Overlap at word boundaries in Russian has been less extensively studied than in English. Barry (1992) used epg to study palatalization assimilation in Russian consonant clusters, but only examined clusters within words. He found that, as with English, coronals are more likely to be overlapped by the palatal gesture of a following segment than are other consonants. Descriptive accounts (e.g., Jones & Ward, 1969; Avenesov, 1984) indicate that final consonants in Russian are generally audibly released. No direct phonetic measurements of overlap at word boundaries are available for this language, however; nor has there been discussion in the literature of how the coordination of different clusters might vary. Here, the measures of percent released and duration ratio are applied to Russian hetero-syllabic sequences, and results for English and Russian are compared.

The second goal of this paper is to investigate the relationship between overlap and palatalization (or lack thereof) at word boundaries. The term palatalization can be used in two different senses: either to refer to the addition of a secondary palatal articulation (as in the Russian series of palatalized consonants, s *vs.* s^j, p *vs.* p^j, etc.) or to a movement of the primary place of articulation toward the palatal region (/t/ *vs.* /tʃ/, /s/ *vs.* /ʃ/). In English, lexical palatalization of the second type occurs within words before /j/-initial suffixes, as revealed in alternations such as *press/pressure*, *fuse/fusion*, *grade/gradual*, *habit/habitual*. An apparent change from an alveolar to an alveopalatal may also take place when alveolars and /j/ abut at word boundaries, as in the phrases "this year", "miss you", or "hit you".

Zsiga (1993, 1995) claims, based on acoustic and electropalatographic evidence, that the perceived change from /s/ to /ʃ/ in American English phrases such as "this year" is in fact due to overlap between the /s/ and /j/ gestures. The claim is that there is no phonological assimilation operating in these cases. Rather, /s/ and /j/ overlap to a large extent, following the general pattern of consonant overlap at word boundaries typical of English. The overlap produces an articulatory configuration which is not identical to /ʃ/, but whose acoustic consequence (lowered frequency of the fricative noise) is similar to /ʃ/. The view that cross-word palatalization can be accounted for simply in terms of consonant overlap is questioned by Scobbie (1995), who draws attention to some design problems in Zsiga (1995): the small number of subjects (3), and problems with the fit of the electropalates. Other doubts are raised by Holst & Nolan (1995), who argue that a related process (/s/ → /ʃ/ /__/ʃ/) can *not* be accounted for gesturally (though *their* interpretation of their results is questioned by Browman, 1995). The present experiment presents further acoustic analysis of American English /s + j/ sequences, replicating Zsiga (1995) using more subjects, who are speaking without the interference of an artificial palate, and adding some new contexts.[1] In this study, findings from a spectrographic analysis of the /s + j/ sequences will be considered in light of the findings on overlap in the stop sequences, to see whether the overlap analysis of Zsiga (1995) is supported. A similar analysis is conducted for Russian /s + j/ sequences. The prediction of Zsiga (1995), that the different patterns in the two languages can be accounted for largely in terms of different articulatory timing (no overlap, no palatalization), will be tested.

The phonemically palatalized consonants in Russian (such as in /des^{j}at/ *ten*) pose a particular problem for an approach linking palatalization and overlap. Russian /s^j/ has been described as an /s/ co-produced with a palatal /j/ gesture (Avanesov, 1984; Keating, 1988*a*; Barry, 1992). Thus, if Zsiga (1995) is correct, both Russian /s^j/ and English /s + j/ involve co-production of an alveolar fricative and palatal approximant; however, the English and Russian "palatalized" fricatives sound very different. The present experiment also examines the spectral and temporal characteristics of Russian /s^j/, and compares them to English /s + j/, with the goal of determining whether the gestural approach can be defended as an adequate account of post-lexical English palatalization.

This leads to the third goal of the paper: to integrate these findings into an overall theory of phonology and phonetics. The relationship between phonology and phonetics has been re-examined as many researchers have moved from rule-based (e.g., Chomsky & Halle, 1968) to constraint-based (e.g., Prince & Smolensky, 1993) models of phonology. In a rule-based model (as elaborated, for example, in Cohn, 1990), the derivation

[1] Further electropalatographic evidence is currently being collected as well.

begins with an underspecified, categorical representation, to which categorical rules apply. At some point late in the derivation, the phonological representation is translated into a phonetic one, where the rules are gradient, manipulating numbers. In a non-derivational constraint-based theory, however, the phonological representation is not changed from one form to another: all the information is present all the time, and all the constraints have access to it. Recently, some linguists (e.g., Steriade, 1997; Kirchner, 1997; Flemming, 1997) have argued that phonological constraints should have direct access not only to contrastive phonological features, but also to quantitative phonetic information. Under this approach, the distinction between phonology and phonetics disappears. This paper will argue against that view, and hold out for the position that, even in a constraint-based model, phonology and phonetics should be kept separate. Further, it will be argued that the difference between them remains unchanged: phonology is categorical, phonetics is gradient (as argued, for example, by Keating, 1988*b*, 1990*a*; Cohn, 1990; Pierrehumbert, 1990; Zsiga, 1993, 1997).

It is not necessary, however, to consider the phonetic component to be composed of a set of rules that fill in numbers as a continuation of a phonological derivation. With Steriade, Kirchner, Flemming, and others, this paper will argue that many phonetic regularities can be expressed in terms of constraints: constraints that look very much like phonological ones, except for making reference to quantitative information. There will be various kinds of phonetic constraints, corresponding to the "markedness" and "faithfulness" constraints of phonological theory (Prince & Smolensky, 1993).

Markedness constraints encode those aspects of linguistic structure that are universally preferred (such as voiced sonorants and voiceless obstruents, syllables with onsets but no codas, binary prosodic feet, heavy syllables bearing stress, etc.) Phonetic markedness constraints, like the phonological ones, would be based on principles such as "ease of articulation" and "perceptual recoverability" (see Kirchner, 1997; Flemming, 1997 for extended discussion). If a "constraint" is defined as any factor that influences or limits possible linguistic forms, the physical capabilities of the speech system (the jaw can open so far and no farther, the tongue tip cannot be in two places at once, pharyngeal stops are impossible) may be considered "top-ranked" or inviolable markedness constraints. A stricter definition that limits the term "constraint" to those factors within the linguistic grammar would exclude such strictly physical determinants.

Faithfulness constraints stipulate that input and output should correspond as closely as possible. On the phonological side, faithfulness constraints prohibit insertion, deletion, or change in featural specification. (When such differences between input and output *are* found, it is only because a more highly ranked markedness constraint compels them.) In the model being developed here, where phonology and phonetics are separate components, it will be assumed that the input to the phonetic grammar is the output of the phonology. Thus, phonetic faithfulness constraints would specify correspondence between a phonological input and a phonetic output. They would, for example, specify exact place and manner of articulation, such as coronal place = 270° (following Browman & Goldstein (1986) in considering place along the vocal tract as an arc). This kind of constraint would need to be parameterized for different languages, to express, for example, the fact that /t/ is dental in Russian but alveolar in English.

A particularly important class of phonetic constraints will be alignment constraints. Advocates of the gestural approach to phonetic or phonological representations have been arguing that overlap — that is, alignment — between articulatory gestures can account for many assimilations, deletions, and insertions in connected speech, as well as phonetic regularities such as degree of aspiration or nasalization, or release of consonants in clusters (Krakow, 1989; Cooper, 1991; Browman & Goldstein, 1986, 1989, 1990, 1992; Byrd, 1992; Zsiga, 1993, 1995, 1997). This paper will propose considering gestural phasing as a phonetic alignment constraint, similar in form to phonological alignment constraints.

In current constraint-based phonological theory, alignment constraints may be used to align phonological and morphological domains, such as affixes to the right or left edges of phonological words, or the edges of harmony domains to words (McCarthy & Prince, 1993*a*, 1993*b*; Cole & Kisseberth, 1995). For example, Cole and Kisseberth account for harmony systems in which all vowels in a word must agree in the feature [± advanced tongue root] by positing constraints that align the left and right edges of the phonological feature domain with the left and right edges of the phonological word. Crucially, to capture the categorical nature of such phonological processes, phonological alignment constraints refer only to domain edges. Gestural phasing rules (such as those proposed by Browman & Goldstein (1990) and throughout the literature on Articulatory Phonology) contain the same sort of information: a point in one constituent is aligned to a point in another constituent. In the phonetic alignment constraints proposed here, the objects being aligned would be gestures or constellations of gestures, and the instants being aligned would be salient points within them, such as onset, offset, closure, or release. The difference between the two types of constraint would be that for phonetic alignment constraints not only edges, but points within the constituent, are available for alignment.[2] It will be argued, as well, that phonological and phonetic constraints must be evaluated differently, and must be kept separate.

The proposed model, then, encompasses two grammatical components, each of which is constraint-based. In the phonological representation place and manner are represented in terms of distinctive features, and relative timing is encoded in terms of association lines, which group features into segments and segments into higher level prosodic constituents. Input–output pairs of representations are evaluated with respect to a strictly-ranked phonological constraint set, and the pairing which violates only lower-ranked constraints is selected. The successful output candidate then serves as input to the phonetic component, where it is paired with a candidate set of phonetic realizations. In the candidate realizations, place and manner are represented in terms of gestural targets, and relative timing is encoded in terms of gestural phasing. Evaluation of input–output pairs with respect to the phonetic constraint set then determines the phonetic realization assigned to each phonological input. The phonetic constraint set includes specification of place and manner, phasing relations that specify alignment, and articulatory and auditory goals such as "be distinct", or "conserve energy". These can, of course, conflict. Which constraint prevails is, as in phonology, a language-specific matter.

The rest of this paper illustrates the kind of work phonetic alignment constraints might do, focusing on new acoustic data on consonant overlap and palatalization in English

[2] A reviewer suggests that exact phasing relations might better be considered as part of the input to a phonetic grammar, rather than as part of the constraint set. In grammars that make no distinction between phonology and phonetics, exact timing might well be part of the input, but such grammars fail to account for the categorical nature of phonological processes. (See Zsiga, 1993, 1997 for extended discussion. Sagey (1988) elaborates the idea that phonological processes make reference only to domain edges, while phonetic processes access internal points.) On the other hand, if phonology and phonetics are considered distinct components, then a specified phasing in the phonetic input requires (1) rules to fill in those numbers and (2) language-specific requirements on possible inputs, both of which are ruled out in constraint-based phonologies.

and Russian. In what follows, Section 2 presents the findings on overlap in stop consonant sequences in the two languages. Section 3 discusses the /s + j/ sequences, and relates the findings there to those in Section 1. Section 4 turns to the palatalized consonants of Russian, presenting data on how they differ from the English sequences. Section 5 concludes with further discussion of phonetic alignment constraints in phonological and phonetic theory.

2. Experiment 1: stop consonant alignment at word boundaries

2.1. *Experimental design*

2.1.1. *Subjects*

Five native speakers of Russian (four women and one man) and five native speakers of American English (also four women and one man) participated. The Russian speakers ranged in age from 25 to 60, all were speakers of the Moscow dialect, and all had lived in the U.S. fewer than 5 years. The English speakers ranged in age from 20 to 30, and were from the American Northeast or upper Midwest. The Russian speakers were all students of English, and the English speakers all students of Russian.

2.1.2. *Materials*

In both Russian and English, /C1#C2/ sequences were studied. A set of two-word phrases containing stop consonant sequences at word boundaries was constructed (Table I). In English, C1 = /p, d, k/ and C2 = /p, t, k/. The voiced coronal stop was used for C1 because many speakers of American English tend to substitute /ʔ/ for final /t/. In each case, C1 is the final consonant of a verb, and C2 is the initial consonant of its object. The verb was monosyllabic or ended on a stressed syllable, and the stress pattern of the object was systematically varied. Finding word pairs in Russian was more difficult. There are few declined verbs that actually end in stops; all are irregular. Because of the difficulty of finding verbs that met the other phonetic criteria, C1 for Russian included both voiced and voiceless stops: /p, b, d, k, g/. C2, as for English, = /p, t, k/. Because the following consonant was voiceless, devoicing applied in these clusters, and C1 was found to be phonetically voiceless in any case. (However, whether the devoicing constitutes a complete neutralization was not directly investigated.) Where possible, verb#object pairs were again used, with stress controlled as for English, though in some cases an adjective phrase or an intransitive verb followed by a locative or other descriptive phrase was used. Each phrase was incorporated into a sentence. (A full list of sentences is provided in Appendix A.) For presentation to the subjects, sentences were printed out onto a set of index cards, one card per sentence.

2.1.3. *Recording procedures*

Subjects were recorded in a quiet room, using a Sennheiser microphone and a Marantz portable tape recorder. The subjects were given instructions in their native language by a research assistant who is fluent in both Russian and English.[3] Subjects were told they

[3] Data were collected by Stefan Kaufmann.

TABLE I. Stop cluster tokens used in experiment 1

Sound	English tokens	Russian tokens	Russian gloss
p#p	stop parts	[grʲop po 'beregu]	rowed along the banks
	stop potatoes	[o'xrip pozli ek'skursi]	got hoarse after the trip
p#t	stop tarts	[grʲop tam]	rowed there
	stop tobacco	[o'xrip tem ni 'meni]	got hoarse nevertheless
p#k	stop carts	[grʲop kak sports'men]	rowed like a sportsman
	stop commercials	[o'xrip kog'da]	got hoarse after
d#p	had parts	[rad 'pasportu]	glad about the passport
	had potatoes	[rad po'silki]	glad about the parcel
d#t	had tarts	[rad 'tapotʃkam]	glad about the sneakers
	had tobacco	[rad ta'baku]	glad about the tobacco
d#k	had carts	[rad 'kamere]	glad about the camera
	had commercials	[rad karanda'ʃu]	glad about the pencil
k#p	make parts	[pʲok 'persik]	baked a peach
	make potato soup	[pʲok pe'tʃenʲa]	baked pastry
k#t	make tarts	[pʲok tort]	baked tarts
	make tomato sauce	[ʒog tele'vizor]	burned the television
k#k	make carts	[pʲok 'kaʃu]	baked kashu
	make commercials	[pʲok kala'tʃi]	baked bagels
p#V	stop Art	[ox'rip ot 'xoloda]	got hoarse because
	stop another	[grʲop o'din]	rowed alone
d#V	had art	[rad 'atlasu]	glad about the atlas
	had another	[rad ar'buzu]	glad about the parcel
k#V	make art	[pʲok 'astru]	baked asters
	make another	[pʲok o'ladi]	baked pancakes
V#p	saw parts	[edʲa 'persik]	eating a peach
	saw potatoes	[edʲa pe'tʃenʲa]	eating pastry
V#t	saw tarts	[edʲa tort]	eating a tart
	saw tobacco	[edʲa tvo'rog]	eating curds
V#k	saw carts	[edʲa 'kaʃu]	eating kashu
	saw commercials	[edʲa kala'tʃi]	eating bagels

were participating in a study comparing Russian and English, but were given no other details until after the recording session. Subjects were given the set of sentence cards, and asked to repeat each sentence three times. They were asked to read "as naturally and smoothly as possible". The cards were shuffled for each subject, with the materials for Russian and English kept separate. (The index cards for the stop sequences were, however, mixed together with a set of cards printed with sentences containing fricatives. The fricative materials are discussed as Experiment 2, below: see Section 3 and Table IV). All subjects read the set of sentences in their native language first, then the set of sentences in the language they were learning.[4] Only the native language productions are analyzed here. (Zsiga (in preparation), analyzes the second-language productions.)

2.1.4. *Analysis*

The acoustic data were digitized at 22 kHz and analyzed using the Signalyze signal analysis software for the Macintosh.

Two acoustic measures of consonant overlap are considered: percent released and duration ratio. A cluster was counted as released if there was evidence in either the waveform or spectrogram of a release burst. Duration ratio, which was computed for each phrase for each subject, was defined as the mean duration of the C1 # C2 cluster (including the intervening release, if any) divided by the sum of the durations of coda C1 and onset C2 occurring intervocalically. That is, duration ratio equals

$$\frac{\text{(average closure duration C1}\#\text{C2)}}{\text{(avg. clos. dur. C1}\#\text{V)} + \text{(avg. clos. dur. V}\#\text{C2)}}$$

Ratios greater than 1 indicate little or no consonant overlap. In order to avoid skewing the data with disfluent tokens, any phrase in which there was a discernible stumble or pause (defined operationally as a period of silence of 350 ms or more) between the two words was excluded (20 tokens, 4% of the total collected, were excluded on this criterion). For the ratio to be meaningful, it must be assumed that the consonants are articulated similarly in clusters and between vowels, so cases where English speakers would normally produce flaps (final /d/ between a stressed and unstressed vowel) are excluded from this measure.

2.2. *Results of experiment* 1

Results for English and Russian speakers were then compared in analysis of variance. For percent released, independent variables were language, C1, C2, and stress. For duration ratio, in the first analysis of variance, only the condition where the second vowel had main stress was considered (due to the effect of flapping, which excludes 'Vd # V, and therefore any ratio where V1 is stressed and C2 is /d/, from the design), so the independent variables were language, C1, and C2. The effect of whether V2 was stressed or unstressed on duration ratio was tested in a second analysis of variance, which included only labials and dorsals as C1. The results of these ANOVA's are shown in Tables II(a) (release) and II(b) (duration ratio) and are considered in detail below.

2.2.1. *Effect of language*

For both percent released and duration ratio, there was a highly significant main effect of language. The means for each language on each measure are given in Table III. The means for Russian are significantly higher than for English on both measures: clusters in

[4] Due to an error, the sentences containing the phrases /p^{j}ok persik/, /p^{j}ok kaʃu/, and /ed^{j}a kaʃu/ were inadvertently excluded from the set of sentence cards for three of the five Russian subjects. These cells, therefore, contain data from only two subjects, and interactions involving these cells should be interpreted cautiously.

TABLE II(a). Results of analysis of variance for % released

	df	*F*-value	*p*-value
Language	1	80.634	0.0001
C1	2	0.062	0.9397
C2	2	9.070	0.0001
Stress	1	6.466	0.0113
Language * C1	2	4.184	0.0158
Language * C2	2	1.170	0.3114
Language * stress	1	0.005	0.9464
C1 * C2	4	42.867	0.0001
C1 * stress	2	0.208	0.8124
C2 * stress	2	1.359	0.2580
Language * C1 * C2	4	15.832	0.0001
Language * C1 * stress	2	0.241	0.7856
Language * C2 * stress	2	3.877	0.0214
C1 * C2 * stress	4	1.413	0.2286
Language * C1 * C2 * stress	2	1.110	0.3510
Error	465		

TABLE II(b). Results of analyses of variance for duration ratio

	Unstressed # stressed only			p # C and k # C only		
	df	*F*-value	*p*-value	df	*F*-value	*p*-value
Language	1	33.970	0.0001	1	61.627	0.0001
C1	2	15.123	0.0001	1	37.716	0.0001
C2	2	0.2863	0.0647	2	2.900	0.0604
Stress				1	3.344	0.0705
Language * C1	2	10.726	0.0001	1	19.839	0.0001
Language * C2	2	4.633	0.0133	2	4.955	0.0092
Language * stress				1	2.256	0.1368
C1 * C2	4	2.411	0.0585	2	3.549	0.0330
C1 * stress				1	8.657	0.0042
C2 * stress				2	6.727	0.0019
Language * C1 * C2	4	0.286	0.8862	2	0.486	0.6165
Language * C1 * stress				1	15.423	0.0002
Language * C2 * stress				2	2.079	0.1313
C1 * C2 * stress				2	6.002	0.0036
Language * C1 * C2 * stress				2	2.338	0.1026
Error	245					

Russian are released more often, and have higher duration ratios. A duration ratio of 0.80 for English indicates that the consonants are overlapped, on average, for 20% of their closure duration. That means, of course, that the movements of the articulators out of C1 closure and into C2 closure will also be overlapped, consistent with the approximately 30–60% overlap in articulatory *contact* measured by Catford (1977), Barry (1991) and Byrd (1996*a*). In contrast, a duration ratio of 0.98 for Russian indicates almost no overlap in closure duration at all.

TABLE III. Main effect of language on stop cluster measures

	% Released	Duration ratio
English	18	0.802
Russian	47	0.977

2.2.2. *Effects of* C1 *and* C2

Fig. 1 shows a graph of percent released and duration ratio for each cluster in each language. The data show both commonalties and differences between Russian and English. Significant effects, as seen in Table II, are discussed in detail below.

First consider percent released. On this measure, there was a large main effect of C2, as well as significant interactions of C1 by C2, language by C1, and language by C1 by C2. The interaction of C2 and language was not significant.

One immediately apparent effect is that of homorganicity: in both languages, homorganic clusters are almost never released. While the effect is seen for both languages, the difference between homorganic and non-homorganic clusters is much more striking for Russian. In Russian, there was not a single example of a /k#k/ cluster with internal release, yet /k#t/ clusters were released 89% of the time, despite the fact that /t/ is articulated further forward than /k/. (Compare these numbers to English, where /k#k/ clusters had a release rate of 7%, and /k#t/ 27%.) Similarly, the /d#t/ cluster in Russian was released 10% of the time, /d#k/ 100% (compare 3 and 30% in English). It has been previously reported (e.g., Ladefoged, 1993), that identical consonants across word boundaries in English will not have an intervening release, though it is interesting to see here that the same effect is found even for /d#t/, when the two consonants differ in voicing. (Catford (1977) notes that homorganic but non-identical clusters will not be released.)

The main effect of C2 is due to the fact that, for both Russian and English, a cluster was less likely to have an audible release if C2 was /p/. Pooled across both languages, clusters were released 20% of the time when C2 was /p/, *vs.* 38% for /t/ and 39% for /k/.

In English, differences in place of articulation seem to account in a straightforward way for the rest of the data: clusters are more likely to have an audible release if C1 is further forward than C2. This is seen in the fact that clusters whose second member is /p/ are less likely to be released than clusters where C2 is /d/ or /k/, and in the fact that /p#t/, /d#k/, and /p#k/ have largest values for percent released.

In Russian, there is much more variability among the hetero-organic clusters, although all are released more often than their English counterparts. The trend, to the extent there is one, is different. Whether or not C1 is further forward than C2 makes little difference. Rather, those clusters that involve one articulation made with the lips and the other with the tongue (/k#p/, /d#p/, /p#t/, and /p#k/) have an audible release less often than /k#t/ and /d#k/, where both C1 and C2 involve tongue closures. Mean percent released for coronal–dorsal combinations in Russian is 94.5%, while for labial–coronal and labial–dorsal it is 55.3%, including those clusters where the labial is first. (The same comparison for English would yield 28 *vs.* 25%.) This finding of a high percent released (and duration ratio) for /d#k/ clusters in Russian, indicating little overlap in these sequences, differs from previous findings for English (Barry, 1991; Byrd, 1996*a*),

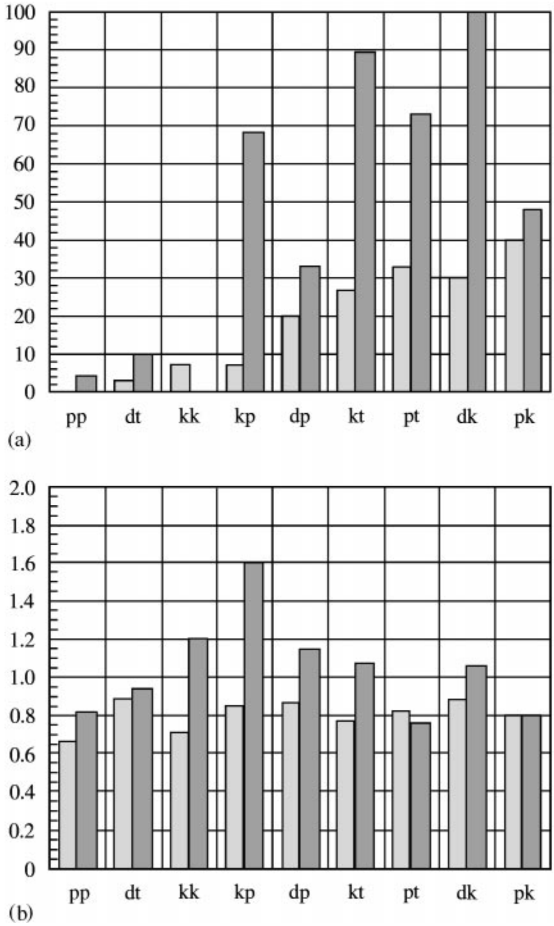

Figure 1. Values for percent released and duration ratio for each cluster in English and Russian. (a) □, English % released; ■, Russian % released. (b) □, English duration ratio; ■, Russian duration ratio.

where coronals were nearly completely overlapped by a following velar stop. Some of the variation in the Russian clusters remains unexplained, particularly why /k # p/ is so high and /p # k/ is so low. However, these results clearly show different articulatory strategies for English and Russian.

Now consider duration ratio, keeping in mind that only data where the second word bears main stress is considered. Statistical analysis (Table II) showed a significant effect of C1, and significant interactions of language by C1 and language by C2. The effect of C1 is that, across both languages, duration ratios are lower when C1 is /p/. The interactions are in line with the fact that Russian is much more variable than English is, as is clear from Fig. 1.

For English, the duration ratio varies little, hovering around 0.8 (that is, the cluster is 80% as long as the sum of C1 and C2 intervocalically). Clusters where C1 is /p/ are slightly lower than average, but not by much. The lowest value, for /p # p/, is 0.659, while the highest, for /d # t/, is 0.881. Variation in duration ratio is not related to variation in percent released. Although /p # p/ is lowest for both percent released and duration ratio, the /d # t/ cluster has the highest duration ratio and one of the lowest percent released, while the /p # k/ cluster has the highest percent released and one of the lower duration ratios. Overall, these data show that, for English, the amount of overlap varies little from cluster to cluster. Whether or not the cluster has an audible release follows from whether C2 is articulated further forward than C1.

The picture is very different for Russian. The highest duration ratio, 1.6 for /k # p/, is twice that of the lowest, 0.75 for /p # t/. Duration ratio more closely mirrors percent released. The greater degree of variation from cluster to cluster is consistent with a speaker strategy in which the two articulations are pulled apart to facilitate an audible release. Although the numbers for /k # p/ should be interpreted cautiously because they represent only two subjects out of five (see note 4 above), it is interesting that duration ratio is highest here, where producing an audible release would be most difficult. (In English, where it is hypothesized there is no reorganization from cluster to cluster, /k # p/ has the lowest percent released of all the hetero-organic clusters.) Conversely in Russian, duration ratio is lowest for clusters where C1 is /p/. In this case, an audible release can be made without drastically pulling apart the two closures. Thus, in Russian, coordination of the stop consonant clusters varies by place of articulation, seemingly in order to facilitate a perceptual goal of audible release.

2.2.3. *Effect of stress*

Previous research (e.g., Hardcastle, 1985; Byrd, 1996*a*; Byrd *et al*., 1999) has shown that the strength of the boundary between two constituents may influence the amount of overlap between them. In the experiment reported here, it was hypothesized that a stressed # unstressed sequence of syllables would form a single prosodic foot, while the unstressed # stressed sequence would form two. Because of the presence of the stronger prosodic boundary, it was hypothesized that both duration ratio and percent released would be higher in the unstressed # stressed condition than in the reverse. In the analysis of variance on percent released, the effect of stress could be tested across all cluster types (Table IIa). Because of flapping of coronals in English, the effect of stress on duration ratio can only be tested for those clusters where C1 is /p/ or /k/ (Table IIb). The result of these analyses is a complex pattern of interactions, as seen in Table II and Fig. 2.

For percent released, the hypothesis is supported: the effect of stress is significant. Across both languages, C1 was significantly more likely to be audibly released when the next word began with a stressed syllable than with an unstressed syllable: 29 *vs.* 36%. The trend held true for almost all cluster types. The exception was Russian clusters where C2 was /t/ (which probably accounts for the language by C2 by stress interaction).

For duration ratio, the main effect of stress was not significant (means were 0.85 in the stressed # unstressed condition, and 0.83 in the unstressed # stressed condition). The language by stress interaction was not significant either, but there are a number of significant interactions of the other factors with stress. Values for the two stress conditions in each language are graphed in Fig. 2. For Russian, there were large differences between stress conditions, but they were not consistent in direction. For English, the effect of stress was smaller, but for five out of six cluster types, duration ratio was higher in the unstressed # stressed condition as predicted. The exception is the /k # t/ cluster, where the effect is reversed.

Overall, then, the effect of stress was not clear in these data.

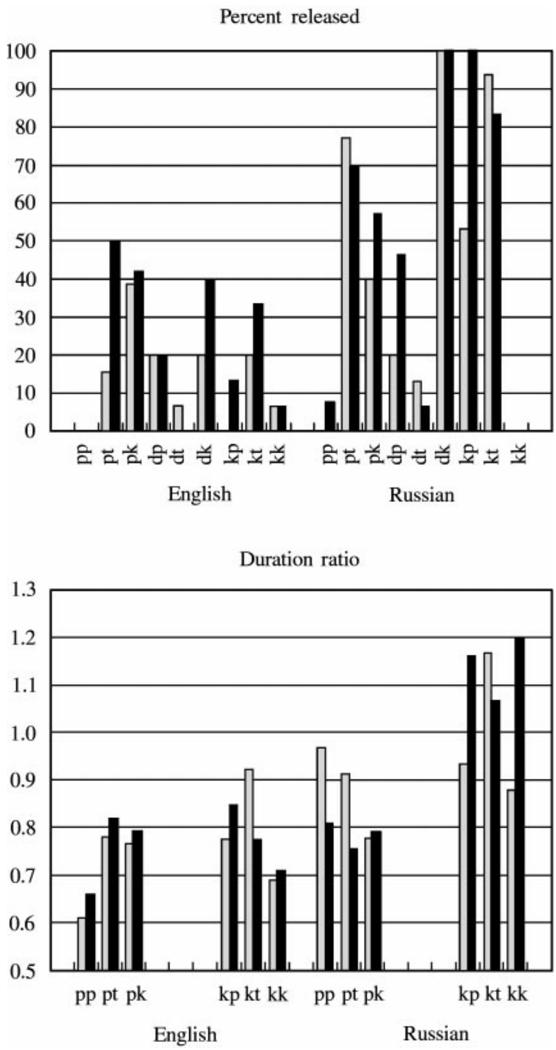

Figure 2. Values for percent released and duration ratio for each cluster in English and Russian, separated by stress condition. □, stress # unstress; ■, unstress # stress.

2.2.4. *Effect of subject*

In each of the above analyses, subjects were pooled. It is worth asking, however, how different the different subjects were, and whether there was a consistent relation for each subject between duration ratio and percent released. A graph of the means for percent released and duration ratio for each subject is given in Fig. 3.

The Russian subjects tend to fall toward the upper right, with higher numbers for both duration ratio and percent released. The English subjects tend to have lower numbers on both, and fall toward the lower left. There is one exception: subject E1 is more Russian-like than English-like on both measures. The anomalous behavior of this subject will be returned to in the discussion of experiment 2.

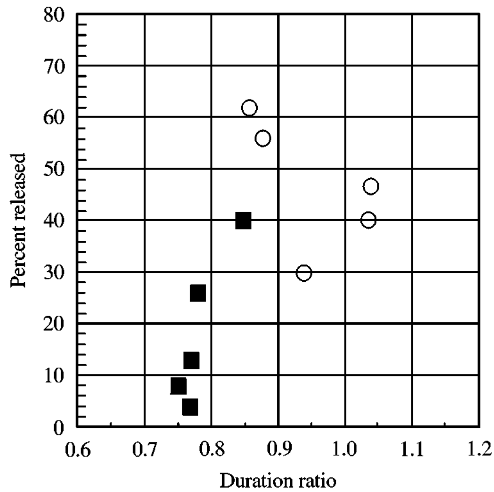

Figure 3. Mean values for percent released and duration ratio for each subject. ■, English; ○, Russian.

2.3. *Discussion of experiment* 1

The strongest finding is the overall difference between English and Russian on the two measures studied. English consonant clusters at word boundaries are more overlapped than their Russian counterparts. A principled statement of consonant alignment in the two languages must account for this consistent difference.

This paper will follow Browman & Goldstein (1986, 1989, 1990, 1992) in assuming that coordination of speech articulations is best expressed as phasing relations between articulatory gestures. Certain points within one gesture (conceived of as a critically damped 360° cycle) are timed to coincide with points in another gesture. Browman & Goldstein (1989) suggest that phasing between hetero-syllabic consonants at a word boundary is not specified, though they leave open the possibility that there is some other word-to-word phasing relationship that would affect overlap in clusters. The data presented here show, however, that different phasing principles of some sort are operative in the two languages. As a first approximation, one might suggest the following.

In English, consonant gestures are aligned such that closure for C2 precedes release of C1. The release then may only be audible if it is articulated further forward than C2. In terms of gestural phasing, where attainment of target position equals 240° and release equals 270°, for English C1(260°) = C2(240°). That is, attainment of target for C2 is timed to occur just before release of C1. In Russian, closure for C2 lags behind release of C1, such that the release is usually audible, regardless of place of articulation. This might be formalized as C1(280°) = C2(240°).

With a slight shift in formalism, these differences are easily expressed as phonetic alignment constraints for the two languages: ALIGN(C1, 260°, C2, 240°) in English and ALIGN(C1, 280°, C2, 240°) in Russian.

These point-to-point alignment constraints, however, make no allowance for variation from cluster to cluster or from speaker to speaker, nor for the differences between English and Russian in terms of how much variation is seen. In English, alignment differed little from cluster to cluster, and whether or not there was an audible release depended most heavily on relative place of articulation. This suggests that a single principle of coordination is operating across all clusters. In Russian, however, there was more variability from cluster to cluster, consistent with a strategy in which the two articulations are pulled apart in non-homorganic clusters in order to facilitate an audible release.

One model which could account well for these facts is Byrd's (1996*b*) phase windows model. Byrd, combining elements of Browman and Goldstein's articulatory phasing and Keating's (1990*b*) window model of coarticulation, suggests that phasing between articulatory gestures be specified not as particular values, but as windows of values within which the phasing relation must fall. (Docherty (1992) proposes a similar model, which also treats inter-articulatory timing in terms of windows, though his model is not couched in terms of Articulatory Phonology.) Different factors, which Byrd terms "influencers" determine where in the window the actual value will fall. These influencers might include rate, formality, and talker idiosyncrasies. They might also include perceptual goals, such as the need for an audible release. The data here suggest a single window for alignment of consonants at word boundaries in English, allowing complete overlap as its lower bound, and audible release as its upper. Incorporating a phase window into the alignment constraints suggested above would mean including a range of values rather than a single point, perhaps ALIGN (C1, [240–270°], C2, 240°) for English. The variability in the Russian data could be accounted for in two different ways, either by specifying different windows for different clusters, or by specifying a single window with a wide range.

The very large differences between homorganic and hetero-organic sequences seen here suggest at least two different windows for these cases. Homorganic clusters would have an alignment constraint similar to English clusters, disallowing audible release. An alternative approach would be to posit a phonological process of gemination for the homorganic clusters, treating them as a single phonologically long stop rather than as a sequence. In that case, no phonetic alignment constraint for homorganic clusters would be necessary. If there were phonological gemination, however, one would expect that the homorganic clusters would *never* have an internal release, and that is not quite the case here.

The alignment constraint for the hetero-organic clusters would specify an alignment in which release was usually audible. A constraint such as ALIGN (C1, [260–320°], C2, 240°) would capture the range of variation found here. Exactly where within the alignment window each cluster fall would be determined in part by pressure to reach the perceptual goal of audible release. Due to this pressure, for example, coronal–dorsal clusters would be pushed toward the outer edges of the window, while labial–dorsal clusters could have greater overlap and still attain the perceptual goal. Thus, by specifying different alignment constraints, but allowing a window of variation within each constraint, the differences both between and within the two languages can be accounted for.

Experiment 2 now turns to the question of whether the patterns of consonant alignment at word boundaries in the two languages can account for the facts of palatalization.

3. Experiment 2: palatalization

3.1. *Experimental design*

3.1.1. *Subjects and recording procedures*

Subjects and recording procedures were the same as for experiment 1. Data on fricatives were recorded in the same session, interspersed with the data on consonant clusters.

3.1.2. *Materials*

In both languages, /s + j/ sequences are contrasted with underlying /s/ and /ʃ/. In Russian, tokens containing /sʲ/ were also recorded. The single /s/, /ʃ/, and /sʲ/ fricatives occurred between vowels, the first of which was stressed, the second unstressed. The fricatives /s/ and /ʃ/ were word-final, /sʲ/ was word-internal. For the /s + j/ sequences, the words used were verb # object pairs (with a few additional pairs, for particular study, noted below). In each case, the verb was monosyllabic or ended on a stressed syllable. The stress pattern of the second word was varied systematically, as was its status as noun or pronoun. The words or phrases containing the fricatives are given in Table IV. The full set of sentences are given in Appendix A.

3.1.3. *Data analysis: centroids*

As in experiment 1, the acoustic data were digitized at 22 kHz and analyzed using the Signalyze signal analysis software for the Macintosh. Of particular interest was how the acoustic quality of the fricatives changed (or did not change) over time. Thus, both duration and spectrographic measures are relevant. The duration of each fricative was measured from the onset to the cessation of aperiodic noise, based on both the waveform and spectrogram. In order to compare the quality of the fricatives, analysis of centroid values was used. Each fricative was divided into thirds, and the mean centroid value (over a range of 500–11 000 Hz) for the beginning, middle, and end was calculated. The centroid is a weighted average (based on amplitude) of all the frequencies present in the spectrum, and gives a measure of the frequency around which the fricative noise tends to be concentrated. A window size of 5 ms was used, then the centroid was calculated based on an averaged spectrum over the third of the fricative.

TABLE IV. Fricative tokens used in experiment 2

Sound	English tokens	Russian tokens	Russian gloss
s	miss another press another	[pas av'tsu] [prɪ'nʲos ar'buz]	tend the sheep brought a watermelon
ʃ	wish another rush another	[vaz'mʲoʃ ar'buz] [u'bʲoʃ a'lenu]	take a watermelon kill Alenu
sʲ		['vosʲəm] ['desʲat]	eight ten
s # j Verb + pronoun	miss you press you press your point	[spas je'vo] [pas je'jo]	rescued him tended it
s # j Verb + noun initial stress	miss yesterday press yards	[spas 'juru] [prɪ'nʲos 'jabloko]	rescued Yuru brought an apple
s # j Verb + noun initial unstress	miss Yolanda press uranium	[spas jev'genʲa] [prɪ'nʲos jit'so]	rescued Yevgeny brought an egg
Other phrases	pressure point boris Yeltsin	['kaʃu] [bo'ris 'jɛltsɪn]	a kind of pastry

3.2. *Results*

3.2.1. *Effects of language and fricative type*

Fig. 4(a) shows the overall results for English. The graph shows the mean centroid value for each fricative type across all five subjects. The /s/ is high-pitched, around 6000 Hz, while /ʃ/ is lower pitched, around 5000 Hz. For the /s + j/ sequences, there is a gradual fall over the course of the fricative, from /s/-like at the beginning to more /ʃ/-like at the end: a gradient palatalization. The gradually falling frequency is clear in Fig. 4(b), which reproduces a representative spectrogram (one token of "press you" from subject E4). Note particularly that the high F3 and F4 formants, indicative of the palatal glide, overlap considerably with the fricative noise.

These overall findings are consistent with the findings of Zsiga (1995): partial assimilation consistent with overlap between the /s/ and /j/ gestures. The gradually falling

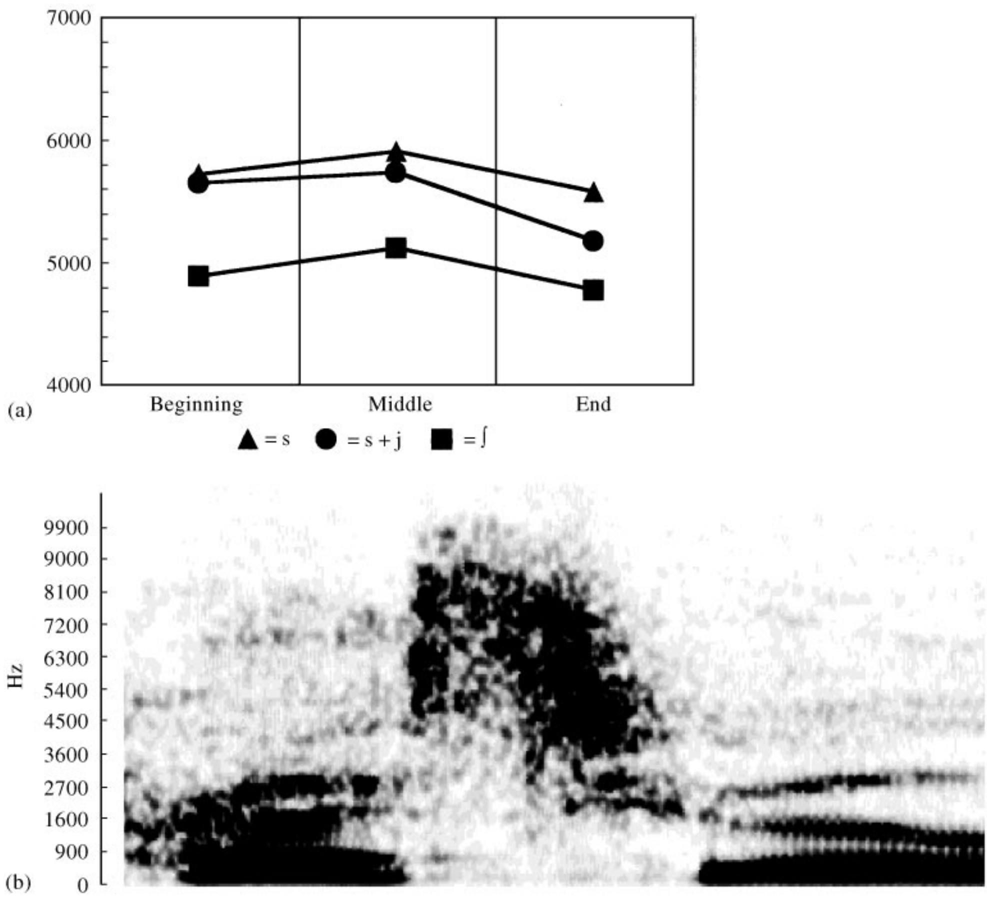

Figure 4. (a) Mean centroid values at beginning, middle, and end of the fricative for /s/, /ʃ/, and /s + j/ in English. (b) A spectrogram of the phrase "press you" spoken by subject E4.

frequency is not, however, the inevitable result of an /s/ followed by a /j/, but the result of the particular pattern of overlap that is typical of English. If a language showed less overlap, we would expect to see less assimilation.

This, in fact, is the case for Russian. As Fig. 5(a) shows, Russian /s + j/ is not distinguished from /s/ between two low vowels. The spectrogram in Fig. 5(b) shows a representative token, /pas jejo/ from speaker R4. Even in the context where English speakers showed the most palatalization — an /s/ followed by an initially unstressed pronoun — Russian speakers showed no assimilation at all. The fricative noise remains steady and high-pitched throughout. The high $F3$ and $F4$ of the palatal glide are clearly visible, but they do not begin until the fricative noise has subsided.

Statistical analysis supports the patterns seen in Figs 4 and 5, and also brings out some interesting further details. A repeated measures analysis of variance was performed on the centroid values for the /s/, /ʃ/, and /s + j/ fricatives, with factors position (beginning, middle, and end), language, subject (within language), and fricative type. Because this

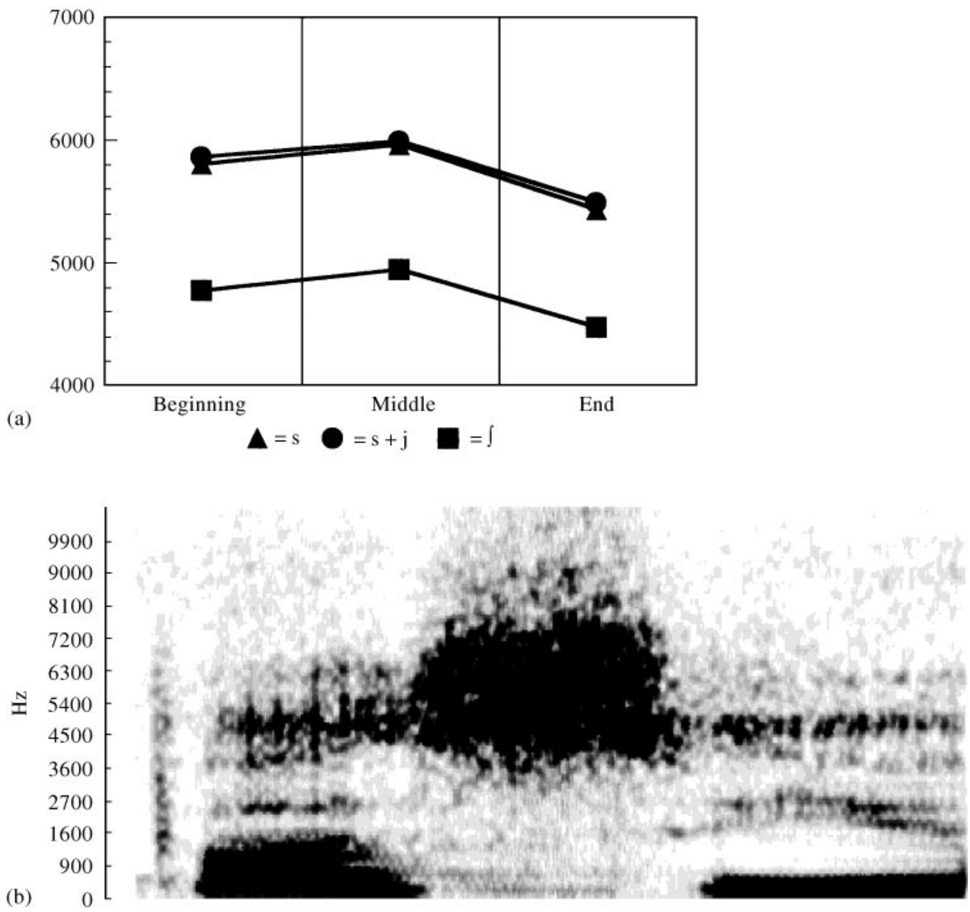

Figure 5. (a) Mean centroid values at beginning, middle, and end of the fricative for /s/, /ʃ/, and /s + j/ in Russian. (b) A spectrogram of the phrase “pas jejo” spoken by subject R4.

TABLE V. Results of analysis of variance for centroid values

	Beginning			Middle			End		
	df	*F*	*p*	df	*F*	*p*	df	*F*	*p*
Language	1	4.896	0.0276	1	3.140	0.0773	1	0.944	0.3320
Fricative type	2	367.8	0.0001	2	295.4	0.0001	2	155.7	0.0001
Subject (within lang)	8	44.56	0.0001	8	45.18	0.0001	8	40.97	0.0001
Fric type * lang	2	11.81	0.0001	2	17.11	0.0001	2	25.52	0.0001
Fric. type * subj (within language)	16	5.613	0.0001	16	4.304	0.0001	16	3.820	0.0001
Error	342			342			342		

analysis showed interactions of position with the other factors, a further analysis of variance was performed for each position. Results are shown in Table V. There was a significant main effect of language only for the initial centroids, where the Russian fricatives were slightly higher-pitched. (By the middle and end of the fricatives, the Russian tokens of /s/ and /s + j/ remained higher than the English, while the /ʃ/ tokens fell lower, so there was no overall effect at those points.) At all three points, however, there were highly significant effects of fricative type and subject, as well as interactions of fricative type with both subject and language. As will be seen below, the Russian tokens consistently showed no assimilation. The patterns in English were more varied. Some tokens, particularly those where the /j/ sound began a content word with an initial stressed syllable, showed little or no assimilation; others, particularly pronouns and words that began with unstressed syllables, showed much more. The English subjects also differed as to how much assimilation they evidenced. These effects and interactions of utterance and subject are explored below.

3.2.2. *Effects of utterance and subject*

To test for differences between the different utterances, a further analysis of variance was performed at each position for each subject, with utterance as the independent variable, and centroid value as the dependent. For all subjects in both languages, the effect of utterance was highly significant. A *post-hoc* Student–Newman–Keuls analysis was used to check which fricatives were significantly different. Of particular interest, of course, are whether the centroid values in the /s + j/ sequences are distinct from either /s/ or /ʃ/. To control for the effect of surrounding vowels in English, the words "press" and "rush" were used as controls for the "press" sequences, and the words "miss" and "wish" were used as controls for the "miss" sequences. In Russian, the /s + j/ sequences were compared with those words containing /s/ and /ʃ/ that are most similar (see Table IV).

Four basic patterns were found: no assimilation, some partial assimilation, greater partial assimilation, and complete assimilation. The tokens that fell into each pattern for English and Russian are listed in Table VI. Example spectrograms from subject E4 illustrating the range of results are shown in Fig. 6.

The results for Russian are very clear: no assimilation. With the exception of only one utterance for one subject, there was no lowering of the fricative noise towards an /ʃ/-like

TABLE VI. Patterns of palatalization for different utterances in English and Russian

(a) No assimilation. /s + j/ is not significantly different from (lower than) /s/ at any point	
English	Russian
E1: press you, press yards, press uranium E3: all /s + j/ sequences E4: miss you, miss yesterday, miss yolanda E5: all /s + j/ sequences, *except* press your	R1: all /s + j/ sequences, except [prin^{j}os jabloko] R2: all /s + j/ sequences R3: all /s + j/ sequences R4: all /s + j/ sequences R5: all /s + j/ sequences
(b) Some assimilation. /s + j/ is significantly different from (lower than) /s/ only at the end	
English	Russian
E1: miss you, miss yesterday, miss yolanda, boris yeltsin E2: miss you, miss yolanda, press you, press yards, press uranium E4: press yards, press uranium E5: press your	No /s + j/ sequences for any subject
(c) Greater assimilation. /s + j/ sequences are like /s/ at onset, but at middle and end are significantly lower than /s/ and the same as /ʃ/	
English	Russian
E1: press your E2: press your, miss yesterday, boris yeltsin E4: press you	R1: [prin^{j}os jabloko]
(d) Complete Assimilation. /s + j/ is significantly lower than /s/, and the same as /ʃ/, at all three points	
English	Russian
E4: press your, boris yeltsin	No /s + j/ sequences for any subject

value in any /s + j/ sequence, regardless of subject or of following word. (In this one anomalous case, /prɪˈn^{j}os ˈjabloko/ for subject R1, the amplitude of the fricative noise was very low, less than a third of the average for this subject.) An example spectrogram showing the lack of assimilation in Russian /s + j/ sequences was shown in Fig. 5. For the Russian speakers, /s/ followed by /j/ does not differ acoustically from /s/ followed by a vowel.

There were also many tokens in English that showed no evidence of assimilation. An example, from the utterance "miss yesterday" from subject E4, is shown in Fig. 6(a). The fricative noise is steady and high-pitched. At the very end of the fricative, formants indicative of the palatal glide become evident, but their effect on the overall pattern of frication is slight and non-significant. Two English subjects, E3 and E5, evidenced this pattern throughout. Subject E3 showed no clear evidence of assimilation at all, and

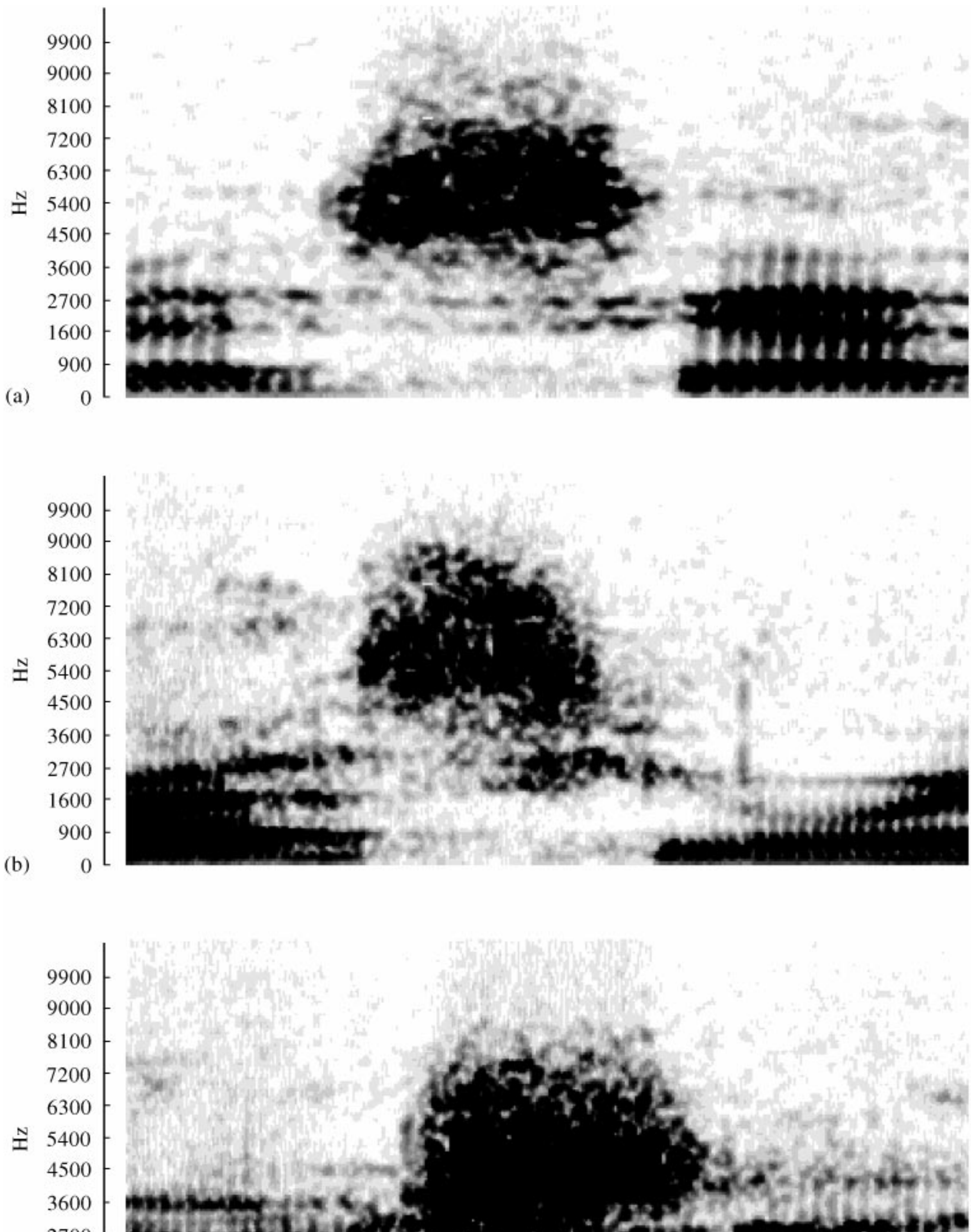

Figure 6. Variation in the amount of palatalization in English. All spectrograms are from subject E4. (a) No assimilation. The phrase "miss yesterday". (b) Some assimilation. The phrase "press uranium". (c) Complete assimilation. The phrase "Boris Yeltsin".

subject E5 showed clear evidence of some assimilation[5] for only one utterance, "press your point". (This utterance was also the only case for which there was even a hint of assimilation for subject E3. For that subject and phrase, centroid values fell in between /s/ and /ʃ/ at all three points, but neither the difference from /s/ nor the difference from /ʃ/ reached significance.) The "press" sequences for subject E1 and the "miss" sequences for subject E4 also fall into the "no evidence of assimilation" category. Here, the centroid values for the control /s/ tokens ("press" for E1 and "miss" for E4) fell considerably at the end of the fricative, so much so that they were not significantly different from /ʃ/ at that point. In the /s + j/ sequences, the centroid value also fell at the end of the fricative but no more so than /s/ did. Thus, the /s + j/ sequences were not, at the end of the fricative, significantly different from either /s/ or /ʃ/, and thus there was no clear statistical evidence of assimilation. Certainly, lowered amplitude of fricative noise may be implicated here. Lowering in the /s + j/ sequences can only be attributed to the following /j/ if the fall in centroid values at the end of the fricative is greater than can be attributed to loss of amplitude alone.

Utterances were counted as showing "some assimilation" if there was significant lowering in the /s + j/ sequence only in the last third of the fricative. That is, /s/ and /s + j/ were indistinguishable for the beginning and middle centroid values, but the end centroid value for the /s + j/ sequence was significantly lower than the comparable value for /s/. No Russian tokens fell into this category. The English cases that showed this pattern are listed in Table VI(b). For the phrase "miss you" for subject E2, the final centroid value was significantly lower than /s/, but also significantly higher than /ʃ/. For all other subjects and phrases listed in Table VI(b), centroid values at the end of the fricative were significantly lower than /s/, and indistinguishable from /ʃ/.[6] (Cases (involving subjects E1 and E4) where the final centroid value is not significantly different from either /s/ or /ʃ/ are counted as "no assimilation", and listed above.) The spectrogram in Fig. 6(b) (a token of "press uranium" from subject E4) is typical of tokens showing "some assimilation". Here, the palatal formants extend further into the fricative, and are excited by noise themselves. Some lowering of the main region of fricative noise toward the palatal formants is also evident.

While this small effect at the end of the fricative should not be ignored, and the fact that no Russian phrases showed this pattern should be noted, the effect must be viewed with caution, especially in light of the fact that a number of tokens also showed some lowering at onset, and some of the /s/ controls also showed lower values at the end. More telling are those cases where there was an effect in the middle of the fricative, where amplitudes are highest, and presumably the target articulatory position has been reached.

[5] For subject E5, the centroid values for "miss" and "wish" were very close together, and the differences between them did not reach significance. The difference between "press" and "rush" was significant only at the middle of the fricative. Again, signal-to-noise ratio may be implicated. Thus, when the /s + j/ sequences are examined, even when they do numerically fall in between /s/ and /ʃ/, they are not significantly different from either. In any case, because the /s + j/ sequences are not significantly different from /s/, this subject does not provide clear evidence to support the hypothesis that there will be partial assimilation in /s + j/ sequences. For the other subjects and tokens in this category, /s/ and /ʃ/ were found to be significantly different, and the /s + j/ sequences were /s/-like at all points.

[6] In some of the phrases in this category, centroid values at the beginning of the fricative also fell between /s/ and /ʃ/, though none were both significantly lower than /s/ and indistinguishable from /ʃ/, as were the values at the end. The phrase "boris yeltsin" for subject E1 and the phrase "press your point" for subject E5, had centroid values at the beginning of the fricative that were in between /s/ and /ʃ/, but not significantly different from either. The phrases "press uranium" and "press yards" for subjects E2 and E4 fell in between /s/ and /ʃ/, and were significantly different from both. This initial lowering is presumably due to lowered amplitude.

Five different English phrases showed such a pattern, listed as Greater Assimilation in Table VI(c). For four of the five phrases in this category, /s + j/ sequences were like /s/ at onset, but at the middle and end of the fricative were significantly lower than /s/ and the same as /ʃ/.[7] For the fifth ("boris yeltsin" for subject E2), there was an additional partial effect at onset. At the onset and middle of this phrase, centroid values fell in between those for /s/ and /ʃ/, and were significantly different from both. At the end, centroid values were significantly lower than those for /s/ and indistinguishable from those for /ʃ/. An example spectrogram ("press you", subject E4) was shown in Fig. 4(b). The pattern is similar to Fig. 6(b), except that evidence of a palatal articulation begins earlier, and there is greater lowering of the main region of fricative noise.

Finally, there were a few cases for subject E4 where the centroid values in the /s + j/ sequences were indistinguishable from the values for /ʃ/ at any point.[8] These are counted as "complete assimilation" and listed in Table VI(d). A spectrogram of one token of the phrase "boris yeltsin" is shown in Fig. 6(c). Palatal formants, excited by noise, extend throughout the fricative, and the noise is uniformly lower pitched.

In this set of data, greater or complete assimilation was the exception rather than the rule. Only two phrases out of a possible 40 looked like complete assimilation. Five more showed a lowering effect as early as the middle of the fricative. Of these seven, four consisted of the verb "press" followed by an unstressed pronoun. Two others were a familiar name. In only one case did a verb followed by a stress-initial content word show an effect earlier than the very end of the fricative. Twelve additional phrases showed a lowering effect (greater than any lowering for /s/ followed by a vowel) at the end of the fricative. Verbs followed by pronouns and by nouns of both stress types fell into this category. Finally, 21 out of the 40 possible cases showed no statistically clear evidence of assimilation at all. Lack of rampant assimilation is perhaps to be expected in read speech recorded in a laboratory setting.

The five English subjects certainly differed in the amount of assimilation in their data. Subject E3 showed no tendency toward assimilation at all; Subject E5 showed some assimilation for only one phrase. E1 tended toward less assimilation rather than more. E2 was the only subject who showed at least some assimilation in all phrases, but E4, the most variable subject, was the only one who provided evidence of apparently complete assimilation, in two different phrases.

It is interesting to note how the subjects' performance in Experiment 2 compared with their performance in Experiment 1. A ranking of the 10 subjects on the three experimental measures — percent released, duration ratio, and amount of palatalization, is given in Table VII. A link between the three measures is supported for most subjects but not for all. Subjects E2 and E4 had the highest percentage of tokens palatalized. They also had the lowest duration ratios and the fewest stop clusters with internal release. Conversely, all of the Russian subjects showed little or no (mostly no) palatalization and had high duration ratios and percentage of clusters released. Subject E5 was intermediate on all three measures. These eight subjects would seem to support a close link between overlap in stop clusters at word boundaries and palatalization in /s + j/ sequences. Subjects E1 and E3 are different, however. Subject E3 has a low percent released, and

[7] Recall, however, that for subject E1 there was a final lowering effect for the /s/ in "press", so that centroid values at the end of the fricative in "press your" were not significantly different from either /s/ or /ʃ/. At the center of the fricative, /s + j/ was significantly lower than /s/ and indistinguishable from /ʃ/.
[8] For "boris yeltsin", the centroid values at the end were not significantly different from either /s/ or /ʃ/.

TABLE VII. Rankings of the 10 subjects on the three experimental measures

Fewest released		Lowest duration ratio		Most /s + j/ tokens palatalized	
E2	4%	E4	0.751	E2	100%
E4	8%	E2	0.769	E4	63%
E3	13%	E3	0.77	E1	63%
E5	26%	E5	0.78	E5	13%
R3	30%	R5	0.798	R1	13% (0%?)
R1	38%	E1	0.847	E3	0%
E1	40%	R4	0.856	R2	0%
R2	47%	R3	0.939	R3	0%
R5	56%	R1	0.985	R4	0%
R4	62%	R2	1.041	R5	0%
Most released		Greatest duration ratio		Fewest /s + j/ tokens palatalized	

a duration ratio as low as that of E2. But she shows no clear evidence of any palatalization. E1, on the other hand, has values for percent released and duration ratio that are Russian-like. She, however, has a high percentage of tokens that show at least some palatalization (albeit usually only in the final third of the fricative, as was seen in Table VI). These subject effects will be returned to in the discussion section.

3.2.3. *Correlations between duration and centroid value*

One further analysis was conducted on the fricative tokens. The duration of noise for each fricative token was measured. Over all the different fricative types, there was no difference in fricative noise duration between English and Russian. Mean duration was 107 ms for Russian, 108 ms for English. This suggests that the Russian and English speakers used similar tempos: it does not seem to be the case that the English speakers were simply talking faster.[9]

For each fricative type in each language, the correlation between duration and the middle centroid value was computed over all the tokens of that fricative type. There was no significant correlation between these two measures for /s/ or /ʃ/ in English, or for any fricative type in Russian. There was thus no general or necessary correlation between short duration and lower centroid value. The only significant correlation between duration and middle centroid value was for /s + j/ in English ($r = 0.291$, $p = 0.0017$). For this sequence, the lower the centroid value, the shorter the fricative tended to be. For the Russian /s + j/ sequences, $r = 0.087$, $p = 0.3911$.

3.3. *Discussion of experiment* 2

Three clear findings stand out from this experiment.

First, English and Russian /s + j/ sequences are different. Except for a single low-amplitude example, there is simply no evidence for /s/ to /j/ assimilation in Russian.

[9] As a reviewer points out, it is possible that an inherent length difference is masking a rate difference. For example, if Russian fricatives were inherently longer than English fricatives, then the finding of equivalent noise durations would mean that the Russian speakers *were* talking faster. Since no independent measures of speaking rate were made, this possibility cannot be completely discounted.

Whatever the cause of the difference between /s/ and /s + j/ in some English sequences, it is not operative for the Russian phrases.

Second, assimilation in English is variable. Some speakers and phrases are more prone to assimilation than others. Subject E2 showed some evidence of assimilation all the time, while subject E3 never did. Subject E4 varied, with some /s + j/ sequences showing no assimilation and others showing apparently complete assimilation. Assimilation was more common with a following pronoun than a content word, but content words could undergo assimilation ("miss yesterday" for subject E2), and pronouns might not ("press you" for subject E1).

Third, assimilation in English is gradient. As can be seen in Figs 4 and 6, the following glide may affect the fricative only slightly at the end, may cause a gradual lowering over the course of the fricative, or may result in lowering throughout. The gradience of the phenomenon is consistent with the findings of Zsiga (1995).

According to a simple overlap account like that given by zZsiga (1995), the three measures in the two experiments reported here — percent released, duration ratio, and lowering of centroid values — can be accounted for by a single parameter: alignment of consonant gestures at word boundaries. There is greater overlap at boundaries in English than in Russian, resulting in lack of release in stop clusters, and gradient palatalization in /s + j/ clusters (and probably a host of other effects, such as casual speech assimilations and deletions of final consonants, as others have argued.) In English, consonant gestures are aligned such that attainment of target constriction for C2 precedes the release of C1. In Russian, attainment of target for C2 follows release of C1. As was discussed above (Section 2.3), these differences may be expressed as different patterns of consonant alignment for the two languages. Importantly, if consonant overlap alone gives rise to the perception of palatalization, as the two combined articulations modify the acoustic result, no separate rule (or constraint) of palatalization is needed. Everything follows from the degree of overlap between the two consonants.

The simple overlap account given in Zsiga (1995) is not fully supported by these data, however. The experiments reported here have shown that English stop consonant clusters at word boundaries have greater overlap than do consonant clusters at word boundaries in Russian, and that gradient palatalization occurs in English, but not in Russian. While there *is* supporting evidence that these two facts are related, the evidence is not unequivocal. Perhaps the clearest and simplest argument in favor of the overlap account is visual evidence from the spectrograms in Figs 4–6. The high F3 and F4 typical of the palatal glide overlap with the higher pitched fricative noise, and are excited by fricative noise themselves. The intrusion of this lower band of noise lowers the overall centroid value into the /ʃ/ range. Another piece of evidence comes from correlations between fricative duration and centroid values. One possible explanation of the fact that only English /s + j/ sequences showed a significant correlation between these two measures is that greater overlap shortens the fricative. The more overlap, the greater the influence of the glide articulation, and the more quickly /s/ would change to /j/. Thus, in English, there would be a correlation between shorter duration and lower centroid value. In Russian, the /s/ and /j/ are separate enough that the following /j/ would not affect the fricative at all. (On the other hand, it is also possible that the correlation arises because palatalization is a characteristic of faster speech in English, so that lower centroid values and shorter fricatives would both be related to faster speech, but not directly related to overlap or to each other.) The subject-by-subject analysis (Table VII) provides some evidence in support of the overlap account, and some against. For eight out of the 10 subjects, a tendency to greater or lesser consonant overlap goes along with a tendency to greater or lesser palatalization. But subject E3 shows that it is possible to have considerable consonant overlap but no palatalization, while subject E1 shows that it is possible to have at least some palatalization in quite a few tokens, with very little consonant overlap. The spectrographic evidence, evidence from the correlations, and evidence from eight of 10 subjects is consistent with an analysis that attributes palatalization to overlap. But subjects E1 and E3 suggest that, at least, something else must be going on.

To examine what this something else might be, we now turn to a fricative type that has not yet been discussed: the palatalized fricative in Russian.

4. Palatalization in Russian

Palatalized fricatives in Russian pose a problem for the overlap approach to English palatalization. X-ray evidence (e.g., Keating, 1988*a*) has shown that palatalized fricatives consist of an /s/ and /j/ articulated at the same time. How is this different from English /s + j/?

Fig. 7 shows a spectrogram of Russian palatalized /s^j/ ("vos^{j}em", subject R5). Consistently for these fricatives, there were two bands of fricative noise visible on the spectrogram: the stronger, higher pitched noise corresponding to the /s/ dental articulation, and lower bands, which seem to be continuations of $F3$ and $F4$, and which can be attributed to the palatal articulation.

This spectrogram differs in two major ways from those of the English /s + j/ sequences in Fig. 6. The first difference is in timing. In Russian, the palatal gesture begins much earlier, and extends throughout. Its effect on $F2$ in the preceding vowel is clear, giving the vowel a definite dipthongal quality. In English the palatal gesture begins, for almost all the tokens in these data, only towards the end of the fricative. The second difference lies in the distinctiveness of the two articulations. The Russian speaker is able to maintain two separate simultaneous articulations throughout the length of the fricative. The different bands of noise remain steady and distinct.

This is in contrast to the gradient palatalization seen in Figs 4 and 6(b), where the English speakers seem to allow the two articulations to blend, such that the /s/ and /j/

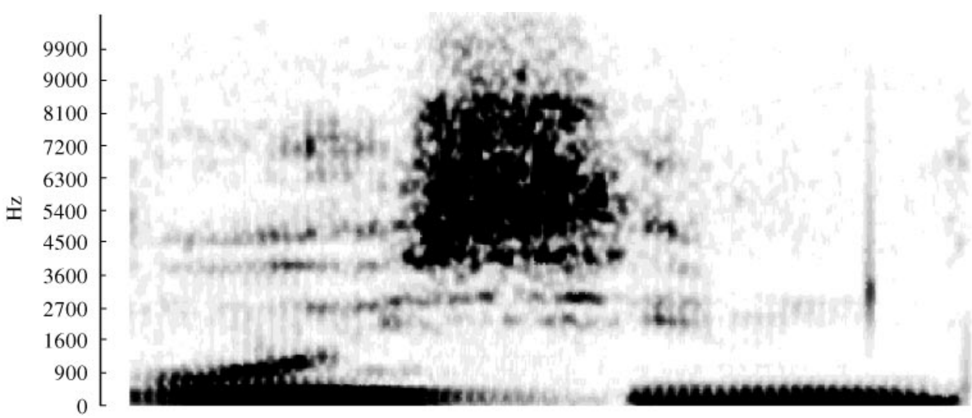

Figure 7. Spectrogram of the word "vos^{j}em", spoken by subject R5.

articulations merge toward the end of the fricative. (Gestural weakening and blending in English /s + j/ sequences was suggested by Scobbie (1995), commenting on Zsiga (1995)). Weakening and blending can be seen even more clearly in Fig. 8, which shows a sequence of spectra, from the beginning, middle and end of one token of "press your" from subject E5. At the beginning and middle of the fricative, there is only a single peak for the /s/. At the end of the fricative, a lower-pitched peak emerges, and the /s/ peak collapses toward it.

Thus, Russian /s^j/ and English /s + j/ both exhibit overlap of the coronal and palatal gestures. They apparently differ in the timing of the two gestures, and in the care taken to keep the two simultaneous articulations separate. Willingness to allow gestural weakening may also help explain the unexpected results for subjects E1 and E3 discussed above (Section 3.2). Subject E3 demonstrated a tendency to greater consonant overlap, but did not evidence much palatalization. This might be attributed to the absence of gestural weakening. Subject E1 had little overlap, but did show some lowering effects at the end of her fricatives. This subject may, on the other hand, have been more prone to weakening, so that a smaller amount of overlap had a greater effect.

Overlap without weakening and blending of the primary /s/ gesture produces a palatalized fricative, as in Russian (Fig. 7). With neither overlap nor blending, there is no effect (Figs 5 and 6(a)). Partial overlap, combined with blending, gives rise to gradient palatalization (Figs 4, 6(b), and 8). Considerable overlap, combined with weakening of the /s/ gesture, produces near-complete collapse of /s/ into an /ʃ/-like fricative (Fig. 6(c)).

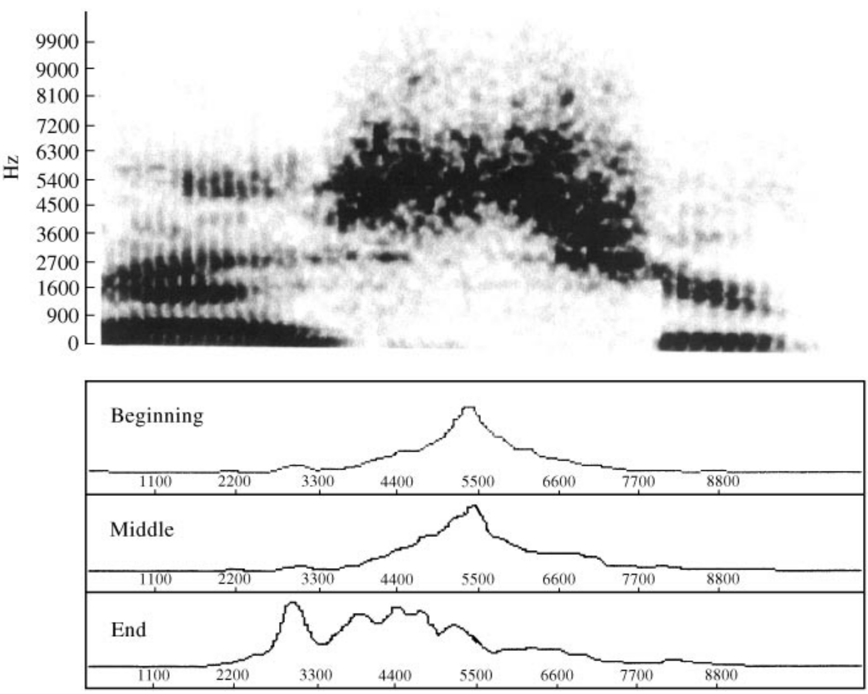

Figure 8. Spectra from the beginning, middle and end of one token of "press your" spoken by subject E5.

5. Conclusion: constraints in phonetics and phonology

Experiment 1 demonstrated that there is greater overlap of stop consonants at word boundaries in English than in Russian. It was suggested that the different patterns of overlap seen in these data be accounted for in terms of phonetic alignment constraints. Such constraints would be similar in form to phonological alignment constraints, but differ in making reference to points internal to the objects being aligned.

Experiment 2 examined /s + j/ sequences in both English and Russian, with an aim toward establishing whether there was a relationship between consonant overlap and palatalization. A gradient and variable palatalization, consistent with the findings of Zsiga (1993, 1995) was found for English, but not for Russian. Close examination of differences between subjects and utterances, and a comparison of /s + j/ sequences to "true" palatalized fricatives in Russian, however, suggested that both gestural overlap and some degree of gestural weakening are necessary for an assimilation to be observed.

The blending seen in English, and the lack of blending in Russian, lead to further consideration of phonetic constraints and their place in the grammar of phonology and phonetics. In the model being presented here, phonological and phonetic constraints belong to different components of the grammar, with the output of the phonological component serving as input to the phonetic component. Phonetic constraints have access to quantitative information; phonological constraints do not. This model contrasts with other constraint-based theories of phonology and phonetics (e.g., Steriade, 1997; Kirchner, 1997; Flemming, 1997) which argue for a single phonological-phonetic component that encompasses constraints referring to both categorical and quantitative information. It has been argued in previous work (Zsiga, 1993, 1995, 1997 and references therein) that a theory which allows quantitative information into the phonological component cannot account for the categorical nature of phonological alternations, such as the true /s/ to /ʃ/ alternation in "press" *vs.* "pressure". The evidence from blending seen here suggests a further reason to keep phonological and phonetic constraints separate: they must be kept separate because they are evaluated in different ways.

Most phonologists pursuing a constraint-based model of phonological alternations (e.g., Prince & Smolensky, 1993) assume that constraints are ranked in strict dominance.[10] If a conflict between constraints arises, the higher-ranked constraint is satisfied, and the lower-ranked constraint is overridden. Thus, if a constraint prohibiting codas is high-ranked and a constraint prohibiting deletion is low-ranked, the coda constraint wins and any consonant that finds itself in the coda will not be pronounced. The deletion is categorical: no trace of the offending consonant remains.

Within the phonetic component, however, conflict resolution is seldom if ever categorical. When phonetic conflicts arise, the two incompatible specifications may be weighted, and a compromise reached. Neither specification wins absolutely; each one contributes something to the final result. This is illustrated in familiar cases such as palatalization of /k/ preceding /i/, dentalization of /t/ before /θ/, and other forms of coarticulation. Browman & Goldstein (1986) modeled such coarticulations as blending between conflicting gestural targets. Later work on gestural models has addressed the

[10] But see Hayes (1997) for a different view. Hayes argues that phonological constraints differ from phonetic constraints in not having access to quantitative information, but that they are like phonetic constraints in undergoing evaluation by weighting rather than strict dominance.

need to assign different weights to the conflicting specifications. Saltzman & Munhall (1989), for example, mathematically model "parameter blending" in coarticulation. Goldstein (in press) discusses weighting of competing constraints in phasing relations. Another model of weighted constraint evaluation is Flemming's (1997) "phonetic optimization", which views coarticulation as arising from the interaction of constraints (such as "achieve targets" and "don't move quickly") which must be given variable weightings.

The data presented here suggest that a weighted constraint evaluation is also at work in the blend between /s/ and /j/ in gradient English palatalization.[11] When overlap occurs between these two articulations, competing demands are placed on the articulatory system. In English, the constraint specifying the place for the coronal articulation (or perhaps the word-final articulation) is given less weight, and the articulation is allowed to weaken and merge with the overlapping /j/, producing an intermediate articulatory configuration, and lower-frequency fricative noise. For Russian, perhaps under pressure to keep three phonologically distinct coronal fricatives also phonetically distinct, the specification of the coronal articulation is given a stronger weight, and persists unchanged even when completely overlapped with a palatal constriction.[12] This weighted evaluation may be considered as a reason to keep phonological and phonetic constraints separate. If strict dominance is to be preserved in the evaluation of phonological constraints, then phonological and phonetic constraints must be evaluated independently, because they must be evaluated differently.

In phonetic theory, the idea of "blending" between two conflicting articulatory demands is certainly not new. The innovation proposed here is that specifications of gestural phasing might fruitfully be considered as phonetic alignment constraints, the correlates of phonological alignment constraints. The shift in formalism allows parallels with phonological theory to be seen more clearly and contributes to the development of an overall model of the phonology–phonetics interface. On the value of formal language, one phonologist affirms his "belief in the explanatory value of formal devices: in many cases, the invention of a good notation has revealed the simplicity behind systems that initially seemed complex. A good formal device takes on a life of its own, revealing previously unseen connections and stimulating further inquiry" (Hayes, 1982, p. 227). It is hoped that the presentation of the data on English and Russian in this paper, and the discussion of the data in terms of phonetic alignment constraints, will indeed suggest connections and stimulate further inquiry, both into cross-linguistic timing patterns, and into the relation between phonetics and phonology.

I would like to thank Stefan Kaufmann, Carrie Finch Weiman and Reneé O'Brien for help in data collection and analysis, and audiences at Georgetown University, Cornell University, the University of Maryland College Park, and the 1998 LSA, as well as the reviewers and editors of *Journal of Phonetics*, for many insightful and helpful comments. Errors, omissions, and failures to take others' insights fully into account are all my own. This work was supported by a grant from the Georgetown University Graduate School.

[11] Again, the model proposed in this paper differs crucially from Flemming's in that it is assumed here that weighted evaluation is appropriate only for the phonetic component, while Flemming assumes that phonology and phonetics are not distinct. The two models agree in arguing that the phonetic component may be fruitfully viewed in terms of constraints, and that phonetic constraints must reference quantitative information and be evaluated in terms of variable weighting rather than strict dominance. The determination of which model better accounts for *phonological* alternations requires further phonological analysis and argumentation beyond the scope of this paper.

[12] Flemming (1997) discusses the role of phonological contrast in determining the language-specific weighting of constraints.

References

Avanesov, R. I. (1984) *Russkoye literaturnoye proiznoshenie*. Moscow: Prosveshchenie.
Barry, M. (1985) A palatographic study of connected speech processes, *Cambridge Papers in Phonetics and Experimental Linguistics*, **4,** 1–16.
Barry, M. (1991) Temporal modelling of gestures in articulatory assimilation. *Proceedings of the XXIIth congress of phonetic sciences*, vol. 4, pp. 14–17.
Barry, M. (1992) Palatalisation, assimilation, and gestural weakening, *Speech Communication*, **11,** 393–400.
Browman, C. P. (1995) Assimilation as gestural overlap: comments on Holst and Nolan. In *Phonology and phonetic evidence*, *papers in laboratory phonology IV* (B. Connell & A. Arvaniti, editors), pp. 334–343. Cambridge: Cambridge University Press.
Browman, C. P. & Goldstein, L. (1986) Towards an articulatory phonology, *Phonology Yearbook*, **3,** 219–252.
Browman, C. P. & Goldstein, L. (1989) Some notes on syllable-structure in articulatory phonology, *Phonetica*, **45,** 140–155.
Browman, C. P. & Goldstein, L. (1990) Tiers in articulatory phonology, with some implications for casual speech. In *Papers in laboratory phonology I*, *between the grammar and physics of speech* (J. Kingston & M. Beckman, editors), pp. 341–376. Cambridge: Cambridge University Press.
Browman, C. P. & Goldstein, L. (1992) Articulatory phonology: an overview, *Phonetica*, **49,** 155–180.
Byrd, D. (1992) Perception of assimilation in consonant clusters: a gestural model, *Phonetica*, **49,** 1–24.
Byrd, D. (1994) Articulatory timing in English consonant sequences. *UCLA Working Papers in Phonetics*, Vol. 86.
Byrd, D. (1996*a*) Influences on articulatory timing in consonant sequences, *Journal of Phonetics*, **24,** 209 –224.
Byrd, D. (1996*b*) A phase window framework for articulatory timing, *Phonology*, **13**(2), 139–169.
Byrd, D., Kaun, A., Narayanan, S. & Saltzman, E. (2000) Phrasal signatures in articulation. In *Papers in laboratory phonology V*, *language acquisition and the lexicon* (M. Broe and J. Pierrehumbert, editors). Cambridge: Cambridge University Press.
Byrd, D. & Tan, C. C. (1996) Saying consonant clusters quickly, *Journal of Phonetics*, **24,** 262–282.
Catford, J. C. (1977) *Fundamental problems in phonetics*. Bloomington: Indiana University Press.
Chomsky, N. & Halle, M. (1968) *The sound pattern of English*. New York: Harper & Row.
Cohn, A. C. (1990) Phonetic and phonological rules of nasalization. *UCLA Working Papers in Phonetics*, Vol. 76.
Cole, J. S. & Kisseberth, C. (1995) Paradoxical strength conditions in harmony systems. *Proceedings of the North East Lingustic Society*, Vol. 25, pp. 17–29.
Cooper, A. (1991) Laryngeal and oral gestures in English /p, t, k/. *Proceedings of the XIIth international congress of phonetic sciences*, Vol. 2, pp. 50–53.
Docherty, G. J. (1992) *The timing of British English obstruents*. Berlin: Foris.
Flemming, E. (1997) Phonetic optimization compromise in speech production. *University of Maryland Working Papers in Linguistics*, Vol. 5. V. Miglio and B. Morén (Eds.), pp. 72–91.
Goldstein, L. (in press) Competing constraints on intergestural coordination and self-organization of phonological structure, *Bulletin de la Communication Parlée*.
Hardcastle, W. J. (1985) Some phonetic and syntactic constraints on lingual coarticulation during /kl/ sequences, *Speech Communication*, **4,** 247–263.
Hardcastle, W. J. & Roach, P. J. (1977) An instrumental investigation of coarticulation in stop consonant seqences. In *Current issues in the phonetic sciences* (H. Hollien & P. Hollien, editors), pp. 531–540. Amsterdam: John Benjamins.
Hayes, B. (1982) Extrametricality and English stress, *Linguistic Inquiry* **13**(2), 227–276.
Holst, T. & Nolan, F. (1995) The influence of syntactic structure on /s/ to /ʃ/ assimilation. In *Phonology and phonetic evidence*: *papers in laboratory phonology IV* (B. Connell & A. Arvaniti, editors), pp. 315–333. Cambridge: Cambridge University Press.
Jones, D. & Ward, D. (1969) *The phonetics of Russian*. London: Cambridge University Press.
Keating, P. A. (1988*a*) Palatals as complex segments: x-ray evidence. *UCLA Working Papers in Phonetics*, Vol. 69, 77–91.
Keating, P. A. (1988*b*) The phonology-phonetics interface. In *Linguistics*: *the Cambridge survey*. *Vol. 1*, *Linguistic theory*: *foundations* (F. J. Newmeyer, editor), pp. 281–302. Cambridge: Cambridge University Press.
Keating, P. A. (1990*a*) Phonetic representations in a generative grammar, *Journal of Phonetics*, **18,** 321–334.
Keating, P. A. (1990*b*) The window model of coarticulation: articulatory evidence. In *Papers in laboratory phonology I*: *between the grammar and physics of speech* (J. Kingston & M. Beckman, editors), pp. 451–475. Cambridge: Cambridge University Press.
Kirchner, R. (1997) Contrastiveness and faithfulness, *Phonology*, **14,** 83–112.
Krakow, R. (1989) *The articulatory organization of syllables*: *a kinematic analysis*. Doctoral dissertation, Yale University.

Ladefoged, P. (1993) *A course in phonetics*. New York: Harcourt, Brace, Jovanovich.
McCarthy, J. J. & Prince, A. (1993*a*) *Prosodic morphology I*: *constraint interaction and satisfaction*. Manuscript, University of Massachusetts at Amherst and Rutgers University.
McCarthy, J. J. & Prince, A. (1993b) Generalized alignment. In *Yearbook of Morphology* 1993 (G. Booij & J. van Merle, editors), pp. 79–153. Dordrecht: Kluwer.
Nolan, F. (1992) The descriptive role of segments: evidence from assimilation. In *Papers in laboratory phonology II*: *gesture*, *segment*, *prosody* (G. J. Docherty & D. R. Ladd, editors), pp. 261–280. London: Cambridge University Press.
Pierrehumbert, J. (1990) Phonological and phonetic representation, *Journal of Phonetics*, **18,** 375–394.
Prince A. & Smolensky, P. (1993) *Optimality theory*: *constraint interaction in generative grammar*. Manuscript, Rutgers University and University of Colorado, Boulder.
Sagey, E. (1988) On the ill-formedness of crossing association lines, *Linguistic Inquiry* **19,** 109–118.
Saltzman, E. & Munhall, K. (1989) A dynamical approach to gestural patterning in speech production. *Haskins Laboratories Status Report on Speech Research*, Vol. 99–100, 38–68.
Steriade, D. (1997) *Phonetics in phonology*: *the case of laryngeal neutralization*. Manuscript, UCLA.
Scobbie, J. (1995) What do we do when phonology is powerful enough to imitate phonetics? Comments on Zsiga. In *Phonology and phonetic evidence*, *papers in laboratory phonology IV* (B. Connell & A. Arvaniti, editors), pp. 303–314. Cambridge: Cambridge University Press.
Zsiga, E. C. (1993) *Features*, *gestures*, *and the temporal aspects of phonological organization*. Doctoral dissertation, Yale University.
Zsiga, E. C. (1994) Acoustic evidence for gestural overlap in consonant sequences, *Journal of Phonetics*, **22,** 121–140.
Zsiga, E. C. (1995) An acoustic and electropalatographic study of lexical and post-lexical palatalization in American English. In *Phonology and phonetic evidence*, *papers in laboratory phonology IV* (B. Connell & A. Arvaniti, editors), pp. 282–302. Cambridge: Cambridge University Press.
Zsiga, E. C. (1997) Features, gestures, and Igbo vowels: an approach to the phonology/phonetics interface, *Language*, **73**(2), 227–274.
Zsiga, E. C. (in preparation) *Relearning consonant timing*. Manuscript, Georgetown University.

Appendix A: Text of sentences used in the experiments

A.1. English sentences for experiment 1 (*stop consonant clusters*)

p # p	The manager wants to stop parts from being stolen.
	The farmers want to stop potatoes from being imported.
p # t	The bakery wants to stop tarts from being stolen.
	The farmers want to stop tobacco from being imported.
p # k	The grocery store wants to stop carts from being stolen.
	The protestors want to stop commercials from being shown.
d # p	The children had parts in the play.
	The farmers had potatoes in the fields.
d # t	The children had tarts after lunch.
	The farmers had tobacco in the fields.
d # k	The children had carts in the race.
	The actors had commercials on TV.
k # p	The machine can make parts by the thousand.
	I'll make potato soup for lunch.
k # t	The baker will make tarts this afternoon.
	I'll make tomato sauce for dinner.
k # k	The manufacturer can make karts as well as bicycles.
	They make commercials for TV.
p # V	The museum wants to stop Art from being stolen.
	The police want to stop another crime.
d # V	The children had Art after lunch.
	I think he had another for dessert.

k # V	The government will make Art a priority.
	I'll make another one tomorrow.
V # p	They saw parts on the table.
	They saw potatoes on the farm.
V # t	They saw tarts at the bakery.
	They saw tobacco in the barns.
V # k	They saw carts in the parking lot.
	They saw commercials on TV.

A.2. Russian sentences for experiment 1 (*stop consonant clusters*)

p # p	Moj drug [gr^{j}op po 'beregu].
	My friend rowed along the banks.
	On [o'xrip pozli ek'skursi].
	He got hoarse after the trip
p # t	Moj drug [gr^{j}op tam], pod derevam.
	My friend rowed there, under the trees.
	On [o'xrip tem ni 'meni].
	He got hoarse nevertheless.
p # k	Moj drug [gr^{j}op kak sports'men].
	My friend rowed like a sportsman.
	On [o'xrip kog'da] my vernulis'.
	He got hoarse after we had returned.
d # p	Ego otec [rad 'pasportu].
	His father is glad about the passport.
	Ego otec [rad po'silki].
	His father is glad about the parcel.
d # t	Ego otec [rad 'tapotʃkam].
	His father is glad about the sneakers.
	Ego otec [rad ta'baku].
	His father is glad about the tobacco.
d # k	Ego otec [rad 'kamere].
	His father is glad about the camera.
	Ego otec [rad karanda'ʃu].
	His father is glad about the pencil.
k # p	Moj deduska [p^{j}ok 'persik].
	My grandfather baked a peach.
	Moj deduska [p^{j}ok pe'tʃen^{j}a].
	My grandfather baked pastry.
k # t	Moj deduska [p^{j}ok tort].
	My grandfather baked tarts.
	Jejo muz [ʒog tele'vizor].
	Her husband burned the television.
k # k	Moj deduska [p^{j}ok 'kaʃu].
	My grandfather baked kashu.
	Moj deduska [p^{j}ok kala'tʃi].
	My grandfather baked bagels.

p#V	On [ox'rip ot 'xoloda].
	He got hoarse because of the cold.
	Moj drug [gr^{j}op o'din].
	My friend rowed alone.
d#V	Ego otec [rad 'atlasu].
	His father is glad about the atlas.
	Ego otec [rad ar'buzu].
	His father is glad about the parcel.
k#V	Moj deduska [p^{j}ok 'astru].
	My grandfather baked asters.
	Moj deduska [p^{j}ok o'ladi].
	My grandfather baked pancakes.
V#p	On sidel u stola, [ed^{j}a 'persik].
	He was sitting at the table, eating a peach.
	On sidel u stola, [ed^{j}a pe'tʃen^{j}a].
	He sat at the table, eating pastry
V#t	On sidel u stola, [ed^{j}a tort].
	He sat at the table, eating a tart
	On sidel u stola, [ed^{j}a tvo'rog].
	He sat at the table, eating curds.
V#k	On sidel u stola, [ed^{j}a 'kaʃu].
	He sat at the table, eating kashu.
	On sidel u stola, [ed^{j}a kala'tʃi].
	He sat at the table, eating bagels.

A.3. English sentences for experiment 2 (*fricatives*)

/s#V/	I don't want to miss another game.
	Can you press another shirt for me?
/ʃ#V/	I wish another bus would come soon.
	We need to rush another part to the factory.
/s#j/	I'm going to miss you if you move away.
	Let me press you a little further on that.
	I'm sorry I had to miss yesterday's class.
	The tailor has to press yards of fabric.
	The students will miss Yolanda when she retires.
	The machine can press uranium into pellets.
	You need to press your point more strongly.
	The President of Russia is Boris Yeltsin
derived /ʃ/	You have a pressure point on your wrist.

A.4. Russian sentences for experiment 2 (*fricatives*)

/s#V/	Papa [pas av'tsu].
	Papa tended the sheep.
	Otec [prɪ'n^{j}os ar'buz].
	Father brought a watermelon.

/S#V/	Ty [vaz'm^{j}oʃ ar'buz] iz skafa
	You take the watermelon out of the cupboard
	Esli ty [u'b^{j}oʃ a'lenu], to ja ub'ju tebja.
	If you kill Alena, I'll kill you.
/s^{j}/	My ozidali ['vos^{j}əm] devusek, no prisli ['d^{j}es^{j}at] malichikov.
	We expected eight girls, but ten boys came.
/s#j/	Milicioner [spas je'vo] iz doma.
	The policeman rescued him from the house.
	Papa [pas je'jo].
	Papa tended it.
	Ljudi gvorojat, cto on [spas 'juru].
	They say he rescued Yura.
	Papa mne [prɪ'n^{j}os 'jabloko].
	Papa brought me an appl.e
	Ljudi gvorojat, cto on [spas jev'gen^{j}a].
	They say he rescued Evgeny.
	Papa mne [prɪ'n^{j}os jit'so].
	Papa brought me an egg.
	On ljubil ['kaʃu], kogda on byl malenki.
	He used to like kasha when he was a boy.
	Prezident Rossi [bo'ris 'jɛltsɪn].
	The president of Russia is Boris Yeltsin.