SSLA, **25**, 399–432. Printed in the United States of America.
DOI: 10.1017.S0272263103000160

ARTICULATORY TIMING IN A SECOND LANGUAGE

Evidence from Russian and English

Elizabeth C. Zsiga
Georgetown University

This study compares patterns of consonant-to-consonant timing at word boundaries in English and Russian and investigates the roles of transfer and the emergence of linguistic universals in second language (L2) articulation. Native Russian speakers learning English and native English speakers learning Russian produced phrases in English and Russian contrasting VC#CV, VC#V, and V#CV sequences. The duration of all stop closures was measured as well as the percentage of consonant sequences in which the first consonant was audibly released. In their native language (L1), Russian speakers had a higher percentage of released final consonants than did English speakers in their L1 as well as a higher ratio of sequence-to-singleton duration. Examination of the timing patterns across different clusters revealed different articulatory strategies for the two languages. The native Russian pattern transferred to L2 English, but the native English pattern did not transfer to L2 Russian. Evidence was found for both articulatory transfer and the emergence of a default pattern of articulation, characteristic of neither L1 nor L2.

This study examines articulatory timing between words in connected speech, comparing native speakers (NSs) with nonnative speakers (NNSs). In particular, it examines the word-to-word timing of NSs of English and Russian speak-

I would like to acknowledge: Stefan Kaufmann, Carrie Finch Weiman, and Renée O'Brien, who helped with data collection and analysis; my colleagues Donna Lardiere and Alison Mackey, who offered invaluable advice; the anonymous *SSLA* reviewers, whose comments focused and strengthened the manuscript; and the influence of Louis Goldstein and Dani Byrd, with whom this experiment began.

Address correspondence to: Elizabeth C. Zsiga, Department of Linguistics, Georgetown University, Washington, DC 20057; e-mail: zsigae@georgetown.edu.

ing both English and Russian. These languages exemplify two very different patterns of word-to-word timing and thus provide an interesting test case for the question as to what happens to phonetic timing in a second language (L2). Previous phonetic research (to be subsequently discussed) has shown that NSs of English tend to produce consonants at word boundaries with a great deal of articulatory overlap and that (as a consequence) they seldom if ever produce an audible release burst between the two consonant closures. Russian NSs, on the other hand, generally produce sequences with little or no overlap and create an audible release burst for final consonants much more often, even in clusters. Because of the acoustic and perceptual effects of these different articulatory patterns, the speech of a learner who employs a nonnative pattern may suffer in both naturalness and intelligibility.

The importance of timing patterns in learners' speech has been highlighted in many studies and texts (e.g., Anderson-Hsieh, Johnson, & Koehler, 1992; Anderson-Hsieh, Riney, & Koehler, 1994; Chen, 1982; Dickerson, 1989; Flege, Munro, & MacKay, 1995; Gilbert, 1993; McNerney & Mendelsohn, 1992; Pennington, 1989; Solé, 1997; Tajima, Port, & Dalby, 1997; Wong, 1987). These publications have emphasized that speaking an L2 with a timing pattern typical of the first language (L1) will in and of itself mark the speech as nonnative and difficult to understand. For example, Tajima et al. demonstrated that nonnative timing patterns in the speech of a Chinese learner of English had a significant, negative effect on intelligibility. When differences in timing were digitally corrected, the learner's speech became significantly more intelligible even though errors in segmental quality remained.

In the case of Russian and English, nonnative timing patterns at word boundaries may result in the inappropriate presence or absence of final release bursts. Because the presence of a release burst enhances the perceptibility of stop consonant contrasts (Liberman, Cooper, Shankweiler, & Studdert-Kennedy, 1967; Stevens & Blumstein, 1978), an English NS learning Russian who failed to produce the audible bursts typical of Russian might sound not only unnatural but also unintelligible to a Russian NS. On the other hand, a Russian speaker who carried over to her L2 English speech the articulatory habit of releasing final consonants might correctly signal contrasts in consonantal place of articulation but at the cost of disruption in prosody. Tajima et al. (1997) noted that the presence of unexpected release bursts on final consonants could be interpreted by native English listeners as extra unstressed syllables, rendering individual words less intelligible and creating garden path effects in the interpretation of running speech.

The English pattern of extensive consonant overlap at word boundaries has further acoustic consequences as well. Not only does articulatory overlap prevent audible release bursts that would more clearly cue contrasts among final consonants, but it also may cause final contrasts to be lost completely. In conversational English speech, it is typical for final consonants, particularly coronals, to be perceived as assimilated to the place of articulation of a following word-initial consonant or even completely deleted (e.g., see Catford, 1977; Gimson, 1962; Ladefoged, 1993; Lass, 1984). For example, the phrase *in part* may be heard as *im part*, *this year* as *thish year*, *that boy* as *tha' boy*, and *last year* as *las' cheer*. In many if not all cases, these connected speech alternations may be caused by the English pattern of producing large amounts of overlap between consonants at word boundaries. When two different consonant constrictions are produced at the same time, such as the /t/ and /b/ in *that boy*, the acoustic cues for the word-initial labial consonant may overwhelm those for the word-final coronal, causing the perception that the final consonant has been changed or dropped (Browman & Goldstein, 1990; Byrd, 1992). Similarly, two constrictions made with the same articulator may "blend" into an intermediate constriction location so that /s/ overlapped with /y/ sounds likes /ʃ/ (Zsiga, 1995a, 2000).

For NNSs of English, such connected speech assimilations and deletions create a challenge both in understanding and producing conversational English speech (Anderson-Hsieh et al., 1994; Brown, 1977; Hieke, 1987; Weinberger, 1994b). In terms of comprehension, many lexical contrasts that the learner might expect, on the basis of dictionary pronunciations, to be realized are not. In terms of production, the learner who does not create the cross-word assimilations so pervasive in English speech will sound overly careful and stilted. Furthermore, phenomena such as assimilations at word boundaries serve to link words together into hierarchical prosodic groupings, which NSs rely on to convey syntactic and discourse relationships (e.g., see Kaisse, 1985; Nespor & Vogel, 1986). Nonnative speech in which these linking phenomena are absent will fail to convey the connections between words that the native prosody signals. Conversely, problems would occur for NSs of English learning other languages if the same assimilations and deletions that occur in native speech surfaced inappropriately in their L2 productions. Zsiga (1995b) noted that American English speakers may pronounce the Russian name *Boris Yeltsin* the same way they pronounce *this year*—with a word-final /s/ that sounds more like /ʃ/. In this case, however, the Russian name was embedded in an English sentence. Would the same pronunciation occur if NSs of English were speaking L2 Russian?

The studies on L2 articulatory timing previously noted (e.g., Anderson-Hsieh et al., 1992; Tajima et al., 1997) have demonstrated the presence of L1 timing patterns in L2 speech, confirming that phonetic timing may transfer from L1 to L2 in the same way that phonological patterns are hypothesized to do (Archibald, 1993; Flege, 1992; Hancin-Bhatt & Bhatt, 1997; Lado, 1957). Other studies (Altenberg & Vago, 1983; Hammarberg, 1990; James, 1988; Rubach, 1984; Weinberger, 1994a) have argued that, when phonological transfer occurs, the postlexical processes of the L1, rather than the lexical alternations, cause the most interference. Following up on these findings, Cebrian (2000) suggested that it is "the difficulty in changing the fossilized articulatory timing habits" in the learners' speech that makes transfer of postlexical processes so common (p. 7); see also Solé (1997). If transfer of articulatory timing occurs in English-Russian interlanguage, as predicted by these studies, we would expect that NSs of English would speak Russian with considerable overlap at word boundaries, which would result in pervasive assimilations and deletions of final consonants. Similarly, NSs of Russian speaking English would produce their English consonant sequences in a manner consistent with their L1 patterns.

Cebrian (2000), however, found an asymmetry between transfer of timing patterns within words and across word boundaries. In his study, NSs of Catalan were significantly more likely to transfer to their English speech the pattern of coordination between oral and glottal gestures that gives rise to word-final devoicing than they were to transfer the pattern of coordination that gives rise to cross-word voicing assimilation.[1] Cebrian attributed this asymmetry to a "word integrity effect" in interlanguage that "treats every word as a separate unit and prevents the articulatory synchronization of sounds belonging to different words" (p. 19). As neither the L1 nor the L2 exhibits a word integrity effect, this effect can be seen as an example of the emergence of unmarked linguistic structure in interlanguage (Broselow, 1987; Broselow, Chen, & Wang, 1998; Eckman, 1977, 1981, 1987; Epstein, Flynn, & Martohardjono, 1996).

Universal preferences for ease of either articulation or perception might make a timing pattern in which each word is separately articulated particularly well suited for learners. Because lack of overlap at word boundaries makes word-final contrasts easier to hear, the preference to keep words separate is consistent with Weinberger's (1994a) recoverability principle—that is, the general tendency of NNSs to preserve contrastive information. That NNSs prefer epenthesis to deletion in difficult clusters—a pattern noted by Weinberger—may be a direct consequence of a pattern of articulation that prefers a lag between closures instead of an overlap.

Other studies of the transfer of cross-word assimilations from L1 to L2 have not found a word integrity effect, however. Altenberg and Vago (1983), Rubach (1984), and Solé (1997) all found L1 rules of voicing assimilation applying across word boundaries in learners' L2 English. Cebrian (2000, p. 20) suggested that these all involved the boundary between function words and content words (as in *this girl* pronounced as *thi*[z] *girl*). If function words cliticize onto content words, forming a larger phonological word in terms of prosodic structure (Nespor & Vogel, 1986), these assimilations might be reanalyzed as prosodically word internal. Kim and Zsiga (2002), however, found that Korean learners of English often applied the Korean rule of lenis stop voicing (a process that Jun, 1995, argued to be the result of articulatory reduction and overlap) in their English speech even at boundaries between two content words. Further research testing the validity of a word integrity effect is clearly called for.

Because word-final obstruents tend to have clear release bursts in native Russian articulation, theories of both transfer and word integrity predict that NSs of Russian will speak English with little overlap of consonants at word boundaries. The two theories make opposite predictions for NSs of English speaking Russian, however. Transfer of the English pattern would result in significant overlap of consonants; a preference for word integrity predicts no overlap. The experiment described in this paper explicitly tests these predictions.

This paper thus investigates the issue as to what happens to word-to-word articulatory timing when NSs of English and Russian produce the L2. Two questions in particular arise:

1. To what extent will the native articulatory timing patterns transfer to the L2?
2. Will any universal tendencies, such as a word integrity effect, appear in the interlanguage articulatory timing?

After a review of the literature on the different articulatory strategies of NSs of English and Russian and the acoustic consequences of these strategies, an acoustic experiment that addresses these questions will be described. Implications for further L2 research will also be noted.

BACKGROUND: ARTICULATORY OVERLAP IN NATIVE ENGLISH AND RUSSIAN

Patterns of Articulation

Many phonetic studies have shown that there is a substantial overlap between two consecutive consonant gestures at word boundaries in English (Barry, 1985, 1991; Browman & Goldstein, 1990; Byrd, 1992, 1996; Byrd & Tan, 1996; Catford, 1977; Hardcastle, 1985; Hardcastle & Roach, 1977; Zsiga, 1994, 1995a, 2000). These researchers have found that in a sequence of two consonants in English, in which the first consonant is word final and the second is word initial, movement of the articulators toward closure for the second consonant begins during the articulation of the first, often resulting in a period of time with two simultaneous closures. The pattern of articulatory overlap at word boundaries is very different in Russian. Articulation of Russian has been less extensively studied than articulation of English, but studies have demonstrated that in general there is a lag between consecutive consonants at word boundaries, not an overlap (Kochetov, 2001; Kochetov & Goldstein, 2001; Zsiga, 2000). The differences between English and Russian found in these studies confirm the generalization that phonetic timing does differ from language to language, and it is therefore part of the grammar that is acquired by the child and that the language student attempts to learn. (For more on phonetics as grammar, see Keating, 1988b, 1990; Liberman & Pierrehumbert, 1984; and Pierrehumbert, 1990.) Furthermore, the two different articulatory strategies lead to very different acoustic and perceptual consequences in the two languages.

In Russian, the lag between consecutive closures allows space for an audible release of the first consonant to be produced before the closure for the second consonant is formed. Descriptive accounts of Russian phonetics (e.g., Avenesov, 1984; Jones & Ward, 1969) and Russian language textbooks (e.g., Kostomarov, 1986) have noted the audible release of word-final consonants as an important aspect of Russian pronunciation. In an acoustic study, Kochetov (2001) found that word-final /t/ and /t^j/ were audibly released more than 90% of the time before word-initial /k/. Russian consonants are usually not released, however, in homorganic clusters (Kochetov; Zsiga, 2000).

English consonants generally do not have an audible release in any cluster (Catford, 1977; Ladefoged, 1993; Lass, 1984). Of course, both stop closures in a cluster must be released in the strict articulatory sense—that is, the contact between active and passive articulators that has occluded the flow of air must be broken. Without release in that sense, speech could not continue. In English, however, this release of contact for the first stop in a sequence often results in no perceptible release burst in the acoustic signal. Although Henderson and Repp (1982) found that a release burst could in fact be seen on an oscilloscope in 58% of the clusters they examined (in words such as *act*, *cactus*, or *pigpen*), their perception experiment revealed that even phonetically trained English-speaking listeners could not reliably hear these bursts. Lack of audible release bursts in clusters is directly related to the English pattern of extensive articulatory overlap. A release burst for the first consonant may be generated when the articulatory seal is broken, but it will be acoustically weak if a second closure in the mouth precludes the free flow of air. The release will not be audible at all if the second closure is made further forward in the mouth than the first, as in a /k#p/ sequence. The lip closure for the /p/ effectively "hides" any acoustic consequence of the release of the /k/ because sound generated at the velum will not propagate past the closed lips (Browman & Goldstein, 1990; Byrd, 1992; Maddieson & Ladefoged, 1989).

Perceptual Consequences in Russian: Preservation of Contrast

When there is a lag between consecutive consonant closures, as in Russian, a release burst for the first consonant is likely to be audible, regardless of the places of articulation of the consonants in sequence. Creating an audible release burst bestows a clear perceptual benefit by increasing the information that is available to the hearer for deciphering the signal. The duration, intensity, and frequency range of the release burst provide important acoustic cues to the identity of a consonant (Hume & Johnson, 2001; Liberman et al., 1967; Mattingly, 1981; Silverman, 1995; Stevens & Blumstein, 1978). Although cues are also provided by closure duration, presence or absence of voicing, and vocalic formant transitions into the closure, perceptual studies (Repp, 1978; Silverman) have shown them to be less salient than the cues provided at release. Steriade (1997) argued that the importance of the information in release bursts is seen in the crosslinguistic generalization that languages will often preserve place and laryngeal contrasts for single consonants in prevocalic position, where a release burst is necessarily heard, whereas they will neutralize contrasts in word-final position or in clusters, where a burst may be absent or unreliable.

Going against this crosslinguistic trend, Russian allows a full range of consonants in coda position, contrasting labials versus coronals versus dorsals, sonorants versus obstruents, and stops versus fricatives.[2] (Voiced obstruents are allowed in Russian codas, but voicing is not contrastive.) Additionally, in both coda and onset positions Russian also contrasts palatalized versus plain (or velarized) consonants, the former articulated with the tongue body high and fronted, the latter with the tongue body retracted—for example, /mat/ "foul language," /m^{j}at/ "crumpled (past part.)," and /mat^{j}/ "mother" (from Padgett, 2001, p. 190). The fact that coda consonants are released in Russian helps to provide the salient perceptual cues that listeners need to distinguish this wide range of contrasts. For example, Kochetov (2001) found that native Russian listeners identified the contrast between final /t/ and /t^{j}/ more reliably before /k/, where the coronal had an audible release burst, than before homorganic /n/ or /s/, where there was no release burst. In fact, the articulatory organization of the language appears to be structured so as to enhance perceptibility of coda contrasts.

Kochetov and Goldstein (2001), in an articulatory study using electromagnetography, and Zsiga (2000), in an acoustic study, independently found not only that there was in general a lag between consecutive consonant closures in Russian but also that this lag increased when the first consonant was articulated further back in the mouth than the second consonant. Russian speakers showed the greatest lag in back-to-front clusters (such as /k#p/) and the shortest lag in front-to-back clusters (such as /p#k/). Zsiga argued that Russian speakers adjust their patterns of articulation from cluster to cluster to ensure that the release of the first consonant in the cluster is heard. When a closure is made in the back of the mouth, as for /k/, the closing gesture for a following stop is delayed so that the release burst for the /k/ will not be blocked behind a closure further forward in the vocal tract. Release of a labial stop is more likely to be perceptible, regardless of whether there is a closure further back in the mouth, so more overlap can be allowed in /p#k/ clusters than in /k#p/ clusters.

Perceptual Consequences in English: Segmental Loss and Prosodic Gain

Unlike Russian speakers, English speakers do not adjust patterns of overlap according to the place of articulation of the sequenced consonants. Zsiga (2000) found that English speakers showed little variation in overlap from cluster to cluster. Sequences across different places of articulation were all produced with two simultaneous closures for about 20% of the total cluster duration.

The consequences of overlap between articulatory gestures have been investigated extensively within the theory of Articulatory Phonology (see Browman & Goldstein, 1992, for an overview). Browman and Goldstein (1990) specifically addressed assimilation and deletion of final alveolar consonants in English, examining phrases such as *perfect memory* [p^{h}ɛrfɛk mɛmri], *seven plus* [sɛvm plʌs], or *hundred pounds* [hʌndrɛb pʰaundz]. In most descriptions of English phonology (e.g., Gimson, 1962; Kenstowicz, 1993; Lass, 1984; Spencer, 1996), assimilations of word-final alveolars to a following word-initial stop are handled in terms of a feature-changing rule. Such a rule, for example, would apply to the /n/ in the phrase *seven plus*, changing its [coronal] feature specification to [labial]. In a constraint-based phonology, such as Prince and Smolensky (1993), the same feature change could be accomplished by high ranking of a markedness constraint against nonhomorganic nasal-stop sequences, coupled with a faithfulness constraint against changing the place of articulation of the onset stop (see Kager, 1999, for specific proposals). Whether through constraint ranking or rule, the result of the phonological feature switch in these approaches is that the speaker's mental representation of the final nasal consonant is [m] instead of [n], and instructions are sent to the articulators to produce the nasal consonant at the bilabial place of articulation rather than the alveolar. Similarly, in cases of deletion in clusters, the final alveolar is not represented in the phonological output, and the articulators receive no instruction to produce an alveolar closure.

Browman and Goldstein (1990) offered evidence that these phonological approaches do not correctly describe the data. They presented X-ray traces showing that even in utterances in which no alveolar consonant is heard, such as those previously listed, there is movement of the tongue tip toward the alveolar ridge, and an alveolar closure may still be made. In these cases, temporal overlap of the consonants at word boundaries masks any acoustic consequences of an alveolar closing gesture, such as formant transitions or audible release, and leads to the perception that the alveolar has been deleted or assimilated. Barry (1985, 1991, 1992) and Nolan (1992) also argued, on the basis of evidence from electropalatography, for the presence of a (sometimes weakened) tongue-tip gesture in phrases in which an alveolar has apparently been deleted. Again, the phonetic data show that the alveolar closure may still be made but that it is hidden behind the closure for a following consonant. Because the alveolar closing gesture can be shown to be phonetically present, if inaudible, in these phrases, phonological accounts that require the alveolar to be categorically deleted or changed cannot be correct.[3]

Thus, the claim of those working within the theory of Articulatory Phonology is that, in these and other cases of apparent deletion and assimilation, there is no phonological alternation. No change in phonological features has taken place; no segments have been deleted. The perception of deletion or assimilation follows from gestural overlap. Gestural overlap, in turn, follows from independently needed principles of articulatory coordination, which differ from language to language. As it happens, the degree of overlap particular to English allows many final contrasts that are preserved in the lexicon or in careful pronunciation to be lost in connected speech.

Although significant overlap of articulations at word boundaries leads to loss of segmental contrast, it also serves to indicate a strong degree of prosodic connection between the two words. Prosodic affiliation has been shown to have a significant effect on timing between consecutive segments (e.g., see Browman & Goldstein, 1989; Byrd, 1996, 2000; Byrd & Saltzman, 1998; Chitoran, 1998; Zsiga, 1997, 2000). Thus, principles of articulatory organization must make reference to specific prosodic categories, although researchers differ as to exactly what these categories should be and how closely they might correspond to the prosodic hierarchy of phonological literature (Inkelas & Zec, 1990; Kaisse, 1985; Nespor & Vogel, 1986; Selkirk, 1986). Browman and Goldstein, for example, argued explicitly for different timing relations between consonant sequences in English onsets and codas, but they found no evidence for the usefulness of other phonological hierarchical structures. Zsiga (2000) suggested that the extensive overlap between consonants at word boundaries seen in native English can be argued to signal the close relationship between two phonological words within a phonological phrase.

Accounting for Variation

In addition to variation depending on prosodic affiliation, variation in articulatory timing is also expected on the basis of rate (Byrd, 1996; Byrd & Tan, 1996; Gay, 1981; Ostry & Munhall, 1985; Zsiga, 1994), degree of formality (Anderson-Hsieh et al., 1994), and idiosyncrasies of individual speakers (Smith, 2000). These speaker- and situation-specific variations, however, have been found to occur within a language-particular range (Byrd; Smith). A theory of articulatory timing thus must account for consistent cross-language differences (as between English and Russian) and intralanguage variation.

The present study adopts *phonetic alignment constraints* (Zsiga, 2000) for the representation of articulatory timing. Phonetic alignment constraints synchronize salient points internal to the realization of articulatory gestures such as onset of movement, achievement of target position, or release of closure. The proposed timing relations are based on the gestural phasing principles suggested by Browman and Goldstein (1990, 1992, 2000) but use a formalism created by McCarthy and Prince (1993) to align edges of phonological and morphological constituents. For the specific cases of Russian and English, Zsiga suggested aligning points of closure and release. In American English, the point at which closure is achieved for the second consonant is timed to coincide with a point just before the articulatory release of the first, resulting in overlap of the two closures. The timing pattern typical of Russian could be expressed as aligning closure for the second consonant with a point just after the release of the first, resulting in a lag.

Variation within a range is accounted for by variable weighting of constraints (or different levels of *bonding strength* in the terminology of Browman & Goldstein, 2000). If a certain timing relation is given a relatively low weight, other exigencies, such as the need to produce an audible release or to speed up the rate of speech, may push the gestures out of perfect alignment, resulting in slightly more or less overlap. In more careful or formal speech, preservation of contrast may be more strongly weighted, resulting in less overlap than is usual for casual speech in English. The alignment constraint is never completely inactive, however, so the variation is limited to a range of values around the specified point. See Anttila (2001) for further discussion of constraint weighting to account for phonological variation, Zsiga (2000) as well as Browman and Goldstein for further discussion of phonetic alignment constraints, and Byrd (1996) for a different approach.

RESEARCH QUESTIONS

If proponents of the theory of Articulatory Phonology are correct in their characterization of connected speech assimilations and deletions as being the result of gestural overlap rather than a switch in phonological features, this changes the target for L2 speakers. If there is no phonological rule or constraint ranking in the grammar of L1 speakers that accounts for final stop assimilation or deletion, then there is no phonological rule or constraint ranking for L2 speakers to learn. Rather, what the L2 speakers need to learn is the target pattern of articulatory overlap—that is, the phonetic alignment constraints.

The same questions arise concerning the L2 acquisition of phonetic constraints as for the acquisition of phonological constraints: Do learners transfer the phonetic constraint ranking of the L1 to the L2, as Hancin-Bhatt and Bhatt (1997) have argued for phonological constraint rankings? Will universal preferences for unmarked linguistic structures result in a timing pattern typical of neither L1 (Broselow et al., 1998; Cebrian, 2000)? A better understanding of how languages differ in terms of their timing patterns, the acoustic consequences of these patterns of articulation, and how or whether such timing patterns transfer should lead to a clearer understanding of the phonetics and phonology of nonnative speech and of crosslinguistic differences and commonalities in general. To that end, an experiment was designed to test the following hypotheses:

1. Consistent with the phonetic studies previously cited, NSs of Russian are predicted to have less overlap of consecutive consonant articulations across word boundaries than NSs of English. Russian NSs will produce audible release bursts of word-final consonants more often than English NSs.
2. Again consistent with the phonetic studies previously cited, NSs of Russian will adjust their patterns of articulation to have less overlap in back-to-front clusters (such as /k#p/) than in front-to-back clusters (such as /p#k/). English NSs will use the same degree of overlap across all clusters.
3. Evidence for transfer will be seen if L2 learners exhibit patterns of articulatory coordination that are different from those of NSs in a direction that is consistent with the learners' L1 patterns. Russian NSs will release final consonants in their L2 English speech, whereas English NSs will overlap final consonants in their L2 Russian speech, producing cross-word deletions and assimilations.
4. Evidence for emergence of the unmarked will be seen if L2 speakers exhibit patterns of articulatory coordination that are different from those of NSs in a direc-

Table 1. Background information on participants

Subject	Sex	Age	L2 fluency (self-described)	Years of L2 instruction	Age at which L2 instruction began	Time spent in Russia/U.S.
English NSs						
EA	F	19	Beginner	1	17	None
EB	F	23	Intermediate	4.5	19	None
EC	F	23	Intermediate	5.5	17	4 months
ED	F	29	Fluent	8	18	1.5 years
EE	F	23	Fluent	9	13	None
EF	F	22	Fluent	8	14	3 months
Russian NSs						
RA	F	45	Beginner	1	44	1 year
RB	M	60	Intermediate	12	11	1 year
RC	F	38	Intermediate	12	12	1 year
RD	F	24	Fluent	8	14	1 year
RE	F	27	Fluent	14	7	6 years
RF	F	30	Fluent	12	18	9 years

tion that is *not* consistent with the learners' L1 patterns but *is* consistent with a default articulatory strategy that favors word integrity and the recoverability of acoustic cues. Both Russian and English NSs will release final consonants in their L2 speech and will not overlap consonants at word boundaries.

EXPERIMENT

Participants

Six NSs of Russian learning English and six NSs of American English learning Russian took part in the experiment. Participants reported the number of years they had studied the L2 and the age at which they began study. As objective measures of proficiency were not available, the participants rated their own L2 fluency. The English speakers were from the American Northeast or upper Midwest (except EF, who was from Hawai'i). They were students at Georgetown University at the time of the study: ED was a graduate student, and the other five were undergraduates. The Russian speakers, who were living in Washington, DC, at the time of the study, spoke the Moscow dialect. RE and RF were graduate students, RB held a diplomatic post at the Russian embassy, and RA, RC, and RD were members of diplomats' families (RA is the mother of RD). Five of them had had 8 or more years of English instruction in Russian schools. For RB and RC, however, their English classes occurred years in the past, and they described their competence in English as intermediate. Table 1 summarizes this information.

Materials

For both Russian and English, a set of two-word phrases was constructed such that the first word ended with a stop consonant (C1) and the second word began with a stop consonant (C2). For comparison to the C1#C2 clusters, stress-matched phrases were also constructed with single consonants, both word-final and word-initial, between vowels. There were no contexts in which American English speakers would produce flaps ([t] or [d] followed by an unstressed vowel). The phrases are listed in Appendix A.[4] Each phrase was incorporated into a sentence; the full list of sentences is provided in Appendix B. For presentation to the participants, a set of individual index cards was created on which each sentence was printed three times.

Twenty-six of the 30 phrases consist of a verb followed by its direct object. The other four phrases (those containing the Russian verb /gr^{j}ob/ "rowed") consist of a verb followed by a locative or descriptive phrase.[5] The final consonant for C1 was /p/, /b/, /d/, or /k/; the first consonant for C2 was /p/, /t/, or /k/. The voiced consonant /d/ was used for C1 instead of /t/ because many speakers of American English tend to substitute /ʔ/ for final /t/.[6]

In both Russian and English some degree of devoicing is found in clusters where an underlying voiced stop is followed by an underlying voiceless one. Devoicing in English is usually partial (Catford, 1977; Ladefoged, 1993), and sources disagree as to whether devoicing in Russian is a gradient phonetic effect or a complete phonological neutralization (Avanesov, 1984; Hayes, 1984; Isachenko, 1955; Jakobson, 1978; Shapiro, 1965; Wells, 1987). Conversely, in both languages, vocal-fold vibration generally extends for a few pitch periods from a vowel into the closure of a final voiceless consonant. For all NNSs in this study, the extent of (de-)voicing was highly variable. The extent and effect of devoicing was not directly considered in this study, however, and the symbols P, T, and K are used as cover terms for the labial, coronal, and dorsal stops, respectively, both as C1 and C2. (Note that Henderson & Repp, 1982, found no effect of voicing on the percentage of English clusters with audible release.)

Recording Procedures

The data were recorded in a quiet room, using a Sennheiser microphone and a Marantz portable tape recorder.[7] The participants were told that they were taking part in a study comparing Russian and English but were given no other details until after the recording session. Participants were given the set of sentence cards and had the opportunity to read through the sentences silently and ask the experimenter about any unfamiliar words or phrases. The cards were shuffled for each speaker, with the materials for Russian and English kept separate. The index cards for the phrases in Appendix A were, however, mixed together with phrases with different stress patterns and with a set of cards printed with sentences containing fricatives. The fricative materials and the effect of stress on the articulation of the consonant clusters are discussed in Zsiga (2000). After they had familiarized themselves with the materials, the participants were asked to repeat each sentence three times, reading "as naturally and smoothly as possible."[8] All participants read the sentences in their L1 first, then in their L2.

Analysis

The recordings were digitized at 22 kHz and analyzed using the Signalyze signal analysis software for the Macintosh. A total of 1,080 sentences were analyzed (12 participants × 2 languages × 15 sentences × 3 repetitions). To avoid skewing the data with disfluent tokens, any phrase in which there was a discernible stumble or pause between C1 and C2 (defined operationally as a period of silence of 350 ms or more) was excluded.[9] In the L1 productions, 13 tokens (2.4% of the total collected) were excluded on the basis of this criterion. In the L2 productions, 30 tokens (5.6%) were excluded. An additional four L2 tokens were excluded because of mispronunciations and an additional three L1 tokens because of extraneous noise in the signal. In five of these instances, a speaker had inordinate difficulty with a particular phrase such that all three repetitions had a pause or misarticulation at the crucial juncture. In these cases in which no fluent token of the phrase could be recorded for a particular subject, the empty cell was filled via mean interpolation for the statistical analysis.

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 between the two closures. The duration of each period of consonant closure and of each release burst was measured from the waveform. Duration ratio, which was computed for each phrase for each speaker, was then defined as the mean duration of the C1#C2 cluster (including both closures and the intervening release, if any) divided by the sum of the mean closure durations of C1 and C2 occurring intervocalically, as shown in equation 1.

$$\frac{M \text{ closure duration C1\#C2}}{(M \text{ closure duration C1\#V}) + (M \text{ closure duration V\#C2})} \qquad (1)$$

Duration ratio gives a measure of the amount of overlap between the two consonant articulations in a cluster such that the smaller the ratio, the greater the overlap. A duration ratio of 1 would indicate that the two consonants in the cluster are exactly sequenced: The duration of the cluster is exactly equal to the sum of the durations of C1 and C2 individually. Ratios less than 1 indicate that there is some overlap between the two closures: The duration of the cluster is less than the sum of its individual parts. A ratio of 0.5, for example, would mean that the two consonants in the cluster were simultaneous. Ratios greater than 1 indicate a lag between the two closures rather than overlap. Duration ratio is an indirect measure. However, if a sequence of two stops has no reliable change in voicing and no detectable internal release, as is generally the case in English and sometimes the case in Russian, there is no way to read directly from the acoustic record exactly how much each consonant contributes to the overall closure duration.

Certainly, experimental means for directly measuring articulatory coordination, including electropalatography and electromagnetic articulography, exist and have been used in previous studies (for discussion of these and other techniques, see Ball, Gracco, & Stone, 2001; Ong & Stone, 1998; Stone, 1996). However, given the drawbacks of direct articulatory measurement (e.g., physically intrusive devices, limits on the places of articulation that can be measured, and sometimes prohibitive expense), indirect techniques for inferring articulatory configurations from the acoustic record may also be found to be useful. The fact that the duration ratios reported here are for the most part consistent with direct articulatory measurements in the literature (see studies noted in the Results section) supports the validity of the calculation.[10]

The measures of duration ratio and percent released were submitted to (separate) repeated measures ANOVAs, with L1 as the between-group variable and language spoken and cluster type (homorganic, front-to-back, and back-to-front) as the within-group variables. A second analysis was performed to test for any effect of proficiency. Within each language context (NSs of English speaking English and Russian; NSs of Russian speaking Russian and English), a one-way ANOVA was conducted; dependent variables were percent released and duration ratio, and the independent variable was proficiency (fluent vs. nonfluent).

RESULTS

Hypothesis 1: Native English and Russian

Hypothesis 1 predicts that in their respective L1s English and Russian speakers will differ in duration ratio and percent released, with Russian NSs showing higher values on both measures. Results of the repeated measures ANOVAs, shown in Table 2, reveal a significant interaction of L1 and language spoken for both measures. This interaction is presented graphically in Figures 1 and 2.

As predicted, NSs of English speaking English have the lowest values on both measures: a duration ratio of 0.797 and 17.1% released. A duration ratio of approximately 0.80 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 Barry, 1991; Byrd, 1996; and Catford, 1977.) With this degree of overlap, the sequences rarely have an audible internal release. In contrast, NSs of Russian speaking Russian have (on average) a significantly higher duration ratio (0.982). Because there is less overlap, clusters in native Russian have an audible internal release more often (on average, 53.6% of the time).

Table 2. Results of the repeated measures ANOVAs

	Duration ratio		Percent released	
Variables	*df*	*F*	*df*	*F*
Between subjects				
Native language	1, 10	0.474	1, 10	4.808
Within subjects				
Language spoken	1, 10	7.401*	1, 10	32.330**
Cluster type	2, 20	4.338*	2, 20	33.691**
Interactions				
L1 × LgSp	1, 10	5.812*	1, 10	18.818**
L1 × ClTyp	2, 20	5.564*	2, 20	3.721*
LgSp × ClTyp	2, 20	0.884	2, 20	2.657
L1 × LgSp × ClTyp	2, 20	0.116	2, 20	0.462

Note. LgSp = language spoken, ClTyp = cluster type.
*$p < .05$. **$p < .001$.

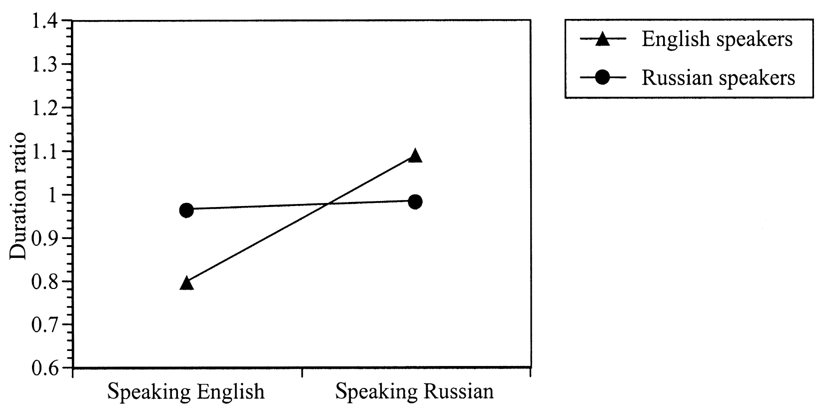

Figure 1. Mean duration ratios for the four language contexts.

Hypothesis 1 is confirmed. Consistent with previous research, the word-to-word articulatory timing patterns of native Russian and English are found to be different.[11] NSs of English speaking English have on average more overlap between consonant closures and fewer clusters with audible release than do NSs of Russian speaking Russian.

Hypothesis 3: Transfer of Timing Patterns

Hypothesis 3 predicts that speakers will carry over their timing patterns when they switch from L1 to L2. At this level of analysis, hypothesis 3 was only partially supported. When NSs of English speak Russian, both duration ratio and

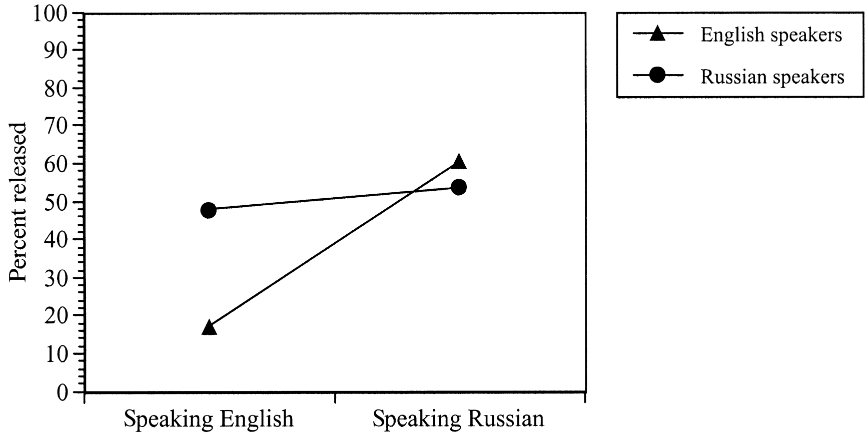

Figure 2. Mean percent released for the four language contexts.

percent released increase, jumping to 1.089 for duration ratio and to 60.6% released. The means for these NNSs are as large as the native Russian means on both measures (the difference between L2 English and L1 Russian speakers was not significant). The means for the Russian NSs, on the other hand, remain essentially unchanged when they switch to speaking English, showing a slight but nonsignificant drop in the English context to a duration ratio of 0.964 and 47.8% released (values that are still significantly higher than those of the English NSs). Considering that, across both groups of speakers the values on both measures are higher for spoken Russian than for spoken English, though only slightly so for the Russian NSs, this then accounts for the main effect of language spoken seen in Table 2.

The fact that NSs of Russian did not produce a nativelike pattern in L2 English, whereas NSs of English are seemingly able to produce a nativelike pattern in L2 Russian, is consistent with the hypothesis that the English articulatory pattern is more marked. According to the Markedness Differential Hypothesis (Eckman, 1977), the difficulty that a language learner encounters in producing some systematic pattern in the target language corresponds to the degree to which that pattern is crosslinguistically dispreferred. Furthermore, the finding that both groups of NNSs preferred on average an articulatory pattern with little overlap between consonant closures, and thus more clearly articulated final consonants, supports the word integrity constraint of Cebrian (2000) and the recoverability principle of Weinberger (1994a). The Discussion section returns to these issues.

The data for the English NSs are not consistent with hypothesis 3: They did not transfer their L1 articulatory coordination to the L2. The data for the Russian NSs are consistent with both hypothesis 3 (transfer) and hypothesis 4 (emergence of a default articulatory pattern). Overall means, however, hide

Table 3. Mean duration ratio (DR) and percent released (PR) for each cluster in each language context

		English NSs				Russian NSs			
		Speaking L1		Speaking L2		Speaking L1		Speaking L2	
Cluster type	Cluster	DR	PR	DR	PR	DR	PR	DR	PR
Homorganic	P#P	0.664	0.0	1.038	33.3	0.803	5.5	0.902	0.0
	T#T	0.904	0.0	1.005	11.0	0.923	5.5	0.871	22.2
	K#K	0.705	5.5	1.130	88.8	1.102	33.2	0.988	30.3
Back-to-front	T#P	0.742	0.0	1.088	44.5	1.086	55.5	1.060	22.2
	K#P	0.800	5.5	1.005	60.0	1.233	77.7	0.967	55.5
	K#T	0.806	11.2	1.185	83.3	1.083	66.7	1.108	69.5
Front-to-back	T#K	0.905	33.3	1.247	88.3	1.041	100.0	1.000	100.0
	P#T	0.862	58.3	1.046	55.5	0.766	77.7	0.852	63.8
	P#K	0.789	40.2	1.061	80.2	0.798	61.0	0.931	66.7

interesting differences that were found between different clusters. The statistical analysis showed both a significant main effect of cluster type and a significant interaction of cluster type and L1 for both measures. The differences between cluster types in the four language contexts bears on the question of transfer versus emergence of unmarked patterns.

Hypothesis 2: Cluster-to-Cluster Variation in L1

Hypothesis 2 predicts that NSs of Russian will adjust their patterns of articulation to have less overlap in back-to-front clusters (such as /k#p/) than in front-to-back clusters (such as /p#k/), whereas NSs of English will use the same degree of overlap across all clusters. This hypothesis was confirmed. The significant interaction of L1 and cluster type for both measures indicates that NSs of Russian and English do indeed time the different clusters differently. Table 3 shows the means for duration ratio and percent released for each individual cluster. Figures 3 and 4 graph the data by language context and cluster type. Data for NSs of English speaking English are graphed with filled triangles, data for NSs of Russian speaking Russian are graphed with filled circles, and data for the nonnative productions are graphed with open symbols and will be subsequently discussed.

As predicted by hypothesis 2, NSs of English speaking English show little variation in duration ratio as a function of cluster type (Figure 3). In this language context, duration ratio remains close to 0.8 for all three cluster types: 0.758 for homorganic, 0.783 for back-to-front, and 0.852 for front-to-back. In contrast, NSs of Russian speaking Russian vary duration ratio by cluster type, with back-to-front clusters having a much higher duration ratio (1.134) than homorganic (0.920) and front-to-back (0.868) clusters. Figure 4 graphs percent

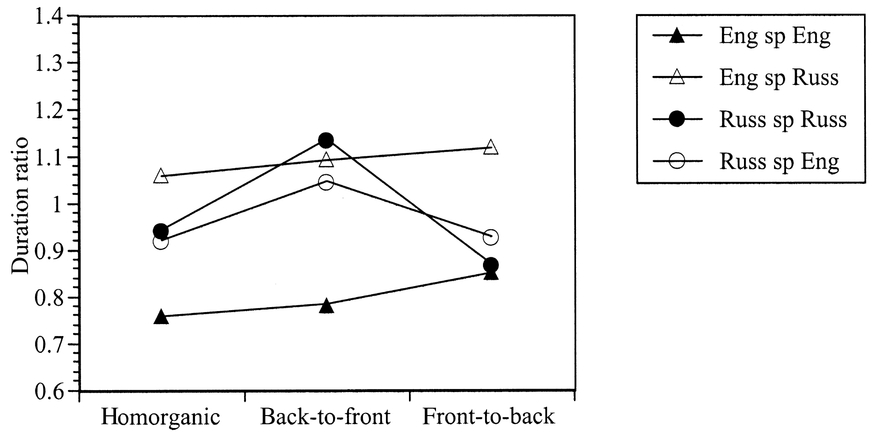

Figure 3. Mean duration ratios for each cluster type in each language context.

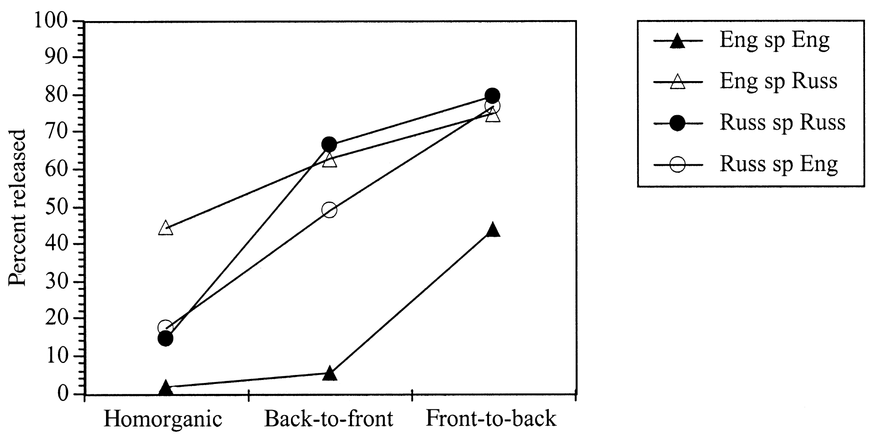

Figure 4. Mean percent released for each cluster type in each language context.

released and shows that this measure is affected both by language-specific, consonant-to-consonant timing and by the language-universal anatomy of the vocal tract.

Homorganic clusters very seldom have an audible release in either L1, consistent with descriptions in the literature (Catford, 1977; Kochetov, 2001; Ladefoged, 1993). When two consecutive consonants are made at the same place of articulation, any amount of overlap between them (duration ratio less than 1) will result in one long closure with no internal release. As illustrated in Table 3, almost all L1 homorganic clusters have mean duration ratios less than 1, so percent released in these clusters remains very low. The only cluster in the L1 productions that has a value for percent released greater than 6% is K#K in Russian (33%), which also has the highest duration ratio (1.102).

In the back-to-front clusters, the large difference in duration ratio between English and Russian in Figure 3 results in a large difference in percent released in Figure 4. For native English, where duration ratio remains less than 1, the closure for C1 is released only after the closure for C2 is made. Because C1 is further back in the mouth than C2, any burst noise generated by the release of C1 is blocked by C2, which renders the burst inaudible. NSs of Russian, however, implement a high duration ratio in these clusters, pushing the two closures far enough apart to allow an audible release burst in nearly 70% of cases—almost as often as in the front-to-back clusters.

The front-to-back clusters, where the release of C1 is never hidden behind the C2 closure regardless of the degree of overlap, are the only consonant sequences in which NSs of English speaking English produce an audible release more than 10% of the time. For example, although the K#P and P#K clusters have nearly identical duration ratios for English NSs, (0.800 and 0.789, respectively, as seen in Table 3), K#P has an audible release only 6% of the time, whereas P#K has an audible release 40% of the time. Front-to-back clusters also have the highest values for percent released for NSs of Russian speaking Russian, though for these speakers duration ratios are lowest in this context. This is the main effect of cluster type on percent released reported in Table 2: In all language contexts, regardless of the amount of overlap, front-to-back clusters favor an audible release.

Despite the fact that release of C1 is never articulatorily hidden in the front-to-back clusters, Figure 4 shows that an audible release burst is not generated 100% of the time even in this cluster type. Factors other than degree of overlap, such as a weakly articulated closure, may prevent the build-up of pressure necessary to generate a strong release burst, particularly in English. In their study of release and nonrelease in English consonant clusters, Henderson and Repp (1982) also found a lower than expected number of audible bursts when the first consonant in a sequence was a labial. They suggested that the dryness of the lips (as compared to surfaces inside the mouth) hinders the formation of a good seal. A weaker labial closure is also implicated in L1 Russian speech: Table 3 shows that heterorganic clusters in which C1 is P have a perceptible release on average 69% of the time, whereas T#K has an audible release in every single token.

In summary, hypothesis 2 is confirmed. In their L1 productions, NSs of English do not adjust overlap patterns to achieve an audible release; audible release occurs only when articulatory conditions happen to favor it. Russian NSs, on the other hand, increase the lag time between closures in back-to-front clusters, consistent with a strategy of reorganizing articulation to facilitate an audible burst. This release burst provides a strong cue to the separate places of articulation of the word-final and word-initial stops in heterorganic clusters. In homorganic clusters, no separate set of cues is necessary because C1 and C2 are identical; and in front-to-back clusters, an audible release burst is likely regardless of the amount of overlap.

Hypothesis 3 versus Hypothesis 4: Cluster-to-Cluster Variation in L2

Now that different patterns of cluster-to-cluster variation have been demonstrated in the L1 speech, the question as to whether these patterns transfer to L2 speech can be addressed. Hypotheses 3 and 4 make different predictions. Hypothesis 3 predicts transfer: NSs of Russian will speak English with the varied Russian pattern, and NSs of English will speak Russian with the flat, significantly overlapped English pattern. Hypothesis 4, on the other hand, predicts emergence of unmarked patterns in L2 speech, derived from a universally unmarked default, rather than transfer of a language-specific pattern. It has been previously suggested that the unmarked pattern favors separation between words (Cebrian, 2000) and clear phonetic cues to underlying contrasts (Weinberger, 1994a). For the present experiment, Figures 3 and 4 show that hypothesis 3 is supported for the Russian NSs, whereas hypothesis 4 is supported for the English NSs, which indicates that both transfer and linguistic universals have a role to play in L2 timing patterns.

When the Russian speakers switch from their L1 to L2 English, there is no significant change in either duration ratio or percent released. The patterns of articulation transfer from L1 to L2. Although some differences can be observed between the open and filled circles in Figures 3 and 4, these differences did not reach significance in this study: Table 2 shows no significant interaction of cluster type and language spoken (though this interaction approached significance for percent released). Further research with more varied tokens might reveal whether the slight flattening of the lines connecting the open circles in Figures 3 and 4 represents a real change in the direction of less variation in L2 speech or whether it is just noise in the data.

English speakers neither maintain their L1 pattern of significant articulatory overlap in their L2 Russian speech nor adopt the native Russian pattern of cluster-to-cluster variation. Rather, they increase duration ratio significantly across the board and keep timing relatively equal across all types of clusters. The result is that NSs of English speaking Russian have higher duration ratios in the homorganic and front-to-back clusters than do NSs of Russian despite the fact that mean values for the two language contexts (averaged across cluster type) are not significantly different (as illustrated in Figures 1 and 2). These differences result in an interesting reversal in the data on percent released. Speaking English with a Russian accent means producing too many audibly released final consonants in heterorganic clusters. Speaking Russian with an English accent means producing too many audibly released final consonants in homorganic clusters.

The articulatory pattern for NSs of English speaking Russian, different from both L1 and L2, is consistent with the emergence of a default pattern of articulation. The default duration ratio is approximately 1.00. Consecutive articulations are sequenced rather than overlapped, and there is no variation depending on place of articulation.[12] This sequenced pattern is consistent with both Cebrian's (2000) word integrity effect and Weinberger's (1994a) recoverability principle. Words are articulated without overlap, consistent with the word integrity effect, and clusters have an audible release more often than not, which facilitates the recoverability of place of articulation of final consonants.

Note, however, that the pattern exhibited by the Russian speakers (in both L1 and L2) is consistent with the recoverability principle but not the word integrity effect. Russian speakers continue to overlap word-final and word-initial consonants when doing so does not compromise perceptual cues. In both L1 and L2 productions, these speakers have low duration ratios in homorganic clusters (where cues to a separate place of articulation for the word-final consonant are not needed) and in front-to-back clusters (where audible release is likely to occur), cuing the place of articulation of the final consonant regardless of degree of overlap.

Proficiency

These effects of L1, language spoken, and cluster type are significant when participants are pooled. However, the NNSs included learners of varying proficiencies. Many studies (e.g., Bradlow, Akahane-Yamada, Pisoni, & Tohkura, 1999; Flege, 1995) have demonstrated that learners can and do improve aspects of their L2 pronunciation with time and appropriate instruction. Did proficiency have any effect on the patterns of articulation studied here?

For each language context (English NSs speaking English and Russian; Russian NSs speaking Russian and English) a one-way ANOVA was conducted; dependent variables were percent released and duration ratio, and the independent variable was proficiency (fluent vs. nonfluent). There was a significant effect in only one case: The proficiency of NSs of English speaking Russian had a significant effect on duration ratio, $F(1, 52) = 5.809$, $p < .05$. It was not the case, however, that the more advanced learners had more Russian-like speech.

To tease out the effect of proficiency, a second ANOVA was run for duration ratio of NSs of English speaking Russian, this time with subject number as the independent variable. This effect was also significant, $F(5, 48) = 6.470$, $p < .001$. A post hoc Student-Newman-Keuls test revealed that the significance was due entirely to one participant, EE, a fluent speaker of Russian. Her duration ratios were significantly different from all the other speakers', and no other significant differences were found. Interestingly, what makes EE stand out is not that her duration ratios in L2 Russian were more Russian-like but rather that they were more English-like. The other five English speakers showed a big increase in duration ratio from English to Russian (on average from 0.770 in English to 1.157 in Russian). For EE, on the other hand, duration ratio actually fell from 0.827 speaking English to 0.754 speaking Russian.[13]

Apart from EE, there were no significant differences within the group of English NSs or within the group of Russian NSs. That is not to say, of course, that other aspects of their pronunciation or proficiency, such as accuracy in reproducing segmental contrasts or the ability to carry on a conversation, were equivalent. Furthermore, none of the participants were true beginners. All were able to understand the (somewhat unusual) sentences presented to them and to produce them fairly fluently: Recall that only 5.6% of the nonnative tokens had to be excluded because of pauses or stumbles between words. In this small group of speakers, however, those who had had the most instruction in the L2 and who judged themselves most proficient did not produce the more nativelike timing patterns.[14]

To the ear of this author (a monolingual speaker of English, who is thus unable to judge the naturalness of the Russian language productions but who has been trained in phonetics), the English speech of the more advanced NSs of Russian was perfectly intelligible and very "fluent" in the sense that these participants spoke smoothly and easily, but their speech was definitely marked by a Russian accent. Part of the accent was due to a difference in segmental realizations, such as dental rather than alveolar /t/, and vowels not perfectly English-like in quality or duration. A large part of the accent, however, was due to the tendency to release final consonants where NSs English never would.

DISCUSSION

Articulatory and Acoustic Goals in Native English and Russian

The first conclusion one can draw from this study is that it supports previous findings that native English and Russian patterns of word-to-word articulation are different. Hypotheses 1 and 2 are confirmed. English NSs display a large amount of overlap between consonants and do not pull consonant closures apart in cases where doing so would facilitate an audible release. This acoustic goal is not important for English NSs. Russian NSs, on the other hand, adopt a pattern of coordination where overlap varies from cluster to cluster in a way that increases the likelihood of audible release. In back-to-front clusters (like /k#p/), Russian NSs increase the lag time between words, allowing time for the release of the first consonant to be heard before the second consonant creates a new vocal tract closure that will block the sound of the burst. In front-to-back clusters, the release burst for C1 is likely to be audible regardless of overlap because its sound will not be blocked by any other constriction, so in these sequences NSs of Russian can increase overlap to the point of values typical for English. The current study thus further emphasizes the language-specific nature of the phonetic grammar.

The Russian results also lend support to those theories of phonology and phonetics that incorporate acoustic goals (e.g., Flemming, 1995; Hume & Johnson, 2001; Steriade, 1997). The findings here are consistent with the hypothesis that Russian speakers adopt their strategy to maintain the acoustic cues that maximize lexical contrasts among the word-final stops. Such a hypothesis would explain the asymmetry in release rates between homorganic and heterorganic clusters in native Russian. In homorganic clusters, no separate cues to separate places of articulation are necessary, so the extra release burst is not produced.

In the English case, however, word-to-word timing is not organized so as to promote the preservation of lexical contrast. English NSs maintain place contrasts between word-final stops in the lexicon and in segmental articulation, only to obliterate them through gestural overlap. In terms of constraint rankings (or weightings), it seems that Russian ranks the constraint requiring maintenance of segmental contrasts in the coda very high, whereas English ranks it very low. For any constraint violation, however, there must be at least one countervailing, higher ranked constraint, which forces the violation of the lower ranked constraint (Prince & Smolensky, 1993). Whenever a contrast is lost, it must be for a reason.

One usual explanation is that the English pattern reflects “ease of articulation” (e.g., Lass, 1984). What counts as easy, however, is seldom quantified. Models for quantifying articulatory effort have been proposed (e.g., Kirchner, 2000; Lindblom, 1983), but in general these address the levels of force or muscular activation needed for specific articulations, not the question of which patterns of interarticulator coordination might be easiest to effect.

Rather than being particularly easy, it may be that the English pattern of overlap sacrifices segmental contrast in order to convey prosodic information. As previously discussed, phonetic research (e.g., Browman & Goldstein, 1989; Byrd, 1996, 2000; Byrd & Saltzman, 1998; Chitoran, 1998; Zsiga, 1997, 2000) has shown that specific timing relations differ according to the prosodic affiliations of the units being coordinated—that is, whether the articulatory gestures are within or cross the boundaries of prosodic categories such as segment, coda, syllable, phonological word, or phonological phrase. Thus, the extensive overlap between consonants at word boundaries seen here in native English may well serve to signal a close prosodic relationship between the two words. According to Selkirk (1986) and Rice (1990), in fact, all cases of assimilations across word boundaries occur when the two words are within a single phonological phrase. In the present experiment, however, syntactic and prosodic relations were kept constant, so the influence of prosodic structure could not be tested directly. (Zsiga, 2000, manipulated stress patterns, and thus foot structure, but the results were not clear.) Further crosslinguistic research is called for, examining how patterns of word-to-word overlap relate to prosodic structure, preservation of segmental contrasts, and sandhi phenomenon. One particularly interesting language to examine might be Korean. Kim and Jongman (1996) showed that Korean final stops are usually audibly released even though Korean neutralizes coda contrasts in voicing and continuancy and displays many cross-word assimilations.

Emergence of the Unmarked

In this study, five of the six NSs of English used a timing pattern in their L2 productions—a duration ratio of approximately 1.1 across all clusters—that was typical of neither English nor Russian. This pattern, in which words are sequenced rather than overlapped, is argued to represent an unmarked crosslinguistic default, and its emergence in the speech of English NSs learning Russian is taken to support hypothesis 4.

The fact that Russian NSs did not use the extensively overlapped English pattern is consistent (following Eckman, 1977) with the proposal that extensive consonant overlap is crosslinguistically a more marked pattern.[15] Interestingly, a situation similar to consonant overlap at word boundaries exists for English NSs learning Russian. The palatalized series of Russian consonants (/p^j/, /t^j/, /k^j/, /s^j/, etc.) is articulated with a secondary palatal constriction of the tongue blade made simultaneously with the primary closure at the lips, teeth, or velum (Keating, 1988a). These palatalized consonants are very difficult for English speakers to learn to produce correctly: There is a strong tendency for English learners of Russian to produce them as consonant-glide sequences (see Zsiga, 1995b).

By allowing the release of word-final consonants to be audible and in general keeping the acoustic cues to the two closures separate, the sequential implementation of consonant closures furthers the recoverability of underlying contrasts. The results of this experiment thus add further support to Weinberger's (1994a) recoverability principle: NNSs prefer configurations in which contrastive information is maintained. As was previously noted, Weinberger's finding that NNSs prefer epenthesis to deletion in difficult clusters may in fact be a direct consequence of a pattern of articulation that prefers a lag between closures rather than overlap. The unexpected release burst may be interpreted by native listeners as insertion of an extra vowel (Tajima et al., 1997). More acoustic research is necessary to determine whether the so-called epenthetic vowels in nonnative speech are in fact phonological insertions or are perhaps more accurately described as nothing more than audible consonant release.

Although this study found ample evidence for Weinberger's (1994a) recoverability principle, the word integrity effect of Cebrian (2000) was less strongly supported. The speech of both English and Russian L2 learners was consistent with the recoverability principle: All learners preferred an articulatory organization that facilitated the audible release of word-final stops in heterorganic clusters, better preserving the underlying place contrasts. Whereas productions of NSs of English speaking Russian (duration ratio greater than 1) were consistent with the word integrity effect, the productions of NSs of Russian speaking English were not. In their L2 speech, Russian learners did not demonstrate a word integrity effect, producing duration ratios less than 1 in homorganic and front-to-back clusters where recoverability was not an issue.

Transfer of Timing Patterns

Whereas NSs of English tended to use a default timing pattern in their L2 speech (consistent with hypothesis 4), NSs of Russian transferred their L1 articulatory timing patterns (consistent with hypothesis 3). Thus, this study supports a role for both transfer and emergence of the unmarked in patterns of L2 articulation.

Only one participant, an English NS fluent in Russian, tended to reproduce the English pattern of extensive overlap in her Russian speech. It may be that the experimental condition of reading sentences predisposed even the fluent speakers to more careful articulation and thus to less overlap. In a more relaxed and conversational task, with the attention of the speakers diverted from their articulation per se (Anderson-Hsieh et al., 1994), more English-like patterns might emerge in L2 speech. It was not the case, however, that the nonnative speech in this experiment was choppy and disfluent. Very few utterances had stumbles or pauses, and these were not included in the statistical analysis. Furthermore, the Russian speakers did produce word-to-word overlap in their L2 speech in those clusters where such overlap was consistent with both the L1 pattern and with recoverability of contrasts.

In this study the asymmetry in transfer effects between English and Russian might be attributed to the recoverability principle: Only patterns that were consistent with recoverability were carried over from L1 to L2. Other studies, however, have found that contrast-neutralizing, cross-word processes of assimilation and deletion can and do transfer (Altenberg & Vago, 1983; Kim & Zsiga, 2002; Rubach, 1984; Solé, 1997). The finding that the recoverability principle is active but violable suggests a solution in terms of constraint ranking or weighting.

Weinberger (1994a) suggested that language learners will tend to rank recoverability high. For Russian NSs, this ranking presents no problem, as recoverability does not conflict with their language-specific phonetic alignment constraints. The native pattern transfers from L1 to L2, and no violations of either the language-specific constraints or of recoverability are incurred. In native English, the language-specific alignment (specifying overlap and consistent with conveying prosodic information) must outrank segmental recoverability, as word-final contrasts are lost in native English. When English NSs switch to an L2, however, the data here suggest that they change their constraint ranking, promoting recoverability over language-specific alignment. When the language-specific native pattern is ruled out, a default pattern consistent with recoverability is used. (The one English NS who did transfer English timing to Russian speech could be argued to have reverted to the original English ranking.) Languages that exhibit transfer of cross-word assimilations from L1 to L2 (e.g., Korean) can be argued to rank cross-word alignment more highly than recoverability in both L1 and L2 speech.

Interestingly, even for English NSs, transfer from L1 to L2 occurs with other articulatory timing patterns—those that govern timing within smaller phonological domains. Intrasegmental timing, such as the coordination of glottal opening with initial consonants that causes English NSs to aspirate initial /p/ in French, or intrasyllabic timing, such as the coordination between velum and oral articulators that causes vowel nasalization in CVn sequences, seems to be very hard to unlearn. This asymmetry is consistent with the hypothesis (Browman & Goldstein, 2000; Zsiga, 2000) that the articulatory constraints linking gestures within smaller constituents are universally more highly ranked (or more strongly weighted) than those linking gestures across word boundaries. When NSs of English switch to an L2, they promote recoverability above the constraints on phonetic alignment across word boundaries but not above constraints on phonetic alignment within words. This ranking—within-word timing constraints above recoverability above cross-word timing constraints—would also account for Cebrian's (2000) findings for Catalan-English interlanguage and for the appearance of a word integrity effect. Under this constraint ranking, within-word articulatory timing patterns will transfer, but cross-word patterns will not. If the hypothesis that smaller prosodic units always display tighter phasing than larger units holds up crosslinguistically, that leads to the prediction that no language learner will exhibit an "anti-word-integrity" preference—a pattern in which cross-word timing transfers from L1 to L2, but intraword timing does not.

Modeling articulatory timing in terms of phonetic alignment constraints provides a framework in which to build a typology of transfer and emergence of the unmarked in L2 articulation. It may also lead to hypotheses about when and how transfer will occur. It is hoped that further study on how variably weighted timing constraints fare in the speech of language learners will shed light on the cognitive representations of articulatory timing patterns both for NNSs and NSs.

CONCLUSION

This study has offered some preliminary evidence from English and Russian on patterns of L2 articulation. The importance of maintaining acoustic cues to underlying contrasts was confirmed in L2 speech and in native Russian, as was the fact that these cues tend to be lost in native English, perhaps being sacrificed to the goal of conveying prosodic information. Some evidence for both transfer of timing patterns and the emergence of an unmarked pattern of articulation was found, and a typology of possible transfer effects, resulting from variable ranking of phonetic alignment constraints and constraints on recoverability, was proposed.

This small study examined only a few speakers of two languages in an experimental setting. More research—more speakers, more languages, and different tasks in different settings—is clearly needed. It is hoped, however, that this study has highlighted the importance of studying patterns of articulation in L2 acquisition and that it may in fact spur further study in this area.

(*Received 19 September 2002*)

NOTES

1. Cebrian (2000, p. 18) argued for a definite role for transfer in devoicing, in addition to a preference for unmarked linguistic structure, because devoicing in Catalan-English interlanguage targeted particular classes of obstruents differentially, following the pattern in native Catalan speech.

2. Both Russian and English also allow a restricted set of consonant clusters in syllable codas. Articulatory timing of tautosyllabic consonants will not be a focus here, although it should not be assumed that timing within syllables will be the same as timing across syllable boundaries. The relation between prosodic categories and articulatory timing will be subsequently discussed.

3. The claim that phonological feature alternations cannot account for these English data does not, of course, entail that categorical feature changes might not be the best way of expressing other connected speech alternations, although it does suggest that many proposed phonological rules deserve further phonetic scrutiny. Furthermore, it does not follow that phonological organization, in terms of prosodic categories, has no role to play in the production of these English phrases. It will in fact be argued that prosodic affiliations largely determine the particular timing relations that are chosen. The point here is not that phonetic explanation should completely supplant phonological explanation, just that in many cases that have been discussed in the literature the feature-changing explanation is not the correct one.

4. Due to an error, the sentences containing the phrases /pʲok 'persik/ and /pʲok 'kaʃu/ were inadvertently excluded from the set of sentence cards for EA, ED, EE, RA, RC, and RD. For these speakers, the phrases /pʲok pe'tʃene/ and /pʲok kala'tʃi/ were substituted and were paired with the stress-matched phrases /e'dʲa pe'tʃene/, /e'dʲa kala'tʃi/, and /pʲok o'ladi/ for the computation of duration ratio.

5. Although every effort was made to control for syntactic and prosodic context, very few declined verbs in Russian end in stop consonants, and no examples of a verb-object phrase that contained a labial-final verb and that met the other phonetic conditions of the study could be devised. An anonymous *SSLA* reviewer pointed out that this difference in syntactic context may have introduced a confounding factor into the experiment, and this possibility cannot be ruled out. Unfortunately, the Russian lexicon offered no better alternative. As it turned out, however (see the Results section), the phrases with labial-final verbs showed the largest amount of overlap of the Russian tokens. This result is consistent with the effect of place of articulation found by Kochetov and Goldstein (2001) but is not consistent with an effect of syntactic or prosodic difference. If there were a stronger prosodic boundary between verb and locative phrase than between verb and direct object, there would be less overlap between C1 and C2 in the locative phrases, exactly the opposite of what was found here.

6. Some speakers of American English tend to add some degree of glottal constriction to all final stops, but this was not noticed to be the case for any of the speakers recorded for this study. Furthermore, none of the speakers substituted /ʔ/ for devoiced Russian /d/.

7. Data for 10 of the 12 participants was collected by Stefan Kaufmann, a research assistant fluent in both Russian and English. Kaufmann gave the participants instructions in their L1. EF and RF were recorded by the author, with instructions given in English.

8. Such an instruction cannot, of course, erase the inherent differences between a reading task and casual conversation. For the purposes of this experiment, in a trade-off inherent in all linguistic phonetic analyses, a degree of informality was sacrificed for the sake of phonetic consistency and a controlled design that would make statistical analysis possible. It is hoped that further research will extend the baseline measurements made here to naturally occurring conversations.

9. An anonymous *SSLA* reviewer noted that, if removing outliers adds any bias to the data, it would be in the direction of making NNSs, who had slightly more disfluencies, more like the NSs. Therefore, any differences between NSs and NNSs are not likely to be due to the elimination of these outliers.

10. Duration of individual consonant closures can be measured in the intervocalic contexts and in clusters with an intervening internal release (236 tokens). In these contexts, pooling across languages, consonant durations were consistent with those reported in the literature on General American phonetics (e.g., Klatt, 1976; Oller, 1973). In word-initial position, the duration of P, T, and K are 104 ms, 87 ms, and 81 ms, respectively. For final consonants, the duration of P, T, and K are 78 ms, 59 ms, and 59 ms, respectively. A two-way ANOVA was conducted on these tokens in which the dependent variable was duration and the independent variables were place of articulation of the stop itself (P, T, or K) and preceding (or following) segment (P, T, K, or vowel). There was no significant main effect of preceding or following segment: Whether the next word began with P, T, K, or a vowel did not significantly affect the duration of a word-final stop, $F(3, 406) = 1.204$, $p = .3079$; and

whether the preceding word ended with P, T, K, or a vowel did not significantly affect the duration of a word-initial stop, $F(3, 426) = 0.885$, $p = .4487$.

11. These results are consistent with but not identical to those presented in Zsiga (2000), which analyzed an overlapping dataset (NS productions of 9 of the 12 participants included here with different stress conditions). Zsiga found a duration ratio of 0.802 and 18% released for NSs of English speaking English and a duration ratio of 0.977 and 47% released for NSs of Russian speaking Russian.

12. Lack of consonant-specific variation in both the L1 and L2 pronunciations of the English NSs might be interpreted as transfer of relative, if not absolute, timing. In the absence of more crosslinguistic data, it is impossible to be certain whether this lack of variation represents a crosslinguistic default or a particularly English configuration. However, treating all clusters alike represents the simplest, most general articulatory strategy and thus argues for the former.

It is also possible that the pattern of across-the-board release in the Russian speech of English NSs might be attributed to hypercorrection, not recoverability. That is, the learners might have been overgeneralizing the Russian pattern of word-final release from heterorganic to homorganic clusters in an attempt to sound more Russian. However, this hypothesis is not consistent with the learners' perceptions of their own speech. After having completed the experiment, participants were asked about the differences they perceived between English and Russian pronunciation, whether they perceived themselves as speaking the L2 with a foreign (Russian or English) accent, and if so whether there was any aspect of their speech they thought sounded particularly nonnative. None mentioned release or nonrelease of final consonants as an aspect of pronunciation to which they had paid particular attention.

13. One can only speculate as to why this speaker was different from the others. It may be that she was proficient enough in Russian that she had ceased to pay careful attention to her articulation and reverted to the pattern typical of English connected speech.

14. The present experiment was not designed to test whether explicit instruction can help learners to change cross-word timing patterns or to investigate what sort of instruction would be most useful. The subjects were all exposed to different classroom experiences, and it is not clear how much or what kind of pronunciation training they received. The learners of Russian had all received some instruction in the Department of Slavic Languages at the same university but from different teachers and at different levels. The NSs of Russian reported in postexperiment discussions that they were sometimes drilled in their advanced English classes on segmental articulation, but they did not receive instruction not to release final consonants. Russian texts and pronunciation drills for English speakers likewise emphasize vowel quality and consonant place of articulation (as well as secondary articulations). Although the college textbook that had been used by the learners of Russian (Kostomarov, 1986) mentioned that final consonants are usually released in Russian, no distinction was made in the text between patterns of release in homorganic and heterorganic clusters, and participants reported that this aspect of pronunciation was not emphasized in their language classes.

15. An anonymous *SSLA* reviewer suggested that the asymmetry in transfer effects might have been due to social factors, not a difference in markedness. It is suggested that the Russian NSs may have retained aspects of their L1 pronunciation to maintain a degree of Russian identity despite immersion in an English-speaking environment. Although no in-depth research into the participants' language attitudes or social networks was conducted, this hypothesis is not consistent with the attitudes the speakers reported in the postexperiment discussions. The participants indicated that they were aware that they spoke with an accent but that they worked very hard to "lose" their accents. When one Russian speaker, who considered herself fluent in English, was shown how her pronunciation of *make parts* (with released /k/) was different from a NS's pronunciation (no release), her response was "Oh, my English teacher will be so disappointed! I'll have to work harder on that!"

REFERENCES

Altenberg, E., & Vago, R. (1983). Theoretical implications of an error analysis of second language phonology production. *Language Learning, 33*, 427–448.

Anderson-Hsieh, J., Johnson, R., & Koehler, K. (1992). The relationship between native speaker judgments of non-native pronunciation and deviance in segmentals, prosody, and syllable structure. *Language Learning, 42*, 529–555.

Anderson-Hsieh, J., Riney, T., & Koehler, K. (1994). Connected speech modifications in the English of Japanese ESL learners. *Issues and Developments in English and Applied Linguistics, 7*, 31–51.

Anttila, A. (2001). Variation and phonological theory. In J. K. Chambers, P. Trudgill, & N. Schilling-Estes (Eds.), *The handbook of variation and change* (pp. 206–243). Oxford: Blackwell.

Archibald, J. (1993). *Language learnability and L2 phonology: The acquisition of metrical parameters*. Dordrecht: Kluwer.

Avanesov, R. I. (1984). *Russkoye literaturnoye proiznoshenie* [Russian literary pronunciation]. Moscow: Prosveshchenie.

Ball, M., Gracco, V., & Stone, M. (2001). A comparison of imaging techniques for the investigation of normal and disordered speech production. *Advances in Speech-Language Pathology, 3*, 13–24.

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 22nd Congress of Phonetic Sciences, 4*, 14–17.

Barry, M. (1992). Palatalisation, assimilation, and gestural weakening. *Speech Communication, 11*, 393–400.

Bradlow, A. R., Akahane-Yamada, R., Pisoni, D. B., & Tohkura, Y. (1999). Training Japanese listeners to identify English /r/ and /l/: Long-term retention of learning in perception and production. *Perception & Psychophysics, 61*, 977–985.

Broselow, E. (1987). An investigation of transfer in second language phonology. In G. Ioup & S. Weinberger (Eds.), *Interlanguage phonology: The acquisition of a second language sound system* (pp. 261–278). Rowley, MA: Newbury House.

Broselow, E., Chen, S.-I., & Wang, C. (1998). The emergence of the unmarked in second language phonology. *Studies in Second Language Acquisition, 20*, 261–280.

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 J. Kingston & M. Beckman (Eds.), *Papers in laboratory phonology I: Between the grammar and physics of speech* (pp. 341–376). New York: Cambridge University Press.

Browman, C. P., & Goldstein, L. (1992). Articulatory phonology: An overview. *Phonetica, 49*, 155–180.

Browman, C. P., & Goldstein, L. (2000). Competing constraints on intergestural coordination and self-organization of phonological structures. *Bulletin de la Communication Parlée, 5*, 25–34.

Brown, G. (1977). *Listening to spoken English*. London: Longman.

Byrd, D. (1992). Perception of assimilation in consonant clusters: A gestural model. *Phonetica, 49*, 1–24.

Byrd, D. (1996). Influences on articulatory timing in consonant sequences. *Journal of Phonetics, 24*, 209–224.

Byrd, D. (2000). Articulatory vowel lengthening and coordination at phrasal junctures. *Phonetica, 57*, 3–16.

Byrd, D., & Saltzman, E. (1998). Intragestural dynamics of multiple prosodic boundaries. *Journal of Phonetics, 26*, 173–199.

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.

Cebrian, J. (2000). Transferability and productivity of L1 rules in Catalan-English interlanguage. *Studies in Second Language Acquisition, 22*, 1–26.

Chen, Y. (1982). Principles and methods of the teaching of pronunciation. *Language Learning and Communication, 2*, 75–76.

Chitoran, I. (1998). Georgian harmonic clusters: Phonetic cues to phonological representation. *Phonology, 15*, 121–141.

Dickerson, W. (1989). *Stress in the stream of speech: The rhythm of spoken English*. Urbana: University of Illinois Press.

Eckman, F. (1977). Markedness and the contrastive analysis hypothesis. *Language Learning, 27*, 315–330.

Eckman, F. (1981). On the naturalness of interlanguage phonological rules. *Language Learning, 31*, 195–216.

Eckman, F. (1987). The reduction of word-final consonant clusters in interlanguage. In A. James & J. Leather (Eds.), *Sound patterns in second language acquisition* (pp. 143–162). Dordrecht: Foris.

Epstein, S., Flynn, S., & Martohardjono, G. (1996). Second-language acquisition: Theoretical and empirical issues in contemporary research. *Behavioral and Brain Sciences, 19*, 677–758.

Flege, J. E. (1992). Speech learning in a second language. In C. Ferguson, L. Menn, & C. Stoel-Gammon (Eds.), *Phonological development: Models, research, implications* (pp. 565–604). Timonium, MD: York Press.

Flege, J. E. (1995). Second language speech learning: Theory, findings, and problems. In W. Strange (Ed.), *Speech perception and linguistic experience: Issues in cross-language research* (pp. 233–272). Timonium, MD: York Press.

Flege, J. E., Munro, M. J., & MacKay, I. R. A. (1995). Factors affecting strength of perceived foreign accent in a second language. *Journal of the Acoustical Society of America, 97*, 3125–3134.

Flemming, E. (1995). *Auditory representations in phonology*. Unpublished doctoral dissertation, University of California, Los Angeles.

Gay, T. (1981). Mechanisms in the control of speech rate. *Phonetica, 38*, 148–158.

Gilbert, J. (1993). *Clear speech: Pronunciation and listening comprehension in North American English*. New York: Cambridge University Press.

Gimson, A. C. (1962). *An introduction to the pronunciation of English*. London: Edward Arnold.

Hammarberg, B. (1990). Conditions on transfer in phonology. In A. James & J. Leather (Eds.), *New sounds 90: Proceedings of the 1990 Symposium on the Acquisition of Second-Language Speech* (pp. 198–215). Dordrecht: Foris.

Hancin-Bhatt, B., & Bhatt, R. (1997). Optimal L2 syllables: Interactions of transfer and developmental effects. *Studies in Second Language Acquisition, 19*, 331–378.

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 sequences. In H. Hollien & P. Hollien (Eds.), *Current issues in the phonetic sciences* (pp. 531–540). Amsterdam: Benjamins.

Hayes, B. (1984). The phonetics and phonology of Russian voicing assimilation. In M. Aronoff & R. T. Oehrle (Eds.), *Language sound structure* (pp. 318–328). Cambridge, MA: MIT Press.

Henderson, J. A., & Repp, B. (1982). Is a stop consonant released when followed by another stop consonant? *Phonetica, 39*, 71–82.

Hieke, A. (1987). Absorption and fluency in native and non-native casual English speech. In A. James & J. Leather (Eds.), *Sound patterns in second language acquisition* (pp. 41–58). Dordrecht: Foris.

Hume, E., & Johnson, K. (Eds.). (2001). *The role of speech perception in phonology*. San Diego, CA: Academic Press.

Inkelas, S., & Zec, D. (Eds.). (1990). *The phonology-syntax connection*. Chicago: University of Chicago Press.

Isachenko, A. V. (1955). [Review of the book *Russiche lautlehre*]. *Zeitschrift für Phonetik, 8*, 415.

Jakobson, R. (1978). Mutual assimilation of Russian voiced and voiceless consonants. *Studia Linguistica, 32*, 107–110.

James, A. R. (1988). *The acquisition of a second-language phonology: A linguistic theory of developing sound structures*. Tübingen: Gunter Narr.

Jones, D., & Ward, D. (1969). *The phonetics of Russian*. New York: Cambridge University Press.

Jun, S.-A. (1995). Asymmetrical prosodic effects on the laryngeal gesture in Korean. In B. Connell & A. Arvaniti (Eds.), *Papers in laboratory phonology IV: Phonology and phonetic evidence* (pp. 235–253). New York: Cambridge University Press.

Kager, R. (1999). *Optimality theory*. New York: Cambridge University Press.

Kaisse, E. (1985). *Connected speech*. San Diego, CA: Academic Press.

Keating, P. A. (1988a). Palatals as complex segments: X-ray evidence. *UCLA Working Papers in Phonetics, 69*, 77–91.

Keating, P. A. (1988b). The phonology-phonetics interface. In F. J. Newmeyer (Ed.), *Linguistics: The Cambridge survey: Vol. 1. Linguistic theory—Foundations* (pp. 281–302). New York: Cambridge University Press.

Keating, P. A. (1990). Phonetic representations in a generative grammar. *Journal of Phonetics, 18*, 321–334.

Kenstowicz, M. (1993). *Phonology in generative grammar*. Oxford: Blackwell.

Kim, H., & Jongman, A. (1996). Acoustic and perceptual evidence for complete neutralization of manner of articulation in Korean. *Journal of Phonetics, 24*, 295–312.

Kim, H.-K., & Zsiga, E. (2002, June). *Assimilation and neutralization in Korean/English interlanguage*. Paper presented at the 8th Conference on Laboratory Phonology, Yale University, New Haven, CT.

Kirchner, R. (2000). Geminate inalterability and lenition. *Language, 76*, 509–545.

Klatt, D. (1976). Linguistic uses of segmental duration in English: Acoustic and perceptual evidence. *Journal of the Acoustical Society of America, 59*, 1208–1221.

Kochetov, A. (2001, January). *Testing licensing by cue: An experimental study of Russian palatalization*. Poster presented at the 75th meeting of the Linguistic Society of America, Washington, DC.
Kochetov, A., & Goldstein, L. (2001, January). *Competing recoverability factors and inter-gestural phasing in Russian stop clusters*. Paper presented at the 75th meeting of the Linguistic Society of America, Washington, DC.
Kostomarov, V. (Ed.). (1986). *Russian for everybody*. Moscow: Russky Yazyk.
Ladefoged, P. (1993). *A course in phonetics* (3rd ed.). New York: Harcourt, Brace, Jovanovich.
Lado, R. (1957). *Linguistics across cultures*. Ann Arbor: University of Michigan Press.
Lass, R. (1984). *Phonology*. New York: Cambridge University Press.
Liberman, A. M., Cooper, F. S., Shankweiler, D., & Studdert-Kennedy, M. (1967). Perception of the speech code. *Psychological Review, 74*, 431–461.
Liberman, M., & Pierrehumbert, J. (1984). Intonational invariance under changes in pitch range and length. In M. Aronoff & R. T. Oehrle (Eds.), *Language sound structure* (pp. 157–233). Cambridge, MA: MIT Press.
Lindblom, B. (1983). Economy of speech gestures. In P. MacNeilage (Ed.), *Speech production* (pp. 217–246). New York: Springer Verlag.
Maddieson, I., & Ladefoged, P. (1989). Multiply-articulated segments and the feature hierarchy. *UCLA Working Papers in Phonetics, 72*, 116–138.
Mattingly, I. G. (1981). Phonetic representation and speech synthesis by rule. In T. Myers, J. Laver, & J. Anderson (Eds.), *Proceedings of the 13th International Congress of Phonetic Sciences* (Vol. 4, pp. 574–577). Stockholm: Royal Institute of Technology and Stockholm University.
McCarthy, J. J., & Prince, A. (1993). Generalized alignment. In G. Booij & J. van Merle (Eds.), *Yearbook of morphology 1993* (pp. 79–153). Dordrecht: Kluwer.
McNerney, M., & Mendelsohn, D. (1992). Suprasegmentals in the pronunciation class: Setting priorities. In P. Avery & S. Ehrlich (Eds.), *Teaching American English pronunciation* (pp. 185–196). Oxford: Oxford University Press.
Nespor, M., & Vogel, I. (1986). *Prosodic phonology*. Dordrecht: Foris.
Nolan, F. (1992). The descriptive role of segments: Evidence from assimilation. In G. J. Docherty & D. R. Ladd (Eds.), *Papers in laboratory phonology II: Gesture, segment, prosody* (pp. 261–280). New York: Cambridge University Press.
Oller, D. K. (1973). The effect of position in utterance on speech segment duration in English. *Journal of the Acoustical Society of America, 54*, 1235–1247.
Ong, D., & Stone, M. (1998). Three-dimensional vocal tract shapes in /r/ and /l/: A study of MRI, ultrasound, electropalatography, and acoustics. *Phonoscope, 1*, 1–13.
Ostry, D. J., & Munhall, K. G. (1985). Control of rate and duration of speech movements. *Journal of the Acoustical Society of America, 77*, 640–648.
Padgett, J. (2001). Contrast dispersion and Russian palatalization. In E. V. Hume & K. Johnson (Eds.), *The role of speech perception in phonology* (pp. 187–218). San Diego, CA: Academic Press.
Pennington, M. (1989). Teaching pronunciation from the top down. *RELC Journal, 20*, 20–38.
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*. Unpublished manuscript, Rutgers University, Newark, NJ, and University of Colorado, Boulder.
Repp, B. (1978). Perceptual integration and differentiation of spectral cues for intervocalic stop consonants. *Perception & Psychophysics, 24*, 471–485.
Rice, K. (1990). Predicting rule domains in the phrasal phonology. In S. Inkelas & D. Zec (Eds.), *The phonology-syntax connection* (pp. 289–312). Chicago: University of Chicago Press.
Rubach, J. (1984). Rule typology and phonological interference. In S. Eliasson (Ed.), *Theoretical issues in contrastive phonology: Studies in descriptive linguistics* (pp. 37–50). Heidelberg, Germany: Julius Groos.
Selkirk, E. (1986). On derived domains in sentence phonology. *Phonology Yearbook, 3*, 371–405.
Shapiro, M. (1965). On non-distinctive voicing in Russian. *Journal of Linguistics, 10*, 189–194.
Silverman, D. (1995). *Phasing and recoverability*. Unpublished doctoral dissertation, University of California, Los Angeles.
Smith, B. (2000). Variations in temporal patterns of speech production among speakers of English. *Journal of the Acoustical Society of America, 108*, 2438–2441.
Solé, M. J. (1997). Timing patterns in the production of a foreign language. In L. Díaz & C. Pérez (Eds.), *Views on the acquisition and use of a second language: Proceedings of the European Second Language Association 7th annual conference* (pp. 539–551). Barcelona: Universitat Pompeu Fabra.
Spencer, A. (1996). *Phonology*. Oxford: Blackwell.

Steriade, D. (1997). *Phonetics in phonology: The case of laryngeal neutralization*. Unpublished manuscript, University of California, Los Angeles.

Stevens, K. N., & Blumstein, S. (1978). Invariant cues for place of articulation in stop consonants. *Journal of the Acoustical Society of America, 64*, 1358–1368.

Stone, M. (1996). Laboratory techniques for investigating speech articulation. In W. Hardcastle & J. Laver (Eds.), *The handbook of phonetic sciences* (pp. 11–32). Oxford: Blackwell.

Tajima, K., Port, R., & Dalby, J. (1997). Effects of temporal correction on intelligibility of foreign-accented English. *Journal of Phonetics, 25*, 1–24.

Weinberger, S. (1994a). Functional and phonetic constraints on second language phonology. In M. S. Yavaş (Ed.), *First and second language phonology* (pp. 283–302). San Diego, CA: Singular Publishing Group.

Weinberger, S. (1994b, October). *Please speak faster, I can't understand you: On the L2 transfer of fast speech processes*. Paper presented at the Second Language Research Forum, Montreal.

Wells, R. (1987). Voicing assimilation across word boundaries in Russian. *UCLA Working Papers in Phonetics, 68*, 170–182.

Wong, R. (1987). *Teaching pronunciation: Focus on English rhythm and intonation*. Englewood Cliffs, NJ: Prentice Hall.

Zsiga, E. C. (1994). Acoustic evidence for gestural overlap in consonant sequences. *Journal of Phonetics, 22*, 121–140.

Zsiga, E. C. (1995a). An acoustic and electropalatographic study of lexical and post-lexical palatalization in American English. In B. Connell & A. Arvaniti (Eds.), *Papers in laboratory phonology IV: Phonology and phonetic evidence* (pp. 282–302). New York: Cambridge University Press.

Zsiga, E. C. (1995b). Phonology and phonetics in the education of second language teachers: The representation of some variable rules of English. In J. E. Alatis, C. Straehle, B. Gallenberger, & M. Ronkin (Eds.), *Proceedings of the 1995 Georgetown University Round Table on Languages and Linguistics* (pp. 575–587). Washington, DC: Georgetown University Press.

Zsiga, E. C. (1997). Features, gestures, and Igbo vowels: An approach to the phonology/phonetics interface. *Language, 73*, 227–274.

Zsiga, E. C. (2000). Phonetic alignment constraints: Consonant overlap and palatalization in English and Russian. *Journal of Phonetics, 28*, 69–102.

APPENDIX A

Table A1. Stop cluster tokens used in the experiment

		Russian tokens	
Cluster	English tokens	Pronunciation	Gloss
P#P	stop parts	/grʲob po 'beregu/	"rowed along the banks"
P#T	stop tarts	/grʲob tam/	"rowed there"
P#K	stop carts	/grʲob kak sports'men/	"rowed like a sportsman"
T#P	had parts	/rad 'pasportu/	"is glad about the passport"
T#T	had tarts	/rad 'tapotʃkam/	"is glad about the gym shoes"
T#K	had carts	/rad 'kamere/	"is glad about the camera"
K#P	make parts	/pʲok 'persik/	"baked a peach" or
		/pʲok pe'tʃene/	"baked pastry"
K#T	make tarts	/pʲok tort/	"baked tarts"
K#K	make carts	/pʲok 'kaʃu/	"baked kasha" or
		/pʲok kala'tʃi/	"baked bagels"
P#V	stop art	/grʲob odin/	"rowed alone"
T#V	had art	/rad 'atlasu/	"is glad about the atlas"
K#V	make art	/pʲok 'astru/	"baked asters" or
		/pʲok o'ladi/	"baked bagels"
V#P	saw parts	/e'dʲa 'persik/	"eating a peach" or
		/e'dʲa pe'tʃene/	"eating pastry"
V#T	saw tarts	/e'dʲa tort/	"eating a tart"
V#K	saw carts	/e'dʲa 'kaʃu/	"eating kasha" or
		/e'dʲa kala'tʃi/	"eating bagels"

APPENDIX B

SENTENCES USED IN THE EXPERIMENT

English

P#P The manager wants to stop parts from being stolen.
P#T The bakery wants to stop tarts from being stolen.
P#K The grocery store wants to stop carts from being stolen.
T#P The children had parts in the play.
T#T The children had tarts after lunch.
T#K The children had carts in the race.
K#P The machine can make parts by the thousand.
K#T The baker will make tarts this afternoon.
K#K The manufacturer can make carts as well as bicycles.
P#V The museum wants to stop art from being stolen.
T#V The children had art after lunch.
K#V The government will make art a priority.

V#P They saw parts on the table.
V#T They saw tarts at the bakery.
V#K They saw carts in the parking lot.

Russian

P#P *Moj drug* /grʲob po 'beregu/.
"My friend rowed along the banks."
P#T *Moj drug* /grʲob tam/, *pod derevam*.
"My friend rowed there, under the trees."
P#K *Moj drug* /grʲob kak sports'men/.
"My friend rowed like a sportsman."
T#P *Ego otec* /rad 'pasportu/.
"His father is glad about the passport."
T#T *Ego otec* /rad 'tapotʃkam/.
"His father is glad about the gym shoes."
T#K *Ego otec* /rad 'kamere/.
"His father is glad about the camera."
K#P *Moj deduska* /pʲok 'persik/.
"My grandfather baked a peach."
Moj deduska /pʲok pe'tʃene/.
"My grandfather baked pastry."
K#T *Moj deduska* /pʲok tort/.
"My grandfather baked tarts."
K#K *Moj deduska* /pʲok 'kaʃu/.
"My grandfather baked kasha."
Moj deduska /pʲok kala'tʃi/.
"My grandfather baked bagels."
P#V *Moj drug* /grʲob odin/.
"My friend rowed alone."
T#V *Ego otec* /rad 'atlasu/.
"His father is glad about the atlas."
K#V *Moj deduska* /pʲok 'astru/.
"My grandfather baked asters."
Moj deduska /pʲok o'ladi/.
"My grandfather baked pancakes."
V#P *On sidel u stola,* /e'dʲa 'persik/.
"He sat at the table, eating a peach."
On sidel u stola, /e'dʲa pe'tʃene/.
"He sat at the table, eating pastry."
V#T *On sidel u stola,* /e'dʲa tort/.
"He sat at the table, eating a tart."
V#K *On sidel u stola,* /e'dʲa 'kaʃu/.
"He sat at the table, eating kasha."
On sidel u stola, /e'dʲa kala'tʃi/.
"He sat at the table, eating bagels."