

Predictions for the Acquisition of American English Vowels by Native Russian Speakers

University of Georgia Working Papers in Linguistics
Sofia A. Ivanova

Abstract: The purpose of this paper is to hypothesize the difficulties native speakers of Russian will have in the acquisition of American English monophthong vowels based on the predictions generated by the Speech Learning Model (SLM) (Flege 1987). The SLM predicts that considerably new L2 phones will be easier to acquire than phones which are similar to or overlap with existing L1 categories. Based on a comparison of the phonological features of the vowel systems of contemporary standard dialects of Russian and American English, for native Russian learners of English, the /i-ɪ/, /u-ʊ/, /ɛ-æ/, and /ɑ-ʌ/ contrasts are anticipated to be most challenging. A brief review of the literature on Russian learners' perception and production of L2 English vowel contrasts supports these predictions, and adds insights into the acquisition of phonological distinctions in an L2. Future research should take care to account for regional variation in English vowels and compare L2 performance with the local norm, not a generalized standard.

0. Introduction

This work compares the vowel systems of Contemporary Standard Russian (**CSR**) and General American (**GA**), emphasizing how phonetic and phonological differences between these two languages' vowel systems might impact the acquisition of English as a second language (**ESL**) by adult monolingual speakers of Russian. Among others, the tense/lax distinction between the high front and high back vowels of English is expected to be particularly difficult to perceive and produce for native Russian (**NR**) learners. Based on predictions from the Speech Learning Model (Flege 1987) and the respective feature inventories of the two languages (Russian has only one vowel phoneme in the high front and one in the high back parts of the vowel space, while English has two in each), NR learners of English are expected to encounter a great degree of difficulty acquiring the /i-ɪ/ and /u-ʊ/ contrasts. Additional problematic contrasts include /ɛ-æ/ and /ɑ-ʌ/, again owing to differences in the respective vowel feature inventories of Russian and English and the difficulty of acquiring novel features in an L2, as well as differences in phonological processes such as vowel reduction.

The idea that comparing the sound systems of the first and second languages can inform hypotheses about Second Language Acquisition (**SLA**) has long facilitated the study of the phenomenon of language acquisition. Contrastive Analysis (Lado 1957), one of the earliest explicit linguistic hypotheses regarding SLA, first sought to describe phonological differences between the first or native language (**L1**) and subsequent or non-native language(s) (**L2**) of adult learners to predict learner difficulties in the target language. Subsequent research has of course shown that the reality of SLA is dramatically more complex than what the strong version of the Contrastive Analysis Hypothesis suggests; by no means can all errors in SLA can be attributed to differences between the L1 and L2. However, a weaker version of this fundamental insight remains at the heart of many theories of L2 phonology today: we know that a speaker's native phonology plays a role in shaping L2 speech, and L1/L2 differences, in addition to other factors including universal markedness, contribute to the pattern of errors observed in SLA.

The notion of *markedness* contributed another fundamental insight. Greenberg's (1966) typological interpretation of this hypothesis utilized the frequency with which linguistic elements appeared cross-linguistically as a measure of their markedness, with less common concepts considered '*marked*' and those more common, '*unmarked*'. This new insight quickly established its place in generative linguistic theory and extended its scope well beyond simple measures of frequency. Chomsky and Halle (1968) devised a series of opposing pairs of *marked/unmarked* features to evaluate segment inventories; a simple inventory that produced the necessary contrasts while relying on a minimal number of total features and few marked features was considered most effective, while grammars with many marked features or redundant feature combinations were regarded as uneconomical. Another response to the fundamental idea of markedness, Eckman's (1977) Markedness Differential Hypothesis posited that universal markedness considerations may also be important to acquisition of L2 structures: when an aspect of the L2 is more marked than what is present in the L1, the learner will experience difficulty learning it. While Eckman provides no explicit method for determining *degree* of markedness, the fundamental idea that typological markedness plays a role in SLA provides a tool for evaluating L1/L2 differences and improves the predictive power of transfer-based approaches. More recently, Optimality Theory (**OT**) (Prince & Smolensky 1993, 2004), a central idea in current L2 phonological theory and research, has attempted to formally develop the relationship between markedness and phonological universals and the contribution of the L1 phonology by positing that transfer from the L1 is a major factor in L2 phonological acquisition. This approach focuses on constraints, positing that learners begin the SLA process with their L1 constraint rankings, and must, over time, acquire the differing rankings of these same constraints in the L2; markedness plays a role, and less marked structures are re-ranked sooner than the more marked.

The role of the L1 in L2 phonological acquisition is well established (Bohn & Best 2012; Brannen 2002; Flege 1987; Flege et al. 1999; Iverson et al. 2003). A number of models and hypotheses attempting to clarify precisely how the L1 and L2 systems interact throughout the acquisition process and what predictions or generalizations can be made about the outcome are built upon this theoretical foundation. This work utilizes one such proposal – the Speech Learning Model – as a starting point for comparing the vowel systems of two distantly related but dissimilar languages to determine what predictions can be made about the relative ease and success with which English vowel contrasts are acquired by adult native speakers of Russian.

0.1 Speech Learning Model

In order to address why adult learners may not achieve fully nativelike pronunciation of all L2 phones, Flege's (1987, 1988, 1991) Speech Learning Model (**SLM**) compares the sound systems of the L1 and L2. This model makes specific predictions about which L2 phones will cause difficulty for learners from specific L1 backgrounds on the basis of how difficult establishing the new phonetic categories is expected to be. The model suggests that "new", or sufficiently different, phones whose categories do not overlap appreciably with existing L1 phones may be difficult initially but are more likely to be eventually mastered than phones that are only *slightly* different from – and thus partially overlap with – existing L1 categories (Flege 2005). Splitting an L1 category to accommodate two or more partially overlapping L2 phones in the same part of the vowel space is argued to be particularly difficult for the adult learner (Flege 2005). This assertion refers to learners' perception of differences: when acoustic distance between exemplars is small, the SLM predicts that the relevant acoustic cues will be difficult for the L2 learner to perceive, and associated features, tough to acquire (Flege 1987, 1988, 1991; Flege & Munro

1994). Therefore, a phonetic and phonological comparison of the L1 and L2 may help predict and elucidate SLA difficulty.

The SLM holds that, even for highly experienced learners, many L2 production errors are perceptual in origin and that the objects of cross-language perception are vowel and consonant segments, as perceived via a set of phonetically relevant features (Flege, Bohn, & Jang 1997). A range of studies from Flege and colleagues has demonstrated that L2 learners utilize and manipulate acoustic cues in acquiring new contrasts. Flege and Port (1981) examined the production of English /p/ by native speakers of Saudi Arabian Arabic, which has the phonemes /b t d k/ in its inventory, but not /p/ or /g/. Based on this inventory, it was reasoned that Arabic must have the features [voicing] and [place] for stops, and the researchers wanted to test if these features could be recombined to achieve a novel English phone, /p/. When the study subjects produced English /p/ inappropriately similarly to (and heard by native listeners as) /b/, the authors reasoned that they had not re-combined abstract features of the L1 to acquire this new segment, and that the failing had been in producing a new speech sound rather than a new phoneme.

Turning to the acquisition of L2 vowels, when McAllister, Flege, and Piske (2002) tested L2 Swedish learners' ability to acquire a new distinctive vowel feature [length], they found that native Spanish and some native English learners who did not have this feature in their L1 tended to rely on spectral cues (or features) that existed in the L1 and showed little sensitivity to length contrasts, while native Estonian learners, whose L1 shows the greatest degree of prominence of the duration feature, performed most like Swedish controls. Some native Spanish and English participants, however, performed well; their performance could be cited as evidence that new L2 features can sometimes be acquired by learners with especially high language learning aptitude. Results were taken to indicate that it is difficult, although not impossible, to acquire a new feature (or sensitivity to a related acoustic phonetic dimension - in this case, duration). Age effects (which lie outside the scope of this work) have also been observed: Flege, Schirru and MacKay (2003) examined the production of rhotic schwa [ə] by early and late Italian learners of English and concluded that late learners have more difficulty with the acquisition of new features (or relevant acoustic phonetic dimensions).

The Speech Learning Model posits that perception and production of L2 vowels depends on their acoustic similarity to L1 vowels: L2 vowels which are more appreciably different from existing L1 vowels are thought to be easier to acquire, and those that partly overlap with L1 vowels, more difficult (Flege 1987, 1988, 1991). The model predicts that it will be difficult for L2 learners to form a new phonetic category very close to or partially overlapping with but nonetheless distinct from one existing in the L1, and relatively easier to form a phonetic category that is appreciably acoustically different from (and thus readily distinguishable from) existing categories (Flege, 1987, 1988, 1991). Vowels perceived by learners as “new” or *quite different* from those of the L1 are argued to be acquired more effectively than those that are *similar* (*neither identical to nor substantially different from*) those of the L1 (Flege 1987, 1988, 1991). In acquiring an L2 category perceptually similar to what exists in the L1, learners are expected to dissimilate the phones from one another by increasing the difference between them (for example, by slightly raising one vowel and slightly lowering the other, even if doing so causes the vowels to diverge somewhat from monolingual production values in either language) (Flege 2005). Where it is not possible to make specific predictions based on the SLM, it is assumed that categories which are more marked (Eckman 1977), and patterns which depend on more features (Moreton & Paton, 2012), are more difficult to acquire.

1. Russian and English Vowel Inventories

1.1 Russian Vowels

Detailed dialectal and regional variation aside, there are two varieties of standard Russian most often written about in the literature: the Moscow and St. Petersburg varieties. Differences between the two varieties were still fairly prominent just over a century ago; more recently, however, these differences have dwindled and surface less and less in younger speakers of Contemporary Standard Russian (**CSR**) (Jones & Ward 1969; Yanushevskaya & Bunčić 2015).

In the majority view among scholars and the view taken here, Russian has a system of five vowel phonemes, all monophthongs, in stressed syllables - /i e a o u/ (Avanesov 1972; Halle 1971; Jones & Ward 1969) (Table 1). Some accounts also attribute phoneme status to /i/ (Bondarko 1998; Halle 1959; Yanushevskaya & Bunčić, 2015). Based on the former and more accepted view held in the literature, as well as the arguments presented in Padgett (2001), [i] is treated here as an environmentally conditioned allophone of /i/. The motivation for this rests on the fact that the two are in near-complementary distribution, with [i] occurring after non-palatalized consonants and [i] elsewhere, and the few instances in which they contrast in identical environments tend to be borrowings or dialectological terms referring to production of the phones themselves (икать ‘to produce the sound и - [i]’; ыкать ‘to produce the sound ы – [i]’). The palatal glide /j/ may follow any of the five vowel phonemes of Russian in coda position to generate falling diphthongs (Jones & Ward 1969).

Table 1. Vowel Phonemes of CSR			
/i/	/kit/	кит	‘whale’
/e/	/net/	нет	‘no’
/u/	/tut/	тут	‘here’
/o/	/kot/	кот	‘cat’
/a/	/skat/	скат	‘stingray’

There is some support for treating [i] and [i] as representations of two underlyingly different phonemes, and advocates of the independent phoneme view (Hale 1959; Scerba 1912; Yanushevskaya & Bunčić 2015) point out that: 1) they are differentiated orthographically; 2) unlike other positional variants such as [æ], a variant of /a/ which occurs between palatalized consonants, [i] is easily produced and identified in isolation by native speakers (Scerba 1912); 3) historical evidence shows that the two were different phonemes in the past; and 4) in a handful of cases /i/ and /i/ appear word-initially in otherwise phonologically identical environments (Chew 2003). However, the dominant allophonic view, and the view adopted in this manuscript, treats [i] and [i] as allophones in complementary distribution, with /i/ surfacing as [i] following palatalized consonants and as [i] following non-palatalized consonants (Avanesov 1972; Chew 2003; Cubberley 2002; Jones & Ward 1969; Padgett 2001, 2003; Timberlake 2004); the few exceptions contrasting the two come from non-native place names (e.g. Ыл [ip] - the name of a river and several villages in the Komi Republic) and dialectological terms referring to the production of the phones themselves (e.g. икать [ikatʲ] ‘to produce the sound и’ – [i]’; ыкать [ikatʲ] ‘to produce the sound ы – [i]’ (Chew 2003).

1.2 English Vowels

General American (**GA**) is very much a generalization in that it attempts to reflect a diverse group of dialects by excluding any salient social features and idiosyncratic elements of the many

regional dialects spoken throughout the U.S., but itself reflects no specific, exemplary dialect (Kretzchmar 2004). If a GA dialect were to be recognized, it would combine features of Canadian, American West, and American Midland dialects (Labov, Ash, & Boberg 2006). Second language learners immersed in an L2 speaking community have been shown to imitate a local variety, rather than a generalized standard, and in studies, ESL speech has aligned with local norms, both in terms of the social group with which learners associate (Adamson & Regan 1991; Anisman 1975; Thompson 1976) and more general parameters like regional pronunciation (Friesner & Dinkin 2006; Wolfram et al. 2004). Together, these studies suggest that social factors such as gender, social class, and peer group can affect the language variety targeted by ESL learners, and that the English pronunciation of ESL speakers should be compared to native speakers of similar social and regional background. As such, this work bases its analyses on the generalization of GA (Table 2) while making reference to relevant dialectal variation.

Table 2. Vowel phonemes of GA					
Monophthongs			Diphthongs		
/i/	‘bead’	/u/	‘booed’	/aɪ/	‘buy’
/ɪ/	‘bid’	/ʊ/	‘good’	/aʊ/	‘bough’
/e-eɪ/	‘bayed’	/o-ou/	‘bode’	/ɔɪ/	‘boy’
/ɛ/	‘bed’	/ɔ/	‘pawed’		
/æ/	‘bad’	/ɑ/	‘pod’		
		/ʌ/	‘bud’		

The vowel phonemes of General American include 11 monophthongs¹ - /i ɪ e ε æ ʌ ɔ o ʊ u ʌ/, and three diphthongs - /aɪ aʊ ɔɪ/ (IPA 1989; Ladefoged 1999). Some accounts analyze as a diphthong the sequence /ju/ or /iu/, as in *you*, *new*, *tune*; in this account, /ju/ is treated as a sequence of an approximant /j/ and a vowel /u/. Despite the ongoing *cot-caught* merger, which has caused many American English speakers to produce /a/ and /ɔ/ as the same sound, speakers in many parts of the U.S. show no sign of the merger (Labov 2006). Since /ɔ/, as in *pawed*, is present in some varieties of American English (Giegerich 1992; Hillenbrand 2003; Ladefoged 1993), it is included in this inventory. Rhoticized vowels such as [ɹ], mentioned in some phonetic descriptions of English (Ladefoged 1993; Ladefoged 1999,) are not included, as they are not usually seen as phonemic categories in English, analyzed instead as an underlying vowel influenced by a following [ɹ] through a co-articulatory effect known as “/r/-coloring” (Giegerich 1992; Ladefoged 1993). In some dialects spoken in the Western and some Mid-Western parts of the U.S., [u] and [ʊ] are reported to be unrounded, with [ʊ] often pronounced with spread lips (Ladefoged 1999).

Several general conventions address environmentally conditioned changes in vowel quality of American English, noteworthy here for their potential to impact the acquisition of vowel contrasts. Vowels are raised before [ŋ] in the same syllable, so the vowel in *sing* /sɪŋ/ is more like the vowel in *seen* than the vowel in *sin* (Ladefoged 1999); before [ɹ], vowels are lowered and centralized (Ladefoged 1999). In some varieties, [u] is fronted after [t, d, n, l], and the preceding consonant acquires a mid-high front glide [ɪ] (Ladefoged 1999). Vowels are longer before voiced than before voiceless obstruents in coda position, and native speakers have been

¹ Despite being formally classified as monophthongs, American English /e/ and /o/ are generally slightly diphthongized (Ladefoged 1999); except when before rhyme /ɪ/, as in *hair* and *short*, they are best represented as diphthongal vowels [eɪ] and [oʊ] (Giegerich 1992; Hillenbrand 2003).

found to utilize vowel duration as a cue for postvocalic contrast voicing (Kondaurova & Francis 2008). Confounding the situation somewhat is the regional variation observed in vowel duration throughout the U.S.: studying speakers from the same six geographical areas as Clopper et al. (2005), Jacewicz, Fox, and Salmons (2007) found differences across all studied vowels, with longest durations in the South and the shortest in the Inland North.

1.3 Russian and English Vowels: Acoustic distance

Speaking very generally about the vowel inventories of CSR and GA, one of the first anticipated difficulties for NR learners of the GA vowel system is subdividing the vowel space to accommodate twice as many vowel phonemes, some perceptually similar to Russian phonemes or allophones, and others showing varying degrees of difference from Russian phones.

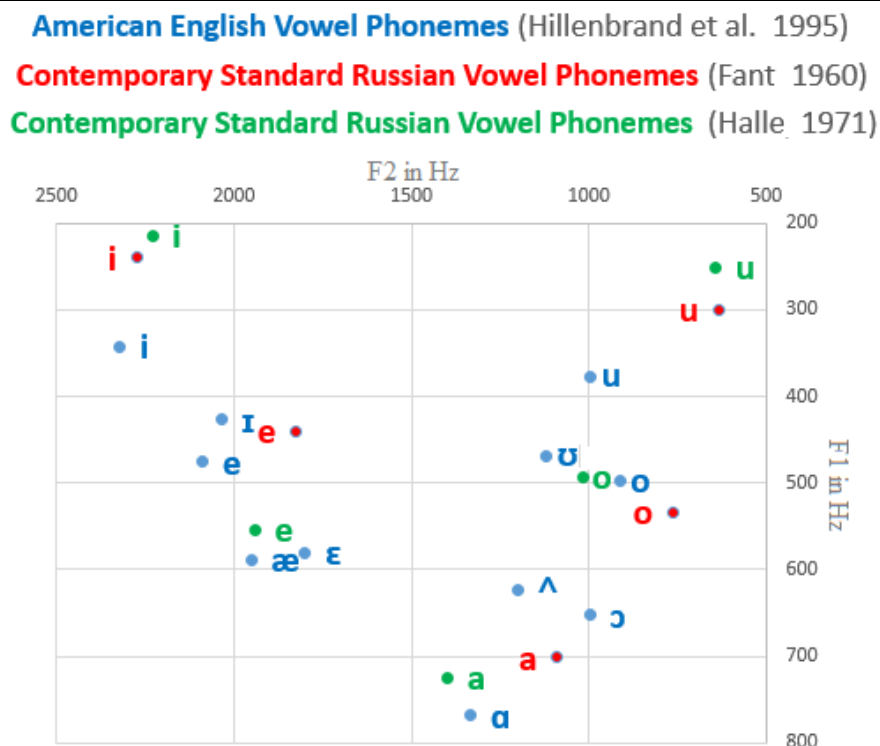
Figure 1 compares the Russian and English vowel systems in terms of acoustic (F1, F2) distance between phonemes. Owing to a paucity of data on formant values of Russian vowels, Russian data (Table 3) are drawn from two studies: one of a single male native speaker of Russian producing vowels in isolation (Fant 1960), and another of three male NR speakers' production of vowels in a variety of CV and VC environments (Halle 1971), selections of which have here been averaged across speakers and presented for [xV] and [Vt] contexts most similar to the English data. English data (Table 4) come from an average of 45 adult male speakers of American English, the majority of whom were raised in Michigan's lower peninsula and who were selected from a larger group of subjects for their production of the /ɑ - ɔ/ distinction; vowels were produced in [hVd] contexts (Hillenbrand 1995). This comparison of the vowel systems of Russian and English in terms of acoustic (F1, F2) distance between vowels (Tables 3, 4; Figure 1) illuminates places where L2 perception and production errors may arise due to overlap of L1 and L2 phonemes in the vowel space.

The formant values for English [ɛ] and [æ], for example, are, at least in some dialects, remarkably similar to one another and most closely approximate those of Russian [e] (Halle 1971). The English speakers' Michigan dialect may play a role, and a greater acoustic distance may indeed be observed between [ɛ] and [æ] in other dialects; nonetheless, the small acoustic difference between these vowels does not generally prevent them from being identified correctly in native production and perception (Hillenbrand 1995). In Figure 1, note the clustering of English /ɪ/, /e/, /æ/, and /ɛ/ near the average F1/F2 values of Russian /e/, and of English /ʊ/ and /o/ with Russian /o/.

Table 3. Formant (in Hz) values of Russian vowels											
Male speaker, vowels spoken in isolation (Fant 1960)						Three speakers, mean of [xV] [Vt] envt. (Halle 1971)					
	i	e	a	o	u		i	e	a	o	u
F1	240	440	700	535	300	F1	221	571	825	492	258
F2	2250	1800	1080	780	625	F2	2250	1933	1408	1013	633
F3	3200	2550	2600	2500	2500	F3	2983	2500	2325	2150	1983

Table 4. Formant (in Hz) values of English vowels											
Mean of 45 male speakers (Michigan), [hVd] environment (Hillenbrand et al. 1995).											
	i	ɪ	e	ɛ	æ	ɑ	ɔ	o	ʌ	ʊ	u
F1	342	427	476	580	588	768	652	497	623	469	378
F2	2322	2034	2089	1799	1952	1333	997	910	1200	1122	997
F3	3000	2684	2691	2605	2601	2522	2533	2459	2550	2434	2343

Figure 1. Russian and English vowel systems



American English vowels exhibit a great range of regional variation (Clopper et al. 2005; Labov et al. 2006). Testing speakers from each of six dialect regions of the U.S. - New England, Mid-Atlantic, North, Midland, South, and West - Clopper et al. (2005) found evidence of the Northern Cities Chain Shift in northern speakers, the Southern Vowel Shift in southern speakers, and an /a-ɔ/ merger in New England, Mid-Atlantic, Midland, and Western speakers, along with other indications of continuing change. Clearly, there is no truly generalized American English, particularly in terms of vowels. Nonetheless, keeping in mind this regional variation and the rich and mutable nature of learner input and experience, it should be possible to compare acoustic distance between what is produced by native and non-native speakers from similar regional and social backgrounds. Moreover, the specific frequencies of a prototypical representation of a given phoneme is not as critical as its relationship to others in the vowel system, which remains very much comparable.

The SLM uses distance between two prototypical vowels on the perceptual (F1-F2) plane as a measure of vowel similarity (Flege & Munro 1994); those L2 vowels which are more similar to one another in both F1 and F2 are more likely to be perceived by L2 learners as members of

one category and, thus, not acquired as separate vowel categories in the L2 (Flege 2005). When acoustic distance between exemplar vowels is small, the SLM predicts that the acoustic cues differentiating the vowels will be difficult for the L2 learner to perceive and acquire, and associated features similarly difficult to acquire. In particular, the tense/lax distinction between the high front and high back vowels of English is expected to pose the greatest overall degree of difficulty in both perception and production for NR learners. Based on the respective feature inventories of the two languages (Russian has only one vowel phoneme in the high front and one in the high back parts of the vowel space, while English has two in each), Russian learners of English are expected based on the SLM to encounter difficulty acquiring the /i-ɪ/ and /u-ʊ/ contrasts. Further difficulty lies in the low vowel space, where Russian has only one phoneme, /a/, while English has /ɑ æ ɔ/. In some varieties of English, /ɑ/ is more central than back and thus slightly more acoustically similar to Russian /a/ (Hillenbrand 2003); NR learners of these varieties may produce a relatively native-like /ɑ/ but struggle with the /ɑ-ʌ/ contrast. The /ʌ-ɑ/ contrast is another area of anticipated perception and production difficulty highlighted by the above comparison of the acoustic distance between Russian and English vowel phonemes.

1.4 Russian and English Vowel Phonemes: Features

L1 categories are thought to serve as a kind of 'magnet' for L2 phones, which map onto L1 categories and trigger a complex process of substitution if new categories do not emerge for the L2 phones (Iverson & Kuhl 1995). Earlier accounts (e.g. Perceptual Assimilation Model – see Best 1993, 1994, 1995) of SLA held that L2 learners, especially at the beginning stages of SLA, do not have (and, with few exceptions, do not generally acquire) access to new L2 features not present in the L1, and that discrimination of L2 contrasts rests on their assimilation to L1 categories. The SLM holds that, even for highly experienced learners, many L2 production errors are perceptual in origin and that the objects of cross-language perception are vowel and consonant segments, as perceived via a set of phonetically relevant features (Flege, Bohn, & Jang 1997). If phonemes are regarded as bundles of distinct features (some associated with specific acoustic or articulatory dimensions, others more abstract), the SLM model may indeed address features in addition to purely perceptual similarities (Flege & Mackay 2004; Flege, Mackay, & Meador 1999; Flege, Munro, & Mackay 1995; Flege, Schirru, & Mackay 2003; Flege, Yeni-komshian, & Liu 1999).

To describe and distinguish the Russian monophthong vowel phonemes with reference to features, only three distinctive features are needed – [High], [Back], and [Round] (Table 5). To describe and distinguish the more numerous English vowel phonemes, at least two additional features are needed – [Low] and [Tense] are typically used (Table 6) (Giegerich 1992). Moreton and Paton (2012) recently added their own experimental data to a review of existing literature in artificial phonology studies to show that patterns which depend on more features are more challenging to acquire. Moving from the Russian system, where only three features are distinctive, a learner of English may be expected to struggle to accurately discern and produce contrasting English vowel phonemes when the distinction rests on these new features, [Low] and [Tense], not present in Russian - /i/ and /ɪ/; /ɛ/ and /æ/;² /ɑ/ and /ʌ/; /u/ and /ʊ/. The /o - ɔ/ distinction is not considered here as the diphthongal properties of English /o/ are thought to make it relatively simple to distinguish from its monophthong neighbors.

The SLM posits that perception and production of L2 vowels depends on their similarity to L1 vowels; the greatest challenge lies in forming new phonetic categories close to or

² /e/ is excluded due to its diphthongal nature, which simplifies its distinction from neighboring monophthongs.

overlapping with existing categories (Flege 1987, 1988, 1991). If phonetic categories are distinguished on the basis of features, then both perception and production of those L2 phonemes which rely on features not present in the L1 are expected to be impacted for L2 learners who cannot access novel L2 features and instead perceive and produce L2 phonemes according to features of the L1. Thus, a NR L2 learner of English has no phonetic basis upon which to accurately perceive or produce the /i - ɪ/, /ɛ - æ/, /ɑ - ʌ/, or /u - ʊ/ distinction without gaining access to the additional phonological features distinctive in English.

Table 5. Phonological features of Russian monophthong vowels					
	i	e	u	o	a
high	+	-	+	-	-
back	-	-	+	+	+
round	-	-	+	+	-

Table 6. Phonological features of GA monophthong vowels (Giegerich 1992)											
	i	ɪ	e	ɛ	æ	ɑ	ɔ	o	ʊ	u	ʌ
high	+	+	-	-	-	-	-	-	+	+	-
back	-	-	-	-	-	+	+	+	+	+	+
round	-	-	-	-	-	-	+	+	+	+	-
low	-	-	-	-	+	+	+	-	-	-	-
tense	+	-	+	-	-	+	+	+	-	+	-

2. Evidence from the literature

2.1 Perception

Perception underlies production, and many production errors are perceptual in origin (Flege, Bohn, & Jang 1997). Few studies have investigated the role of perception in the acquisition of specific GA vowel contrasts by NR speakers. In one such study, native listeners used predominantly spectral differences and relied only somewhat on duration cues in distinguishing English vowels along the /i - ɪ/ continuum, while NR learners did not appear to have access to the relevant features and relied entirely on duration, employing it as a 'default' contrast despite its absence from the L1 (Kondaurova & Francis 2004, 2008).

A recent dissertation was one of the first perceptual studies to examine the acquisition of three English vowel categories by adult speakers of Russian (Makarova 2010) and look specifically at cue weighting in the acquisition of the new vowel contrasts. The new categories formed in the English vowel system by the addition of the features [Low] and [Tense] are the very distinctions the perception of which Makarova's (2010) dissertation seeks to investigate. This study examined the effect of vowel duration and spectral differences on categorization of the high front, high back, and mid/low front lax (/ɛ/ and /æ/) English vowels by adult NR learners. Makarova found that the distinctions between /i - ɪ/, /u - ʊ/, /ɛ - æ/ are indeed difficult for Russian learners to acquire. Moreover, at least in perception, learners initially display overreliance on duration³ (most for the /ɛ - æ/ pair; least for the /u - ʊ/ pair) (Makarova 2010). This is unsurprising given the correlation between tenseness and length in English vowel phonemes, but perhaps somewhat surprising given that duration is not distinctive in Russian.

³ Native speakers of GA rely mostly on spectral cues to distinguish the high tense/lax vowel pairs (Hillenbrand, Clark, & Houde 2000).

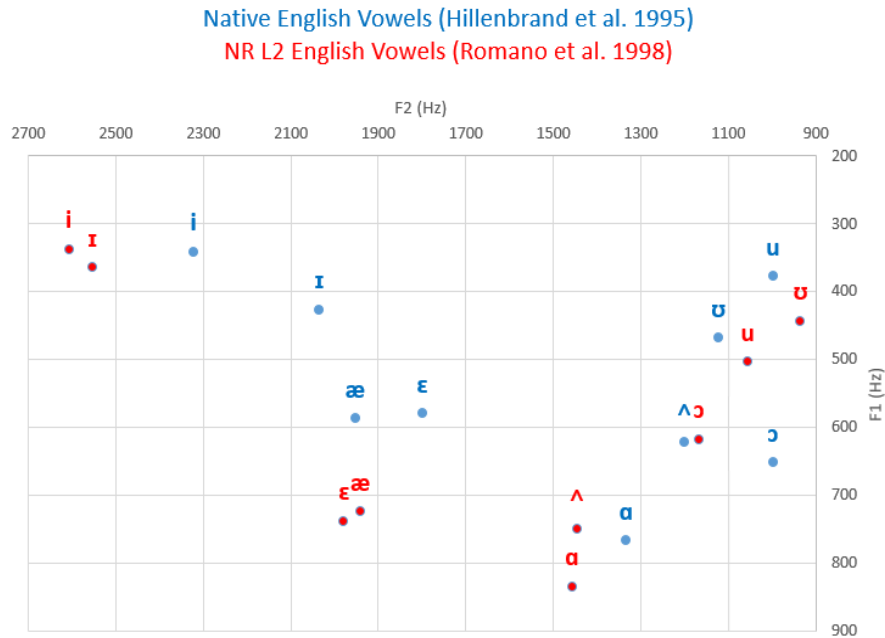
Other authors, however, also report overreliance on temporal cues in the acquisition of these contrasts by NR ESL learners and learners from other language backgrounds which, like Russian, have no vowel duration contrast (Cebrian 2006 - Catalan; Flege et al. 1997 - Mandarin, Korean, Spanish; Kondaurova & Francis 2009 – Russian, Spanish). Particularly for the /i - ɪ/ distinction, overreliance on duration has been shown to remain even with increased experience in the L2 (Cebrian 2006.) NR learners of English in Makarova's study did not confuse low *front* /æ/ with low *back* /ɑ/ or /ɔ/; however, /æ/ was frequently confused with its higher front neighbor /ɛ/, and both /æ/ and /ɛ/ tended to be mapped onto Russian /e/ (Makarova 2010). This tendency may be at least partially explained by the acoustic distance between one American English dialect's /ɛ æ/ and Russian /e/ (particularly the figures given in Halle 1971). Although formant values show that Russian /ɑ/ is more back than front, and thus closer to the low back vowels of English, Russian /ɑ/ is somewhat less back than its perceptual English equivalent, unrounded back /ʌ/, most often confused by NR learners with unrounded mid /ʌ/.

The suggestion that L2 learners utilize contrasts familiar from the L1 to hone their L2 perception and production is not without problems, particularly with reference to features. Results from Flege and Port (1981) suggest that most L2 learners – those who have difficulty perceiving and identifying features that are not phonemic in their L1 system – may not be able to re-combine abstract features already present in the L1 to access feature combinations unique to the L2 and produce a new L2 segment natively. If abstract features are more challenging for learners to transfer and manipulate, then perhaps those which have easily identifiable visual correlates (e.g. [round]) have an advantage in this respect.

Additionally, Makarova (2010) finds that the /u - ʊ/ contrast is mastered before the other two contrasts tested. This is somewhat surprising, as the two are not differentiated well in English orthography, appear in different phonotactic environments, and have relatively few minimal pairs. Makarova's explanation is that Russian speakers are extra sensitive to variation in vowel quality in the high back corner of the vowel space because they are accustomed to consciously subdividing it to accommodate the presence of [ɨ], a high central allophone that is so prominent in the system that it is represented orthographically and is considered by some linguists to be a distinct phoneme. Indeed, at least in some dialects of English, including that found in California, [ʊ] is unrounded and pronounced with spread lips (Ladefoged 1999: 43), making it even more similar to Russian [ɨ]; NR learners getting their input in one such dialect may transfer their native /u/, which is rounded, onto English /u/, and utilize the spread lip articulation of [ʊ] to create an additional salient cue to distinguish it from /u/ for the L2 learner.

2.2 Production

Turning now to production, Figure 2 compares the first and second formants of the English vowels of native and NR L2 English speakers (Hillenbrand et al. 1995; Romano et al. 1998). Romano et al. (1998) examined the L2 English pronunciation of NR learners and found that, fitting with the predictions made by the SLM, NR learners do not make as great a spectral distinction between /i - ɪ/, /ɛ - æ/, and /ɑ - ʌ/ as do native speakers, suggesting both phonemes in each pair have been mapped onto a single L1 category and the features necessary for their distinction have not been acquired.

Figure 2. Native and NR L2 English vowels⁴

The NR learner's challenge in acquisition of GA vowel categories, then, lies in subdividing the vowel space to account for a greater number of vowel contrasts in the L2 than the L1. This is accomplished by learning to pay attention to vowel quality distinctions allophonic in the L1, but phonemic in the L2, and producing these phones in the L2 as contrasting phonemes, as well as cuing in to the fine gradients of English vowel quality as they pertain to phonological environment and processes (e.g. vowel reduction). Russian speakers have both [ɪ] and [ʊ], English phonemes considered to be some of the most problematic elements of Russian-accented English, in their L1 inventory as reduced allophones of the high vowels (Jones & Ward 1969). Between two soft consonants, Russian /a/ is raised to [æ], another particularly troublesome English phoneme for NR learners, and word-initially or between a hard consonant and /l/, the retracted allophone [ɑ] occurs (Jones & Ward 1969), yielding a set of phones that resembles the inventory of low vowel phonemes of at least some GA dialects (Ladefoged 1999). Suppressing native phonological rules that control vowel quality in the L1 to utilize L1 allophones as contrasting phonemes in the L2 is thought to be a difficult task; the learner must become proficient at identifying and producing the relevant distinctions in the L2 in various environments. The fact that these phones are environmentally conditioned allophones in Russian, and not under conscious control, makes it more difficult for the NR learner to acquire these phonemic distinctions in English. Since patterns which depend on more features are more challenging to acquire (Moreton & Paton 2012), the more complex vowel phoneme system of English is predicted to pose difficulty for NR learners.

It is worth noting that ESL learners learn a local variety rather than a generalized standard of English (Friesner & Dinkin 2006; Wolfram et al. 2004). Given the variation noted in the vowel systems of several distinct dialects of GA, the acquisition of certain contrasts may

⁴ The F1-F2 acoustic plane positions for L2 English /u/ and /ʊ/ appear reversed due to one speaker's nonstandard performance: /u/ only slightly higher and more back than the same speaker's /ʊ/. Removing this speaker's outlier values for /u/ causes the average L2 English /u/ to plot directly above NR /ʊ/.

have as much to do with social and geographical factors, and whether the imitated variety maintains those contrasts with sufficiently different prototypical variants (e.g. /ɑ-ɔ/ or /ɛ-æ/), as with any transfer-based explanations.

2.2 Future Research

Additional research may help elucidate what factors impact the differential rate or final attainment of acquisition of L2 contrasts predicted to be equally challenging (e.g. /i - ɪ/, /u - ʊ/) and why some cues (e.g. duration) appear to be ‘defaulted’ to for discrimination of L2 contrasts in the absence of salient spectral cues. A greater focus is needed on dialectal variety; while generalizations about CSR vowels may be specific enough, GA shows such variation, particularly in vowels, that study subjects must be chosen carefully based on the dialect they are acquiring and compared to native speakers of the same variety (as well as sociological background).

3. Conclusion

This work has compared the sound systems of two distantly related but dissimilar languages, Russian and English, in terms of their vowel inventories, based on predictions from the Speech Learning Model. This model predicts that L2 phones appreciably different from existing L1 phones are more likely to be mastered than L2 phones which are only slightly different from and partially overlap with existing L1 categories. It is argued to be difficult for the adult learner to split an L1 category to accommodate two or more partially overlapping L2 phones (Flege 2005).

Several GA phonemes are predicted, based on acoustic similarities and differences with the L1 as well as L1 and L2 feature inventories, to pose significant problems for NR learners of English due to errors caused by L1 transfer and substitution. Most notably, these are the /i - ɪ/, /u - ʊ/, /ɛ - æ/, /ɑ - ʌ/ distinctions, which rely on features and contrasts not present in the L1 and require the learner to subdivide the vowel space to accommodate multiple new L2 phonemes in perceptual space formerly occupied by considerably fewer L1 vowels. Perceiving these new contrasts and acquiring new features, or re-combining abstract features of the L1 to master L2 contrasts, is considered challenging for the adult learner. The learner’s task is to form new categories by learning to perceive and produce the relevant features and combinations in the L2. The GA vowel system, which subdivides the vowel space more so than that of CSR and requires more features to establish all necessary vowel contrasts, is predicted to pose difficulty for NR learners, specifically in subdividing the L2 vowel space and attuning to relevant contrasts.

4. References

- Adamson, H. Douglas & Vera M. Regan. 1991. The acquisition of community speech norms by Asian immigrants learning English as a second language. *Stud Second Lang Acquis* 13. 1-22.
- Anisman, Paul H. 1975. Some aspects of code switching in New York Puerto Rican English. *Biling Rev* 2. 56-85.
- Avanesov, Ruben Ivanovich. 1972. *Russkoe Literaturnoe Proiznoshenie*. [Russian literary pronunciation] Fifth edition. Moskva: Prosvezchenie.
- Best, Catherine T. 1993. Emergence of language-specific constraints in perception of nonnative speech: A window on early phonological development. In B. de BoyssonBardies, S. de Schonen, P. Jusczyk, P. MacNeilage, and J. Morton (eds.), *Developmental Neurocognition: Speech and Face Processing in the First Year of Life*, 289–304. Dordrecht, The Netherlands: Kluwer Academic Publishers.

- Best, Catherine T. 1994. The emergence of native-language phonological influences in infants: A perceptual assimilation model. In J. C. Goodman and H. C. Nusbaum (eds.), *The Development of Speech Perception: The Transition from Speech Sounds to Spoken Words*, 167–224. Cambridge, MA: MIT Press.
- Best, Catherine T. 1995. A direct realist view of cross-language speech perception. In W. Strange (ed.), *Speech Perception and Linguistic Experience: Issues in CrossLanguage Research*, 171–204. Baltimore, MD: York Press.
- Bohn, Ocke-Schwen & Catherine T. Best. 2012. Native-language phonetic and phonological influences on perception of American English approximants by Danish and German listeners. *J Phon* 40. 109-128.
- Bondarko, Liya V. 1998. *Fonetika sovremennogo russkogoazyka*. St. Petersburg.
- Brannen, Kathleen. 2002. The role of perception in differential substitution. *Canadian Journal of Linguistics* 47. 1-46.
- Chew, Peter A. 2003. *A Computational Phonology of Russian*. Dissertation. Parkland, FL: Dissertation.com.
- Chomsky, Noam & Morris Halle. 1968. *The sound pattern of English*. Cambridge, Massachusetts: The MIT Press.
- Clopper, Cynthia G., David B. Pisoni & Ken de Jong. 2005. Acoustic characteristics of the vowel systems of six regional varieties of American English. *J Acoust Soc Am* 118(3). 1661-1676.
- Cubberley, Paul. 2002. *Russian: A linguistic introduction*. Cambridge: Cambridge University Press.
- Eckman, Fred R. 1977. Markedness and the Contrastive Analysis Hypothesis. *Lang Learn* 27(2). 315-330.
- Fant, Gunnar. 1960. *Acoustic Theory of Speech Production*. The Hague: Mouton.
- Flege, James E. 1987. The production of 'new' and 'similar' phones in a foreign language: evidence for the effect of equivalence classification. *J Phon* 15. 47-65.
- Flege, James E. 1988. Factors affecting degree of perceived foreign accent in English sentences. *J Acoust Soc Am* 84. 70-79.
- Flege, James E. 1991. The interlingual identification of Spanish and English vowels: Orthographic evidence. *Q J Exp Psychology Section A* 43(3). 701-731.
- Flege, James E. 2005. *Origins and development of the Speech Learning Model*. Keynote lecture presented at the 1st ASA Workshop on L2 Speech Learning, Simon Fraser Univ., Vancouver, BC April 14-15, 2005.
- Flege, James E., Ian R. MacKay & Diane Meador. 1999. Native Italian speakers' perception and production of English vowels. *J Acoust Soc Am* 106: 2973-2987.
- Flege, James E. & Ian R. MacKay. 2004. Perceiving vowels in a second language. *Stud Second Lang Acquis* 26. 1-34.
- Flege, James E. & Murray J. Munro. 1994. Auditory and categorical effects on cross-language vowel perception. *J Acoust Soc Am* 95. 3623-3641.
- Flege, James E., Murray J. Munro & Ian R. MacKay. 1995. Effects of age of second-language learning on the production of English consonants. *Speech Commun* 16. 1-26.
- Flege, James E., Ocke-Schwen Bohn & Sunyoung Jang. 1997. Effects of experience on non-native speaker production and perception of English vowels. *J Phon* 25. 437-70.
- Flege, James E. & Robert Port. 1981. Cross-language phonetic interference: Arabic to English. *Lang and Speech* 24. 25-146.
- Flege, James E., Carlo Schirru & Ian R. MacKay. 2003. Interaction between the native and second language phonetic subsystems. *Speech Commun* 40. 467-491.
- Flege, James E., Grace H. Yeni-Komshian, & Serena Liu. 1999. Age constraints on second-language acquisition. *J Mem Lang* 41. 78-104.
- Friesner, Michael L. & Aaron J. Dinkin. 2006. The acquisition of native and local phonology by Russian immigrants in Philadelphia. *University of Pennsylvania Working Papers in Linguistics* 12(2). 91-104.
- Giegerich, Heinz J. 1992. *English phonology*. Cambridge: Cambridge University Press.
- Greenberg, Joseph H. 1966. *Language Universals*. The Hague: Mouton.
- Halle, Morris. 1971. *The Sound Pattern of Russian*. The Hague, Paris: Mouton.
- Hillenbrand, James M. 2003. American English: Southern Michigan, *JIPA* 33(1). 121–126.
- Hillenbrand, James M., Laura A. Getty, Michael J. Clark, & Kimberlee Wheeler. 1995. Acoustic characteristics of American English vowels. *J Acoust Soc Am* 97. 3099-3111.
- Hillenbrand, James M., Michael J. Clark, & Robert A. Houde. 2000. Some effect of duration on vowel recognition. *J Acoust Soc Am* 108. 3013-3022.
- IPA. 1989. Report on the 1989 Kiel Convention. *JIPA*. 19. 67-80.
- Iverson, Paul & Patricia K. Kuhl. 1995. Mapping the perceptual magnet effect for speech using signal detection theory and multidimensional scaling. *J Acoust Soc Am* 97(1). 553–562.
- Iverson, Paul, Patricia K. Kuhl, Reiko Akahane-Yamada, Eugen Diesch, Yoh'ich Tohkura, Andreas Kettermann, & Claudia Siebert. 2003. A perceptual interference account of acquisition difficulties for non-native phonemes. *Cognition* 87. 47-57.

- Jacewicz, Ewa, Robert A. Fox & Joseph Salmons. 2007. Vowel duration in three American English dialects. *Am Speech* 82(4). 367-385.
- Jones, Daniel & Dennis Ward. 1969. *The Phonetics of Russian*. Cambridge: Cambridge University Press.
- Kondaurova, Maria V. & Alexander L. Francis. 2004. Perception of the English tense/lax vowel contrast by native speakers of Russian. *J Acoust Soc Am* 116. 2572.
- Kondaurova, Maria V. & Alexander L. Francis. 2008. The relationship between native allophonic experience with vowel duration and perception of the English tense/lax vowel contrast by Spanish and Russian listeners. *J Acoust Soc Am* 124(6). 3959-3971.
- Kretzschmar, William A. 2004. Standard American English pronunciation. In E.W. Schneider, K. Burridge, B. Kortmann, R. Mesthrie, and C. Upton (ed.), *A Handbook of Varieties of English: A Multimedia Reference Tool*, Berlin: Mouton de Gruyter.
- Labov, William, Sharon Ash & Charles Boberg. 2006. *The Atlas of North American English*. Berlin: Mouton-de Gruyter.
- Ladefoged, Peter. 1993. *A Course in Phonetics*, 3rd ed. Fort Worth TX: Harcourt Brace & Company.
- Ladefoged, Peter. 1999. Illustrations of the IPA: American English. *JIPA* 19(2). 77–80
- Lado, Robert. 1957. *Linguistics across cultures: Applied linguistics for language teachers*. Ann Arbor: University of Michigan Press.
- Makarova, Aleksandra Olegovna. 2009. *Acquisition of three vowel contrasts by Russian Speakers of American English*. PhD Dissertation. Cambridge, MA: Harvard University dissertation.
- McAllister, Robert, James E. Flege & Thorsten Piske. 2002. The influence of the L1 on the acquisition of Swedish vowel quantity by native speakers of Spanish, English and Estonian. *J Phon* 30. 229-258.
- Moreton, Elliott & Joe Pater. 2012. Structure and substance in artificial-phonology learning, part I: Structure. *Lang Linguist Compass* 6(11).686-701.
- Padgett, Jaye. 2001. Contrast dispersion and Russian palatalization. In E. Hume & K. Johnson (eds.), *The role of speech perception in phonology*, 187-218. Cambridge, MA: Academic Press.
- Padgett, Jaye. 2003. Contrast and post-velar fronting in Russian. *Nat Lang Linguist Theory* 21. 39-87.
- Padgett, Jaye. 2004. Russian vowel reduction and Dispersion Theory. *Phonological Studies* 7. 81-96.
- Padgett, Jaye & Marija Tabain. 2005. Adaptive Dispersion Theory and phonological vowel reduction in Russian. *Phonetica* 62. 14-54.
- Prince, Alan & Paul Smolensky. 2004. *Optimality Theory: Constraint Interaction in Generative Grammar*. Malden, MA/Oxford, England: Blackwell Publishers.
- Romano, Lisa Jayne, Fredericka Bell-Berti & Eugenia Lorin. 1998. *J Acoust Soc Am* 103(5). 3092-3094.
- Tabain, Marija. 1998. Non-sibilant fricatives in English: spectral information above 10 kHz. *Phonetica* 55. 107-130.
- Thompson, Roger M. 1976. Mexican-American English: Social correlates of regional pronunciation. *Am Speech* 50. 18-24.
- Timberlake, Alan. 2004. *A reference grammar of Russian*. Cambridge: Cambridge University Press.
- Wolfram, Walt, Phillip Carter & Beckie Moriello. 2004. Emerging Hispanic English: New dialect formation in the American South. *J Socioling* 8. 339-358.
- Yanushevskaya, Irena & Daniel Bunčić D. 2015. Illustrations of the IPA: Russian. *JIPA* 45(2). 221-228.