CUE WEIGHTING IN THE ACQUISITION OF FOUR AMERICAN ENGLISH VOWEL CONTRASTS BY NATIVE SPEAKERS OF RUSSIAN

by

SOFIA ALEXANDROVNA IVANOVA

(Under the Direction of Victoria Hasko and Keith Langston)

ABSTRACT

This dissertation investigates the differential weighting of auditory cues in perception and production of four American English vowel contrasts – /i/ and /ɪ/, /u/ and /ʊ/, /ɛ/ and /æ/, and /ɑ/ and /ʌ/ – by Russian-English bilinguals. Its primary goal is to investigate the role of native (Russian) phonology in the acquisition of new acoustic cues – specifically vowel duration and spectral differences – in a second language (English). Literature suggests it is more difficult to acquire more complex vowel systems in a second language; for second language learners of English, the tense/lax contrast, as well as other contrasts between vowels adjacent in the vowel space and distinguished via both spectral and duration cues, are particularly challenging, and non-native speakers, including native speakers of Russian, differ from natives by attuning more to duration than spectral cues. However, many unanswered questions remain about the role of cue weighting in adults' acquisition of vowel systems with larger inventories in second languages. Despite evidence that acquisition patterns vary not only by first language but also by contrast, few studies that examine this issue study more than one vowel pair (typically /i - ɪ/), and very few study native-Russian learners of English. The present dissertation addresses this understudied language pairing and examines both perception and production of four vowel contrasts, focusing on one specific regional variety of American English. While native English talkers in the present study made a perceptible duration distinction between all four contrasts tested, the Russian-English bilinguals did so only between the high vowel pairs, and their accuracy varied by vowel, contrast, and individual in both perception and production. Perception and production ability correlated on some vowels and contrasts, with perception ability generally exceeding production ability; however, due to the difficulty of comparing perception and production accuracy, it is not possible to conclude that these data offer strong support for the Perception-Production link. The Russian-English bilinguals did not appear to be acquiring a local variety, although it is possible that some talkers with low AOA, high LOR, and extensive cultural experience (including education) in Georgia may have acquired a local rather than generalized standard.

INDEX WORDS: Second language acquisition, phonetics, phonology, Russian-English bilinguals, acquisition of vowels, southern English

CUE WEIGHTING IN THE ACQUISITION OF FOUR AMERICAN ENGLISH VOWEL CONTRASTS BY NATIVE SPEAKERS OF RUSSIAN

by

SOFIA ALEXANDROVNA IVANOVA

BA, University of North Carolina at Greensboro, 2006

MA, The Ohio State University, 2010

A Dissertation Submitted to the Graduate Faculty of The University of Georgia in Partial Fulfillment of the Requirements for the Degree

DOCTOR OF PHILOSOPHY

ATHENS, GEORGIA

2018

CUE WEIGHTING IN THE ACQUISITION OF FOUR AMERICAN ENGLISH VOWEL CONTRASTS BY NATIVE SPEAKERS OF RUSSIAN

by

SOFIA ALEXANDROVNA IVANOVA

Major Professor: Victoria Hasko
Keith Langston

Committee: Margaret Renwick
Linda Harklau

Electronic Version Approved:

Suzanne Barbour
Dean of the Graduate School
The University of Georgia
August 2018

ACKNOWLEDGEMENTS

First and foremost I wish to thank my dissertation committee for their guidance and encouragement; this undertaking would have been impossible without their support. I thank Lisa Lipani and Joey Stanley for their invaluable assistance with PRAAT and statistical analysis, respectively. I am grateful to Kim Waters, Vanessa Swenson, and Lindsey Antonini for their insightful, kind, and well-timed advice on everything from waveforms to life. I owe my gratitude also to the individuals who participated in and lent their time and experience to helping me carry out my research. Finally, I am especially thankful for the unending support that I have received from my mother and from Michael Greenfield through the ups and downs of my doctoral studies, and for the uplifting influence of two very special cats, Dean and Nina.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1

INTRODUCTION

1.1 Introduction

This dissertation deals with the acquisition of non-native vowel categories. More specifically, it investigates the differential weighting of auditory cues in perception and production of four American English vowel contrasts – the high front vowel pair /i/ and /ɪ/ (as in English *heat* and *hit*), the high back pair /u/ and /ʊ/ (as in *hoot* and *hook*), the mid and low front /ɛ/ and /æ/ (as in *head* and *had*), and the low back and low-mid central /ɑ/ and /ʌ/ (as in *hot* and *hut*) – by Russian-English bilinguals who began the study of English after age 13 and who currently live in the same major city in central Georgia, USA. The primary goal of this dissertation is to investigate the role of native (Russian) phonology in the acquisition of new acoustic cues – specifically vowel duration and spectral differences – in a second language (English). Secondary goals include: 1) to test the link between perception and production ability in a non-native language among the same group of bilinguals, 2) to investigate whether non-native learners acquire local language norms or a more generalized standard variety, 3) to test whether early age of arrival to the U.S. is linked with more nativelike performance, and 4) to test whether longer residence in an English-speaking community is associated with more nativelike performance.

Despite decades of inquiry into the subject of second language acquisition, many unanswered questions remain about the process of language acquisition in adulthood in general, as well as about the role of cue weighting, defined as differential treatment of acoustic cues in categorization of speech sounds (Holt & Lotto 2006), in adults' acquisition of vowel systems with larger inventories in second languages. Very generally speaking, it is thought to be more difficult to acquire more complex vowel systems in a second language because their inventories rely on additional contrasts not present in the first or native language, and gaining access to non-native cues is argued to be challenging, particularly in adulthood, when language learning ability tends to decline. To address why adult learners generally do not achieve fully nativelike pronunciation of all second language phones, the Speech Learning Model (SLM) (Flege 1987, 1988, 1991), the theoretical framework utilized by the present work, posits that perception and production of non-native vowels depends on their acoustic similarity to one another and to the native language's vowels and that therefore acoustic-phonetic similarity between first and second language sounds can predict their acquisition by adult learners. Splitting one native language category to accommodate two or more nonnative phones in that same part of the vowel space is proposed to be particularly difficult for the adult learner (Flege 2005). Per the SLM, second language vowels occupying overlapping or closely adjacent space on the perceptual plane are more likely to be perceived by learners as members of one category and not acquired as separate categories (Flege 2005). The SLM uses distance between vowels on the F1-F2 plane as a measure of vowel similarity (Flege and Munro 1994); when acoustic distance between exemplar vowels is small, the SLM predicts that the acoustic cues and the associated features differentiating the vowels will be difficult for learners to perceive and acquire (Flege 1987, 1988, 1991; Flege and Munro 1994).

Most studies that have examined the question of how adult acquire a more complex vowel system in a second language have found that learners attempting to acquire the tense/lax contrast, or other contrasts that rely both on spectral and duration cues, over-rely on duration, weighting duration more heavily than do native speakers, who rely more on spectral cues (Escudero 2000;

Kondaurova and Francis 2004; 2006; 2008; Tsukada et al. 2005; Cebrian 2006; Makarova 2010). However, despite evidence that acquisition patterns vary by contrast (Kim, Clayards, & Goad 2017), most studies that examine this issue study only one vowel pair (typically /i - ɪ/), and very few have studied native-Russian learners of English. One pilot study with just four Russian-English bilingual participants and a limited number of tasks examined native-Russian perception and production performance on ten vowels – /i ɪ ɛ æ u ʊ ɔ ɑ ʌ ɝ/ (Bell Berti et al. 1998; Romano et al. 1998) but did not address the weighting of acoustic cues directly. Makarova (2010) – arguably the most in-depth study currently on the subject of cue weighting in native-Russians' acquisition of English vowels – reports on listening performance only, on just three vowel pairs, /i - ɪ/, /u - ʊ/, and /ɛ - æ/.

Additional research is needed to clarify what factors impact the differential rate of acquisition and final attainment of second language contrasts that are predicted to be equally challenging based only on the associated features. For example, the English high front and high back tense-lax pairs /i - ɪ/ and /u - ʊ/ might be predicted to be equally challenging for speakers whose native language (e.g. Russian) distinguishes only one high front and one high back vowel phoneme; however, Makarova's (2010) recent perception study of the acquisition of three American English vowel contrasts by native Russians in northern U.S. has determined that these two contrasts are acquired at different paces and with differing levels of reliance on duration cues (potentially linked to contrast-specific differences in native reliance on duration cues) through the acquisition process. The finding that acquisition patterns may vary both by learner and by vowel contrast is not surprising (Kim, Clayards, & Goad 2017); however, relatively few studies explore this intriguing topic, and even fewer do so with native-Russian learners of English, who face unique challenges in the acquisition of the English vowel system due to their L1's vowel inventory;

to my knowledge, Makarova's (2010) dissertation was the first in-depth study on this subject dealing with this specific first- and second-language pairing. Examining performance on multiple English vowel pairs by the same participants whose native language contrasts fewer vowels than English should be a primary goal for research attempting to address the acquisition of vowel systems with larger inventories in second languages, and while Makarova's (2010) dissertation represents an important first step in this direction, unanswered questions remain about whether Makarova's findings can be generalized to other regional varieties of American English, to production as well as to perception, and to vowel contrasts not reported on in this study. Further research is also needed to address why some cues (e.g. duration) appear to serve as a default for discriminating these novel vowel contrasts in the absence of salient spectral cues, to examine the relationship between native and non-native use of duration cues, and to further explore the connection between first language and the acquisition of non-native vowel contrasts.

1.2 Statement of the Problem

Perception and production studies previously conducted with Russian-English bilinguals (Bell-Berti, Romano, and Lorin 1998; Romano, Bell-Berti, and Lorin 1998; Kondaurova and Francis 2004; 2006; 2008; Makarova 2010), as well as the casual observation of the English pronunciation of Russian-English bilinguals, suggest that native speakers of Russian who began acquiring English in late childhood or adulthood struggle with the acquisition of several specific English vowel contrasts, namely the contrasts between vowels adjacent in the vowel space. Russian-English bilinguals have also been reported by some sources to rely more than do native listeners on duration in perception of some vowel contrasts known to differ in both spectrum and duration in native performance (Kondaurova and Francis 2004; 2006; 2008; Makarova 2010), while other sources note no such excessive reliance on duration in production (Romano et al. 1998). This is not surprising, as it is well established in the fields of linguistics and second language acquisition that nativelike second language pronunciation is quite challenging (Bongearts 1999; Hyltenstam and Abrahamson 2000, 2003), and many adult language learners experience difficulty perceiving and producing non-native contrasts (Strange 1995; Bohn and Munro 2007) often relying on different acoustic cues than do native listeners to make the distinctions (Bohn 1995; Flege, Bohn, and Jang 1997; McAllister, Flege, and Piske 2002). However, much remains unknown about the specific mechanisms behind the process of acquisition of a new phonology, in particular differences pertaining to first language phonology, differences in perception and production of problematic contrasts, and individual differences in acquisition. Indeed, no study to date has rigorously tested all four contrasts known to be problematic for native-Russian learners of English from the perspective of both perception and production within the same large test population. The present study aims to address the research gap introduced above and discussed at length below by adding a fourth vowel contrast to the three contrasts reported on by Makarova (2010), combining perception and production data from the same individuals in order to address multiple aspects of the second language acquisition process, and focusing on a different regional variety of English, spoken in the southern U.S. state of Georgia, in its investigation of cue weighting in the acquisition of four English vowel contrasts predicted to be problematic for native speakers of Russian acquiring English as a second language in late childhood or early adulthood.

1.3 Research Questions

To address the literature gap briefly outlined above and discussed in detail in Chapter 2, this dissertation poses the following seven research questions:

Question 1 – Duration Distinction: Do Russian-English bilinguals produce the four English contrasts tested with a ***different duration ratio*** between the typically longer and typically shorter vowel than do native speakers of English?

Question 2 – Spectral Distinction: Do Russian-English bilinguals produce the four English contrasts tested with a ***different spectral distance*** between the vowels in each contrast than do native speakers of English?

Question 3 – Overall Accuracy: Are Russian-English bilinguals ***more accurate*** in perceiving or ***more nativelike*** in producing certain English vowels than others?

Question 4 – Spectrum vs. Duration: Do native-English listeners more accurately identify English vowels produced by native-Russian talkers when they match native values more in ***spectral*** or ***duration*** cues; that is, which cues contribute more to overall comprehensibility?

Question 5 – Perception vs. Production: Do native-Russian listeners who more accurately ***discriminate*** certain vowels also ***perform*** those contrasts better?

Question 6 – Local Variety: Are Russian-English bilinguals acquiring a ***local variety*** of American English or a generalized standard?

Question 7 – AOA and LOR: Do age of arrival to and length of residence in an English-speaking community correlate with performance?

Questions 1 - 3 address the core goals of the present study – to investigate the differential weighting of acoustic cues by native Russian speakers of American English in their acquisition of novel vowel contrasts not present in their native language – while Questions 4 - 7 address the secondary goals of testing the perception/production link, investigating whether non-native learners acquire local language norms or a more generalized standard variety, verifying the cues utilized by native listeners, and investigating the effect of age of arrival and length of residence on second-language performance.

1.4 Potential Contributions of the Study

The present study has the potential to fill gaps in the existing literature on cue weighting in the process of acquisition of a second language in general and of second language vowels specifically, using evidence from an infrequently studied native (Russian) and second (English) language pairing. The question of cue weighting in second language acquisition is complex, and more research is needed examining various contrasts and different native- and second-language pairings in order to be able to describe general trends in the acquisition of non-native cues.

Additionally, the present study aims to provide further evidence in support of 1) the Perception-Production link and 2) the roles of age of arrival and 3) length of residence in second language performance. Age of arrival and length of residence have fairly extensive histories of being studied in connection with second language acquisition; the Perception-Production link is a somewhat newer area of inquiry, and recently aspects of the intuitive connection between listening and speaking ability in a second language have been questioned, and the present study is expected to provide evidence in support of its key ideas.

Finally, the present study has the potential to contribute additional insights to the growing literature on whether non-native learners acquire a local variety or a generalized standard in the process of acquisition of a second language.

1.5 Conclusion

In this section, I briefly introduced the research topic and questions. Before moving on to the present study's research design and findings, it is necessary to address the previous research conducted on the topics addressed here, as well as theoretical approaches and models. I accomplish this in Chapter 2, by first presenting background information on the relevant languages, then discussing recent and relevant literature on the acquisition of second languages, with a special focus on the acquisition of phonology and vowel contrasts and particularly the acquisition of English vowels by native speakers of Russian; I end Chapter 2 by discussing the link between perception and production in second languages and introducing the Speech Learning Model, the theoretical foundation of the present work. In Chapter 3, I introduce the research participants and provide information about the experiments the results of which are reported in Chapter 4. In Chapter 5, I discuss how these findings inform the present work's research questions and compare with the findings of previous research.

CHAPTER 2

BACKGROUND AND LITERATURE REVIEW

2.1 Introduction

The existing literature on phonological acquisition in second languages is extensive, and a considerable number of studies report on the acquisition of novel vowel contrasts by second language learners, as well as that of specific problematic consonants (for English /ɹ/ and /l/, see: Bradlow, Pisoni, Akahane-Yamada, and Tohkura 1997; Iverson et al. 2003; Aoyama et al. 2004; Bradlow 2008; for English /θ/ and /ð/, see: Miller & Nicely 1955; Hancin-Bhatt 1994; Brannen 2002; Lombardi 2003; Cutler, Weber, Smits, & Cooper 2004; Kinney 2013). However, among the studies investigating the acquisition of novel vowel contrasts in a second language, relatively few studies report on native-Russian learners of English, as compared to English learners from other language backgrounds. If, as proposed by the Speech Learning Model (SLM, discussed in detail in §2.8; see Flege 1987, 1988, 1991), the difficulties faced by adult learners in acquiring second language phones can be attributed to similarities and differences between first and second language sounds, and perception and production of non-native vowels depends on their acoustic similarity to one another and to the native language's vowels, then it becomes critical to thoroughly investigate as many native/non-native language pairings as possible. Even fewer publications address the question of cue weighting – the differential treatment of acoustic cues in categorization of speech sounds (Holt & Lotto 2006) – in this process. This chapter first provides background information on the Russian and English vowel systems in terms of their inventory, features, and acoustic distance (§2.2); continues with a review of relevant findings regarding the acquisition of a second language phonology (§2.3) and second language vowel contrasts (§2.4); offers a review of the handful of perception (§2.5) and production (§2.6) studies from recent decades which report on the acquisition of specific English vowel contrasts by native Russians; and concludes with a discussion of theoretical approaches utilized in the present study (§2.7 and §2.8).

2.2 Background on the Languages

This section compares the vowel systems of Russian and English in terms of inventory, phonological description, features, and acoustic distance as a background for subsequent discussion of the existing literature on the acquisition of English vowels by native Russians.

2.2.1 Russian Vowel Phonemes

In the majority view among scholars and the view taken in this manuscript, Russian has a system of five vowel phonemes, all monophthongs, in stressed syllables: /i e a o u/[1] (Jones & Ward 1969; Halle 1971; Avanesov 1972; see Table 2.1). The palatal glide /j/ may follow any of the five vowel phonemes in coda position to generate falling diphthongs (Jones & Ward 1969: 74-75). The prominent allophone [ɨ], as well as details regarding the quality and articulation of all five vowel phonemes (especially /a/, /e/, and /o/), merit additional discussion.

Table 2.1 Vowel Phonemes of Russian

i *as in*	kit	кит	'whale'
e	net	нет	'no'
u	tut	тут	'here'
o	kot	кот	'cat'
a	skat	скат	'stingray'

[1] These do not represent IPA notation; see §2.2.2 below for an articulatory description of these vowels.

Some authors also attribute phoneme status to /ɨ/. Advocates of this **independent phoneme view** include Ščerba (1912), Halle (1971), Bondarko (1998), Yanushevskaya & Bunčić (2015), and Canepari and Vitali (2018). There is indeed some support for treating [i] and [ɨ] as representations of two underlyingly different phonemes:

1) they are differentiated orthographically, with [i] represented by the letter и and [ɨ] represented by the letter ы, as in бить [bʲitʲ] 'to hit' and быть [bɨtʲ] 'to be'
2) historical evidence shows that the two were different phonemes in the past
3) unlike other positional variants such as [æ], a variant of /a/ which occurs between palatalized consonants as in пять [pʲætʲ] 'five', [ɨ] is easily produced and identified in isolation by native speakers (Chew 2003: 61)
4) in a handful of cases /i/ and /ɨ/ appear word-initially in otherwise phonologically identical environments (икать [ˈikətʲ] 'to produce the sound и' – [i] – and ыкать [ˈɨkətʲ] 'to produce the sound ы' – [ɨ] (Chew 2003: 61).

However, the dominant **allophonic view** treats [i] and [ɨ] as allophones of one phoneme in complementary distribution, with /i/ surfacing as [i] following palatalized consonants and as a more central allophone [ɨ] following hard consonants (Jones & Ward 1969; Avanesov 1972; Padgett 2001; 2003; Cubberley 2002; Chew 2003; Timberlake 2004). The motivations for this position, explained briefly[2], include:

1) the two are in near-complementary distribution, with [ɨ] occurring after hard consonants and [i] elsewhere, a process that may be attributed to the velarization of the preceding consonant (Padgett 2001: 321).

[2] An in-depth exploration of the phonemic status of Russian [ɨ] is beyond the scope of this manuscript.

Table 2.2 Alternation of Russian [i] and [ɨ]

пил [pʲil] ‘he was drinking’	пыл [pɨl] ‘ardor/fervor’
бить [bʲitʲ] ‘to hit’	быть [bɨtʲ] ‘to be’
тикать [ˈtʲikətʲ] ‘to tick’	тыкать [ˈtɨkətʲ] ‘to poke’
Дима [ˈdʲimə] ‘Dima’ (name)	дыма [ˈdɨmə] ‘of smoke’
мило [ˈmʲilə] ‘cute’	мыло [ˈmɨlə] ‘soap’
вить [vʲitʲ] ‘to wreathe, twist’	выть [vʲitʲ] ‘to howl’
гори [gəˈrʲi] ‘burn’ (imperative)	горы [gəˈrɨ] ‘of the mountain’

2) the few instances in which [i] and [ɨ] contrast in identical environments tend to be borrowings such as non-native place names or dialectological terms referring to production of the phones themselves (икать [ˈikətʲ] ‘to produce the sound и’ – [i] – and ыкать [ˈɨkətʲ] ‘to produce the sound ы’ – [ɨ] (Chew 2003: 61).

Based on the more widely accepted allophonic view, as well as the arguments presented in Padgett (2001), [ɨ] is treated here as an environmentally conditioned allophone of /i/.

2.2.2 Russian Vowels: Articulatory Description

Turning now to a more detailed articulatory description of the five vowel phonemes of Russian, the two **high vowels**, [i] and [u], are articulated with a relatively high position of the lower jaw. For [i], the tongue is forward in the mouth, with the tongue tip against the bottom incisors and the tongue blade rising toward the hard palate; the sides of the tongue are pressed against the inner sides of the lower dental arch; the lips are “passive” (not rounded) and additionally pulled slightly to the sides (Avanesov 1972: 30). For [u], the tongue is held relatively back in the oral cavity, with the tongue back and root rising toward the soft palate and the tongue tip pulled away from the bottom incisors; the lips are rounded and protrude far forward.

The two **mid vowels**, [e] and [o], are articulated with the mouth open more than for [i] and [u] but less than for [a]. For [e], the tongue is forward in the mouth, with the blade rising toward the palate, but to a lesser degree than for [i]; the sides of the tongue are pressed against the inner sides of the lower dental arch, the tongue tip is behind the bottom front teeth, and the lips are "passive" and unrounded (Avanesov 1972: 31). For [o], the tongue is relatively back in the mouth, with the tongue back rising toward the soft palate, though to a lesser degree than for [u]; the tongue tip is retracted away from the bottom incisors and the lips are rounded and protrude slightly forward, again, to a lesser degree than for [u]. Word-initially, word-finally, and between a palatalized and a non-palatalized consonant in the case of [e] and two non-palatalized consonants for [o], these two vowels are somewhat diphthongized to include a slight /j/ glide at the onset of [e] and a slight /w/ glide at the onset of [o] (Avanesov 1972: 31; Yanushevskaya & Bunčić 2015: 225).

The Standard Russian **low vowel**, [a], is relatively central in front/backness (Avanesov 1972: 29). In its articulation, the bottom jaw is dropped lower than for any other vowel, the tongue takes on a slightly concave shape, so that the tongue back is lifted and the tongue tip is resting behind the bottom teeth; the sides of the tongue are lightly touching the inner sides of the lower dental arch, and the lips are not rounded and "passive" (Avanesov 1972: 29).

Because allophonic variation in stressed Russian vowels is affected largely by the palatalization of neighboring consonants, and much but – importantly – not all of the consonant inventory is characterized by palatalization contrasts, it is necessary to present an overview of the Russian consonant system before addressing allophonic variation in the Russian vowel system.

2.2.3 Russian Consonants

The following 33 consonant phonemes (Table 2.3) are included in most accounts of Common Standard Russian (Avanesov 1972; Jones & Ward 1969; Halle 1971; Yanushevskaya & Bunčić 2015). This phonemic inventory is largely characterized by contrasts in both voicing and palatalization (hard/non-palatalized, henceforth *hard* vs. soft/palatalized, henceforth *soft* – e.g. /p/ vs. /pʲ/) with the exception of always-hard /t͡s, ʃ, ʒ/ (which do not have soft variants in native words) and always-soft /t͡ʃ, ʃʲː, j/ (which do not have hard variants in native words). Note that /ʃ/ and /ʃʲ/ are analyzed as underlyingly separate phonemes rather than a hard and soft pair.

Table 2.3 Consonant phonemes of Russian. Adapted with modification from Yanushevskaya & Buncic 2015: 222.

p	[ˈpalʲt͡sɨ]	пальцы	‘fingers’	l	[ˈlot]	лот	‘plummet’ (tool)
pʲ	[ˈpʲalʲt͡sɨ]	пяльцы	‘embroidery hoop’	lʲ	[ˈlʲot]	лед	‘ice’ (n)
b	[ˈbas]	бас	‘bass’	r	[ˈrat]	рад	‘glad’ (adj.M)
bʲ	[ˈbʲasʲ]	бязь	‘calico’ (fabric)	rʲ	[ˈrʲat]	ряд	‘row’ (n)
m	[ˈmatʲ]	мать	‘mother’ (n)	s	[ˈsat]	сад	‘garden’ (n)
mʲ	[ˈmʲatʲ]	мять	‘crumple’ (inf)	sʲ	[ˈsʲatʲ]	сядь	‘sit’ (imper.)
f	[ˈfota]	фото	‘photo’	z	[ˈzal]	зал	‘hall’
fʲ	[ˈfʲodar]	Фёдор	‘Fedor’ (name)	zʲ	[ˈvzʲal]	взял	‘(he) took’
v	[ˈvalit]	валит	‘(s/he) fells’	t͡s	[ˈt͡sarʲ]	царь	‘tsar’ (n)
vʲ	[ˈvʲalit]	вялит	‘(s/he) dry-cures’	t͡ʃ	[ˈt͡ʃarɨ]	чары	‘charms’ (n)
t	[ˈtok]	ток	‘current’ (noun)	ʃ	[ˈʃar]	шар	‘sphere’
tʲ	[ˈtʲok]	тёк	‘flowed’	ʃʲː	[ˈʃʲːuka]	щука	‘pike’ (fish)
d	[ˈdom]	дом	‘house’	ʒ	[ˈʒar]	жар	‘heat’ (n)
dʲ	[ˈdʲorn]	дёрн	‘turf’	j	[ˈjat]	яд	‘poison’ (n)
n	[ˈnos]	нос	‘nose’	k	[ˈkot]	кот	‘tomcat’
nʲ	[ˈnʲos]	нес	‘(he) carried’	g	[ˈgot]	год	‘year’
				x	[ˈxam]	хам	‘cad’

There is lack of consensus over the phonemic status of palatalized velars in Russian. Some accounts give the velar phonemes as /k g x/ only, with /kʲ gʲ xʲ/ described as possible environmentally conditioned allophones (Avanesov 1972; Cubberley 2002); others attribute phonemic status to a combination of hard and soft velars: /k kʲ g x/ (Jones & Ward 1969; Halle

1971) or /k k^j g g^j x x^j/ (Kavitskaya 1999; Padgett 2001, 2003; Yanushevskaya and Bunčić 2015). Jones and Ward (1969: 114) consider [k] and [k^j] to be distinct phonemes, as both occur preceding /i e a o u/ while the appearance of /g^j/ before phonemes other than [i e] is quite rare and restricted to loanwords. Admittedly, the occurrence of [k^j] before [a o u] is limited to very few words, nearly all of foreign origin; one common example of /kɨ/ is /kɨʃ/ (shooing or animal-driving noise), while another occurs across a proclitic boundary as in /k ɨgoru/ ‘to Igor’ (Cubberley 2002: 65).Yanushevskaya & Bunčić (2015) attribute phoneme status to all three palatalized velars but acknowledge that their distribution is limited to certain phonological contexts. The phonemic status of [$ʒ^j$] and [$ʒ^j$:] is also contested, with some (typically older) sources treating these phones as phonemes of Russian, and others treating them as allophones of nonpalatalized [ʒ]. Table 2.4 indicates with “+” which soft variants of uncontested hard consonants in the leftmost column are included in the Russian consonant phonemic inventory by the authors at the top of each column. The present work treats soft [k^j g^j x^j] as allophones of /k g x/, in complementary distribution with their plain counterparts, and [$ʒ^j$] along with the marginal phone [$ʒ^j$:] as allophones of /ʒ/, occurring largely in foreign borrowings and older pronunciations.

Table 2.4 Marginal phonemes of Russian

	Vinogradov (1960)	Jones & Ward (1969)	Shvedova (1970)	Halle (1971)	Avanesov (1972)	Chew (2003)	Yanushevskaya & Bunčić (2015)
k^j	+	+		+			+
g^j	+						+
x^j	+						+
$ʒ^j$			+				
$ʒ^j$:					+		

2.2.3 Russian Vowels: Allophonic Variation

Allophonic variation in stressed Russian vowels is affected largely by the palatalization, rather than the place or manner of articulation, of neighboring consonants (as in English, below), although see notes on [ɨ], [e], and [o] below. Since the English consonant system does not feature the palatalization contrast, I do not hypothesize that this potential influence from the L1 will pose a problem in the acquisition of the English vowel system; however, the quality of the five monophthongs themselves merits additional discussion, as it relates to a comparison of the Russian and English vowel systems and how native Russians assimilate and acquire English vowels.

While the Russian sound system is limited to just five vowel phonemes, the number of Russian vowel phones is significantly greater when environmentally conditioned allophones are considered. Broadly speaking, each phoneme has at least four variants, described in detail below, affected by stress (which triggers a complex process of vowel reduction) and palatalization (which encourages various articulatory adjustments) of neighboring consonants. Vowels are more fronted before, after, and especially between soft consonants; that is, progressively more front in CVC^j, C^jVC, and C^jVC^j contexts (Timberlake 2004; Yanushevskaya & Bunčić 2015), which can have the effect of diphthongizing vowels in CVC^j and C^jVC contexts, as the vowel accommodates both the velarization and palatalization of its neighbors (Yanushevskaya & Bunčić 2015). This paradigm extends throughout the CSR vowel system, as demonstrated in Timberlake's (2004: 38) representation of midpoints of stressed Russian vowels in four different palatalization contexts (Figure 2.1). In this figure, the leftmost point for each vowel represents its positioning on the F1/F2 plane in C^jVC^j context; that is, between two soft consonants. Moving left to right, the next point

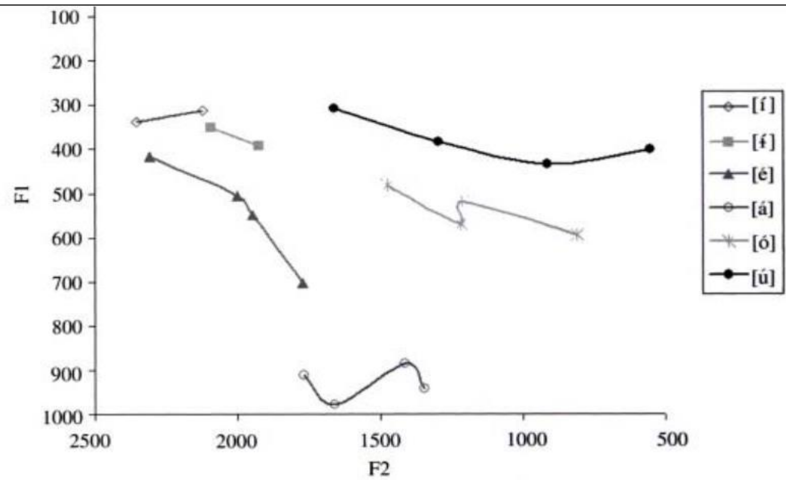


Fig. 2.9 Midpoints of stressed vowels, contexts *ÇVÇ / ÇVC° / C°VÇ / C°VC°*

Figure 2.1 Midpoints of stressed Russian vowels (Timberlake 2004: 38)

represents the vowel in a C^jVC context; the next, in a CVC^j context, and the rightmost point represents the vowel in a CVC context – between two hard consonants.

Vowel reduction, on the other hand, neutralizes several vowel contrasts and shortens vowels (Jones & Ward 1969; Avanesov 1972; Crosswhite 2000; Timberlake 2004). Reduction patterns are described for individual vowels and summarized below for the sake of completeness; however, the quality of reduced vowels in English or Russian is not the focus of the present work.

The principal allophone of Russian /u/ is [u]; it occurs only in stressed syllables and never between two soft consonants (Jones & Ward 1969). In stressed syllables between soft consonants, /u/ is centralized to [ʉ] (Jones & Ward, 1969). Considerable variation is reported in F2 values for stressed Russian /u/ due to articulatory adjustments to neighboring palatals (see Figure 2.1). In unstressed syllables and never between two soft consonants, a near-close near-back allophone [ʊ] occurs; a half-close, central allophone [ʊ̈] occurs in unstressed syllables between two soft consonants (Jones & Ward 1969). These positional variants of /u/ are summarized in Table 2.5.

The Russian [i] occurs only in stressed syllables and can only be preceded by soft consonants. Between soft consonants, both stressed and unstressed /i/ are raised (Jones & Ward

Table 2.5 Positional variants of Russian /u/

Phoneme	*Allophone*	*Stress*	*Context*
/u/	[u]	stressed	all but C^jVC^j
/u/	[ʉ]	stressed	C^jVC^j
/u/	[ʊ]	unstressed	all but C^jVC^j
/u/	[ʊ̈]	unstressed	C^jVC^j

1969). This raised, stressed allophone [i̝] is closest to Cardinal Vowel 1 (Jones & Ward 1969). After a hard consonant, /i/ surfaces as the more central [ɨ]; it may be argued that this [ɨ] is slightly diphthongized from the velarization of the preceding consonant (Padgett 2001; 2003). [ɨ] is advanced further forward in the mouth to [ɨ̟] between hard coronals or the always-hard palato-alveolar [ʃ] or [ʒ], and retracted to [ɨ̠] and slightly diphthongized after hard labial consonants and /l/ (Jones & Ward 1969). Unstressed /i/ lowers to a reduced variant, the near close [ɪ], in most environments, and a lowered and more centralized [ɨ̞] following a hard consonant (Jones & Ward 1969). These positional variants of /i/ are summarized in Table 2.6.

Table 2.6 Positional variants of Russian /i/

Phoneme	*Allophone*	*Stress*	*Context*
/i/	[i]	stressed	most cases after a soft consonant or #V
/i/	[i̝]	stressed	C^jVC^j
/i/	[ɨ]	stressed	CVC, CVC^j, CV#, #V[3]
/i/	[ɨ̟]	stressed	between hard coronals, [ʃ], or [ʒ]
/i/	[ɨ̠]	stressed	after hard labial consonants and [l]
/i/	[ɪ]	unstressed	all but CV
/i/	[ɪ̝]	unstressed	C^jVC^j
/i/	[ɨ̞]	unstressed	CV

[3] Recall that word-initial [ɨ] is possible but rare in Russian, occurring only in borrowings such as non-native place names or dialectological terms referring to production of the phones themselves (Chew 2003; Padgett 2001; 2003)

The Russian /e/ and its ɛ-type allophones , summarized in Table 2.7, occur only in stressed syllables - unstressed /e/ surfaces as [ɪ] word-initially and after soft consonants and the always-soft /j/, and as [ɨ] after hard consonants[4] and the always-hard /ʃ/, /ʒ/, and /t͡s/ (Jones & Ward 1969).

Table 2.7 Positional variants of Russian /e/

Phoneme	*Allophone*	*Stress*	*Context*
/e/	[e]	stressed	VC^j
/e/	[e̠]	stressed	between [ʃ]/[ʒ] and C^j
/e/	[ɛ]	stressed	#VC, CV, never VC^j
/e/	[ɛ̝]	stressed	C^jV, never VC^j
/e/	[ɪ]	unstressed	all but CV
/e/	[ɨ]	unstressed	CV

[e], similar to Cardinal Vowel 2, is always followed by a soft consonant; it is slightly retracted following the consonants [ʃ] or [ʒ] (Jones & Ward 1969). [ɛ], similar to Cardinal Vowel 3, occurs word-initially or after a hard consonant (recall that /e/ following hard consonants is typically restricted to borrowings.), and is never followed by a soft consonant (Jones & Ward 1969). [ɛ̝], a front, unrounded vowel midway between half-close and half-open, occurs after soft consonants, but is never followed by a soft consonant (Jones & Ward 1969). Thus, it would be most accurate to describe the Russian e-vowel (as in нет [nʲɛ̝t] 'no') as characterized by considerable variation in both F1 and F2 and falling along a continuum between Cardinal Vowels 2 and 3 (/e/ and /ɛ/) depending on its environment. Having acknowledged the variability in the quality of this phoneme, the present study will henceforth refer to it as /e/ but refer back to the more nuanced descriptions of its quality as necessary for its analysis.

[4] Always-hard /ʃ/, /ʒ/, and /t͡s/ excluded, /e/ typically only occurs after hard consonants in borrowings.

The Russian /o/ (Table 2.8) appears in most stressed environments as [o̩], a back vowel, slightly less than half-open, positioned between Cardinal Vowels 6 and 7, with moderate lip rounding (Jones & Ward 1969). In stressed positions between two soft consonants the half-close,

Table 2.8 Positional variants of Russian /o/

Phoneme	*Allophone*	*Stress*	*Context*
/o/	[o̩]	stressed	all but C^jVC^j
/o/	[ö]/ [ɵ̩]	stressed	C^jVC^j
/o/	[ɪ]	unstressed	pretonic positions only after soft consonants, /j/, /ʃ/, or /ʒ/
/o/	[ɐ]	unstressed	word-initially or in pretonic position when not preceded by a soft consonant
/o/	[ə]	unstressed	pre-pretonic positions after a hard consonant, or in any post-tonic positions

central allophone [ö] occurs (also transcribed [ɵ̩]). After a soft consonant, /o/ has a slight on-glide similar to an i-sound, and the beginning of the /o/ may be slightly fronted. Before a soft consonant, the end of the vowel may be slightly fronted, and there occurs a slight off-glide similar to schwa ([oᵊ]; this off-glide can become more i-like before a soft consonant) (Jones & Ward 1969). Thus,the Russian o-vowel (as in кот [ko̩t] 'cat') is perhaps most accurately described as falling somewhere between Cardinal Vowels 6 and 7 (/ɔ/ and /o/) in terms of height and characterized by considerable centralization in some environments. Having acknowledged the variability in the quality of this phoneme, the present study will henceforth refer to it as /o/ but refer back to the more nuanced descriptions of its quality as necessary for its analysis. In unstressed positions, /o/ may surface as [ə], [ɐ], or [ɪ]. [ɪ] surfaces in pretonic positions only after soft consonants, /j/, /ʃ/, or /ʒ/ (Jones & Ward 1969). For /o/ (as well as for /a/ – see below for merger of unstressed /o/ and /a/), the reduced variant [ɐ] occurs word-initially or in immediate pretonic position when it is not preceded by a soft consonant; in other pretonic positions after a hard consonant, or in any post-tonic positions, an allophone [ə] occurs (Jones & Ward, 1969).

The principal allophone of the Russian /a/ (Table 2.9) is [a], an open, relatively front unrounded vowel that occurs only in stressed syllables (Jones & Ward 1969). Between two soft consonants, /a/ is raised to [æ]; word-initially or between a hard consonant and /l/, the retracted allophone [ɑ] occurs (Jones & Ward 1969). Thus, the Russian a-vowel (as in скат [skat] 'stingray') is perhaps most accurately described as falling somewhere between Cardinal Vowels 4 and 5 (/a/ and /ɑ/) in terms of front/backness and raised to a vowel similar to English /æ/ in some environments. Having acknowledged the variability in the quality of this phoneme, the present

Table 2.9 Positional variants of Russian /a/

Phoneme	*Allophone*	*Stress*	*Context*
/a/	[a]	stressed	most environments
/a/	[æ]	stressed	C^jVC^j
/a/	[ɑ]	stressed	word-initially or between a hard consonant and /l/
/a/	[ɐ]	unstressed	word-initially or in pretonic position when not preceded by a soft consonant
/a/	[ə]	unstressed	pre-pretonic positions after a hard consonant, or in any post-tonic positions
/a/	[ɪ]	unstressed	pretonic positions only after soft consonants and /j/

study will henceforth refer to it as /a/ but refer back to the more nuanced descriptions of its quality as necessary for its analysis. In unstressed environments, /a/ can surface as [ə], [ɐ], or [ɪ]. [ɐ] occurs word-initially or in immediate pretonic position when it is not preceded by a soft consonant (Jones & Ward 1969). Additionally, [ɐ] can occur in a sequence of two vowels where the letters 'a' or 'o' are used, although some speakers produce an even further reduced schwa here (Jones & Ward

1969). [ə] generally occurs in syllables other than the stressed or first pretonic syllable (see exception above); in pre-pretonic syllables it occurs only after a hard consonant; in posttonic syllables, there is no such restriction (Jones & Ward 1969). In pretonic positions only after soft consonants and /j/, unstressed /a/ reduces to [ɪ] (Jones & Ward 1969).

To summarize the range of allophones attested for the five Russian monophthongs depending on their environment, in **stressed** syllables, Russian /i/ can surface as (potentially raised or centralized) [i] or as (potentially fronted or retracted) [ɨ]; /u/ can surface as (centralized) [u]; /e/ can surface as (retracted) [e] or (lowered) [ɛ]; /o/ can surface as (centralized) [o]; and /a/ can surface as [æ] or [ɑ]. In **unstressed** positions, /i/ and /e/ may surface as (centralized) [ɪ] or [ɨ]; /u/ may surface as (centralized) [ʊ]; and /o/ and /a/ may surface as [ə], [ɐ], or [ɪ]. Russian vowels in unstressed positions tend to be shorter and reduced, and some merge in their surface representations. Very generally, it can be said that the vowel immediately preceding the stressed syllable occupies a 'moderate' reduction position (Crosswhite 2000) and shows a lesser degree of reduction than the vowels of all other unstressed syllables, which occupy a 'radical' reduction position (Crosswhite 2000) characterized by a greater degree of reduction (Jones & Ward 1969; Avanesov 1972; Crosswhite 2000) in terms of both quality and length (Timberlake 2004). Ignoring the gradient changes in vowel quality pertaining to palatalization of neighboring consonants described in more detail above, it can be said that in stressed positions, the Russian vowels are comprised of /i e a o u/; in the pretonic syllable, /e/ merges with /i/ and /o/ with /a/ to yield the set [i a u]; and in all other unstressed syllables, /o/ and /a/ reduce even further to yield the set [i ə u].

2.2.4 English Vowel Phonemes

The vowel phonemes of General American English include 11 monophthongs - /i ɪ e ɛ æ ɑ ɔ o ʊ u ʌ/ and three diphthongs - /aɪ aʊ ɔɪ/ (IPA 1989; Ladefoged 1999; see Table 2.10). Note that, despite being formally classified as monophthongs, American English /e/ and /o/ are generally slightly diphthongized (Ladefoged 1999); except when before rhyme-/ɹ/, as in 'hair' and 'short', they are best represented as diphthongal vowels [eɪ] and [oʊ] (Giegerich 1992).

Some accounts analyze as a diphthong the sequence /ju/ or /iu/, as in *you, new, tune*; in this account, /ju/ is treated not as a diphthong, but as a sequence of an approximant /j/ and a vowel /u/. However, since no data elicitation words used in the present study contain this sequence of coronal before /u/, the precise classification of this segment/sequence is not critical to the analysis. Diphthongs may be monophthongized in typically southern speech (Clopper et al. 2005; Labov et al. 2006); in these cases, length is preserved to maintain underlying contrasts with monophthongs; this study, however, examines only underlying monophthongs, and thus is not concerned with this aspect of southern speech.

Despite the ongoing *cot-caught* merger, which has caused some American English speakers to lose the /a–ɔ/ distinction, other speakers in many parts of the U.S. show no sign of the merger (Labov, Ash, & Boberg 2006). Since /ɔ/, as in '*pawed'*, is present in some varieties of American English (Giegerich 1992; Ladefoged 1993; Hillenbrand 2003), and particularly because it is present in Southern English (Clopper, Pisoni, & de Jong 2005), it is included in this inventory. Rhoticized vowels such as [ɚ], mentioned in some phonetic descriptions of English (Ladefoged 1993; 1999), are not included, as they are not typically seen as phonemic categories in English and are analyzed instead as an underlying vowel influenced by a following [ɹ] through a co-articulatory effect known as "/r/-coloring" (Giegerich 1992; Ladefoged 1993).

Table 2.10 Vowel Phonemes of American English

Monophthongs				Diphthongs	
i *as in*	'bead'	u	'booed'	aɪ as in	'buy'
ɪ	'bid'	ʊ	'good'	aʊ	'bough'
e/eɪ	'bayed'	o/oʊ	'bode'	ɔɪ	'boy'
ɛ	'bed'	ɔ	'pawed'		
æ	'bad'	ɑ	'pod'		
		ʌ	'bud'		

2.2.5 English Vowels: Articulatory Description

The 11 English monophthongs listed above correspond generally with Cardinal Vowels 1, 2, 3, 4, 5, 6, 7, and 8, with the addition of high front and high back lax vowels /ɪ/ and /ʊ/ and an open-mid back unrounded /ʌ/. The back vowels /u/, /ʊ/, /o/, and /ɔ/ are generally rounded, although in some dialects, /u/ and especially /ʊ/ are unrounded and produced with spread lips (Hagiwara 1997; Ladefoged 1999).

2.2.6 English Vowels: Allophonic Variation

Lexical stress affects the pronunciation of English vowels in that unstressed vowels are often reduced to a centralized [ə] or other reduced versions of vowels. Unstressed, lax vowels reduce to schwa; vowels in prevocalic positions and non-low vowels in word-final positions escape this process by becoming tense (Chomsky & Halle 1968). In some varieties of English [ɨ] occurs as a reduced vowel, in addition to schwa (Ladefoged 2001; 2005). Schwa occurs in unstressed word-final positions (e.g. the unstressed vowel in *Rosa*) and inflected forms of such words (e.g. the unstressed vowel in *Rosa's*), where the word-final [ə] is preserved, while [ɨ] occurs in most other unstressed positions (e.g. in *roses*; see Flemming & Johnson 2007).

English vowels lengthen before voiced consonants in coda position, and thus most vowels are longer before voiced obstruents; native speakers of American English utilize vowel duration as a cue for postvocalic contrast voicing (Kondaurova and Francis 2008). English vowels are raised before [ŋ] in the same syllable (Ladefoged 1999); speaking generally, English also features some nasalization of most vowels before nasal consonants. English vowels are lowered and centralized before [ɹ], and [u] is fronted after [t, d, n] (Ladefoged 1999).

2.2.7 Russian and English vowel inventories compared

While the Russian phonemic system is limited to just five monophthongs, the English vowel inventory includes 11 monophthongs (keeping in mind that /e/ and /o/, despite being classified as intrinsically monophthongal, are typically realized as diphthongs) and three diphthong phonemes. Only the monophthongs are considered here.

Several monophthongs – Russian /i a u/ and English /i ɑ u/ – appear in relatively similar forms in both languages' phoneme inventories in that they occupy the same part of the vowel space (Figure 2.2) and are both acoustically similar and readily assimilated to the corresponding native categories even by monolingual Russian speakers (Gilichinskaya and Strange 2010 – see more extensive discussion of this study in 2.5, and the mean formant values for Russian and English vowel phonemes in §2.2.9). This does not, however, mean that they are easy for native Russians to acquire, as perceptually and featurally similar neighbors – /ɪ ʌ ʊ/ – occupy adjacent, in some cases overlapping, parts of the vowel space and present a challenge to native-Russian learners. A considerably larger number of English vowel phonemes - /ɪ ɛ æ ɔ ʊ/ - are rather different from any Russian phonemes, although they may resemble environmentally conditioned allophones of Russian phonemes ([ɛ], [æ], and [ɔ] as allophones of /e/, /a/, and /o/, respectively); and [ɪ] and [ʊ] as unstressed variants of /i/ and /e/, and of /u/, respectively). Thus, speaking very generally about the vowel inventories of Russian and English, the first anticipated challenge for NR learners of the

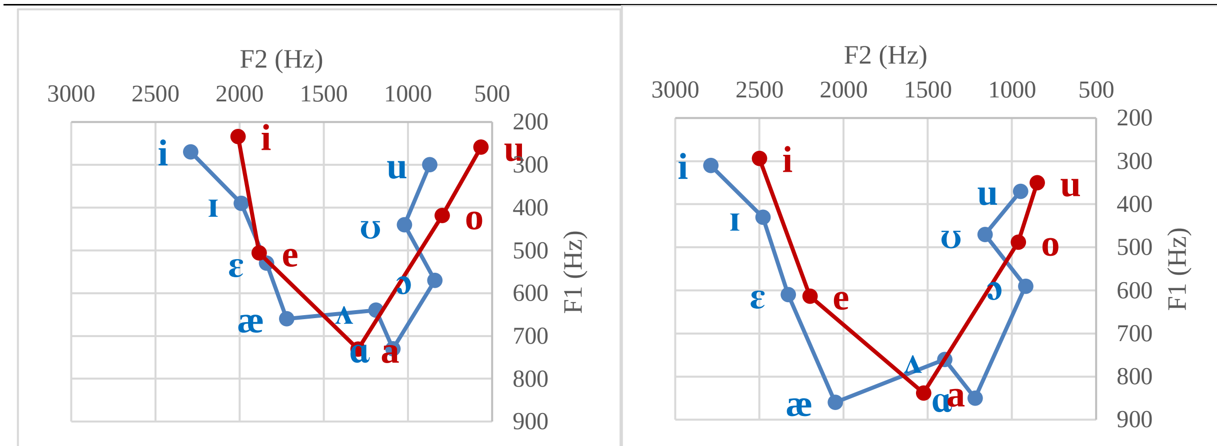


Figure 2.2 American English and Russian vowel inventories. English blue (Peterson & Barney 1952), Russian red (Halle 1971); males left (33 English, 2 Russian); females right (28 E, 1 R).

English vowel system is subdividing the vowel space to accommodate twice as many English vowel phonemes, some of which are quite perceptually similar to Russian phonemes (/i u ɑ/), but most of which are only comparable to environmentally conditioned and reduced allophones Russian (/ɪ ɛ æ ɔ ʊ/)[5]. §2.2.8 and §2.2.9 compare the two vowel systems in terms of features and acoustic distance.

2.2.8 Russian and English Vowel Phonemes: Features

To describe and distinguish the Russian monophthong vowel phonemes with reference to features, only three distinctive features are needed – [High], [Back], and [Round] (Table 2.11). Of course, other feature combinations are possible which would adequately differentiate the five vowel phonemes of Russian, but these would lead to difficult to support predictions about the acquisition of the English contrasts in question; for example, assuming [high], [low], and [back] would suggest that native Russians might recombine their native features [-high], [+low], and [-

[5] See Chapter 5 for a more extended discussion of how experience with Russian vowel allophones may impact the acquisition of L2 English phonemes.

back] and to distinguish English /æ/ but would need to acquire the non-native feature [tense] in order to distinguish /i - ɪ/, /u - ʊ/, and /o - ɔ/ contrasts.

To describe and distinguish the vowels of the larger English inventory, at least two additional features are needed – [Low] and [Tense] are typically used (Table 2.12) (Giegerich 1992). There is some disagreement about the phonological features of English /ɔ/, specifically regarding the features [Low] and [Tense]. Chomsky and Halle (1968) describe /ɔ/ as [+back], [+round], and [+low], but [-tense]; Hall (2007) describes it as [+back] and [+round] only. Following Hall (2007), this account treats /ɔ/ as having the features [+back] and [+round] only, while acknowledging that it is possible to motivate [+tense] for this phoneme due to attested English words such as *claw*, *flaw*, *jaw*, *law*, *maw*, *paw*, *raw*, *saw*, and *straw* which demonstrate that /ɔ/ can occur in a word-final stressed open syllable, a typical feature of tense vowels[6].

Table 2.11 Phonological features of Russian monophthongs

	i	e	u	o	a
high	+	-	+	-	-
back	-	-	+	+	+
round	-	-	+	+	-

Table 2.12 Phonological features of English monophthongs. Vowels not considered in this study are shaded in gray.

	i	ɪ	e	ɛ	æ	ʌ	ɑ	ɔ	o	ʊ	u
high	+	+	-	-	-	-	-	-	-	+	+
back	-	-	-	-	-	+	+	+	+	+	+
round	-	-	-	-	-	-	-	+	+	+	+
low	-	-	-	-	+	-	+	-	-	-	-
tense	+	-	+	-	-	-	-	-	+	-	+

[6] A full discussion of this issue is beyond the scope of this work, particularly as this phoneme is not considered in the final analysis.

Moreton and Pater (2012) recently contributed experimental data to a review of existing literature in artificial phonology studies to show that patterns that depend on more features are more challenging to acquire. Moving from the Russian system, where only three features are distinctive, a native-Russian 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 distinctive in Russian – /i/-/ɪ/ and /u/-/ʊ/, which rely on the feature [Tense], and /ɛ/-/æ/ and /ɑ/-/ʌ/, which rely on the feature [Low]. The o/ɔ and e/ɛ distinctions are not considered here, as the diphthongal properties of English /o/ and /e/ are perceptually salient and facilitate the native Russian listener's task in distinguishing these segments from their monophthong neighbors in perception. Of course, additional considerations arise in the production of these phonemes. Recall that, in Russian, the palatal glide /j/ may follow /e/ in coda position to generate a falling diphthong similar to English [eɪ], but neither /ʊ/ nor /w/ appear in the Russian sound system as candidates for the second half of the diphthong /oʊ/ (Jones & Ward 1969). Historically, borrowings from English into Russian which feature the diphthong /oʊ/ have been produced with a sequence of the two vowels /o/ and /u/ in place of the diphthong (ex: шоу [ʃo.ʊ] 'show'). Note that this post-tonic /u/ would be realized as [ʊ]. Additionally, /o/ and /e/, typically diphthongs in English unless before coda [ɹ], occur in Russian as both monophthongs and diphthongs with a palatal glide, occasionally contrasting in identical environments (e.g. сто /sto/ 'hundred' vs. стой /stoʲ/ 'stand.IMPV.2SG'; все /vsʲe/ 'all.NOM.PL' vs. всей /vsʲeʲ/ 'all.F.GEN-DAT-LOC-INST.SG'). This should serve to facilitate identification of these segments in perception, but may interfere with nativelike production, particularly for [e/eɪ].

The tense/lax distinction between the high front and high back vowel pairs is expected to be particularly difficult for native Russians to perceive and produce since the critical feature

[Tense] is not distinctive in the L1. Further anticipated difficulty lies in the low vowel space, where Russian has only one phoneme, /a/, while English has /ɑ æ/ and utilizes the feature [Low], not distinctive in Russian, to create the distinction between /ɛ/ and /æ/ in the front vowel space and between /ɑ/ and /ʌ/ in the mid-back vowel space.

Thus, based on the respective feature inventories of the two languages, Russian learners of English are expected to encounter difficulty acquiring the /i - ɪ/, /u - ʊ/, /ɛ - æ/, and /ɑ - ʌ/ contrasts. A review of recent literature offers support for this hypothesis – see Kondaurova and Francis 2008 for /i - ɪ/; Makarova 2010 for /i - ɪ/, /u - ʊ/, and /ɛ - æ/; Bell-Berti et al. 1998 and Romano et al. 1998 for all four contrasts.

Although vowel duration is not contrastive in English, it is accepted that the distinctions between tense and lax vowels, as well as the distinctions between /ɛ/ - /æ/ and /ɑ/ - /ʌ/, are typically signaled not only through spectral differences but through the use of duration as a secondary cue (Bohn & Flege 1990; Flege, Bohn, & Jang 1997; Hillenbrand, Clark, & Houde 2000), with tense vowels being longer than lax vowels, and /æ/ and /ɑ/ longer than /ɛ/ and /ʌ/, respectively. Few sources, however, actually report detailed acoustic data such as duration ratios for tense/lax vowel pairs in the same part of the vowel space and for the pairs /æ/-/ɛ/ and /ɑ/-/ʌ/. Peterson and Barney's (1952) benchmarks for general American pronunciation include frequency but not duration data; Clopper et al. (2005) compare duration among six dialect regions but do not provide the durations themselves. One study (Hillenbrand, Getty, Clark, and Wheeler 1995) reports data on the ratio of duration of adjacent vowels (Table 2.13) in a regional variant of English spoken in the north and upper Midwest (Michigan, Illinois, Wisconsin, Minnesota, northern Ohio, and northern Indiana), a geographical area correlating with the North dialect region in Clopper et al. (2005), a key study in the area of regional variation within American English vowels.

Manipulating materials from the production database created by Hillenbrand et al. (1995), Hillenbrand et al. (2000) tested the ability of native listeners to identify American English vowels when their duration was altered. Results suggested that, overall, altering the vowels' duration did not have a great effect on the ability of native listeners to identify them. However, the modest effect that was detected was greater for /a/ - /ʌ/ and /æ/ - /ɛ/ – the same pairs that were found to have the greatest duration difference in Hillenbrand et al. (1995) – than other vowel pairs tested, suggesting native listeners rely more on duration to distinguish these pairs than other pairs known to differ in duration in native speech.

Table 2.13 Ratios between four American English vowel pairs (Hillenbrand et al. 1995: 3103)

Vowels compared	Ratio of durations (vowel 1: vowel 2)
u/ʊ	1.19:1
i/ɪ	1.25:1
æ/ɛ	1.37:1
ɑ/ʌ	1.39:1

There is evidence that in Southern English, lax vowels are produced with longer duration than in other regional dialects of the US (see Clopper et al. 2005 for all 11 monophthongs of American English in six regional varieties; see Jacewicz, Fox, and Salmons 2007 for durations of /ɪ, ɛ, æ/ in three dialect areas), which decreases the contrast with their longer, tense counterparts. If this assertion is true, then duration may be an even less reliable cue to those contrasts in Southern English, which in turn complicates the perception and production by learners relying excessively or solely on duration.

2.2.9 Russian and English Vowel Phonemes: Acoustic distance

Table 2.14 presents formant frequencies for the five Russian vowels, and Table 2.15 presents the formant values for nine[7] American English vowels from a large group of speakers representing a general American dialect of English. A comparison of acoustic (F1, F2) distance between Russian /i, e, a, o, u/ and English /i, ɪ, ɛ, æ, ɑ, ʌ, ʊ, u/ illuminates places where L2 perception and production errors are likely to arise due to acoustic similarity between pairs of adjacent L2 vowels /i, ɪ/, /ɛ, æ/, /ɑ, ʌ/, and /ʊ, u/ and the relative proximity of those vowel pairs to L1 categories: English /i, ɪ/ are acoustically similar to one another and to Russian /i/; English /ɛ,

Table 2.14 Formant (Hz) frequencies of Russian vowels[8] (Halle 1971)

Mean of two male speakers (1971: 169-171)					
	i	e	a	o	u
F1	234	506	731	419	259
F2	2009	1884	1297	797	566
F3	2741	2453	2231	2197	2206
One female speaker (1971: 174)					
	i	e	a	o	u
F1	294	613	838	488	350
F2	2500	2200	1525	963	850
F3	3050	3006	2550	2513	2888

Table 2.15 Formant (Hz) frequencies of English vowels[9] (Peterson & Barney 1952)[10]

Mean of 33 male speakers (1952: 183)									
	i	ɪ	ɛ	æ	ʌ	ɑ	ɔ	ʊ	u
F1	270	390	530	660	640	730	570	440	300
F2	2290	1990	1840	1720	1190	1090	840	1020	870
F3	3010	2550	2480	2410	2390	2440	2410	2240	2240
Mean of 28 female speakers (1952: 183)									
F1	310	430	610	860	760	850	590	470	370
F2	2790	2480	2330	2050	1400	1220	920	1160	950
F3	3310	3070	2990	2850	2780	2810	2710	2680	2670

[7] Values for /e/ and /o/ were not provided.
[8] Vowels produced in pV# and bV# environments.
[9] Northern variety of American English; vowels produced in hVd environment.
[10] Vowels not considered in this study are shaded in gray.

æ/ to one another and to Russian /e/; English /ɑ, ʌ/ to one another and to Russian /a/; and English /u, ʊ/ to one another and to Russian /u/.

Note, however, that American English vowels exhibit a considerable range of regional variation (Clopper et al. 2005; Labov et al. 2006), and General American English is more a term used for convenience rather than a reference to any specific unified variety of English. Particularly in terms of its vowel system, American English exhibits considerable variation among its many regional dialects spoken throughout the US; thus the term “General American English” is a generalization that itself reflects no specific variety but rather refers to a continuum of varieties by attempting to highlight overarching similarities and exclude salient social features and idiosyncratic elements of its dialects (Thomas 2001; Labov et al. 2006). If a General American dialect were recognized at all, it would combine features of Canadian, American West, and American Midland dialects (Labov et al. 2006) and would be unlikely to accurately represent the vowels of many native speakers, particularly those from the north south.

An important early benchmark for acoustic characteristics of American English vowels comes from Peterson and Barney’s (1952) paper (see Table 2.15), which reported on an acoustic analysis of 10 vowels produced by 75 speakers, who varied considerably in terms of native dialect; not surprisingly, the results showed a high degree of variability. Since then, other scholars have sought to describe vowel variation (Thomas 2001), to detail vowel characteristics and variation in specific regional variants (Hillenbrand et al. 1995; Hagiwara 1997; Clopper et al. 2005), and to define regional varieties of American English (Labov 1998; Labov et al. 2006).

Hillenbrand et al. (1995) replicated the work of Peterson and Barney (1952) with speakers from the north and upper Midwest and found considerable differences from the earlier study, including evidence of the Northern Cities Chain Shift, a clockwise rotation of the low and low-mid vowels, also noted by Labov (1998) and Clopper et al. (2005) in Northern dialects. When Hagiwara (1997) examined the mean formant frequency of vowels of speakers from southern California, a pattern of fronted high back vowels (/u/ and /ʊ/) emerged, later also reported, at least for /u/, for Western speakers by Thomas (2001), Clopper et al. (2005), and Labov et al. (2006).

Clopper et al. (2005) attempted to fill a void in the acoustic-phonetic description of the major regional varieties of American English by reporting on vowel variation between six US dialect regions, based on variation described in Labov et al. (2006). Clopper et al. (2005) found consistent variation associated with region. Most notable were a merger of low back vowels in New England, Midland, and Western speakers (as previously reported by Labov (1998)); evidence of the Northern Cities Chain Shift in Northern speakers (previously noted by Hillenbrand et al. (1995) and Labov (1998)); and evidence of the Southern Vowel Shift (fronted back vowels /u/ and /o/) in Southern speakers (reported also by Labov (1998)). In addition to spectral differences, Clopper et al. (2005) noted duration differences, with southern speakers producing longer vowels, and in particular, longer lax vowels. Labov and colleagues additionally report a reversal of the high and mid front vowels in southern speech: laxing and centralization of /i/ and /e/ paired with a tensing and peripheralization (consisting of raising and fronting) of /ɪ/ and /ɛ/, often with an accompanying development of an onglide (Labov 1998; Labov et al. 2006).

As such, it is important to acknowledge that acoustic distance between Russian and English vowel phonemes is necessarily influenced by the variety of English used for the comparison. Moreover, studies show that second language learners actively seek out and imitate a local variety, not a generalized standard, and ESL speech aligns with local norms both in terms of the social group with which learners associate (Anisman 1975; Thompson 1976; Adamson & Regan 1991) and more general parameters like regional pronunciation (Wolfram et al. 2004; Friesner & Dinkin

2006). These studies suggest that, where possible, the English pronunciation of ESL speakers ought to be compared to native speakers of similar regional background.

Therefore, the present work, which aims to investigate the acquisition of English vowel categories by native-Russian speakers of English living in the south, will rely on existing Southern English benchmarks (Labov 1998; Clopper et al. 2005; Stanley et al. 2018) for its comparisons of the Russian and English vowels systems, which will inform the experiment design and preliminary hypotheses, before comparing the perception and production performance of native-English and native-Russian speakers of English living in the same major southern city of the US.

2.2.10 Southern English

For Southern English[11], the variety considered in this work, Clopper et al. (2005) reported consistent spectral differences from the Peterson and Barney (1952) benchmarks, including fronting of /u/ and /o/, centralization of /e/, and raising of /u/ and /ʊ/ among both male and female speakers (Figure 2.3). Among male speakers only, Clopper et al. (2005) additionally noted raising of /ɛ/ and advancing and raising of /æ/. In addition to spectral differences, increased lax vowel duration was reported in Southern speech. Although Clopper et al. (2005) do not provide formant frequencies for Southern English, Figure 2.3, which represents the vowel ellipses of male and female southern speakers, helps demonstrate the spectral features described above and additionally underscore the fact that adjacent vowel pairs /i, ɪ/, /ɛ, æ/, /ɑ, ʌ/, and /u, ʊ/ vary in how much the

[11] Importantly, Clopper et al. (2005) based measurements for Southern speech on only four males (one from Alabama, one from South Carolina, and two from Indiana) and four female speakers (two each from Kentucky and Texas). The hometowns of half of this group - the two male speakers from Indiana and the two females from Kentucky - are on the periphery of the 'Southern' region (defined based on Labov et al. (2006)) and quite close to the 'Midland' region – geographically closer, in fact, to the hometowns of all eight 'Midland' speakers than to the hometowns of the other four 'Southern' speakers in the study (TX, AL, and SC).

vowels of each pair overlap with one another in this regional variety, variation which may impact the relative ease with which the four vowel pairs are acquired by non-native learners.

Most recently, Stanley, Renwick, Kretzschmar, Olsen, and Olsen (2018) have contributed to the growing body of research on Southern speech by creating a detailed and flexible web resource to explore acoustic data on vowels from the Digital Archive of Southern Speech audio corpus (Kretzschmar et al. 2013). Figure 2.4 presents a small part of data available via this project, selected to best match the phonological environment used by Clopper et al. (2005) and the present study. These data are largely in line with the observations of Clopper et al. (2005): for males, both (?) show at least some degree of fronted /u/, raised /ʊ/, raised /ɛ/ and /æ/. Taken together, these southern features noted in Clopper et al. (2005) and Stanley et al. (2018) may impact the of specific English vowel contrasts by native speakers of Russian by reducing acoustic distance between English phonemes in some cases, and changing the acoustic similarity to Russian vowels in others. Southern /u/ and /ʊ/ may be less distinct from one another than in other variants of English due to fronting of /u/ and raising of /ʊ/; furthermore, these two shifts in the vowel space may serve to make English /u/ (which is typically assimilated easily to Russian /u/ - see §2.5 below) perceptually less like Russian /u/ (which is quite back) and English /ʊ/ more like Russian /u/ (compare Figures 2.2, 2.3, and 2.4), making the two English categories even more difficult to discern for a native speaker of Russian acquiring a southern variety of English rather than one where these shifts have not occurred. Raised Southern /ɛ/ and /æ/ may become more perceptually similar to Russian /e/, further confusing the mid and low front vowel space for the native-Russian speaker of English (see Figure 2.5 and §2.5).

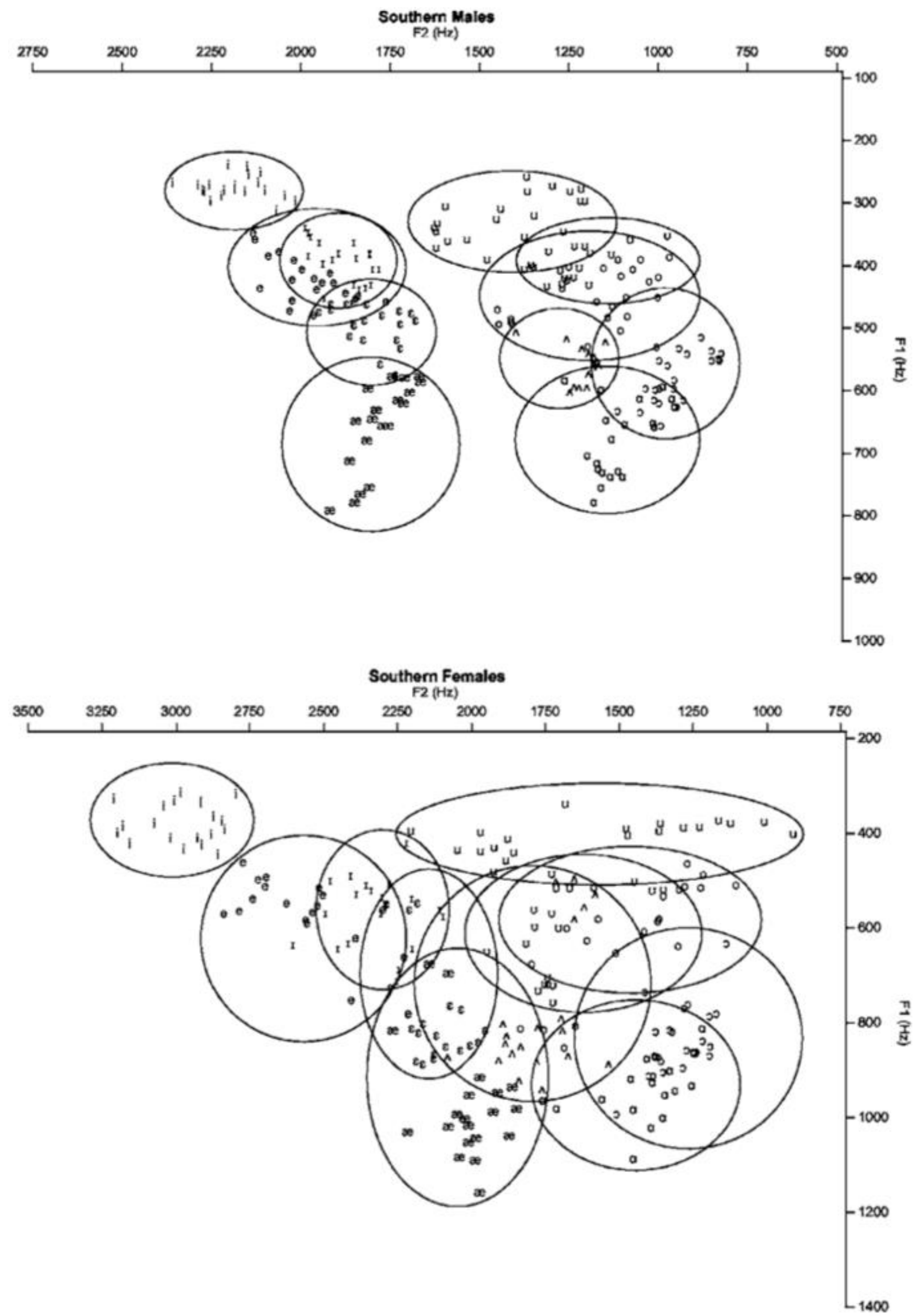


Figure 2.3 11 Southern English vowels - AL, SC, IN, KY, TX (Clopper et al. 2005: 1672)[12]

[12] Vowels /i ɪ e ɛ æ ɑ ɔ ʌ o ʊ u/ in hVd context; produced by four male (top) and four female (bottom) talkers; all tokens, ellipses hand-drawn to include every token for each vowel.

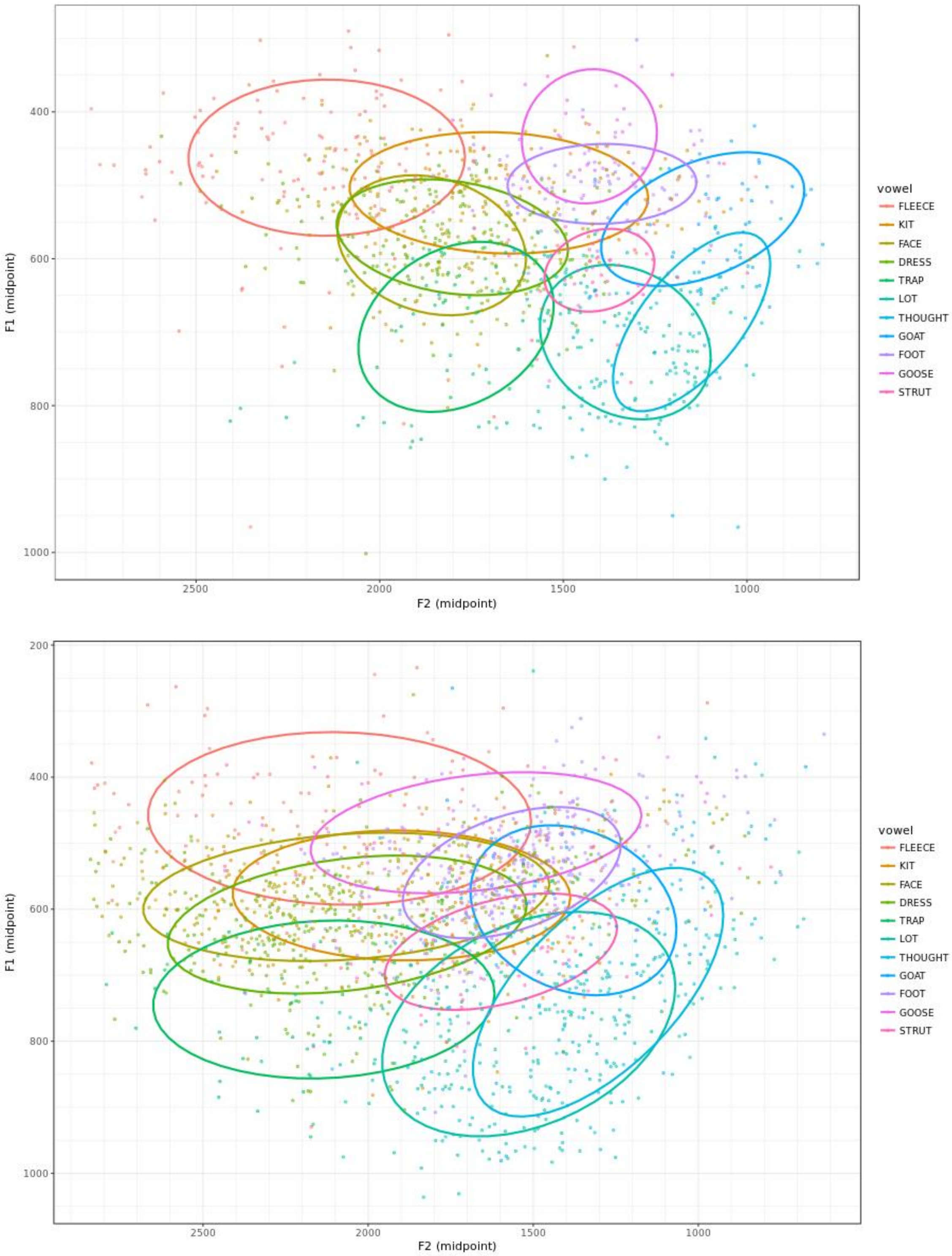


Figure 2.4 11 Southern English vowels - GA (Stanley, Renwick, Kretzschmar, Olsen, & Olsen 2018)[13]

[13] Vowels /i ɪ e ɛ æ ɑ ɔ ʌ o ʊ u/ in Vt and Vd contexts; produced by three male (top) and three female (bottom) talkers; all tokens, ellipses at 67% of data.

Finally, it is helpful to consider how the acquisition of English vowels by native speakers of Russian might be influenced by their native dialect. The next section addresses some trends in variation in Russian and how these differences in the L1 might impact perception and production of L2 English vowels.

2.2.11 On regional variation in Russian

Detailed dialectal and regional variation aside, there are two versions of standard Russian: 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). It is worth noting that some sources report differences between the vowels of regional dialects of Russian. Holden and Nearey (1982) provide acoustic and perceptual data on stressed vowels in three dialects of Russian, spoken in Moscow, Russia; Kiev, Ukraine; and Minsk, Belarus. Both in a characterization task of synthetic vowels and in production, listeners from Kiev differed considerably from those from Minsk and Moscow in their performance with the high vowels, their most salient differences indicating that the vowel [ɨ][14] is positioned differently in the Kiev dialect of Russian – high front, rather than high central (i.e. like the high front /ɪ/ of Ukrainian), approximating English /ɪ/.

Figure 2.5 represents visually the results of a categorization task with synthetic vowels by native speakers of three Russian dialects (Holden and Nearey 1986) – listeners heard synthetic vowel stimuli and identified them as one of six ([ɨ] was treated as a phoneme) Russian vowels. A visual comparison of Figure 2.5 with Figures 2.3 and 2.4 illuminates places where native dialect

[14] Holden and Nearey (1982) treat [ɨ] as an independent phoneme of Russian. It is represented by “y” in Figure 2.5.

may impact how native speakers of Russian perceive English vowels. In particular, the high, front English /ɪ/ is acoustically most similar to Russian [i] for native speakers of Russian in Moscow and Minsk, but for speakers from Kiev, whose [ɨ] is positioned differently than for the other two groups, English [ɪ] may overlap with their native [ɨ] instead[15].

A recent highly detailed analysis of Russian pronunciation (Canepari & Vitali 2018) partially supports this finding (mainly with respect to the Moscow and Minsk position of [ɨ]): in the more recent analysis, the Moscow [ɨ] (argued by these authors also to be a phoneme rather than allophone in Russian) is high central, perhaps even a little more back than truly central (Canepari& Vitali 2018: 162), while [ɨ] in Ukraine (Canepari & Vitali 2018: 183) and Belarus (Canepari & Vitali 2018: 185) is slightly more front than central.

Figure 2.5 Categorization of synthetic vowels in three Russian dialects (Holden and Nearey 1986: 9)

An additional perception task (Holden and Nearey 1982), in which listeners from the three dialects were asked to identify tokens of English /ʌ/ as either Russian /a/, /o/, or /e/, revealed that

[15] Recall that while the dominant view considers [ɨ] to be an allophone of Russian /i/, [ɨ] is nonetheless easily distinguished from [i] and easily produced in isolation by native speakers, which may cause this phone to have a stronger impact on these speakers' L2 English performance than other, less prominent, allophones.

Moscow listeners were most likely to categorize /ʌ/ as /o/, while Kiev listeners overwhelmingly chose /a/. Minsk listeners chose /o/ and /a/ equally frequently and furthermore selected /e/ more often than either of the other two groups. Holden and Nearey (1982) attributed these differences to influence from minor differences in native vowel (/a/, /o/, and /e/) quality among dialects. Gilichinskaya and Strange (2010) tested monolingual Russian listeners' assimilation of English vowels, recruiting participants from a pool of current and former researchers, employees, and students at an education and research institution in Pushchino, Russia, some 100km south of Moscow. These participants most commonly assimilated English /ʌ/ to Russian /a/ (69% of total responses) and less commonly with /o/ (20% of total responses); the remaining 11% of responses were distributed among other Russian vowels and were not reported by the study (see detailed discussion in §2.5).

The areas of realization outlined by Canepari & Vitali (2018) for Russian /a/ and /o/ suggest that both /a/ or /o/ would be an equally good fit perceptually for American English /ʌ/. In contrast with the findings of Holden and Nearey (1982), Canepari & Vitali's (2018) reported data on the variety of Russian spoken in Moscow, Ukraine, and Belarus does not reveal considerable differences in the allophones of /a/ other than perhaps a tendency of stressed Moscow /a/ to vary in F2 rather than height due to the influence of neighboring soft consonants (which may explain why Moscow listeners in Holden and Nearey 1982 hesitated to categorize English /ʌ/ as /a/), while for Ukrainian and Belarusian speakers of Russian, stressed /a/ is both fronted *and* raised in some contexts. Little variation is described by Canepari & Vitali (2018) in the quality of allophones of /o/ for the three dialects, outside of minor, general tendencies, such as that of Moscow /o/ to centralize more than in the other two varieties and Ukrainian Russian /o/ to be positioned slightly lower on the F1/F2 plane, with a higher F1 (which should make it perceptually more similar to

English /ʌ/ and does not help explain why Ukrainian listeners overwhelmingly chose /a/ for English /ʌ/). Based on the additional perspectives of Gilichinskaya and Strange (2010) and Canepari & Vitali (2018), it is unclear whether the performance reported in Holden and Nearey (1982) for assimilation of English /ʌ/ and quality of Russian /ɨ/ by Moscow, Minsk, and Kiev natives is generalizable outside of the earlier study.

Loanwords are an additional source of insights about how the English /ʌ/ is perceived and assimilated by native Russians. Table 2.16 presents the typical pattern of substitution of stressed English vowels in English borrowings into Russian (Benson 1959). In some cases (e.g. that of words like *dollar* and *club*) the observed pattern of substitution may be attributed to orthography; in fact, more recent borrowings of *club* as part of cultural terms (i.e. Club Med, Comedy Club) that do not utilize the existing Russian word *клуб* [klup] show substitution with /a/ rather than /u/.

Table 2.16 Substitution of English vowels in Russian borrowings

English phoneme/s	Russian substitutions in loanwords
/i/ (e.g. *meeting)*	/i/ - *митинг* [mʲitʲɪnk]
/ɪ/ (e.g. *business)*	/i/ - *бизнес* [bʲiznɨs]
/ɛ/ (e.g. *check)*	/e/ - *чек* [ʧek]
/æ/ (e.g. *slang, tanker)*	/e/ - *сленг* [slʲenk]
	/a/ - *танкер* [tanʲkʲɪr]
/u/ (e.g. *boom)*	/u/ - *бум* [bum]
/ʊ/ (e.g. *bulldozer)*	/u/ - *бульдозер* [bulʲdoʷzʲɪr]
/ɑ/ (e.g. *dollar, log)*	/o/ - *доллар* [dolər]
	/a/ - *лаг* [lak]
/ʌ/ (e.g. *bluff, bumper, club)*	/e/ - *блеф* [blʲeʲf]
	/a/ - *бампер* [bamʲpʲɪr]
	/u/ - *клуб* [klup]

It is clear that /ʌ/ exhibits the most variation in possible substitutions, which may be impacted (in addition to phonology) not only by orthography, as pointed out above, but also by language variety and the precise source of the borrowing (e.g. British English may utilize [ʊ] in some words pronounced with [ʌ] in American English); however, it should be noted that substitution with /a/ is by far the most common. Additional support for this observation comes from more recent borrowings in the fields of technology, where /ʌ/ is overwhelmingly substituted with /a/ (e.g., upgrade – апгрейд [apgrejt], update – апдейт [apdejt], bug – баг [bak], touchpad – тачпад [taʧpat]) and occasionally with /u/ (e.g., ultrabook – ультрабук [ulʲtrɐbuk]) (Janulienė and Andriulaitytė 2016: 89). Additional recent borrowings of cultural terms (e.g., lunch – ланч [lanʲʧ], bucks – бакс [baks]) and brand names (Starbucks – Старбакс [starbəks], Pizza Hut - Пицца Хат [pʲiʦə xat], Subway – Сабвей [sabvej], Dunkin' Donuts –Данкин Донатс [dankʲɪn doʷnəʦ]) demonstrate the same trend. This evidence from loanwords suggests that English /ʌ/ would be assimilated most consistently by native Russians to Russian /a/ rather than /o/ or /u/.

A critical comparison of the vowel inventories of English and Russian, as well as evidence from attested substitutions in loanwords adopted into Russian from English, suggest some potential patterns for how native Russian phonology might impact the acquisition of the English vowel system in adulthood. A growing body of experimental and theoretical research on perception and production in the acquisition process offers an additional perspective.

Having established key differences between the Russian and English vowels systems in terms of inventory, features, and acoustic distance, this work moves next into a review of canonical and contemporary research on the acquisition of L2 phonology and vowel contrasts, then turns to a review of recent publications reporting on the perception and production of English vowels by native Russian speakers.

2.3 Acquisition of L2 phonology

Many factors have been found to impact, to varying degrees, success in second language acquisition (**SLA**), particularly pertaining to differential ultimate attainment in second or non-native language (**L2**) pronunciation and accent among L2 learners. These have included age of arrival in an L2-speaking community (typically negatively correlated with L2 performance), length of residence in an L2 speaking community (typically positively correlated with L2 performance), and interaction between the first language (**L1**) and L2 (typically portrayed as influence, as well as affective factors, gender, intelligence, speech sound mimicry, L1 ability, and quality and quantity of input (Table 2.17). The former three most strongly implicated factors – AOA, LOR, and L1 transfer – are considered in the present study in connection with perception and production ability and are discussed in more detail in turn below.

Table 2.17 Factors implicated in success in second language acquisition

Age of arrival (**AOA**) in the L2-speaking community	(Asher and Garcia 1969; Seliger et al. 1975; Oyama 1976; Suter 1976; Tahta, Wood, and Lowenthal 1981; Flege 1988; Thompson 1991; Flege and Fletcher 1992; Flege et al. 1995; 1999b; Moyer 1999; Piske, MacKay, and Flege 2001; Aoyama, Flege, Guion, Akahane-Yamada, and Yamada 2004; Flege and MacKay 2004; Tsukada, Birdsong, Bialystok, Mack, Sung, and Flege 2005; Oh et al. 2011)
Length of residence (**LOR**) in the L2 speaking community	(Thompson 1991; Bialystok 1997; Piske et al. 2001)
Interaction and transfer between the first or native (**L1**) and second or non-native (**L2**) languages	(Best 1995; Flege 1995; Bialystok 1997; Flege et al. 1997; Jarvis and Pavlenko 2008; Kondaurova and Francis 2008)
Affective factors	(Oyama 1976; Suter 1976; Purcell and Suter 1980; Busch 1982; Thompson 1991; Elliott 1995; Bongearts, van Summeren, Planket, and Schils 1997; Moyer 1999; Muñoz and Singleton 2007)

Gender	(Asher and Garcia 1969; Suter 1976; Purcell and Suter 1980; Thompson 1991; Piske et al. 2001)
Language aptitude and intelligence	(Ioup, Boustagui, El Tigi, and Moselle 1994; Harley and Hart 1997; DeKeyser 2000; Abrahamsson and Hyltenstam 2008)
Speech sound mimicry and phonological short-term memory	(Suter 1976; Purcell and Suter 1980; Thompson 1991; MacKay, Meador, and Flege 2001)
L1 ability	(Piske et al. 2001)
Quantity and quality of input	(Suter 1976; Piske, Purcell and Suter 1980; Tahta et al. 1981; Thompson 1991; Flege and Fletcher 1992; Elliott 1995; Flege et al. 1995; Bialystok 1997; Bongearts et al. 1997; Bongearts 1999; Moyer 1999; Guion, Flege, and Loftin 2000; MacKay, and Flege 2001; Flege and MacKay 2004)

Pronunciation stands out among other domains of linguistic competence in that it appears to be the rarest domain for nativelike L2 performance (Bongearts 1999), frequently lagging behind the native benchmark even in speakers whose morphosyntactic, lexical, semantic, and pragmatic intuitions are indistinguishable from native speakers (Hyltenstam and Abrahamson 2000, 2003). However, while it is well established that language learners also experience difficulty perceiving non-native contrasts in the L2 (Strange 1995; Bohn and Munro 2007) and may come to rely on different acoustic cues than native listeners to make the distinctions (Bohn 1995; Flege, Bohn, and Jang 1997; McAllister, Flege, and Piske 2002), much remains unknown about the specific mechanisms behind the process of acquisition of a new phonology. Of the factors listed above that are implicated in ultimate attainment in SLA and L2 phonological acquisition, AOA is perhaps the most highly investigated, and the strongest single predictor of ultimate attainment in L2 phonological acquisition. It is generally accepted that language learning ability, including the ability to acquire a new pronunciation system, declines after maturity is reached and deteriorates further still, particularly if left unexercised, with increasing age (DeGroot 2011). Most studies investigating the link between L2 phonological proficiency and AOA show an increase in the degree of non-native accent with increasing AOA (Asher and Garcia 1969; Oyama 1976; Tahta et al. 1981; Munro, Flege, and MacKay 1996; Flege, Schirru and MacKay 2003; Aoyama et al. 2004; Tsukada et al. 2005; Oh et al. 2011; see Birdsong 2006 for overview). Notable exceptions (Thompson 1991; Tsukada et al. 2005) may be attributed to the testing methods used (e.g. precise AOA of study participants, tasks through which they are tested). Importantly, some of these studies that do report a correlation with AOA find not an abrupt drop-off (e.g. Flege et al. 1995) but a linear decrease in judgments of nativelike pronunciation (Bialystok & Hakuta 1999; Hakuta 2003; Munro & Mann 2005), suggesting that either the role of age in L2 phonological acquisition is limited to a sensitive rather than critical period or that AOA effects are mediated by other factors.

Noting this strong connection between AOA and L2 phonological acquisition, some authors have suggested a critical or sensitive period for phonology but not for other domains (Flege, Yeni-Komshian, and Liu 1999), or multiple age-related sensitive or critical periods for L2 acquisition, each affecting different linguistic domains, with the sensitive period for phonology ending sooner than that of other domains (Seliger 1978; Johnson and Newport 1989; Singleton 1989; Long 1990; Hyltenstam 1992; Ioup 2005). For L2 phonological acquisition, Long (1990) proposed that an L2 learned before age 6 is generally spoken without an accent, an L2 learned after age 12 is generally accented, and variability is expected in nativelikeness of accent in an L2 learned between ages 6 and 12. Higher age limits are suggested for nativelike acquisition of morphology and syntax (Johnson and Newport 1989; Hyltenstam 1992; Long 1990). Flege et al. (1999) found that the effect of AOA on grammaticality judgment scores in their test group was erased by controlling for variables confounded with AOA, such as education and language use, while the effect on foreign accent ratings remained; the researchers concluded that a critical period stemming purely from maturational effects may exist for phonology but not morphosyntax.

However, even a very low AOA does not guarantee nativelike acquisition (see Thompson 1991 for the role of L1 use, and Hyltenstam 1992 for the roles of identity and quality/quantity of input even paired with a very early start of acquisition). On the other hand, overall nativelikeness in late L2 acquisition, while atypical, is also not exceptionally rare (Ioup et al. 1994; Bongearts 1999; Flege, Yeni-Komshian, and Liu 1999; Marinova-Todd 2003; Gass and Selinker 2008; Abrahamsson and Hyltenstam 2009). Nativelikeness following a late start of acquisition has been associated with certain L1-L2 pairings and extensive L2 use (Flege, Yeni-Komshian, and Liu 1999) as well as with high levels of practice in the L2, motivation to sound native-like, and L2 phonetic training in pronunciation (Bongearts 1999).

Perhaps the second most frequently cited variable in studies of L2 foreign accent is amount of L2 experience, operationalized in many studies as length of residence in an L2-speaking community (LOR). As with AOA, there is no consensus regarding the precise strength of this variable's influence on L2 pronunciation and its interaction with other variables; nonetheless, most studies of L2 foreign accent tend to find overall better L2 pronunciation in subjects with longer LOR in an L2 speaking community (Asher & Garcia 1969; Flege & Fletcher 1992; Flege et al. 1995; 1999b). In one canonical example, Flege & Fletcher (1992) found significantly stronger accent in late Spanish-English bilinguals who had resided in the U.S. for 0.7 years compared those with an LOR of 14.3 years. Interestingly, however, a multiple regression analysis later identified AOA, not LOR, as the most important predictor of degree of L2 foreign accent (Piske, MacKay, & Flege 2001), leading the researchers to conclude that LOR is a significant but overall less important determiner of degree of L2 foreign accent than AOA. L2 perception, too, has been demonstrated to improve with increased LOR: in one study, Spanish-speaking learners of English who had more experience with the L2 showed greater ability to distinguish L2 from L1 vowels in a vowel categorizing experiment (Flege 1991).

Despite strong evidence such as outlined above, the role played by LOR in language acquisition is neither simple nor undisputed, and some studies have failed to find an effect of LOR on L2 foreign accent (Oyama 1976; Flege 1988; Thompson 1991; Munro 1993; Moyer 1999; Flege et al., 2006). For example, Flege (1988) observed no significant difference in accent ratings between two groups of adult native Taiwanese learners of English differing in LOR (1.1 vs 5.1 yr); these results, however, were not taken to minimize the role of LOR but instead interpreted as support for the hypothesis that, following an initial period of rapid learning, additional unaided language experience (extended LOR) does not significantly affect degree of L2 foreign accent in adulthood, although it may certainly significantly impact the pronunciation of new learners (as in Oyama 1976). It has been suggested (Flege & Fletcher 1992) that studies such as Flege's (1988) fail to find a significant LOR effect on accent because they examine too narrow a range of LOR values. Flege & Fletcher (1992) reasoned that their Spanish-English bilingual participants, but not the native Taiwanese subjects studied by Flege (1988), showed an effect of LOR because those inexperienced subjects had lived for a much shorter period of time in the U.S. than the inexperienced Taiwanese subjects *(0.7 yr, S.D. = 0.3 vs. 1.1 yr, S.D. = 0.7).*

When analyzed from this perspective, the results from a number of studies failing to demonstrate effect of LOR become less detrimental to the LOR hypothesis, such Aoyama et al.'s (2004) study which found that L1-Japanese L2-English adults were judged to be no less accented at 1.5 years of US residence than at 0.5 and Flege et al.'s (2006**)** study which found that accent did not significantly improve for native-Korean speakers with two additional years of exposure to English (LOR of 3 and 5 years). On the other hand, the subjects analyzed by Munro (1993), Meador et al. (2000), and Cebrian (2006) may have resided in their respective L2-speaking communities for too long at the time of study for LOR effects to emerge.

To identify precisely when in the first year of residence the effects of LOR on adult learners begin to taper off, Munro and Derwing (2008) analyzed the English vowel production of recent arrivals to Canada at two-month intervals. They found that intelligibility increased rapidly but plateaued within a few months. The authors did admit that this result is complicated by different learning trajectories among the tested contrasts, differences in transfer-related influence from differing L1s, and possible effects of word frequency; nonetheless, this finding underscores that the benefit of LOR, especially for older learners, begins to wane within the first year, and adds support to the developing explanation (Krashen, Long, and Scarcella 1979; Flege & Fletcher 1992) for why LOR does not have a significant effect on pronunciation in some studies. Taken together, these studies demonstrate that

1) LOR is more likely to affect pronunciation of early, rather than late, L2 learners, particularly those with LOR of <1 year for adults (Flege & Fletcher, 1992),
2) LOR effects are more likely to emerge when comparing groups with greatly different LOR (*ex:* Spanish-English bilinguals with LOR of 0.7 and 14.3 years in Flege & Fletcher, 1992), and studies that have failed to find an effect of LOR were likely comparing groups with overly similar LORs, studying learners too advanced to be significantly impacted by LOR, or both (Piske, MacKay, Flege 2001),
3) when LOR effects are found, they tend to be a less important predictor of degree of L2 foreign accent that AOA (Flege & Fletcher, 1992; Piske, MacKay, Flege 2001).

Clearly, L2 pronunciation is a special case within the field of SLA research, requiring special treatment to elucidate the role of factors beyond AOA and LOR in the successful acquisition of a non-native phonology. One such factor is influence from the native language: it is well known that the L1 affects perception of L2 phones (Polka, Colantoni, and Sundara 2001; Iverson et al. 2003; Escudero, Benders, and Lipski 2009). One of the earliest explicit linguistic hypotheses regarding L2 acquisition, Contrastive Analysis (Lado 1957), first sought to describe the phonological differences between the L1 and 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. 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.

The L1 is thought to exert a powerful influence on the acquisition of L2 phonology due to the very nature of L1 acquisition: infants are born with an innate ability to discriminate most speech sounds but lose sensitivity to most non-native contrasts within the first year of life (Best, McRoberts, and Sithole 1988; Best 1991; Kuhl, Williams, Lacerda, Stevens, and Lindblom 1992). That is, early in the process of acquiring their L1, infants experience a shift in their perception of speech sounds, causing them to filter subsequent input through the lens of their L1 phonology (which is beneficial for L1 processing but potentially problematic for processing input in other languages). Thus, as part of the very process of L1 acquisition, infants begin to attune to a set of L1-specific acoustic cues that indicate the phonological category membership of speech sounds in the L1.

Acoustic dimensions do not contribute equally to categorization of speech sounds; this differential treatment of acoustic cues in categorization of sounds is referred to as cue weighting (Holt & Lotto 2006). Since the use of acoustic cues in speech perception and production is language-specific (Crowther & Mann 1992; Bohn 1995; Escudero, Benders, & Lipiski 2009), experience with specific acoustic cues in the L1 is hypothesized to impact the perception of L2 sounds (Repp 1982; Iverson et al. 2003). Language learners may not readily detect or implement in production of the L2 acoustic cues or features not present in the L1 (Flege 1995 1998; McAllister et al. 2002). There are indications that even preattentive cue weighting may depend on language experience (Lipski, Escudero, and Benders 2012). Indeed, L1 and L2 listeners have been found to weight individual acoustic cues differently (Repp 1982), and cues treated as primary in distinguishing a particular contrast by native speakers may play a secondary role for L2 learners (Holt and Lotto 2006; Ylinen, Uther, Latvala, Vepsäläinen, Iverson, Akahane-Yamada, and Näätänen 2010). Thus, these language-specific differences in cue weighting can affect the salience of acoustic cues in the L2 and interfere with the acquisition of L2 categories (Iverson et al. 2003). One well-known example of this is the case of the acquisition of the English /ɹ/-/l/ distinction by native Japanese speakers: where native English speakers and listeners rely on differences in third formant (F3) values to distinguish these two contrasting phones, native Japanese speakers and listeners instead rely on differences in the second format (F2) (Bradlow, Pisoni, Akahane-Yamada, and Tohkura 1997; Iverson et al. 2003; Aoyama et al. 2004; Bradlow 2008). The next section narrows in scope to address the acquisition of non-native vowel contrasts, particularly as it relates to the differential weighting of cues in the L1 and L2.

2.4 Acquisition of L2 vowel contrasts

The process described in the previous section, by which early language experience influences subsequent ability to attune to speech sound contrasts in the L2, is generalizable for both consonants and vowels, and in the case of vowels, this phenomenon is termed the Perceptual Magnet Effect (Kuhl 1991, 2000): early experience with the vowels of the L1 warps a listener's perception of vowel sounds toward prototypical members of native vowel categories.

The tense/lax contrast of English – exemplified by /i - ɪ/ and /u - ʊ/ – is one of the more commonly studied contrasts in the area of acquisition of L2 vowels, although many studies focus only on the /i - ɪ/ contrast or choose instead to examine the front vowel /i - ɪ - ɛ - æ/ continuum. Native speakers of English rely mostly on spectral differences – formant frequency and formant movement – and only marginally on duration differences to distinguish the tense and lax vowels in the same part of the vowel space (Peterson and Barney 1952; Hillenbrand et al. 1995; Hillenbrand et al. 2000). The majority of L2 learners of English, however, have been found to attune more to temporal than spectral cues in making the distinction, both in perception and production (Flege et al. 1997; Ingram and Park 1997; Escudero 2000; Tsukada et al. 2005; Cebrian 2006; Morrison 2006; Kondaurova & Francis 2008, 2010; Makarova 2010).

It is unsurprising that L2 learners of English who have experience with contrastive duration in their L1 (*Arabic* – Munro 1993; *Japanese* – Ingram and Park 1997, Minnick-Fox and Maeda 1999; *Hungarian* – Altenberg and Vago 1987) might rely on duration to contrast adjacent English vowels which differ primarily in terms of spectral values but also feature some measurable duration difference (e.g. /i - ɪ/, /u - ʊ/, /ɛ - æ/, and /ɑ - ʌ/). Interestingly, however, this over-reliance on temporal cues has been reported among even those learners whose L1s have no vowel duration contrast (*Catalan* - Cebrian 2006; *Korean* - Flege et al. 1997, Ingram and Park 1997, Tsukada et al. 2005; *Mandarin* - Flege et al. 1997, *Spanish* - Escudero 2000, Flege et al. 1997, Kondaurova & Francis 2008, 2010, Morrison 2006; *Russian* - Kondaurova & Francis 2008, Makarova 2010). Particularly for the /i - ɪ/ distinction, over-reliance on duration has been shown to remain even with increased experience in the L2 (Cebrian 2006). Since over-reliance on duration in making the tense/lax distinction has been observed both in learners whose L1 utilizes temporal contrasts and in those whose L1 does not, some researchers have proposed that duration is psychoacoustically more salient than fine spectral differences in a part of the vowel space that is not subdivided in the L1 (Bohn 1995), and thus potentially available as a cue even to learners whose L1 does not utilize it contrastively.

Although many details about the acquisition of a new vowel system remain unclear, it is now well established that subdividing the vowel space in the L2 further than what is familiar in the L1, learning to attune to new cues such as vowel duration, or acquiring new distinctive features, such as [low] or [tense], in the L2 is a particularly challenging aspect of L2 phonological acquisition (for the *tense/lax distinction*, see Flege et al. 1997; Cebrian 2006; Kondaurova & Francis 2008; for *length*, see Dupoux and Peperkamp 2002; McAllister, Flege and Piske 2002). And while this general trend is assumed – and has been shown – to hold true for native Russian learners of English, even less is known specifically about the acquisition of English vowel contrasts by native speakers of Russian (for *perception*: Bell-Berti et al. (1998), Kondaurova and Francis (2004), (2008), Makarova (2010); for *production*: Romano et al. (1998)). The case of native Russian learners of English is particularly interesting because, as discussed in 2.2, Russian contrasts considerably fewer vowels in stressed syllables than English. As a consequence of their native phonology, Russian learners, like others whose native languages contrast a relatively smaller number of vowels, are well known to struggle to perceive and produce English vowel contrasts - namely /i - ɪ/, /u - ʊ/, /ɛ - æ/, /ɑ - ʌ/ - which occupy parts of the vowel space that are either unused entirely or occupied by a single category in the L1 and which rely on new L2 features not present in the L1. The native Russian English-learner's challenge, then, lies in subdividing the vowel space to account for nearly twice as many vowel contrasts in the L2 than the L1. This is presumably accomplished by attuning to new vowel quality distinctions; that is, forming new vowel categories appropriate to the L2 by attuning to fine spectral differences in a previously not subdivided part of the vowel space or, from a featural perspective, learning to attune to new features or a new combination of familiar features to distinguish new vowel contrasts.

However, little is known about the details of this process among native-Russian learners of English: the strategies employed by learners, whether these strategies vary by vowel pair, whether these strategies are best described with reference to acoustic distance or features, which contrasts pose the greatest challenges, and whether the considerable regional variation of American English plays a role. Studies addressing these questions typically examine one vowel pair or compare non-native performance with vowels in the /i/ to /æ/ continuum; to my knowledge, no study has simultaneously examined the acquisition of all four vowel contrasts by native Russian learners. Kim, Clayards, and Goad (2017) report different acquisition patterns for two Canadian English vowel pairs (/i - ɪ/ and /ɛ - æ/) by native-Korean learners, including the use of both spectral and duration cues to distinguish the /i - ɪ/ contrast but largely only duration to distinguish /ɛ - æ/ among the studied groups of learners, and different developmental trajectories of individual learners according to vowel contrast. Additionally, examining cue weights over time, the researchers were able to partially confirm the developmental stages proposed by Escudero (2000) for native-Spanish learners of English. To my knowledge, Makarova's (2010) dissertation has been the only study to investigate this question with native-Russian learners of English, finding that they, too, relied more on duration to distinguish /ɛ - æ/ than /i - ɪ/ and demonstrated a pattern similar to that reported by Escudero (2000). Barrios, Jiang, and Isardi (2016) compared spatial and featural approaches to similarity in L2 phonology by examining the perception of English vowel contrasts /ɑ/-/æ/ and /i/-/ɪ/ among native-Spanish late learners of English; to my knowledge, no such investigation has been undertaken with native-Russian late learners of English despite the fact that such a study would present a fascinating comparison, as Russian and Spanish share a similar vowel inventory. In terms of which vowel contrasts of English present the greatest challenge to the native-Russian learner, Makarova (2010) reported most nativelike listening performance on /u - ʊ/, intermediate performance on /i - ɪ/, and poorest performance on /ɛ - æ/; however, these results may be connected to the treatment of duration cues in the particular regional variety examined by this study (see details below); a 1998 pilot study of Russian-English bilinguals in the same part of the country demonstrated poorest perception performance on /ɑ/, /ʌ/, /æ/, /ɔ/, /u/ (Bell-Berti et al. 1998) and least nativelike production on /ɪ ɛ ɔ ʊ ʌ/ and the /i/-/ɪ/ and /ɛ/-/æ/ pairs in particular (Romano et al. 1998). However, to my knowledge, no study has examined the performance of Russian-English bilinguals across different dialect areas or even compared performance of Russian-English bilinguals in other parts of the U.S. to the Bell-Berti et al. (1998), Romano et al. (1998), or Makarova (2010) findings.

The next section addresses some of the above questions in more detail with a review of literature on the perception and production of English vowel contrasts by native Russian speakers. The gaps that remain, which shape the research questions and hypotheses of the present work, are addressed at the end of the present chapter, after a discussion of the theoretical orientation adopted by the present study.

2.5 Perception of English vowels by Native Russians

Perception underlies production, and, a few notable exceptions aside, many L2 production errors are thought to be mainly perceptual in origin, as a speaker must first perceive the difference between two phones before reliably producing it (Flege, Bohn, and Jang 1997)[16]. It is generally accepted that both naïve listeners and inexperienced L2 learners assimilate L2 phones to L1 phonemes because our early and extensive experience with those native phonemes causes them to serve as 'sieves' (Trubetskoy 1969) or 'magnets' (Iverson and Kuhl 1995; Kuhl and Iverson 1995) for the perception of L2 phones (see also Best 1995). In one example, Grimaldi et al. (2014) found that three studied groups of native Italian speakers - naïve listeners and 1st and 5th year academic learners of English – all assimilated L2 (English) phones to L1 (Italian) categories. This section presents a survey of studies (summarized in Table 2.18) investigating the perception of English vowels by native speakers of Russian – both monolinguals and Russian-English bilinguals – and the next section, 2.6, addresses the production of English vowels by Russian-English bilinguals.

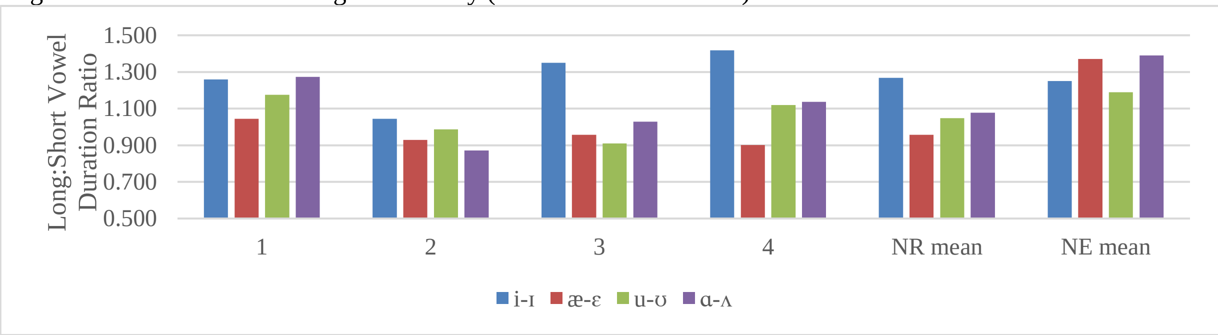
Table 2.18 Major findings of SLA studies investigating Russian-English L1-L2 pairing

study group	*source*	*finding*	*tasks*
monolingual Russian listeners	Gilichinskaya and Strange (2010)	English /i/ assimilated to Russian /i/ English /ɪ, ɛ/ assimilated to Russian /e/ English /æ, ɑ, ʌ/ assimilated to Russian /a/ English /ʊ/ assimilated to Russian /o/ English /u/ assimilated to Russian /u/ Least variation on /i, u, ɑ/	listening – assimilation of English vowels to NR categories
monolingual English, monolingual Russian talkers	Gilichinskaya and Strange (2010)	English /i/ classified most similar to Russian /i/ English /ɪ/ classified most similar to Russian /e/ English /ɛ, æ, ɑ, ʌ, ʊ/ most similar to Russian /a/ English /u/ classified most similar to Russian /u/	acoustic analysis – comparison of English and Russian vowels on F1/F2 plane
Russian-English bilinguals, North	Bell-Berti et al. (1998)	Most confusions with vowels adjacent in the vowel space Worst performance on /ɑ/, /ʌ/, /æ/, /ɔ/, /u/.	listening – identification of English vowels in NE speech
Russian-English bilinguals, North	Kondaurova and Francis (2004; 2008)	Reliance entirely on duration in distinguishing English vowels along the i-ɪ continuum	listening – identification of i-ɪ tokens varied in spectrum and duration

[16] See §2.7 below for a discussion of evidence against the perception-production link.

Russian-English bilinguals, North	Makarova (2010)	Most nativelike performance, least reliance on duration for /u/-/ʊ/ contrast Intermediate performance and reliance on duration on /i - ɪ/ Least nativelike performance, most reliance on duration for on /ɛ - æ/	listening – identification of tokens varied in spectrum and duration, /i/-/ɪ/, /u/-/ʊ/, /ɛ/-/æ/
Russian-English bilingual talker, monolingual English listeners	Bell-Berti et al. (1998)	Most confusions with vowels adjacent in the vowel space Worst performance on /ɪ ɛ ɔ ʊ ʌ/ among adjacent pairs, /i/-/ɪ/ and /ɛ/-/æ/ Most often confused with each other	listening – identification of English vowels in NR speech
Russian-English bilinguals, North	Romano et al. (1998)	Smaller spectral distinction for NR talkers on all four adjacent pairs NR participant who was least proficient at identifying vowels in the speech of a native speaker was also the talker whose vowels were least successfully identified by the NE listeners Duration contrast not acquired (see chart below)	acoustic analysis – comparison of NR and NE production

Duration Ratios of Four English Vowel Pairs for Four Russian-English Bilinguals (Romano et al. 1998) and English benchmark for same regional variety (Hillenbrand et al. 1995)



One recent study investigated the perception of English vowels[17] by monolingual Russian listeners (Gilichinskaya and Strange 2010). The majority of these listeners had academic experience with English in Russian high schools and some, at the college level, but none were Russian-English bilinguals. Only three had spent any time – between 8 months and 3 years – in the U.S. as adults, and none had lived in the U.S. as children. The results of the vowel assimilation part of the study showed the least variation in the perception of English /i/, /u/, and /ɑ/. These

[17] The native speaker who produced listening stimuli for Gilichinskaya and Strange (2010) was a New York native; stimuli consisted of /Vpə/ disyllables containing the vowels [i, ɪ, ɛ, æ, ɑ, ʌ, ʊ, u].

phonemes were assimilated most consistently, with at least 15 of the 16 Russian listeners assimilating them to Russian /i/, /u/, and /a/ at least 7 times out of 9 presentations. These responses constituted 100%, 93%, and 94% of total responses, with high mean goodness ratings of 8.0, 7.0, and 8.0 of 9. Unsurprisingly, acoustic analysis showed that these vowels were also most acoustically similar to Russian /i/, /u/, and /a/, respectively.

More variation was observed, with respect to the Russian vowel selected, the percent of total responses represented by each vowel, the number of participants who consistently chose that vowel, goodness ratings, or acoustic similarity to the chosen vowel, in the assimilation of the other tested vowels. English /ɪ/ was consistently assimilated to Russian /e/ by 14 participants (84% of total responses), with a mean goodness rating of 8.0, and acoustic classification was consistent with this assimilation pattern; however, one participant consistently (and confidently – with a mean goodness rating of 9.0) chose Russian /ɨ/ for English /ɪ/ (these responses comprised less than 10% of total responses and thus the percent value is not provided; they are unlikely due to the influence of phonological context used in the stimuli, as native Russian speakers rarely hear or use [ɨ] word-initially). English /ɛ/ was consistently assimilated to Russian /e/ by only 9 participants (66% of total responses) with a mean goodness rating of 7.0; acoustically, it was classified as most similar to Russian /a/, which no participants chose 7 or more times out of 9 presentations but which did receive 14% of the total responses, with a mean goodness rating of only 4.5, indicating that it was selected infrequently (6 or fewer times out of 9 presentations) by at least three participants. English /æ/ was consistently assimilated to Russian /a/ by only 6 participants (62% of total responses) with a mean goodness rating of 8.0; acoustically, it was also most similar to Russian /a/. English /ʌ/ was consistently assimilated to Russian /a/ by 9 participants (69% of total responses) and to Russian /o/ by one participant (20% of total responses, indicating that many others also infrequently chose

/o/) with a mean goodness rating of 6.0; it was also most similar to Russian /a/. English /ʊ/ was consistently assimilated to Russian /o/ by 12 participants (81% of total responses) with a mean goodness rating of 7.0; it was also most similar to Russian /a/, which did not receive enough responses in this category to be included in the results. The surprising determination that all English vowels but /i/, /u/, and /ɪ/ are most acoustically similar to Russian /a/ was made based on a linear discriminant analysis comparing of the centers of gravity of six Russian vowel categories (using a mean of the values of 3 speakers) and tokens of the eight tested English vowels (one speaker) in terms of F1/F2/F3 Bark values at the 50% point in each vowel. Importantly, the Russian listeners had a choice of nine Russian vowel sounds in the assimilation experiment (a choice probably made based on orthography – Russian has ten vowel letters), but only six vowels were utilized in the acoustic classification part of the study (likely motivated by the Russian vowel phoneme inventory, although [ɨ], which the authors acknowledge to be an allophone of /i/, is included)[18].

To summarize these results for the assimilation of English vowels by monolingual Russian listeners and their acoustic similarity to Russian phonemes, English /æ ɑ ʌ/ were most often assimilated to Russian /a/, English /ɪ ɛ/ to Russian /e/, English /ʊ/ to Russian /o/, English /u/ to Russian /u/, and English /i/ to Russian /i/ (Gilichinskaya and Strange 2010). The highest level of consistency was observed in the assimilation of English /i ɑ u/ (the same response was selected in 93% of cases or more) and the lowest level of consistency in the assimilation of /ɛ æ ʌ/ (the most common response was selected between 62% and 69% of cases); the vowels /ɪ/ and /ʊ/ had a moderate degree of consistency (84% and 81% agreement on the most common response, respectively). The most frequently chosen Russian vowel in the assimilation experiment also

[18] I was unable to reach Y.D. Gilichinskaya regarding this matter.

tended to be the closest in acoustic comparison, with the exception of English /ɛ/, which was assimilated to /e/ but was most acoustically similar to /a/, and English /ʊ/, which was assimilated to /o/ but was most acoustically similar to /a/.

Relatively few studies have investigated the role of perception in the acquisition of specific English vowel contrasts by NR speakers. In one pilot study, Bell-Berti et al. (1998) had four bilingual English/Russian participants identify 10 monosyllabic English words *heed, hid, head, had, hod, hawed, hood, who'd, hud, heard* with the vowels /i ɪ ɛ æ u ʊ ɔ ɑ ʌ ɝ/ as produced by a native speaker of American English. Results indicated that these Russian/English bilinguals varied in their ability to identify native English vowels in the speech of a native speaker[19], with most confusions occurring with vowels adjacent in the vowel space; exceptionally poor performance was noted for the /ɑ - ʌ/ pair, as well as /æ/, /ɔ/, and /u/.

Another study asked native Russian, Spanish, and American English listeners to identify stimuli from a *beat/bit* continuum (as well as in the native language for the non-native English participants) in order to examine the role of native language vowel duration use in the use of duration cues in non-native English (Kondaurova and Francis 2004; 2008). The researchers found that NR learners of English do not appear to have access to the relevant features used by native speakers[20] in distinguishing English vowels along the i/ɪ continuum and instead rely entirely on duration as a default contrast, despite its absence from the L1 for both non-native groups tested. (Kondaurova & Francis 2004; 2008). Some reliance on duration is expected and desirable, given the correlation between tenseness and length in English vowel phonemes. However, any reliance on duration at all, and especially overreliance, is perhaps somewhat surprising among native

[19] The regional variety spoken by the native speaker in Bell-Berti, Romano, and Lorin (1998) is not reported.
[20] Recall that native English speakers rely mostly on spectral cues to distinguish the high tense/lax vowel pairs and treat duration cues as secondary - see Hillenbrand, Clark, and Houde (2000)

speakers of both Russian and Spanish, given that duration is not distinctive in either language. Thus, such a strategy could not be explained as influence from the L1 and may instead represent a language universal. Indeed, it has been suggested that duration is more salient than fine spectral differences, even to learners without previous L1 experience with duration cues or variability in that specific part of the vowel space (Bohn 1995).

A recent dissertation (Makarova 2010) was one of the first perceptual studies to examine cue weighting in the acquisition of English vowel contrasts by native Russians, investigating new categories formed by the addition of the features [Low] and [Tense]. This study reported on the reliance on vowel duration and spectral cues in the categorization of the high front (/i - ɪ/), high back (/u - ʊ/), and mid and low front (/ɛ - æ/) English vowels by NR learners. Makarova compared the performance of 15 monolingual native speakers of English and 38 Russian-English bilinguals on a listening task in which listeners identified manipulated stimuli varying along five-point spectral and duration continua.

Like Bell-Berti et al. (1998), who found that most confusions occur with vowels adjacent in the vowel space, Makarova (2010) found that the distinctions between English /i - ɪ/, /u - ʊ/, and /ɛ - æ/ are indeed difficult for Russian learners. Moreover, Makarova's (2010) perception results partially aligned with those of Kondaurova and Francis (2004, 2008) in that learners *initially* displayed overreliance on duration in the perception of all three contrasts tested.[21] However, as a group, Makarova's (2010) Russian-English bilinguals demonstrated a pattern of acquisition more similar to that reported by Escudero (2000) for native-Spanish learners of English: a set of developmental stages beginning with (0) no distinction between /i/ and /ɪ/, progressing to (1) a

[21] Overreliance on duration was greatest for the /ɛ/-/æ/ pair and least for the /u/-/ʊ/ pair (Makarova 2010). It may be noteworthy that duration differences themselves are also greatest for the /ɛ/-/æ/ pair and least for the /u/-/ʊ/ pair in Makarova's (2010) stimuli (duration of stimuli was adjusted to match values in Hillenbrand *et al.* 1995) – 1.47 to 1.0 for /æ/ - /ɛ/, 1:29 to 1.0 for /i/ - /ɪ/, 1.23 to 1.0 for /u/ - /ʊ/.

stage where distinction based entirely on duration (which most closely matches the findings of Kondaurova & Francis 2008), then to (2) a stage where distinction is based on both spectrum and duration with greater reliance on duration, and ending with 3) distinction based on both spectrum and duration with greater reliance on spectrum (most nativelike weighting of acoustic cues).

It is surprising that Makarova (2010) and Kondaurova and Francis (2008) would report such different results with listeners from the same L1 background and with similar LOR in an L2 speaking community (mean LORs of 2 years and 3.8 years, respectively; see discussion of the role of LOR in **§**2.3). One possible reason for the difference in findings between these two studies is the contrasts under investigation: while Makarova (2010) examined listening performance on three vowel contrasts, Kondaurova and Francis (2008) focused on only one, /i - ɪ/, which was associated with intermediate results in Makarova's (2010) findings: more reliance on duration than in the /u - ʊ/ pair but less than the /ɛ - æ/ pair. Perhaps the differential performance on the three vowel contrasts was enough to tease out the role of LOR in Makarova's study. Makarova's findings receive some additional conceptual support from the results of Kim et al. (2017), who also found stages of acquisition similar to that reported by Escudero (2000), with performance on the /i - ɪ/ and /ɛ - æ/ contrasts among native-Korean learners of English in their first year of residence in an L2 speaking community, tested on arrival to Canada, and then at 4 months, 8 months, and 12 months LOR. Whatever the reason for the dissimilar results of these studies, I hypothesize that the participants in the present study, due to their higher LOR (see Chapter 3) than the participants of Kondaurova and Francis' (2008), Makarova's (2010), or Kim et al.'s (2017) studies[22] are **less** likely to be in Escudero's (2000) earlier stages of acquisition (no distinction or distinction based entirely on duration) or to show effect of LOR at all (recall that LOR is thought by some scholars

[22] Escudero (2000) does not provide LOR for her participants.

to lose its predictive power for L2 foreign accent after one year of residence, particularly when not controlling for factors such as amount and quality of input: see Krashen, Long, & Scarcella 1979; Long 1990; Flege et al. 1992; Flege, Munro, & Fox 1993; Munro 1993), and **more** likely to rely both on spectrum and duration and to be in Escudero's (2000) later stages (reliance on both spectral and duration cues to varying degrees; see **§**2.9 for present study's research questions and hypotheses).

Additionally (and somewhat surprisingly, as the two are not differentiated well in English orthography, appear in different phonotactic environments, and have relatively few minimal pairs), Makarova (2010) reports that her NR listeners performed best on the /u - ʊ/ contrast; they were less nativelike in perceiving the /i - ɪ/ contrast, and least nativelike with /ɛ - æ/.[23] Makarova (2010) suggests that her native Russian participants performed best on the /u - ʊ/ contrast because Russian speakers are especially sensitive to variation in vowel quality in the high back corner of the vowel space due to consciously subdividing it to accommodate the presence of [ɨ], a high central allophone that is so prominent in the Russian sound system that it is represented orthographically and is considered by some linguists to be a distinct phoneme (Bondarko 1998; Yanushevskaya and Bunčić 2015 - see §2.2.1). Indeed, in at least some dialects of English (e.g., California), [ʊ] is unrounded and pronounced with spread lips (Ladefoged 1999: 43), making it more perceptually similar in terms of articulatory gestures to Russian [ɨ], which could be argued to be acoustically similar in at least some Russian dialects (e.g. Moscow and Minsk, as opposed to [ɨ] in Kiev, which is more acoustically similar to English [ɪ] – see Holden and Nearey 1986), and which, unlike other allophones, is easily identified and produced in isolation by native Russians (Ščerba 1912). Some NR learners, particularly those whose native [ɨ] is indeed central and who are getting their English

[23] It is important to note that Makarova (2010) looked only at L2 perception, not production.

input in a variety of English that produces such an unrounded [ʊ], may transfer their native /u/, which is rounded, onto English /u/, and utilize the spread lip articulation of [ʊ] to assimilate it to [ɨ], creating an additional salient cue to distinguish it from /u/.

2.6 Production of English vowels by Native Russians

Compared to the number of *perception* studies, even fewer sources attempt to provide a detailed acoustic analysis of L2 English vowels *produced* by native Russians. In one small pilot study referenced above in the discussion of the existing literature on NR perception, Bell-Berti et al. (1998) also found that native English listeners had difficulty identifying L2 English vowels produced by native-Russian/English bilinguals, particularly /ɪ ɛ ɔ ʊ ʌ/. Most confusions occurred between vowels adjacent in the vowel space rather than more distant vowels separated by other vowel phonemes. Among the adjacent pairs, the /i/-/ɪ/ and /ɛ/-/æ/ pairs were most often confused with each other. The most problematic individual phoneme was /ʌ/, identified by native listeners as /ɑ/, /æ/, and even /ɛ/.

As part of the same pilot study described above, Bell-Berti et al. (1998) analyzed four native Russian talkers' production of 10 monosyllabic English words *(heed, hid, head, had, hod, hawed, hood, who'd, hud, heard)* with the vowels /i ɪ ɛ æ u ʊ ɔ ɑ ʌ ɝ/ via 14 native English listeners' judgments. In general, /ɪ ɛ ɔ ʊ ʌ/ were poorly identified by native listeners, and most confusions, as in the NR perception task, occurred with vowels adjacent in the vowel space. Acoustic analysis of the same native Russian speakers' production (Romano et al. 1998) showed that these subjects did not make as great a spectral distinction between members of the vowel pairs /i - ɪ/, /ɛ - æ/, /ɑ - ʌ/, and /u - ʊ/ as do native speakers[24], and produced the four pairs with considerable F1/F2 overlap.

[24] This study used formant values reported in Peterson and Barney (1952) as a native baseline.

Finally, Bell-Berti et al. (1998) report that the NR participant who was least proficient at identifying vowels in the speech of a native speaker was also the talker whose vowels were least successfully identified by the NE listeners (see Perception-Production link addressed next in §2.7). These results are consistent with the findings of perception studies reviewed in the previous section.

Furthermore, the native Russians as a group appeared not to have acquired the secondary duration cue to the tested contrasts, a finding that is somewhat surprising given that duration is often argued to be perceptually more salient than spectral cues (Bohn 1995). However, with no biographical data (such as AOA and LOR) on the four Russian-English bilinguals in Bell-Berti et al.'s (1998) study, it is impossible to determine whether these results are truly incongruent with previously published data. Additionally, the four talkers in this pilot study exhibited a considerable range of individual variation. All speakers produced at least a slightly longer /i/ than /ɪ/; however, for one speaker, /i/ was only 4% longer than /ɪ/, which is not a duration difference that could be detected in normal speech and most likely represents a lack of duration distinction for this contrast by this talker, while for the other three, /i/ was 26%, 35%, and even 42% longer. /ɛ/ and /æ/ were much closer in duration for all four speakers (one never more than 11% longer than the other), suggesting no intended duration difference for this pair. Interestingly, /æ/ was longer than /ɛ/ for only one speaker; for the other three /ɛ/ was just slightly longer (again, most likely not a duration distinction intended by the speaker or discernible in regular speech) than /æ/, the typically longer vowel. For the /ɑ - ʌ/ pair, /ɑ/ was longer for three of the four speakers (by 27%, 14%, and 3% respectively, with the latter and possibly two latter distinctions likely imperceptible in regular speech and not indicative of an intended duration distinction), while /ʌ/ was 15% longer than /ɑ/ for the fourth speaker, a surprising outcome. For the /u - ʊ/ pair, /u/ was longer for two speakers

(18% and 12%), while /ʊ/ was 11% and 1% longer than /u/ for the other two (likely imperceptible and not indicative of an intended duration distinction).

The literature suggests that differences of at least 10% and perhaps even 20% are the smallest that can be perceived in natural speech (Lehiste 1970; Klatt 1976). Thus, these very small duration differences of 4% (one talker's /i - ɪ/), 3% (one talker's /ɑ - ʌ/), and 1% and 11% (two talkers' /u - ʊ/), as well as all observed duration differences between /ɛ/ and /æ/ in this study (≤11%) are almost certainly imperceptible in normal speech; slightly greater duration differences of 14% (one talker's /ɑ - ʌ/) and 12% (one talker's /u - ʊ/) may also fail to be perceived outside the context of an acoustic study. These most likely imperceptible differences, which probably indicate that the speakers are not making an intentional duration distinction between the paired vowels at all, constitute ten of 16 possible cases (four contrasts produced by four talkers), leaving only six cases where it might be reliably said that the talkers are making an intentional duration distinction in the vowel pair: three for /i - ɪ/, two for /ɑ - ʌ/, and one for /u - ʊ/.

These results are conceptually inconsistent with other existing perception studies; for example, based on Makarova's (2010) finding that native Russian listeners demonstrated greatest overreliance on duration to distinguish the /ɛ - æ/ pair compared to other contrasts tested, we might expect Bell-Berti et al. (1998) to have found considerable duration differences rather than very similar duration values between /ɛ/ and /æ/ in production. However, since this pilot study included only four participants, individual differences may be obscuring a larger pattern. Another important consideration stems from the very different methodologies of these two studies and a critical review of the perception-production link – the rather intuitive but not unquestionable idea that accurate perception of a contrast is necessary (and, in some interpretations of the hypothesis, sufficient) for the successful production of that contrast. That is, it is impossible to say how the four talkers in Bell-Berti et al.'s (1998) study might have performed on a listening task such as used by Makarova's (2010) study, and impossible to say whether an acoustic analysis of vowels produced by Makarova's (2010) listener would have yielded the same results for vowel pair duration. Reliance on duration cues in perception is not guaranteed to translate to comparable duration differences in production of the same contrasts. This idea is addressed in the next section.

2.7 The Perception-Production Link

Having discussed connections between the perception and production of English vowels by native Russian speakers in the previous section, it is important to acknowledge the recent criticisms brought against the Perception-Production Link – the idea that perception necessarily precedes and is a requisite for production.

The perception-production link is well-supported by a considerable number of studies investigating L2 phonological acquisition, which find: 1) links between perception and production accuracy (Flege et al. 1999a; Jia et al. 2006); 2) improvements in perception followed by improvements in production (Bradlow et al. 1997; Aoyama et al. 2004), and 3) perception exceeding production or production lagging somewhat behind perception (Baker and Trofimovich 2006; Cardoso 2007, 2011).

However, a handful of attested cases of accurate L2 perception paired with considerable production difficulties (Chan 2014 – English /ð/; Derwing and Munro 2015 – Spanish /r/) may be interpreted as a challenge to the idea that accurate perception of a contrast supports its production. On the one hand, these findings may indicate that perception and production are controlled by fundamentally different mechanisms and are not as closely connected as posited by the perception-production link (Baker and Trofimovich 2006). On the other hand, these findings may simply suggest that accurate production of some sounds, while still relying on their accurate perception, may involve *additional* articulatory considerations not involved in their perception – that is, learners may accurately perceive an L2 sound but be unable to produce it if doing so requires the implementation of articulatory gestures to which they are unaccustomed. In the second interpretation, these findings are not devastating to the perception-production link, as it may be argued that, additional articulatory production requirements aside, production of these sounds still relies on a foundation of their successful perception, but "accurate production requires *something more than* accurate perception" (Isbell 2016, p.61, emphasis added). One possibility, proposed by Baker and Trofimovich (2006), is that for at least some learners, accurate *self-perception* underlies successful L2 production.

An even smaller subset of studies potentially pose a greater challenge to the fundamental idea that perception precedes and supports production. In one well-known example of production without perception, Sheldon and Strange (1982) reported more accurate production than perception of English /r/ and /l/ by native Japanese learners; however, the small sample size and the possibility of a ceiling effect in the test group may at least partly explain these idiosyncratic findings (Isbell 2016). Tsukada et al. (2005) and Darcy and Kruger (2012) found more nativelike L2 production than perception among child bilinguals; however, these findings may be better evidence of the unique status of early bilinguals and heritage learners as compared to monolingual peers and a need for more congruent perception and production testing tasks than evidence against the perception-production link (Isbell 2016).

Overall, the perception-production link, while certainly not undisputed or undeserving of reevaluation as new findings come to light, enjoys both strong support from empirical research and a strong theoretical foundation within the field of SLA, including Flege's (1987, 1988, 1991)

Speech Learning Model (discussed next, in §2.8) and Best and Tyler's (2007) Perceptual Assimilation Model. Thus, the current work assumes that perception must precede production and, idiosyncratic articulatory challenges aside, that the degree of perceptual accuracy influences the accuracy of production.

2.8 Theoretical Foundations

The following section outlines the theoretical foundation adopted in this work: the Speech Learning Model **(SLM)** (Flege 1987, 1988, 1991), which is rooted in the idea that acoustic-phonetic similarity between L1 and L2 sounds can predict their acquisition by L2 learners. The SLM was selected as the theoretical framework for the present work for two reasons: 1) it is uniquely well-suited to an analysis of the acquisition of L2 vowels in that it makes specific predictions about their acquisition based on acoustic distance between prototypical exemplars of the L1 and L2 vowel systems, and 2) unlike other models which address early stages of SLA (e.g. PAM, see Best 1995; Best & Tyler 2007), the SLM was developed to address L2 performance at advanced stages of acquisition.

In order to address why adult learners generally do not achieve fully nativelike pronunciation of all L2 phones, Flege's (1987, 1988, 1991) SLM compares the sound systems of the L1 and L2. Flege argues that perception and production of L2 vowels depends on their acoustic similarity to one another and to L1 vowels. The SLM uses distance in Hz between two prototypical vowels on the perceptual (F1-F2) plane as a direct and objective measure of vowel similarity (Flege and Munro 1994); when acoustic distance between exemplar vowels is small, the SLM predicts that the acoustic cues (and presumably the associated features) differentiating the vowels will be difficult for L2 learners to perceive and acquire (Flege 1987, 1988, 1991; Flege and Munro 1994).

Thus, per the SLM, L2 vowels occupying overlapping or closely adjacent space on the perceptual plane are more likely to be perceived by L2 learners as members of one category and not acquired as separate L2 categories (Flege 2005).

The model also posits that it is difficult for L2 learners to form a new (L2) phonetic category that overlaps with but remains distinct from an existing (L1) category, 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, but only partly overlapping with what exists in the L1, learners may dissimilate the phones from one another by increasing the difference between them (e.g., by raising one vowel and lowering the other, even if doing so causes each vowel to diverge from monolingual production in either language) (Flege 2005). Splitting one L1 category to accommodate two or more L2 phones in that same part of the vowel space is proposed to be particularly difficult for the adult learner (Flege 2005), as learners are not accustomed to subdividing that part of the vowel space in the L1.

A common criticism of the SLM is that it provides no precise measure of similarity for phones. In the case of vowels, this work assumes that measurable differences in first and second formant values and duration may be used as a measure of similarity of phones. Euclidean distance (Fox and Jacewicz 2009) is used here to assess distance between vowels on the F1/F2 plane; duration ratios between vowels known to differ in duration in native speech are used to assess how those duration relationships are treated by non-natives. Additionally, it may be possible to generalize predictions of the SLM to include features. 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 et al. 1997). If phonemes are regarded as bundles of distinct features (some associated with specific acoustic or articulatory dimensions, others more abstract), this model may indeed address features in addition to purely perceptual similarities in order to predict and explain L2 phonological acquisition (Flege et al. 1995; 1999a; 199b; 2003; Flege & Mackay 2004).

It is worth noting that the suggestion that L2 learners utilize familiar L1 contrasts to hone their L2 perception and production has not gone without criticism, particularly with reference to features. Earlier accounts of SLA held that L2 learners do not have (and cannot acquire) access to new L2 features not present in the L1. Flege and Port (1981), for example, suggest that most L2 learners have difficulty perceiving and identifying L2 features not phonemic in their L1 system and may not be able to re-combine abstract features already present in the L1 to access feature combinations unique to the L2 in order to produce a new L2 segment natively. Native-Arabic speakers of English in Flege and Port's (1981) study of cross-language phonetic interference in Saudi Arabians' production of English stops grasped that 'the phonological nature of /p/' (1981: 125) was 'equivalent to a "voiceless /b/" or a "labial /t/"' (143) based on feature combinations present in Arabic but were nonetheless 'unable to control all articulatory dimensions by which this sound is produced' (125). Note that Flege and Port's (1981) participants were not entirely unaware of the featural basis of the /p - b/ contrast; rather, they made some distinction between the two but were unable to produce /p/ in a native-like manner, suggesting to me that they had some access to the relevant features but were prevented by additional factors such as coordination and phonotactic constraints from producing the contrasts in the given environments.

Other studies appear to show evidence suggesting that L2 learners can successfully redeploy L1 features to build new segmental representations in the L2 (Brown 1998; 2000; Archibald 2005). Based on the ability of native speakers of Japanese and Chinese (neither of which has the /ɹ - l/ contrast) to distinguish English /ɹ/ from /l/, Brown (1998) concluded that the Chinese speakers were relying on the feature [CORONAL] (distinctive in Chinese to distinguish some dental/alveolar liquids, fricatives, and affricates from retroflex variants) while native Japanese speakers, whose L1 only contrasts coronal stops and fricatives via voicing and does not utilize the feature [CORONAL] as distinctive (see Clements 2001: 124), had no access to this contrastive feature at all. Importantly, Brown's (1998) observation comes from onset and cluster experimental conditions; in coda position, native speakers of Japanese were quite successful at discriminating /ɹ/ from /l/, suggesting that additional factors are involved, such as the strong effect of English /ɹ/ on the preceding vowel (Brown 1998). Additionally, it is also possible to interpret Brown's (1998) results as evidence that Chinese listeners map English /l/ and /ɹ/ onto two different L1 categories – /l/ and /ɻ/ – while native-Japanese listeners map them onto a single category – /r/ (Escudero 2007).

Yet other accounts offer limited evidence that at least some L2 learners can implement some L1 allophones as phonemes in an L2 or develop access to new features in the L2 after all (Escudero and Boersma 2004; Vokic 2010, Barrios et al. 2016). Escudero and Boersma (2004) concluded based on the distinct patterns observed in the acquisition of the /i - ɪ/ contrast by native-Spanish learners of two varieties of English that it is possible even for adult learners to fully access their language learning capacity and re-rank constraints based on incoming L2 data to acquire L2 contrasts in a manner similar to L1 acquisition. Vokic's (2010) results with native English Spanish learners' [ɾ] and [ð] (whose phonological relationship and distributional properties differ between

English and Spanish) suggest that access to L1 allophones is limited but possible, particularly given favorable conditions in L1 orthography, stage of L2 acquisition, functional load, and motivation. Barrios et al. (2016) argue against the position that L2 learners are constrained in their acquisition of an L2 phonology by the phonological features made available by the L1. In their study of native-Spanish learners of English, features available in the L1 were not redeployed to acquire L2 vowel contrasts, consistent with the earlier findings of Flege and Port (1981); however, L2 features not available in the L1 were, it seems, acquired by some learners. Based on these findings, Barrios et al. (2016) conclude that "feature availability is neither a necessary, nor a sufficient, condition to predict the observed learning outcomes" (p. 367).

The present work assumes that if it is indeed possible for L2 learners to redeploy L1 features or acquire new L2 features, the acquisition of a nonnative phonological system is nonetheless a challenging process, and rates of successful development of L2 phonological features are expected to vary. This may explain why some studies (e.g. Flege and Port 1981) failed to find evidence of successful development of L2 segments via successful acquisition of novel L2 features or from a recombination or redeployment of existing L1 features. Even Flege and Port's (1981) results suggest some awareness on the part of the participants of the relevant features.

Thus, the position adopted here is one that combines the more recent accounts of feature redeployment and development: that it may indeed be possible for at least some learners to redeploy features from the L1 or develop access to new features in the L2, particularly given favorable conditions. These include two conditions cited by Vokic (2010), specifically L2 orthography and functional load[25]. English orthography least reliably distinguishes /u - ʊ/ among the four vowel contrasts tested in the present study, and functional load for this contrast is also

[25]The present study did not measure learner motivation or stage of acquisition other than to ensure that participants had lived in the desired location for at least five years – see Chapter 3 for a description of study participants.

relatively low compared to the other three contrasts in that there are very few minimal pairs. Based on this position, it is possible that, of the four contrasts tested, acquiring the features necessary to perceive and produce the /u - ʊ/ contrast might be the most challenging for non-native speakers, particularly if the local variety minimizes the distance (and thus perceptual difference) between these two phonemes on the F1-F2 plane or in terms of duration, as compared to a more generalized standard. On the other hand, if abstract features are more challenging for learners to transfer and manipulate, then perhaps those which have easily identifiable visual correlates – e.g. [round], as in the case of /u - ʊ/ in varieties where one is unrounded – offer some advantage to this pair. Particularly in the case of vowels, where distance between prototypical members of categories can be objectively measured based on differences between the F1, F2, and duration of members of the vowel pair, perceptual differences in spectral and duration cues may be better predictors of ease of acquisition than features, which can be quite abstract.

2.9 Research Gap

It is clear from a review of the literature that unanswered questions remain about the process of L2 phonological acquisition, specifically the acquisition of English vowels by native Russian learners, who face considerable challenges due to the differences between the two languages' sound systems. As demonstrated above, little is known about cue weighting in the acquisition of American English vowel contrasts by native Russian learners. It is known that Russian-English bilinguals may struggle to discriminate vowels adjacent in the vowel space both in perception and in production (Bell-Berti et al. 1998; Romano et al. 1998), may over-rely on duration in perception of some contrasts that rely on both spectral and duration cues (Makarova 2010; Kondaurova & Francis 2004; 2008), and may acquire contrasts predicted to be equally challenging based on associated features at different rates (Makarova 2010). However, few studies investigate all problematic vowels and contrasts (e.g. three contrasts in Makarova 2010; /i - ɪ/ only in Kondaurova & Francis 2004; 2008) and few examine both perception and production in the same research subjects. One pilot study that attempts this does so with only four participants (Bell-Berti et al. 1998; Romano et al. 1998), and the most in-depth study of this topic to date (Makarova 2010), which implements a rigorous methodology to investigate the acquisition of three specific vowel contrasts among a large group of native Russian speakers of English living in northern U.S., deals with perception only. This gap in the literature can only be addressed by studies that examine both perception and production of multiple challenging vowel contrasts by the same research subjects.

The present study attempts to help fill this gap in the fields of SLA, phonetics, phonology, and Russian and English linguistics by conceptually replicating the two existing studies on the subject of acquisition of English vowels by native speakers of Russian. Makarova's (2010) perception findings are replicated here with both perception and production data from the same NR participants, perception of those participants' production by NE listeners, and the addition of a fourth vowel contrast not tested by Makarova (2010). The pilot study reported on by Bell-Berti et al. (1998) and Romano et al. (1998) is replicated by the present study with considerably more participants and a methodology targeted at testing specific vowel contrasts. The primary aims of replicating Makarova's (2010) study are to test, with a different methodology, the reliability of these earlier perception findings (do Russian-English bilinguals perform best on the /u/ - /ʊ/ contrast compared to /i/ - /ɪ/ and /ɛ/ - /æ/?) as well as their generalizability (does the same result hold for production as well as perception, and can findings be generalized to native Russians acquiring southern English, as opposed to a northern variety?). The urgent need for replication of existing studies, be it “exact,” “approximate,” or “conceptual” replication (Porte 2012: 8) is widely acknowledged in the field of linguistics (Plonsky 2012); the subfield of SLA is particularly “ripe for replication” (Mackey 2012: 29), including questions pertaining to L2 processing, acquisition, and development (Mackey 2012).

Additionally, this dissertation contributes to the understanding of the relationship between perception and production in second language acquisition and to the literature on salience of acoustic cues (vowel quality vs. quantity) for native and non-native listeners. Finally, this work also contributes to the literature on native pronunciation in Southern English and addresses the question of whether non-native speakers appear to acquire a local variety rather than a generalized standard. To accomplish this, I conduct three experiments (see Chapter 3 for experiment design) to address the following seven research questions, previewed in Chapter 1:

Question 1 – Duration Distinction: Do Russian-English bilinguals produce the American English contrasts /i - ɪ/, /u - ʊ/, /ɛ - æ/, and /ɑ - ʌ/ with a ***different duration ratio*** between the typically longer and typically shorter vowel of each pair than do monolingual native speakers of American English? If so, does the effect vary in magnitude among the four vowel pairs?

Recall from the literature review above that Romano et al.’s (1998) Russian-English bilinguals made a slightly greater duration distinction than did natives only for the high front contrast (where they also had the most nativelike performance in terms of duration ratios between typically longer and shorter adjacent vowels). Kondaurova and Francis’ (2004; 2008) NR listeners relied entirely on duration in discriminating the /i - ɪ/ contrast, and Makarova’s (2010) native Russian listeners over-relied, to varying degrees, on duration in discriminating the three English contrasts /u - ʊ/, /i - ɪ/, and /ɛ - æ/; if *reliance on duration* in *perception* of vowel contrasts is related to the *duration* *ratio* between the long and short vowels of those contrasts in *production* (that is, those contrasts the perception of which relies most on duration will also have the greatest duration difference in their production), then it is possible that Russian-English bilinguals may produce a greater duration distinction between the vowels of the four tested contrasts than do native talkers.

Question 2 – Spectral Distinction: Do Russian-English bilinguals produce the American English contrasts /i - ɪ/, /u - ʊ/, /ɛ - æ/, and /ɑ - ʌ/ with a ***different spectral distance*** between the vowels in each contrast than do monolingual native speakers of American English? If so, does the effect vary in magnitude among the four vowel pairs?

All four of Romano et al.'s (1998) Russian-English bilingual talkers had a smaller spectral distance between the vowels of the four vowel pairs tested by the present study. Additionally, casual observation of the speech of accented Russian-English bilinguals suggests these talkers may make a smaller spectral distinction between these vowels on the F1/F2 plane.

Question 3 – Overall Accuracy: Are Russian-English bilinguals ***more successful*** at perceiving or producing certain American English vowels than others, as measured by nativelikeness in quality (F1/F2) and quantity (duration), as well as by the ability of native English listeners to identify vowels in their speech?

Makarova's (2010) Russian-English bilingual listeners performed most like natives on the u/-/ʊ/ contrast, had intermediate performance on /i - ɪ/, and were least nativelike in their use of duration cues, displaying greatest reliance on duration, for on /ɛ - æ/ contrast. Bell-Berti's (1998) native-Russian listeners performed worst at discriminating the vowels /ɑ/, /ʌ/, /æ/, /ɔ/, /u/ in native speech, while the native-English listeners performed worst at discriminating the vowels /ɪ ɛ ɔ ʊ ʌ/ and particularly the /i/-/ɪ/ and /ɛ/-/æ/ contrasts in the speech of Russian-English bilinguals.

Question 4 – Spectrum vs. Duration: Do native-English listeners more accurately identify English vowels produced by Russian-English bilinguals when they match native values more in ***spectral*** or ***duration*** cues; that is, which cues contribute more to overall comprehensibility?

Native, monolingual speakers of English typically rely more on spectral cues and only marginally on duration cues to distinguish the English vowel contrasts characterized by both spectral and duration differences, such as the tense and lax vowel pairs in the same part of the vowel space (Peterson and Barney 1952; Hillenbrand et al. 1995; Hillenbrand et al. 2000).

Question 5 – Perception vs. Production: Are native-Russian listeners more successful at distinguishing some vowel contrasts than others in native speech, and do the native-Russian listeners who more accurately discriminate certain vowels in native speech also ***perform*** those contrasts better?

The perception-production link is supported by studies which report: 1) links between perception and production accuracy (Flege et al. 1999a; Jia et al. 2006); 2) improvements in perception followed by improvements in production (Bradlow et al. 1997; Aoyama et al. 2004), and 3) perception exceeding production or production lagging somewhat behind perception (Baker and Trofimovich 2006; Cardoso 2007, 2011).

Question 6 – Local Variety: Based on spectral and duration analysis of their speech compared with the present study's native speakers and previously published data from northern and western varieties, do Russian-English bilinguals appear to be acquiring a ***local variety*** of American English or a generalized standard?

Second language learners have been demonstrated to actively seek out and imitate a local variety, not a generalized standard, and ESL speech has been shown to align with local norms both in terms of the social group with which learners associate (Anisman 1975; Thompson 1976; Adamson & Regan 1991) and more general parameters like regional pronunciation (Wolfram et al. 2004; Friesner & Dinkin 2006).

Question 7 – AOA and LOR: Do age of arrival to and length of residence in an English-speaking community correlate with Russian-English bilinguals' performance, either in perception, production, or both?

AOA has frequently, but not universally, been demonstrated to correlate negatively and LOR to correlate positively with accuracy of second language performance (Aoyama et al. 2004; Flege and MacKay 2004; Tsukada et al. 2005; Oh et al. 2011; Bialystok 1997; Piske et al. 2001).

2.10 Conclusion

This chapter provided background information – including a comparison of the languages in question, a review of relevant literature, and the main theoretical approaches employed – as well as a discussion of the research questions employed by the present study. The next chapter presents the methodology of the current work, detailing how the present study aims to address the questions formulated in Chapter 2 by examining perception and production of four American English vowel contrasts among native speakers of Russian.

CHAPTER 3

RESEARCH DESIGN

3.1 Introduction

This chapter describes the research design implemented in the present study. It begins with a description of the study participants, outlines the motivations based on which the study was designed, and describes the stimuli and prompts, procedure, and analysis for each experiment. Data collections methods were approved in 2017 by the Human Research Committee at the University of Georgia (STUDY00004934).

3.2 Participants

Twenty-nine participants took part in the experiments: 20 Russian-English bilinguals whose native language was Russian (this group henceforth **NR**) and 9 monolingual native speakers of English (henceforth **NE**)[26]. NR participants were recruited through internet advertising and word of mouth and screened via an eligibility questionnaire to ensure they met the following criteria:

1. 18 years of age or older at the time of data collection
2. Native speaker of Russian
3. Arrived in the US (AOA) and began interacting with native English speakers after age 12
4. Not native or fluent in any languages other than Russian or English

[26] A tenth native English speaker completed the listening portion of the study but was unable to complete the speaking portion due to having lost her voice just prior to data collection; she was excluded from the analysis.

5. Did not intensively study or fluently speak a language other than Russian before age 12
6. Lived in the Atlanta, GA area for at least 5 years immediately prior to data collection

NE participants were also recruited through internet advertising and word of mouth and screened via an eligibility questionnaire to ensure they met the following criteria:

1. 18 years of age or older at the time of data collection
2. Native speaker of English
3. Not native or fluent in any languages other than English
4. Did not intensively study or fluently speak a language other than English before age 12
5. Lived in the Atlanta, GA area for at least 5 years immediately prior to data collection

3.2.1 Russian-English Bilinguals (NR)

Background information on the NR participants is summarized in Table 3.1. This group was assigned random letter identifiers; the letters 'i' and 'o' were avoided to eliminate confusion with the numbers '1' and '0'. NR participants ranged in age from 21 to 52 years at the time of data collection (mean = 38.7 years; SD = 7.7 years) and included 16 females and 4 males. Participants had arrived in the United States between the ages of 13 and 32 (AOA mean = 22.7 years; SD = 4.9 years) and had spent between 5 and 27 total years in the U.S. (LOR U.S. mean = 16.0 years; SD = 4.5 years) and between 5 and 27 years in Atlanta, Georgia immediately prior to data collection (LOR GA mean = 14.4 years; SD = 5.7 years). A few had lived in other U.S. states prior to Georgia, and most were exposed to other languages academically or naturalistically but none were native or fluent in any language other than Russian or English.

Table 3.1 Russian-English Bilingual Participants (NR)

ID	*Age*	*Sex*	*Born*	*Academic L*	*Natural L*	*AOA*	*LOR US*	*LOR GA*	*Before GA*
A	42	M	Russia	Uzbek	Uzbek	23	19	19	
B	44	F	Russia	French		26	18	12	NY
C	45	F	Ukraine	German Georgian	Georgian Armenian Yiddish	21	24	24	
D	42	F	Russia		Bengali	26	16	16	
E	36	F	Ukraine	French	Ukrainian	19	17	14	KS, TN
F	44	F	Belarus	German Belarussian	Belarussian	28	15	15	
G	41	F	Ukraine	Ukrainian	Ukrainian	27	14	11	KY
H	34	F	Russia	German		21	13	10	FL
J	24	F	Russia	Spanish	Armenian	14	10	10	
K	49	F	Moldova	Romanian	Romanian	32	17	17	
L	44	M	Russia	Ukrainian	German Polish Portuguese	27	17	17	
M	37	F	Russia	French German Spanish Latin		21	17	17	
N	42	F	Uzbekistan	Uzbek	Uzbek Spanish	29	13	5	TX, NJ, WA
P	21	M	Russia	French German		17	5	5	
Q	39	F	Crimea	Ukrainian	Ukrainian	22	15	15	
R	35	M	Crimea			21	14	5	MD, CA
S	36	F	Latvia	Latvian	Moldovan	21	15	15	
T	29	F	Moldova	Spanish Moldovan		13	16	16	
U	38	F	Crimea	Spanish	Ukrainian	21	17	17	
V	52	F	Russia	German		25	27	27	

Explanation of column names:

“ID”: unique identifier assigned to each participant
“Age”: participant age, in years, at time of data collection
“Sex”: participant sex
“Born”: participant country of birth
“Academic L”: languages participants had studied academically, if any, other than English or Russian
“Natural L”: languages participants had been exposed to in a naturalistic setting, if any, other than English or Russian
“AOA”: age, in years, at which participants arrived in the United States
“LOR in US”: number of years participants lived in the USA
“LOR in GA”: number of years participants lived in Atlanta, GA immediately prior to data collection
“Before GA”: other states in the US where participants had lived before Georgia, if any

The majority of NR participants (C, D, E, G, J, K, M, N, P, Q, R, S, U) started academic study of English around grade 5 (ages 9-12) of elementary school[27] in their countries of origin, as is typical in schools in many Russian-speaking countries; however, they were not exposed to native English speakers until their arrival in the US. One participant (T) started academic study of English in school at age 5, but was also not exposed to native speakers of English until her arrival in the US. Participants A, B, F, H, L, and V started their acquisition of English considerably later (at 17 to 27 years of age, average = 22.8), in a mix of academic and naturalistic settings. Three (A, L, V) did not study English at all before arriving in the US as adults (at ages 23, 27, and 25, respectively). As a group, the Russian-English bilinguals who studied English academically did so between ages 5 and 27 (mean age 13.9).

Nine of the NR group were from parts of Russia, including the western cities Moscow and Kaliningrad, and eastern areas including the Ural region and as far east as Novosibirsk. Six NR participants were from Ukraine[28], including three participants from the Crimean Peninsula, which has historically been more strongly associated with Russian, rather than Ukrainian, language and culture, which may impact their native dialect and its influence on their performance in English. Two of these participants self-identified as being from Russia while specifying the Crimean region (and reported little to no exposure to or study of the Ukrainian language in the time they lived there) and one self-identified as being from Ukraine while specifying the same region (and reported both academic and naturalistic exposure to the Ukrainian language); all may exhibit influence from dialectal variation associated with Russian spoken in Ukraine (see Holden and Nearey 1986), but

[27] Students in many Russian-speaking countries begin academic study of English through school, without access to native speakers, in the 5th grade, around age 9-11.

[28] Recall that for at least some speakers, the variety of Russian spoken in Ukraine differs from other varieties in that the vowel [ɨ] is high front, rather than high central (i.e. like the high front /ɪ/ of Ukrainian), similar to English /ɪ/. See §2.2.11

only the latter is likely to exhibit influence on her English performance from Ukrainian itself, due to exposure to it prior to leaving Ukraine. Two participants were from Moldova, one from Belarus, one from Latvia, and one from Uzbekistan.

3.2.2 Native-English Monolinguals (NE)

Background information on the NE participants is summarized in Table 3.2. NE participants were assigned random number identifiers from 1 to 9. This group ranged in age from 18 to 32 years at time of data collection (mean = 23.1 years; SD = 4.8 years) and included 5 females and 4 males. Participants had spent between 7 and 21 years in Atlanta, Georgia immediately prior to data collection (mean = 15.0 years; SD = 4.9 years). A few had lived in other US states (and one in Canada) prior to Georgia, and all had been exposed to other languages academically or naturalistically but none were native or fluent in any language other than English.

Table 3.2 Monolingual English Participants (NE)

ID	*Age*	*Sex*	*Born*	*Academic L*	*Natural L*	*LOR GA*	*Years other*	*Before GA*
1	26	F	GA	Spanish French		18	8	SC, NC
2	22	F	GA		Italian	20	2	CA
3	32	M	UT	French	Hungarian	14	16	CA, Canada
4	21	F	WA	Spanish French Arabic		18	3	WA
5	19	F	NJ	Spanish German		15	4	NJ
6	18	M	SC	German	French Mandarin Spanish Italian	14	5	SC, KS
7	20	F	SC	Spanish ASL	Spanish	8	12	SC, NC, TX, VA
8	29	M	OH	Spanish German		7	22	OH
9	21	M	TX	Spanish Russian	German	21	0	

Explanation of column names:

“ID”:	unique identifier assigned to each participant
“Age”:	participant age, in years, at time of data collection
“Sex”:	participant sex
“Born”:	participant state of birth (all USA)
“Academic L”:	languages participants had studied academically, if any, other than English
“Natural L”:	languages participants had been exposed to in a naturalistic setting, if any, other than English
“LOR GA”:	number of years participants lived in Atlanta, GA immediately prior to data collection
“Years other”:	number of years participants had lived in other US states (in one case Canada) before GA
“Before GA”:	other states in the US where participants had lived before Georgia, if any

To address the research questions posed in Chapter 2 and fill the gap in existing literature surveyed in Chapter 2, the present work will conduct the following three experiments.

3.3 Experiment Design: Experiment 1 – Production

Experiment 1 collects production data from Russian-English bilingual and from monolingual English speaking participants in order to address, in full or in part, all seven research questions (see §2.9 for research questions and motivations) with a focus on production. Questions 3, 4, and 5 additionally rely on Experiments 2 and 3 (see §3.4 and §3.5) with a focus on perception.

3.3.1 Experiment 1 Prompts

Production data elicitation words:

Thirty-two common English words (see Table 3.3) were selected to provide four example words for each of the tested vowels; this yielded four minimal (where possible) or near-minimal pairs for each tested vowel pair - /i - ɪ/, /ɛ - æ/, /u - ʊ/, /ɑ - ʌ/. Bilabial or laryngeal onsets were chosen where possible to minimize the effect of coarticulation on target vowels, as the tongue is not active in the production of these onsets.

Table 3.3 Data elicitation words

i	ɪ	ɛ	æ	u	ʊ	ɑ	ʌ
heat	hit	pet	pat	hoot	put	hot	hut
heed	hid	head	had	who'd	hood	pod	bud
beat	bit	bet	bat	boot	foot*	pot	but
bead	bid	bed	bad	booed	good*	spot**	spud**

Two exceptions (*) had to be made in selection of elicitation words for the high, back, lax vowel due to its low frequency in the English lexicon; a labiodental /f/ onset was considered fairly neutral in terms of its coarticulatory effect, and its use (in *foot*) justified; a velar /g/ onset was considered less desirable, but again due to the limited appearance of the high, back, lax vowel in the English lexicon, its use (in *good*) was deemed necessary. Two complex onsets (**) (appended /s/) were necessary to demonstrate the /ɑ - ʌ/ contrast.

Carrier phrases:

Each target word was embedded at random in one of six carrier phrases (see Table 3.4), all with the same rhythm; all carrier phrases featured the same word (‘say’) before the elicitation word and the same word (‘again’) after the elicitation word.

Table 3.4 Carrier phrases for data elicitation words

In wrath I say _____ again.
Then Matt will say _____ again.
They pass and say _____ again.
The theme is say _____ again.
The team will say ____ again.
They seem to say ____ again.

It was predicted that some speakers would flap the coda t/d of the elicitation word due to its phonological environment; however, having a vowel immediately after the elicitation word (and thus possibly triggering flapping) was deemed necessary in order to ensure that native Russians did not assimilate voicing of the elicitation word coda to the onset of the following word if that word began with a consonant, potentially affecting the length of the elicitation word vowel (Russian voicing assimilation is known to operate across word and morpheme boundaries – see Avanesov 1972; Kulikov 2013). Although, neither flapping nor word-final devoicing (Chen 1970; Kulikov 2012) have been demonstrated to significantly impact the underlying length of a preceding vowel; thus, having a vowel immediately after the elicitation word was deemed safest.

3.3.2 Experiment 1 Procedure

Participants (both NR and NE groups) were recorded on a Lenovo laptop running Windows 10, using the computer's internal microphone. A PowerPoint presentation was loaded on the screen; participants sat at a comfortable distance in front of the computer, at a distance of about 1' from the microphone, and saw one elicitation sentence per slide as they progressed through the presentation slides. Participants were instructed to read the sentence on the screen, as normally and clearly as possible, and press the space bar to go on to the next sentence. Speech samples were recorded and digitized via PRAAT (Boersma and Weenink 2017) at a 44.1 kHz sampling rate directly onto the laptop's hard disk drive. After 32 sentences, participants were offered a short break and instructed to press the space bar again when they were ready to begin the next block of sentences. When participants noticed that they had misspoken, they were permitted to repeat the sentence on the screen (and only the second attempt at the target word was used in analysis), but no backtracking through slides was allowed.

Each participant produced each target word (Figure 3.3) three times over the course of the experiment, for a total of 96 target words and 12 instances of each target vowel produced by each participant, totaling 2784 total target words (1920 by the NR group and 864 by the NE group) and

348 tokens of each of the 8 test vowels (240 by the NR group and 108 by the NE group). In the NR group, 50 tokens total were excluded from analysis (37 due to a participant's misreading of a word; 7 due to a participant's skipping a word, and 6 due to background noise, totaling 2.6% of all NR speaking data collected). In the NE group, 7 tokens were excluded from analysis due to participant error (misreading a word, totaling 0.8% of all NE speaking data collected) – see Appendix A.

3.3.3 Experiment 1 Analytical Design:

The start and end of each test vowel and word was labeled by hand in a PRAAT Text Grid. For consonant onsets, segment boundaries were assigned at start of frication for fricatives (typically /h/ but some /f/ in 'foot' and /s/ in 'spot'); at start of prevoicing, or – if no prevoicing was noted – at the start of the burst, for stops (typically /b/ but some /g/ in 'good'); and at the burst for /p/. Vowel segment boundaries were assigned based on waveform and spectrogram displays. Vowel onset was measured at the zero crossing closest to the start of visible formants and increase in amplitude following the end of the release burst or frication of the preceding consonant. Vowel offset was measured at the point when amplitude dropped to or close to (in cases where the coda consonant was voiced) zero. For consonant coda offsets (all alveolar stops, which were in some cases affricated, flapped, or unreleased), segment boundaries were assigned at the end of frication for affricated codas, at the end of the voicing bar if it was present, or at the transition to next vowel (the [ə] of 'again').

A PRAAT script was used to obtain measurements from the sound files and corresponding text grids, including duration for each test vowel and the frequencies of the first (F1), second (F2), and third (F3) formants at the 25%, 50%, and 75% point of each test vowel (vowel formants at the

50% point were used in analysis). Formant values were then spot-checked by hand. For analyses combining male and female speakers, Lobanov transformation/z-scoring (Lobanov 1971) was used to normalize formant values to control for gender related pitch differences via The Vowel Normalization and Plotting Suite, Version 1.1 (Thomas & Tyler 2007). Vowel duration was not normalized to word duration because pre-voicing of onset plosives p/b varied greatly both between and within (especially for NR) L1 backgrounds and some NR talkers affricated coda plosives t/d (which extended the word duration), while some NE talkers did not release final consonants, shortening word duration; this variation in onset and coda consonant durations would have distorted the relative vowel duration.

NR and NE performance, including 1) the F1 and F2 values for all eight tested vowels, 2) the distances between the vowels of the four tested vowel contrasts on the F1/F2 plane, and 3) the duration ratios between the vowels of the four tested vowel pairs were compared with the goal of identifying statistically significant differences in vowel quality and quantity between the two groups, as well as correlations between NR performance and biographical factors such as AOA and LOR. Euclidean distance (Fox and Jacewicz 2009) was used as a measure of distance between vowels on the F1/F2 plane (see below).

Calculating Euclidean Distance – Vowel (EDV)

EDV is defined as the Euclidean distance (in Hertz) between each talker's vowels and the NE sex-matched means for those vowels on the F1 by F2 plane and was calculated for every instance of all eight tested vowels for all 29 talkers, for a maximum total of 96 (minimum total of 88 for the talker with the most excluded tokens) EDV calculations per talker[29]. It is treated as a measure of similarity of each talker's production to the NE mean, or *nativelikeness* in the case of

[29] Recall that some tokens had to be excluded from analysis, decreasing the number of EDV calculations for that speaker. See Appendix A for a list of all excluded tokens.

NR talkers. NE talkers are expected to have a smaller EDV for all tested vowels than NR talkers. EDV for each vowel is calculated as

$$EDV = \sqrt{((F1_{talker} - F1_{NEmean})^2 + (F2_{talker} - F2_{NEmean})^2)}.$$

Calculating Euclidean Distance – Pair (EDP)

EDP is defined as the Euclidean distance (in Hertz) between the two members of each talker's four tested vowel pairs (/i/ and /ɪ/, /ɛ/ and /æ/, /u/ and /ʊ/, /ʌ/ and /ɑ/) and was calculated for each talker's mean values for all four tested vowel pairs for all 29 talkers, for a total of 8 EDP calculations per talker. It is treated as a measure of dissimilarity between adjacent vowels and predicted to be dissimilated to a lesser degree by NR talkers. NE talkers are expected to have a greater EDP for all vowel pairs than NR talkers. EDP for the four vowel pairs is calculated as

$$EDP = \sqrt{((F1_{i} - F1_{ɪ})^2 + (F2_{i} - F2_{ɪ})^2)}$$

$$EDP = \sqrt{((F1_{ɛ} - F1_{æ})^2 + (F2_{ɛ} - F2_{æ})^2)}$$

$$EDP = \sqrt{((F1_{u} - F1_{ʊ})^2 + (F2_{u} - F2_{ʊ})^2)}$$

$$EDP = \sqrt{((F1_{ʌ} - F1_{ɑ})^2 + (F2_{ʌ} - F2_{ɑ})^2)}.$$

Calculating Nativelikeness of EDP (EDPN)

Nativelikeness of NR talkers' EDP (EDPN) is defined as the difference between each talker's Euclidean distance for each contrast and the NE mean for that contrast, and was calculated per contrast as

$$EDPN = |\,(EDP_{NRtalker} / EDP_{NEmean}) - 1|$$

This calculation yielded a comparable numerical score of how far from the NE mean EDP each NR EDP was regardless of whether the NR EDP was greater or smaller than the NE mean. Note that for this measure, small values indicate performance that is more nativelike (having a smaller difference from the NE mean), while large values indicate a greater divergence from the NE mean (i.e. less nativelike performance).

Calculating Duration Ratio (DR)

Duration ratios (DRs) between the typically long and short members of the four tested vowel pairs for each pair for each speaker were calculated as

$$DR = Duration_{LongVowel} / Duration_{ShortVowel}$$

Calculating Nativelikeness of DR (DRN)

Nativelikeness of NR talkers' DRs (DRN) is defined as the difference between each talker's duration ratio for each contrast and the NE mean for that contrast, and was calculated for each contrast as

$$DRN = |\ (DR_{NRtalker} / DR_{NEmean}) - 1|$$

This calculation yielded a comparable numerical score of how far from the NE mean each NR duration difference was regardless of whether the NR duration difference was greater or smaller than the NE mean. Note that for this measure, small values indicate more nativelike performance, and large values indicate a greater divergence from the NE mean (i.e. less nativelike performance).

Finally, spectral and duration results from Experiment 1 were compared with Peterson and Barney's (1952) data for general American English, Hillenbrand et al.'s (1995) data for Michigan English, Hagiwara's (1997) data for California English, and Clopper et al.'s (2005) results (although actual formant values were not available).

3.4 Experiment Design: Experiment 2 – NR Listeners' Perception of NE Talker

Experiment 2 was conducted to collect listening data from NR participants to address, in conjunction with data from Experiments 1 and 3, Questions 3 and 5, summarized below.

Question 3 – Overall Accuracy: Are Russian-English bilinguals ***more accurate*** in perceiving or ***more nativelike*** in producing certain English vowels than others?

Question 5 – Perception vs. Production: Do native-Russian listeners who more accurately ***discriminate*** certain vowels also ***perform*** those contrasts better?

3.4.1 Experiment 2 Stimuli

To create perception stimuli for the NR group, one female native speaker of American English who had lived in the Atlanta, GA area for 5+ years was recorded on the same Lenovo laptop running Windows 10, using the computer's internal microphone, reading aloud from a list the same 32 elicitation words used for Experiment 1 (Table 3.3) embedded in carrier phrases (Figure II.4); the list this speaker used is provided in Table 3.5. The speaker sat at a comfortable distance in front of the computer, at a distance of about 1' from the microphone. Speech samples were recorded and digitized via PRAAT (Boersma and Weenink 2017) at a 44.1 kHz sampling rate directly onto the laptop's hard disk drive. In anticipation of participant or recording errors, the list was read twice to ensure at least one good token of each word was recorded; since no errors occurred, the first recording was used. The experimenter then used PRAAT (Boersma and Weenink 2017) to extract the 32 target words from their carrier phrases and create an identification experiment (procedure described in 3.5.2) for the NR group. 30 milliseconds of silence were added to the beginning and end of each token to prevent stimuli from being presented too suddenly.

3.4.2 Experiment 2 Procedure

In Experiment 2, NR participants performed a multiple forced choice task with stimulus-dependent response buttons, carried out via PRAAT (Boersma and Weenink 2017), which recorded their responses and response times and saved the results directly onto the laptop's hard disk drive. In this task, NR participants sat before the same laptop computer used for Experiment 1, wearing Maxell max-95dB (model KFITHPNNO) over-the-ear headphones, and were instructed

Table 3.5 Elicitation sentences for single native talker

In wrath I say heed again.	The theme is say bat again.
Then Matt will say hoot again.	They seem to say hot again.
In wrath I say beat again.	The team will say pet again.
In wrath I say bud again.	The theme is say but again.
The team will say bet again.	They seem to say hid again.
Then Matt will say spud again.	They pass and say boo'ed again.
The theme is say boot again.	They seem to say pod again.
The theme is say put again.	They pass and say pot again.
The team will say foot again.	Then Matt will say who'd again.
Then Matt will say heat again.	In wrath I say bit again.
They seem to say bid again.	In wrath I say bed again.
They pass and say hut again.	The team will say pat again.
They seem to say hood again.	The theme is say bad again.
In wrath I say bead again.	The team will say head again.
Then Matt will say good again.	They pass and say had again.
They pass and say hit again.	Then Matt will say spot again.

to place their left index finger on the "F" key and their right index finger on the "J" key of the computer's keyboard. Test subjects were instructed to adjust the volume to a comfortable listening level while hearing several random test words from the experiment; the experiment was then re-started (this preview of experimental stimuli is not thought to have affected performance, as stimuli were randomized on each running of the experiment).

Once the MFC experiment began, participants heard one repetition of a target word at a time, randomized, while seeing two real English words (a minimal pair composed of the target word and another word with the same onset and coda but the 'other' vowel in the four vowel pairs tested - /i/ and /ɪ/, /ɛ/ and /æ/, /u/ and /ʊ/, and /ɑ/ and /ʌ/) on the screen. The word choices on the screen were always a minimal pair contrasting the vowel pairs under investigation; for example, when the auditory stimulus was 'heat', testing perception of /i/, the words on the screen were 'heat' and ‘hit’, a minimal pair contrasting /i/ and /ɪ/ (see Figure 3.6). The visual words were always in the same order, as presented in Figure 3.6, to avoid influencing the participant visually. Participants were instructed to press the “F” key if they thought they heard the word on the left of the screen and the “J” key if they thought they heard the word on the right of the screen. The program recorded participant responses and response times.

For the front vowels, each stimulus word from Figure 3.3 (left side of Figure 3.6) was heard three times by each NR participant throughout the experiment, for a total of 240 judgments of each front vowel (4 example words x 3 repetitions x 20 participants) and 480 judgments of each vowel pair /i/ - /ɪ/ and /ɛ/ - /æ/.

For the back vowels, however, a problem arose in that some target words had no plausible English variant containing the ‘other’ vowel from that pair (e.g. there is no real English word pronounced [hʊt] to pair with [hut] to test the /u/ - /ʊ/ contrast). Those words that had no plausible pair containing the adjacent vowel were excluded as perception experiment stimuli, and cells referring to these words are shaded gray in Figure 3.6. To make up for the paucity of listening tokens containing back vowels, the number of times that some of the remaining tokens were repeated in the experiment was increased: *good, put, pot, bud,* and *but* were each heard 6 times by each participant (instead of 3 times each for each word with front vowels), but *who’d, hood, hot,* and *hut,* being members of true minimal pairs, were heard 3 times each, as the front vowels, in order to keep the number of visual stimuli balanced. Due to investigator error, 120 judgments of ‘bud’ (bold and gray in 3.6) had to be excluded from analysis: when ‘bud’ was the auditory stimulus, the screen showed ‘bod’ and ‘but’ as the possible choices – neither option was correct (‘bod’ had the wrong vowel, and ‘but’ had the wrong coda).

This yielded, for the entire NR test group, 60 judgments where the stimuli were *who'd*, 60 of *hood*, 120 of *good*, 120 of *put*, 60 of *hot*, 120 of *pot*, 60 of *hut*, and 120 of *but*, for a total of 120 judgements of the *who'd/hood* pair (60 of the auditory stimuli were *who'd* and 60 were *hood*), 120 of *goo'ed/good* (all auditory stimuli were *good*), 120 of *poot/put* (all auditory stimuli were *put*),

Table 3.6 Audio and visual stimuli for NR perception task

Target vowel	Stimuli heard	Words seen on screen	Target vowel	Stimuli heard	Words seen on screen
i	heed	heed hid	u	who'd	who'd hood
i	heat	heat hit	u	*hoot*	NA
i	bead	bead bid	u	*boot*	NA
i	beat	beat bit	u	*booed*	NA
ɪ	hid	heed hid	ʊ	hood	who'd hood
ɪ	hit	heat hit	ʊ	put	poot put
ɪ	bid	bead bid	ʊ	*foot*	NA
ɪ	bit	beat bit	ʊ	good	gooed good
ɛ	head	head had	ɑ	hot	hot hut
ɛ	pet	pet pat	ɑ	*pod*	NA
ɛ	bed	bed bad	ɑ	pot	pot putt
ɛ	bet	bet bat	ɑ	*spot*	NA
æ	had	head had	ʌ	hut	hot hut
æ	pat	pet pat	**ʌ**	**bud**	**bod but**
æ	bad	bed bad	ʌ	but	bot but
æ	bat	bet bat	ʌ	*spud*	NA

120 of *hot/hut* (60 of the auditory stimuli were *hot* and 60 were *hut*), 120 of *pot/putt* (all auditory stimuli were *pot*), and 120 of *but/bot* (all auditory stimuli were *but*) by the NR group, creating a data set somewhat smaller than that for the front vowels, but nonetheless balanced in terms of contrasts tested (360 total judgments from the NR group testing /u/ - /ʊ/ and 360 testing /a/ - /ʌ/ vs. 480 each for /i/ - /ɪ/ and /ɛ/ - /æ/).

3.4.3 Experiment 2 Analytical Design

Accuracy of NR listeners' perception of vowels produced by a native English speaker was analyzed using Microsoft Excel and the statistical software JMP. Accuracy was compared individually per vowel and per listener; means were calculated for each NR listener's accuracy on each vowel and for each NR accuracy on perception of all eight vowels pooled together. Vowel pair accuracy was calculated as a mean of the two vowels making up the contrast and compared between contrasts. Both individual vowel and vowel pair accuracy for individual talkers and the group were analyzed for correlations with talkers' biographical factors (e.g. AOA, LOR).

3.5 Experimental Design: Experiment 3 – NE Listeners' Perception of NR Talkers

Experiment 3 was conducted in order to collect listening data from NE participants to address, in conjunction with data from Experiments 1 and 2, Questions 3, 4, and 5, provided for the reader's reference below.

Question 3 – Overall Accuracy: Are Russian-English bilinguals ***more accurate*** in perceiving or ***more nativelike*** in producing certain English vowels than others?

Question 4 – Spectrum vs. Duration: Do native-English listeners more accurately identify English vowels produced by native-Russian talkers when they match native values more in ***spectral*** or ***duration*** cues; that is, which cues contribute more to overall comprehensibility?

Question 5 – Perception vs. Production: Do native-Russian listeners who more accurately ***discriminate*** certain vowels also ***perform*** those contrasts better?

3.5.1 Experiment 3 Stimuli

To create perception stimuli for the NE group, I used the NR production from Experiment 1 (see §3.4). As part of Experiment 1, each NR participant had been recorded reading aloud three instances each of 32 data elicitation words (Figure 3.3) in carrier phrases (Figure 3.4). The experimenter used PRAAT (Boersma and Weenink 2017) to extract the target words from their carrier phrases to create an identification experiment task for the NE group.

The experimenter's goal was to use a selection of eight words from, rather than all of, each NR speaker's production – ideally the h_t/d minimal pairs (*heat/hit*, *head/had*, *who'd/hood*, and *hot/hut*) – to create listening tokens for the NE group's perception task (using all 32 words from all 20 NR speakers, even with only one repetition of each word, would have resulted in 640 listening tokens, raising concerns of listener fatigue). However, several problems arose with the production of two NR participants.

- Participant E produced 'head' as 'heed' in all three repetitions; thus, this participant's minimal pair *head/had* actually sounded like heed/had, and could not be used for the NE perception task, as /i/ and /æ/ are not adjacent in the vowel space, and it would not be informative to investigate their discrimination by NE listeners. The researcher determined this to be a case of misreading rather than mispronunciation, due to orthographically similar words which can be pronounced with /i/ (e.g. 'bead'), and this participant's /ɛ - æ/ pair *head/had* was replaced with her production of bed/bad in the NE perception task.
- Participant N pronounced 'hut' as 'hoot' in all three repetitions; thus, this participant's minimal pair *hut/hot* actually sounded like *hoot/hot*, and could not be used for the NE perception task, as /u/ and /ɑ/ are not adjacent in the vowel space, and it would not be informative to investigate their discrimination by NE listeners. The researcher determined

this to be a case of misreading rather than mispronunciation (due perhaps to lack of familiarity with the word), and this participant's /ɑ - ʌ/ pair *hot/hut* was replaced with her *pot* (to contrast with *putt*) and her *but* (to contrast with *bot*) in the NE perception task.

An attempt was made to extract tokens for the NE perception task from words produced in Block 2 of each NR participant's production; Block 1 was avoided to give participants time to grow accustomed to the task, and Block 3 was avoided to minimize the effects of participant fatigue. Thus, for most participants, all words but 'head' were pulled from Block 2; as 'head' did not occur in Block 2 due to the randomization of words in the production task, the closest instance (2/3 of the way through Block 1) was used. Where it was not possible to use a word produced in Block 2 due to participant or recording error, another occurrence of the same word (or another word or word pair containing the same vowel – as for E12 and N09 above) was used, with preference for instances at the end of Block 1 or the beginning of Block 3 (closest to Block 2).

The recording obtained from participant H was considerably louder than the others, while the recording obtained from participant N was considerably quieter than the others, so much so that their production stood noticeably out from the other 27 talkers in the production experiment. To adjust this, intensity was scaled down to 50dB for H and up to 70dB for N using PRAAT.

3.5.2 Experiment 3 Procedure

In Experiment 3, NE participants performed a multiple forced choice task with stimulus-dependent response buttons, administered via PRAAT (Boersma and Weenink 2017), which recorded their responses and response times and saved the results directly onto the laptop's hard disk drive. In this task, NE participants sat before the same Lenovo laptop running Windows 10 that was used in Experiments 1 and 2, wearing Maxell max-95dB (model KFITHPNNO) over-the-ear headphones, and were instructed to place their left index finger on the "F" key and their right index finger on the "J" key of the computer's keyboard. As in Experiment 2, test subjects were instructed to adjust the volume to a comfortable listening level while hearing several random test words from the experiment; the experiment was then re-started (this preview of experimental stimuli is not thought to have affected performance, as stimuli were randomized on each running of the experiment).

Once the MFC experiment began, participants heard one repetition of a target word at a time, randomized, while seeing two real English words (a minimal pair composed of the target word and another word with the same onset and coda but the 'other' vowel in the four vowel pairs tested - /i/ and /ɪ/, /ɛ/ and /æ/, /u/ and /ʊ/, and /ɑ/ and /ʌ/) on the screen. The word choices on the screen were always a minimal pair contrasting the vowel pairs under investigation; for example, when the auditory stimulus was 'heat', testing perception of /i/, the words on the screen were 'heat' and 'hit', a minimal pair contrasting /i/ and /ɪ/ (see Figure 3.7). The visual words were always in the same order, as presented in Figure 3.7, to avoid influencing the participant visually. Participants were instructed to press the "F" key if they thought they heard the word on the left of the screen and the "J" key if they thought they heard the word on the right of the screen. The program recorded participant responses and response times.

Table 3.7 Audio stimuli for NE perception experiment

	i	ɪ	ɛ	æ	u	ʊ	ɑ	ʌ
All talkers but those noted below	*heat*	*hit*	*head*	*had*	*who'd*	*hood*	*hot*	*hut*
Variations for E12			*bed*	*bad*				
Variations for N09							(*putt*) vs. *pot* [30]	*but* vs. (*bot*)[31]

[30] *Pot* was the word produced by the NR talker and heard by the NE listeners; *putt* was the 'other' option presented visually in the experiment.
[31] *But* was the word produced by the NR talker and heard by the NE listeners; *bot* was the 'other' option presented visually in the experiment.

Unlike the NR listeners, who heard only one talker's production, the NE listeners would hear the production of all 20 NR talkers, and the number of words heard from each talker was reduced in order to eliminate listener fatigue. Thus, each NE listener heard one repetition each of a set of eight words (Figure 3.7) from each NR talker, for a total of 9 judgments of each test vowel, and 18 of each test vowel pair, for each NR talker.

3.5.3 Experiment 3 Analytical Design

Accuracy of NE listeners' perception of vowels produced by NR talkers was analyzed using Microsoft Excel and the statistical software JMP. Accuracy was compared individually per vowel and per talker; also, means were calculated for NE listener accuracy on each vowel spoken by all 20 NR talkers pooled together and for NE accuracy on perception of individual talkers' eight vowels pooled together. Vowel pair accuracy was calculated as a mean of the two vowels making up the contrast and compared between contrasts. Both individual vowel and vowel pair accuracy for individual talkers and the group were analyzed for correlations with talkers' biographical factors (e.g. AOA, LOR) and spectral and duration data for those NR talkers' production from Experiment 1.

3.6 Conclusion

This chapter described the research design implemented in the present study, including a description of two groups of participants and the experiments which test those hypotheses. The next chapter presents the results of these experiments and addresses research questions one by one.

CHAPTER 4

RESULTS

4.1 Introduction

This chapter presents the results obtained from Experiments 1, 2, and 3. It is organized most broadly by experiment – first production (§4.2), followed by NR perception of NE speech (§4.3), and lastly NE perception of NR talkers (§4.4) – and provides both raw and normalized formant frequencies when reporting production data, addresses calculated Euclidean distance and duration differences between tested vowels, and compares perception accuracy on individual vowels as well as vowel contrasts. Throughout the chapter, I provide descriptive and statistical analyses between individuals and language groups as well as between the present study's findings and previously published benchmarks for spectral and duration data.

4.2 Experiment 1 – Production

This section presents results obtained from Experiment 1. §4.2.1 addresses spectral data from NR and NE production results for the eight vowels tested, summarized in Tables 4.1-4.4 and Figures 4.1-4.7 below. Formant information is addressed first in terms of raw formant frequencies, then Lobanov normalized formant values. In each case, individual speaker results are discussed first, then summarized in group means. Figures also represent first raw formant frequencies plotted on the F1/F2 plane (Figures 4.1 and 4.3 present the results of the current study only, while Figures 4.2 and 4.4 compare those findings with previously published data on Russian vowels), then values for normalized vowels (Figures 4.5-4.7). Euclidean distance is addressed next, in §4.2.2 and §4.2.3, calculated first between the vowels of each NR talker and the NE mean for that vowel (EDV), then between the individual members of the four tested vowel pairs (EDP), along with results of statistical analyses measuring the significance of the differences between the tested groups and correlations with biographical factors. Finally, duration measurements for the eight tested vowels among all NR and NE talkers are presented in §4.2.4; duration differences between members of the four tested vowel pairs are calculated and compared between the NR and NE groups; and correlations between these values and factors such as AOA and LOR are addressed.

4.2.1 Spectral data

This section presents spectral information about the eight tested vowels produced by the NR and NE groups. Spectral data are presented first for male talkers (NR, then NE), then female talkers (NR, then NE). Next, frequency data are normalized and presented for NR and NE talkers (male and female combined). Finally, calculations of Euclidean distance between NR and NE production on individual vowels and within vowel pairs are presented.

Table 4.1 Mean F1 and F2 (Hz; standard deviations in parentheses) of eight American English vowels (4 male native Russian talkers)

		i	ɪ	ɛ	æ	u	ʊ	ɑ	ʌ
A	F1	311 (23)	295 (15)	451 (29)	467 (24)	346 (24)	348 (27)	627 (43)	601 (63)
	F2	2372 (287)	2251 (206)	1857 (104)	1889 (92)	891 (163)	1055 (429)	1097 (76)	1050 (211)
L	F1	293 (13)	296 (15)	527 (71)	566 (46)	341 (17)	409 (134)	671 (79)	650 (37)
	F2	2255 (124)	2243 (125)	1792 (165)	1698 (249)	1184 (266)	1110 (236)	1057 (89)	1059 (59)
P	F1	319 (28)	338 (17)	562 (52)	646 (37)	340 (12)	343 (11)	659 (60)	589 (44)
	F2	2410 (206)	2264 (105)	1769 (194)	1612 (311)	1008 (247)	1107 (235)	1132 (83)	1201 (44)
R	F1	274 (40)	293 (30)	559 (72)	617 (72)	357 (14)	362 (38)	714 (56)	688 (94)
	F2	2295 (65)	2202 (98)	1606 (223)	1648 (178)	945 (314)	909 (179)	1181 (66)	1067 (200)
Mean	F1	299	306	525	574	346	366	668	632
	F2	2333	2240	1756	1712	1007	1045	1117	1094
Min	F1	274	293	451	467	340	343	627	589
	F2	2255	2202	1606	1612	891	909	1057	1050
Max	F1	319	338	562	646	357	409	714	688
	F2	2410	2264	1857	1889	1184	1110	1181	1201
SD	F1	20	22	52	79	8	30	36	46
	F2	71	27	107	123	127	94	53	72

Table 4.2 Mean F1 and F2 (Hz; standard deviations in parentheses) of eight American English vowels (4 male native English talkers)

		i	**ɪ**	**ɛ**	**æ**	**u**	**ʊ**	**ɑ**	**ʌ**
3	F1	271 (16)	426 (15)	586 (34)	697 (40)	330 (26)	445 (32)	686 (45)	593 (24)
	F2	2363 (82)	1871 (61)	1716 (58)	1697 (62)	1581 (174)	1470 (190)	1130 (71)	1346 (108)
6	F1	262 (6)	470 (29)	613 (30)	724 (52)	323 (23)	534 (35)	878 (81)	678 (34)
	F2	2547 (60)	2047 (67)	1887 (49)	1765 (41)	1300 (202)	1446 (124)	1127 (125)	1396 (53)
8	F1	276 (13)	431 (30)	683 (65)	852 (42)	335 (19)	480 (17)	777 (41)	681 (63)
	F2	2481 (67)	2017 (89)	1695 (55)	1548 (57)	1608 (182)	1337 (227)	1103 (24)	1262 (203)
9	F1	267 (15)	391 (25)	534 (42)	673 (25)	325 (20)	393 (40)	680 (30)	571 (38)
	F2	2471 (50)	1955 (82)	1783 (85)	1711 (50)	1759 (98)	1578 (168)	1137 (28)	1339 (51)
Mean	F1	269	426	604	737	328	463	755	631
	F2	2466	1973	1770	1680	1562	1458	1124	1336
Min	F1	262	470	534	673	323	393	680	571
	F2	2363	1871	1695	1548	1300	1337	1103	1262
Max	F1	276	470	683	852	335	534	878	681
	F2	2547	2047	1887	1765	1759	1578	1137	1396
SD	F1	5.9	32.3	62.0	79.8	5.4	59.3	93.1	57.0
	F2	76.2	77.8	86.4	92.9	191.4	98.9	14.8	55.3

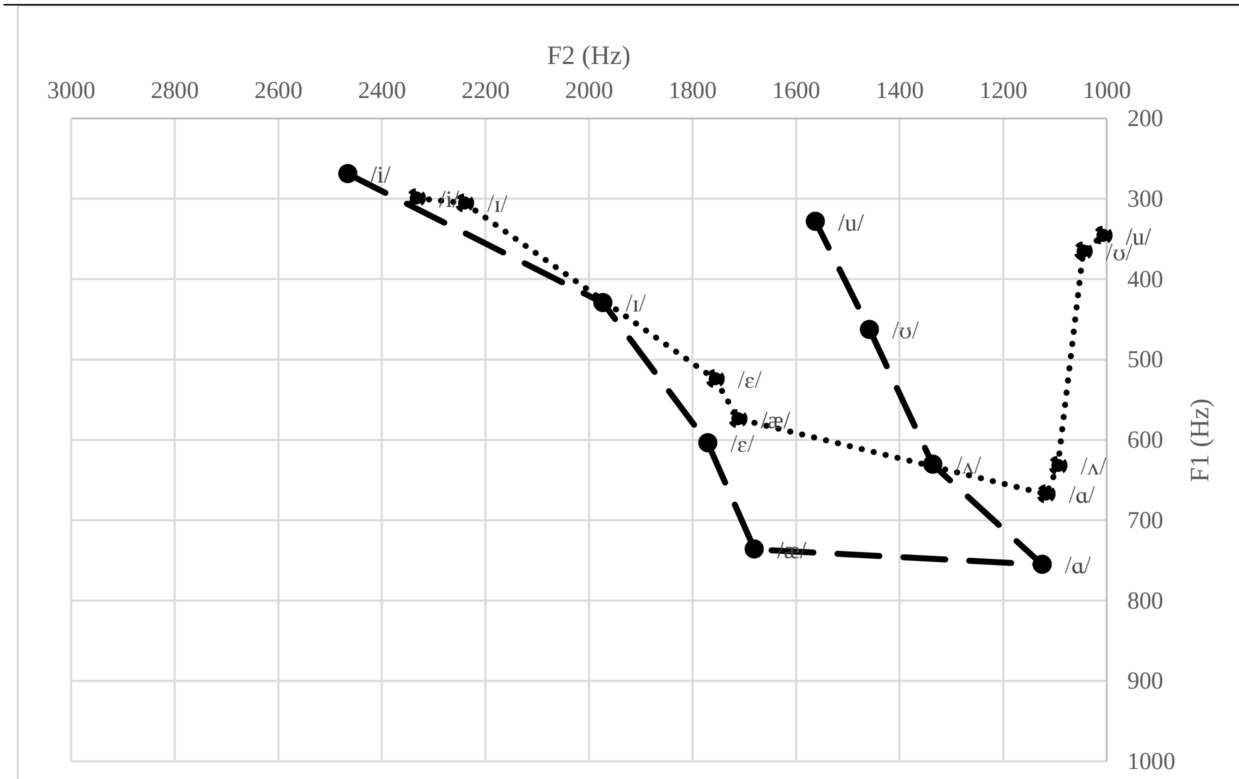


Figure 4.1 Mean F1/F2 (Hz) of eight American English vowels, NR (4 speakers; dots) and NE (4 speakers; dashes) males. Vowels in p/b/h_t/d environments.

Results of spectral analysis for NR and NE male takers (N=4 for each language group) are presented in Tables 4.1 and 4.2 and Figure 4.1. NR male talkers' F1 and F2 values for /i/ were rather similar to those of NE male talkers: only 30 Hz higher in F1 and 133 Hz lower in F2. However, mean values for /ɪ/ among the NR males were considerably further away from the NE males' mean: 120 Hz lower in F1 and 267 Hz higher in F2. Recall from Chapter 2 (see also Figure 4.2) that F1/F2 values for male talkers' Russian /i/, to which, based on previous reports from both monolingual Russians (Gilichinskaya and Strange 2010) and Russian-English bilinguals (Bell-Berti et al. 1998; Kondaurova and Francis 2004; 2008; Romano et al. 1998) we would expect NR speakers of English to assimilate both English /i/ and /ɪ/, were 234 Hz for F1 and 2009 Hz for F2 (Halle 1971: 169-171.) The NR males produced /i/ and /ɪ/ closer to one another on the F1/F2 plane than did the NE males, and both vowels in the production of NR males fell between the NE males' /i/ and /ɪ/, closer to /i/. This and the above-referenced values for Russian /i/ suggests that NR males as a group fail to reliably distinguish these two vowels and produce both as /i/-like, likely due to influence from the L1. Euclidean distances between these and the three other tested vowel pairs illuminate the subject further, especially with respect to individual performance, and are provided below, after a comparison of results from female talkers.

As for the high front pair discussed above, male NR talkers' F1 and F2 values for /ɛ/ and /æ/ are also closer to one another than those of male NE talkers. The difference between NR males' mean F1 for /ɛ/ and /æ/ is only 47 Hz, and the difference in F2 is only 44 Hz. NE males make a greater distinction: 133 Hz in F1 and 90 in F2. For the NR male talkers as a group, both /ɛ/ and /æ/ are positioned very close to but slightly higher and more centralized than NE male talkers' mean /ɛ/ on the F1/F2 plane. This, too, can be interpreted as failure to – as a group – make a strong distinction between these two vowels, likely due to influence from Russian /e/ (which has an F1 of 506 and an F2 of 1884 based on values provided in Halle 1971: 169-171), to which we might expect NR speakers of English to assimilate both /ɛ/ and /æ/ (Gilichinskaya and Strange 2010; Bell-Berti et al. 1998; Romano et al. 1998).

While NR males' /i/ and /ɪ/ were close to NE males' /i/ on the F1/F2 plane, and NR males' /ɛ/ and /æ/ were close to NE males' /ɛ/, the positioning of NR males' /u/ and /ʊ/ shows the most divergence from the NE male mean for either phoneme. NR males' /u/ and /ʊ/ are quite close to one another on the F1/F2 plane, suggesting that as a group, the NR males do not make a strong distinction between these two phonemes, but they are quite far from the NE males /u/ and /ʊ/, which are centralized, with higher F2 values than the NE males' /ʌ/. Rather than approximating English or Russian /u/ (which has an approximate F1 of 259 Hz and F1 of 566 Hz per Halle 1971: 169-171), NR male talkers' /u/ and /ʊ/ appear to fall somewhere between the two native values.

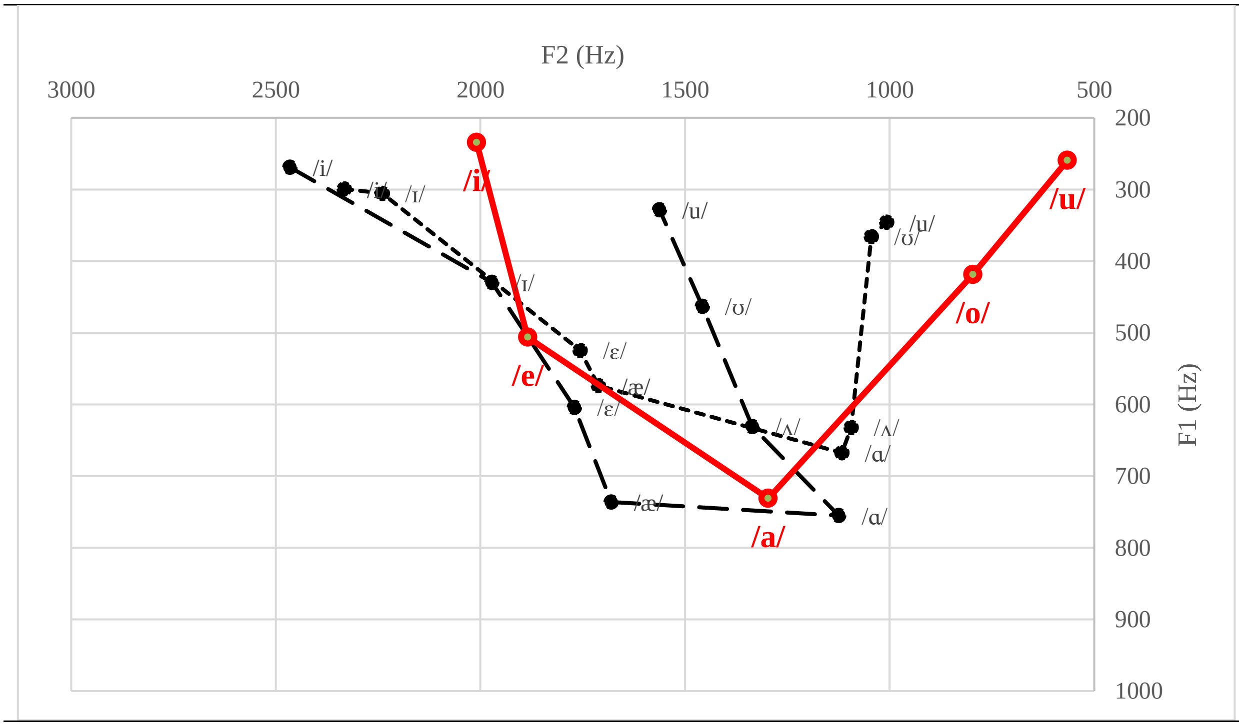


Figure 4.2 Mean F1/F2 (Hz) of eight American English vowels, NR (4 male speakers; dots) and NE (4 male speakers; dashes) (Figure 4.1) compared with five Russian vowels (2 male native speakers; red). English vowels in p/b/h_t/d environments; Russian vowels in p/bV contexts (Halle 1971) (F2 axis modified to include lower Russian values)

The extreme fronting observed in NE males' high back vowels may be connected to dialect (fronting of /u/ and /ʊ/ is a an acknowledged feature of southern English); additionally, the small sample size (only four male speakers per language group) may play a role.

Turning to the /ɑ - ʌ/ pair, NR males' F1 and F2 values for these two vowels are once again much closer to one another than are those of the NE males, suggesting that these two phonemes are also not well differentiated. Both vowels in NR male production are more back and have lower F2 values than NE /ɑ/, and fall between values for Russian /a/ and /o/, which have an F1 of 731 Hz and an F2 of 1297 Hz, and an F1 of 419 Hz and an F2 of 797 Hz, respectively (Halle 1971: 169-171).

To summarize results from the male participants, NR talkers made a smaller distinction between all four tested contrasts than did NE talkers. For the high vowels, NR talkers' F1 and F2 values fall between NE /i/ and /u/ and Russian /i/ and /u/; this could reflect an adjustment to the L2 phonology or simply be a consequence of small sample size, particularly since the Russian phoneme values do not come from the NR speakers in the present study. Next, we turn to production data for female talkers.

Like the male NR talkers, female NR talkers (Table 4.3) made a smaller spectral distinction between the members of the four tested vowel pairs than did NE female talkers (Table 4.4). The spectral distinction between members of all four tested vowel pairs was greater, on average, for NR females than NR males; however, this difference could be due to a larger sample size (N=16) in the female NR group, and individual performance will be addressed in the discussion on Euclidean distance between vowels below. The general pattern of NR females' production of the eight tested vowels follows that of the male NR talkers in that the members of the four tested vowel pairs are not as dispersed as they are in NE production and, at least in some cases, the pairs are

Table 4.3 Mean F1 and F2 (Hz; standard deviations in parentheses) of eight American English vowels (16 female native Russian talkers)

		i	ɪ	ɛ	æ	u	ʊ	ɑ	ʌ
B	F1	315 (27)	312 (31)	744 (72)	780 (139)	385 (57)	405 (78)	811 (272)	954 (52)
	F2	2960 (210)	2530 (307)	2090 (97)	1920 (315)	868 (281)	898 (178)	1248 (290)	1426 (70)
C	F1	348 (59)	530 (87)	704 (67)	1062 (114)	406 (63)	448 (63)	812 (183)	907 (123)
	F2	2624 (163)	2440 (137)	2018 (306)	1937 (314)	1503 (342)	1349 (332)	1281 (206)	1731[32] (247)
D	F1	298 (7)	307 (17)	610 (49)	628 (114)	353 (45)	347 (30)	753 (205)	799 (134)
	F2	2697 (50)	2690 (43)	1706 (87)	1658 (109)	790 (166)	744 (79)	1200 (247)	1244 (118)
E	F1	402 (36)	538 (54)	866 (57)	1063 (63)	414 (30)	539 (95)	1000 (78)	840 (44)
	F2	2953 (33)	2477 (82)	1857 (179)	1672 (116)	1562 (296)	1680 (301)	1245 (68)	1499 (141)
F	F1	379 (18)	377 (35)	685 (41)	694 (56)	390 (35)	459 (126)	821 (119)	782 (38)
	F2	2832 (56)	2649 (223)	1983 (197)	1984 (94)	1141 (175)	1234 (315)	1224 (96)	1207 (87)
G	F1	345 (13)	344 (17)	610 (34)	616 (71)	354 (16)	369 (26)	556[33] (140)	758 (20)
	F2	2670 (60)	2614 (78)	1996 (75)	1897 (248)	789 (85)	811 (42)	929 (167)	1038 (115)
H	F1	418 (40)	483 (63)	847 (71)	903 (58)	438 (39)	480 (58)	821 (80)	850 (51)
	F2	2848 (64)	2508 (175)	1724 (258)	1493 (260)	1109 (144)	1239 (280)	1172 (43)	1269 (83)
J	F1	329 (23)	567 (33)	802 (46)	980 (48)	362 (27)	545 (117)	944 (53)	800 (56)
	F2	2957 (55)	2267 (62)	2048 (60)	1803 (110)	1358 (174)	1583 (162)	1403 (106)	1697[34] (81)
K	F1	356 (56)	470 (84)	801 (71)	944 (56)	404 (47)	471 (163)	812 (66)	772 (68)
	F2	2786 (172)	2175 (111)	1891 (81)	1794 (133)	1483 (365)	1468 (258)	1337 (175)	1455 (133)
M	F1	328 (35)	367 (33)	714 (59)	805 (69)	362 (33)	356 (28)	763 (51)	791 (40)
	F2	2703 (49)	2679 (84)	1964 (51)	1952 (144)	820 (147)	849 (138)	1323 (141)	1320 (111)
N	F1	412 (42)	410 (29)	694 (76)	741 (37)	367 (39)	391 (35)	698 (120)	571 (89)
	F2	2864 (83)	2813 (91)	2081 (83)	2092 (311)	882 (133)	1078 (317)	1084 (141)	1112 (96)
Q	F1	403 (36)	406 (35)	716 (48)	741 (50)	409 (35)	422 (40)	616 (63)	650 (73)
	F2	2670 (47)	2712 (53)	2066 (138)	2004 (106)	691 (103)	734 (94)	948 (76)	1026 (109)
S	F1	379 (87)	505 (80)	864 (228)	976 (146)	397 (49)	501 (103)	910 (190)	877 (97)
	F2	2810 (259)	2418 (157)	1603 (342)	1754 (259	1137 (147)	1322 (150)	1288 (128)	1350 (187)
T	F1	409 (59)	527 (34)	663 (44)	953 (54)	426 (75)	589 (61)	994 (65)	691 (74)
	F2	3021 (143)	2340 (112)	1917 (114)	1956 (175)	1641 (229)	1892 (222)	1436 (126)	1829[35] (177)
U	F1	354 (28)	441 (30)	827 (106)	997 (37)	385 (18)	513 (84)	1010 (44)	864 (80)
	F2	2794 (45)	2346 (88)	1961 (91)	1931 (173)	1344 (148)	1554 (128)	1418 (123)	1488 (85)
V	F1	366 (20)	373 (12)	712 (37)	700 (20)	382 (17)	389 (20)	879 (55)	864 (56)
	F2	2826 (30)	2810 (29)	2078 (94)	2056 (64)	872 (54)	903 (95)	1352 (83)	1402 (84)
Mean	F1	365	435	741	849	390	452	825	798
	F2	2662	2529	1936	1869	1124	1209	1243	1381
Min	F1	298	307	601	616	353	347	556	571
	F2	2624	2175	1603	1493	691	734	929	1026
Max	F1	418	567	866	1063	438	589	1010	954
	F2	3021	2813	2090	2092	1641	1892	1436	1829
SD	F1	37	84	84	152	26	73	132	98
	F2	119	191	147	161	319	364	151	236

[32] Speaker C's /ʌ/ tokens were very fronted, unrounded but approximating the positioning of /œ/ on the F1/F2 plane; this may be influence from German, which has this vowel in its inventory, and which this participant studied academically, or from Armenian, which has /ə/ rather than /ʌ/ in its inventory, and to which this participant was exposed in a naturalistic setting.

[33] Speaker G pronounced some of her /a/ tokens as [o], probably due to orthography; this is particularly strong in *pot* and *pod* , which she produces as [pot] and [pod], sequences permitted by Russian phonotactics. These tokens were not excluded from the analysis, as they are not a clear case of misread word (there are no English words pronounced [pot] and [pod]). There is a considerable difference in F1 between these very low /a/ tokens (in the 400's Hz) and her more typical ones (700's Hz); however, the tokens in the 400's comprise 2/3 of her production.

[34] Speaker J's /ʌ/ tokens were quite fronted; this may be influence from Armenian, which has /ə/ rather than /ʌ/ in its inventory, and to which this participant was exposed in a naturalistic setting for the first 14 years of her life.

[35] Speaker T's /ʌ/ tokens were high and fronted to a central position; this may be influence from Moldovan, which has /ə/ rather than /ʌ/ in its inventory, and which this participant studied academically; she does not report consistent informal exposure to Moldovan in childhood but reports living the first 13 years of her life in Moldova.

Table 4.4 Mean F1 and F2 (Hz; standard deviations in parentheses) of eight American English vowels (5 female native English talkers)

		i	**ɪ**	**ɛ**	**æ**	**u**	**ʊ**	**ɑ**	**ʌ**
1	F1	360 (16)	491 (33)	718 (15)	840 (50)	375 (23)	517 (55)	790 (85)	731 (20)
	F2	3010 (52)	2300 (55)	2079 (60)	1803 (71)	1548 (95)	1445 (244)	1150 (60)	1554 (147)
2	F1	413 (44)	620[36] (19)	809 (32)	974 (64)	397 (18)	642 (15)	856 (96)	802 (36)
	F2	3165 (67)	2450 (52)	2178 (49)	2015 (92)	1986 (160)	1921 (124)	1267 (178)	1872 (135)
4	F1	345 (15)	493 (19)	746 (26)	971 (30)	347 (13)	504 (26)	904 (45)	679 (47)
	F2	3003 (93)	2368 (55)	2133 (41)	1902 (54)	1533 (271)	1886 (137)	1384 (281)	1845 (60)
5	F1	379 (28)	480 (34)	706 (71)	848 (138)	403 (16)	510 (74)	774 (141)	712 (68)
	F2	2487 (243)	2098 (116)	1915 (51)	1798 (79)	1547 (202)	1760 (142)	1273 (92)	1684 (74)
7	F1	355 (22)	540 (26)	792 (57)	1037 (50)	388 (11)	584 (18)	895 (97)	800 (56)
	F2	2788 (33)	2102 (52)	1906 (75)	1559 (104)	1705 (157)	1723 (110)	1225 (44)	1651 (124)
Mean	F1	360	501	754	934	382	551	844	745
	F2	2891	2264	2042	1815	1664	1747	1260	1721
Min	F1	345	480	706	840	347	504	774	679
	F2	2487	2098	1906	1559	1533	1445	1150	1554
Max	F1	413	540	809	1037	403	642	904	802
	F2	3165	2450	2178	2015	1986	1921	1384	1872
SD	F1	26.8	26.6	45.1	86.3	22.2	60.0	59.5	54.6
	F2	262.5	158.5	125.3	168.5	193.4	188.1	85.0	134.5

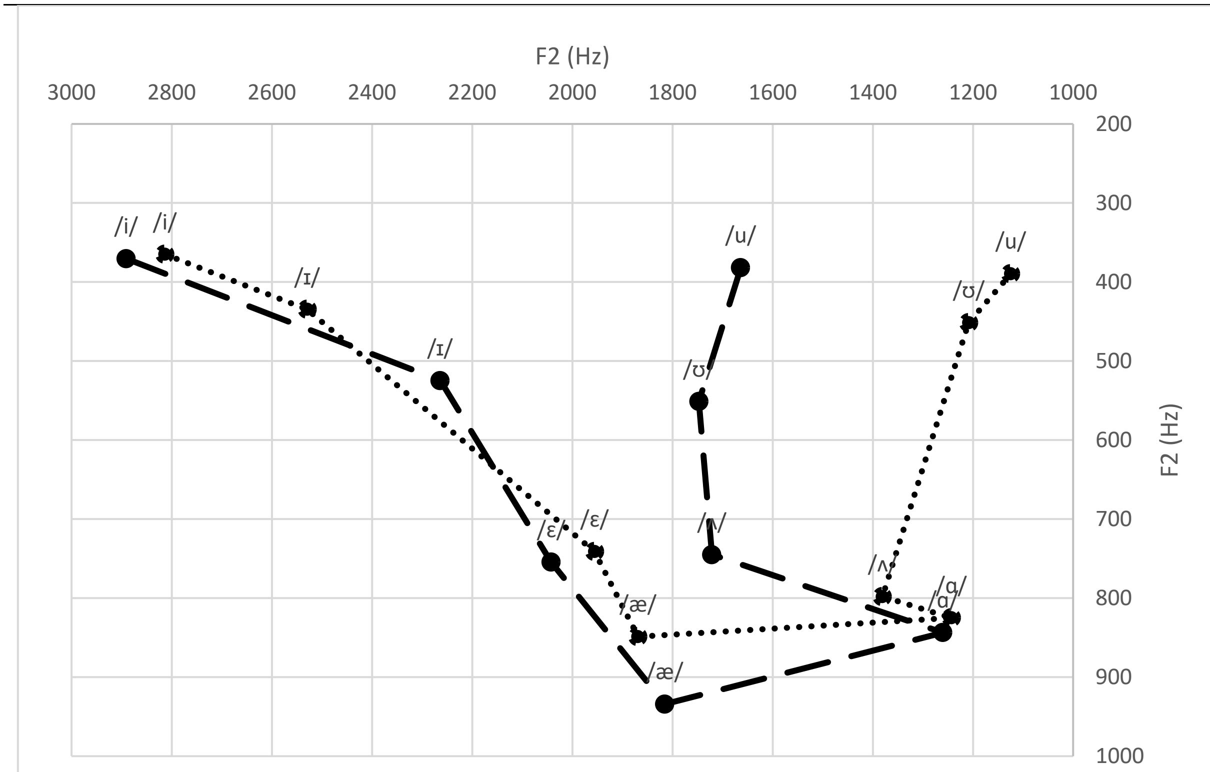


Figure 4.3 Mean F1/F2 (Hz) of eight American English vowels, NR (16 speakers; dots) and NE (5 speakers; dashes) females. Vowels in p/b/h_t/d environments

[36] Speaker 2 has a very high F1 for /ɪ/; this speaker was born in California but relocated go Georgia at the age of two and has spent the last 20 years in the Atlanta, GA area. Her high F1 values may be explained by a combination of her high pitch and southern speech.

positioned near what may be the perceptual magnet to which both members of the pair are assimilated (Figure 4.3). One notable exception is /ɪ/, produced by NR females approximately halfway between the NE /i/ and /ɪ/ values (recall that for NR males, /i/ and /ɪ/ were clustered close together and close to the NE /i/), suggesting that the female group may be further along or more successful at acquiring the /i - ɪ/ contrast (but could also, again, be a consequence of sample size). Another exception is /æ/, which NR females produced much closer on the F1/F2 plane to the NE /æ/ than what NR males produced compared to their sex-matched NE group. Finally, /ʌ/ is slightly fronted in NR female production, not nearly as much as observed for NE females but potentially following the NE female pattern (recall that for NR males, /ʌ/ had a slightly lower F2 than their /ɑ/). Figure 4.4 compares eight American English vowels in the production of NR and NE females.

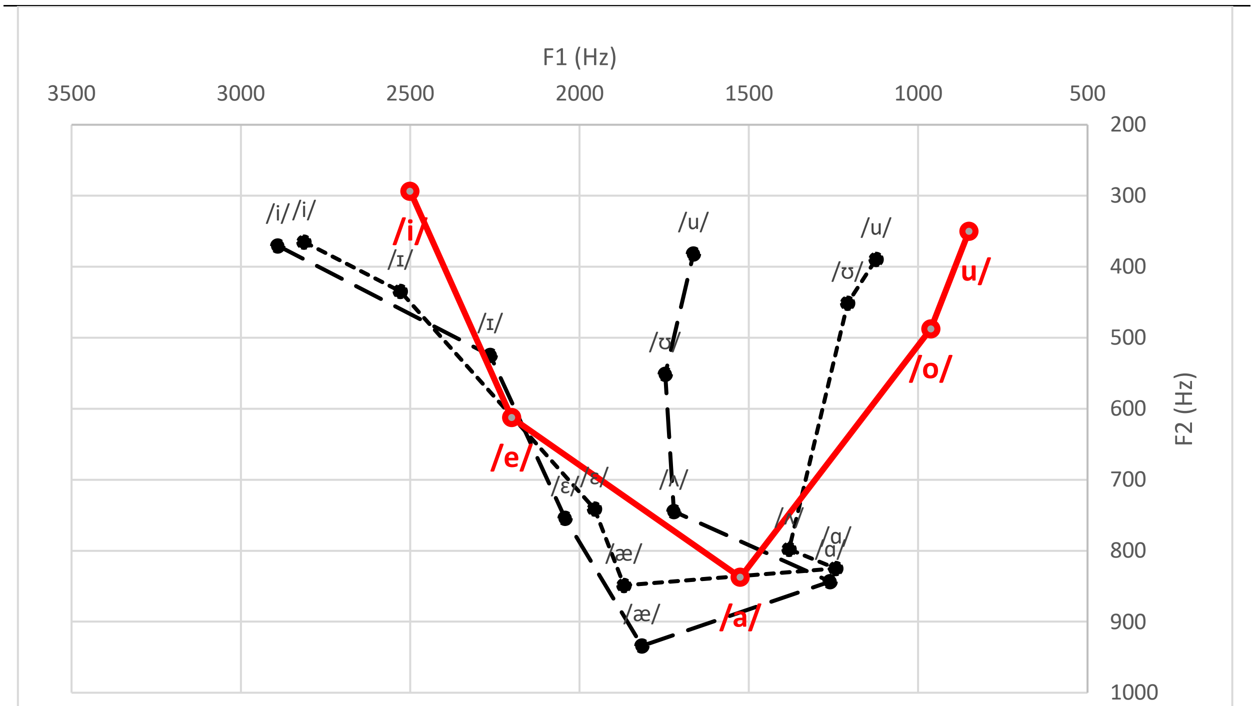
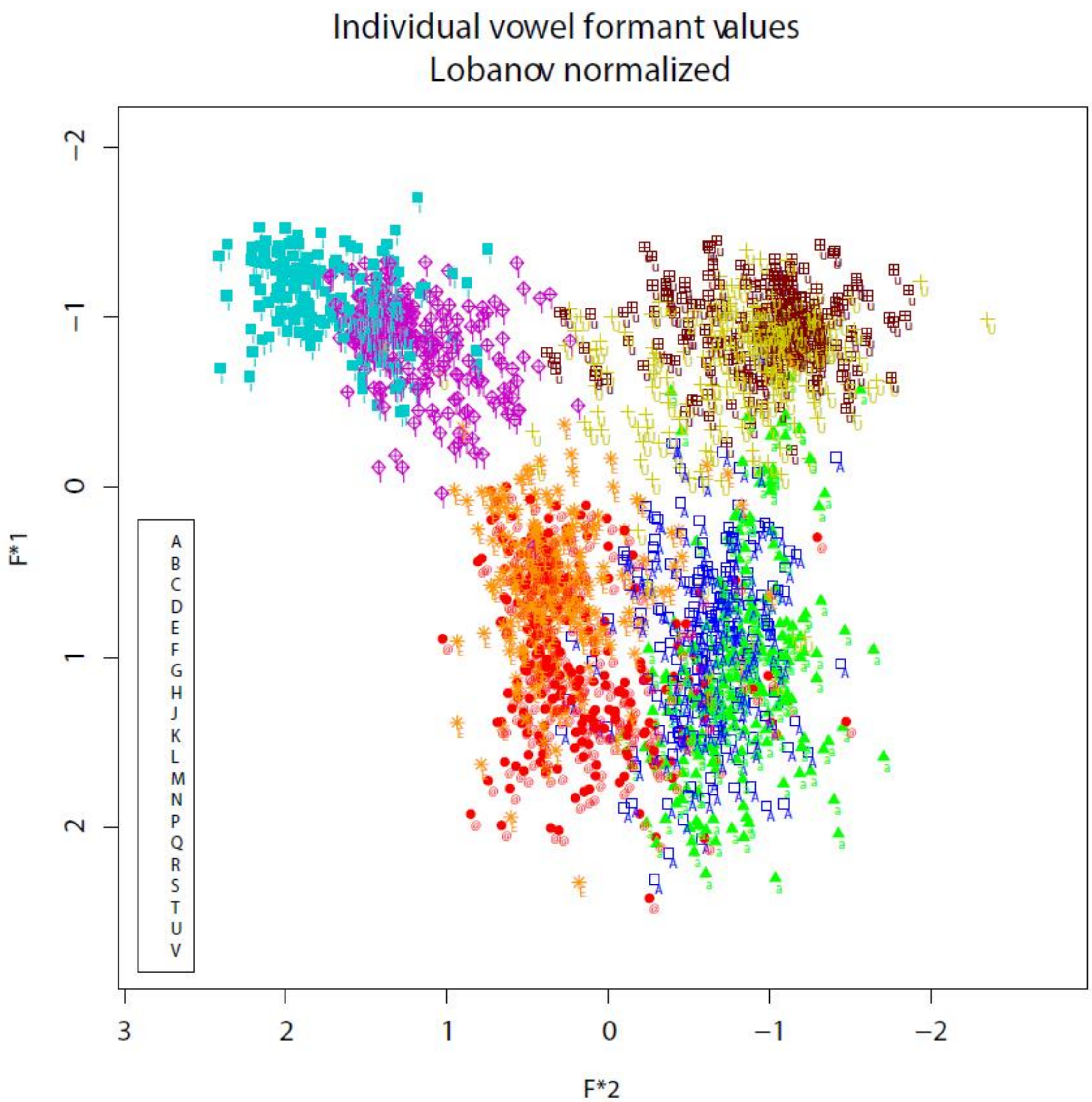


Figure 4.4 Mean F1/F2 (Hz) of eight American English vowels, NR (16 female speakers; dots) and NE (5 female speakers; dashes) (Figure 4.3) compared with five Russian vowels (1 female native speaker; red). English vowels in p/b/h_t/d environments; Russian vowels in p/bV contexts (Halle 1971) (F2 axis modified to include lower Russian values)
with Russian vowels produced by a female talker.

Unlike the male NR talkers, female NR talkers appear to make a greater distinction between the tested vowel pairs and more successfully subdivide the vowel space to accommodate the American English vowel inventory, although they still make a smaller spectral distinction between members of tested pairs than do native speakers.

In several cases, it is possible that the overall shape of both individual and group NR data is disproportionately influenced by either a handful of atypical tokens within one individual's production or the exceptional performance of just one or two speakers in the group. To shed light on this possibility, Figure 4.5 presents all tokens for all NR talkers (both male and female) in a normalized vowel space, helping to visualize these outliers. In particular, Speaker G's idiosyncratic high pronunciation of /a/ stands out, as well as J's and T's exceptionally fronted /ʌ/. A number of NR /u/ and /ʊ/ tokens venture into the high central F1/F2 space occupied by the present study's NE /u/ and /ʊ/ (seen Figures 4.1 and 4.3), although the NR mean is nonetheless clustered at much lower F2 values, somewhere between the present study's central NE /u/ and /ʊ/ and the more back /u/ and /ʊ/ of other varieties of English and Russian /u/.

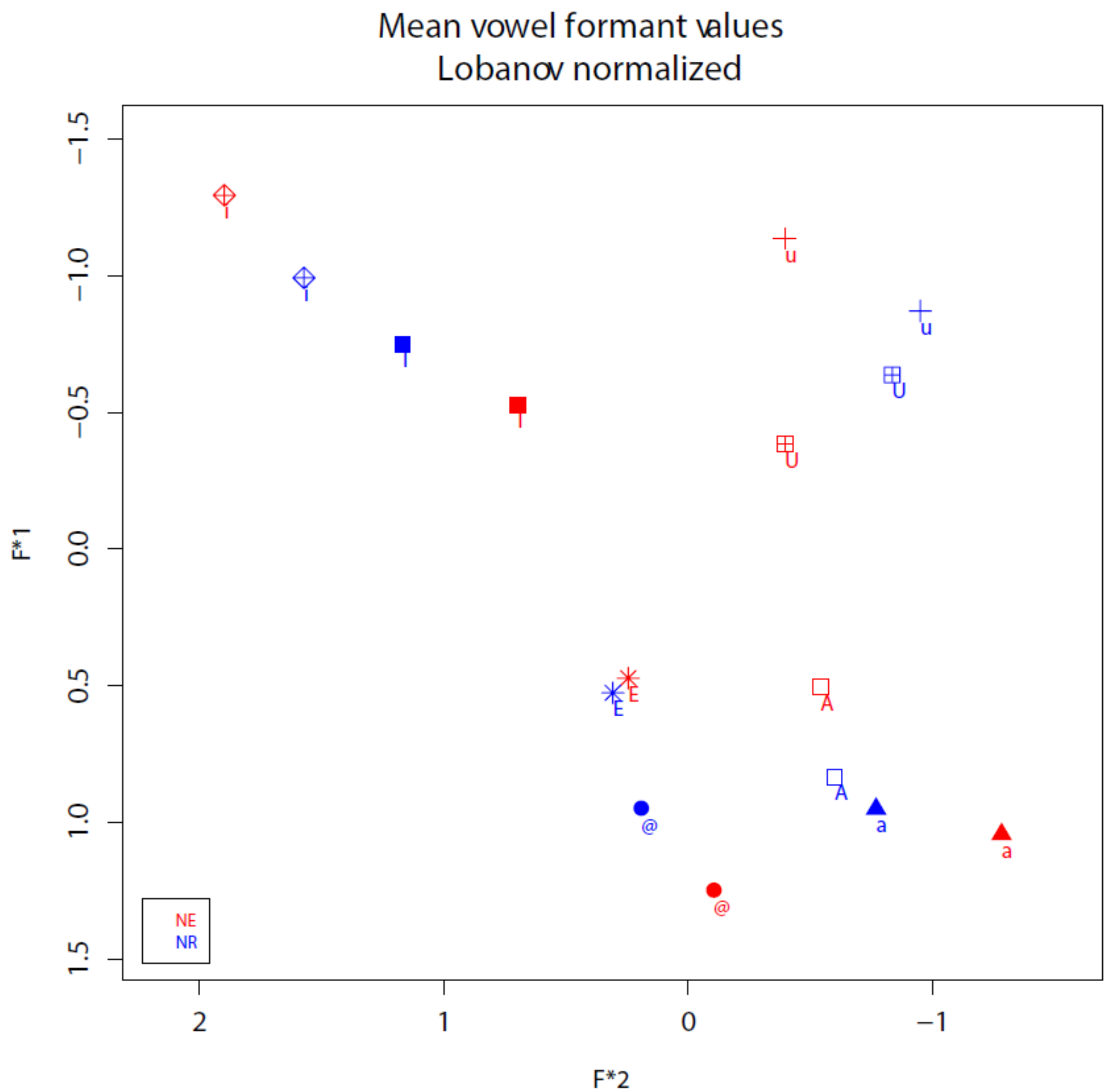
See Appendix B for normalized individual and group mean F1,F2 values of the eight tested vowels in the production of all 20 NR talkers (Tables A.2, A.3), and for normalized F1 and F2 values for the eight tested vowels in the production of all nine NE talkers (Tables A.4, A.5). These normalized data facilitate comparison between the test groups while minimizing the effect of pitch from individual talkers. These data are compared in Figures 4.6 and 4.7, which plot group mean values from Tables A.3 and A.5 on an F1/F2 plane to visually represent differences in production of the eight tested vowels between the NR and NE groups (male and female talkers combined).



Symbol	Phoneme Represented		Symbol	Phoneme Represented	
i	i *as in*	'bead'	u	u *as in*	'booed'
I	ɪ	'bid'	U	ʊ	'good'
E	ɛ	'bed'	a	ɑ	'pod'
@	æ	'bad'	A	ʌ	'bud'

Figure 4.5 F1 and F2 of vowel tokens for all eight tested vowels, NR talkers (4 male speakers; 16 female; normalized data)

Figure 4.6, which represents the production of NR (blue) and NE (red) talkers of both sexes in normalized values on the F1/F2 plane, underscores the results observed individually in the male and female group comparisons above and serves as an excellent summary of production results



Symbol	Phoneme Represented		Symbol	Phoneme Represented	
i	i *as in*	'bead'	u	u *as in*	'booed'
I	ɪ	'bid'	U	ʊ	'good'
E	ɛ	'bed'	a	ɑ	'pod'
@	æ	'bad'	A	ʌ	'bud'

Figure 4.6 Mean F1 and F2 of all eight tested vowels, NR (blue) and NE (red) talkers (4 NR male; 16 NR female; 4 NE male; 5 NE female; normalized data)

reported so far (see Appendix C for individual NR plots). The NR talkers made a smaller spectral distinction between the members of all four tested vowel pairs. As a group, NR talkers produced the /i - ɪ/ pair with F1/F2 values intermediate to NE /i/ and /ɪ/, suggesting that, for at least some speakers, these two phonemes are not well distinguished and treated as members of one category, likely assimilated to the Russian /i/; the /ɛ - æ/ pair is produced closer to NE values for /ɛ/, and the /ɑ - ʌ/ pair is produced between NE /ɑ - ʌ/ on the F1/F2 plane, suggesting that these two pairs are also not well distinguished, with /ɛ/ and /æ/ produced in the part of the vowel space that might (see Chapter 2 for a discussion of Russian allophones) be occupied by Russian /e/, and /ɑ - ʌ/ occupying part of the typical vowel space for Russian /a/. The high back /u/ and /ʊ/ are perhaps the most different from the NE mean in terms of placement in the vowel space – instead of being positioned between NE /u/ and /ʊ/, the NR /u/ and /ʊ/ are considerably more back, likely due to influence from Russian /u/, as well as to the fact that the NE talkers' /u/ and /ʊ/ are fronted. Overall, these results, particularly when taking into account differences in standard deviations between the NR and NE groups (presented in Figure 4.7), suggest that all four tested vowel pairs are not well distinguished by NR talkers as a group (however, see Appendix C for individual plots

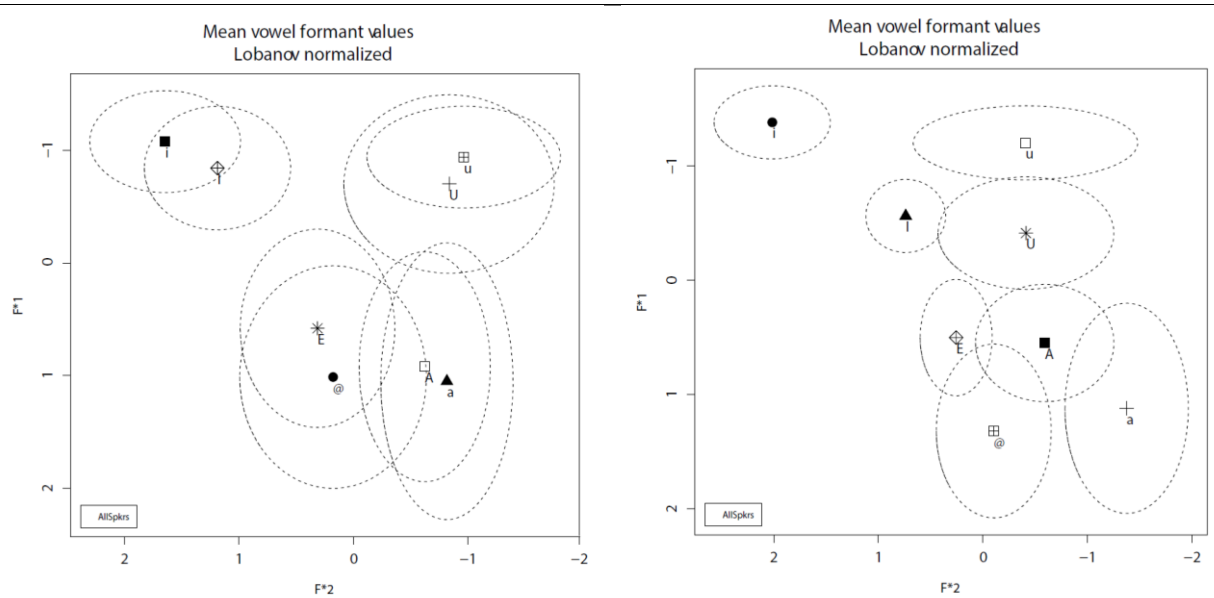


Figure 4.7 NR (left) and NE (right) group means (points) and 2.5 standard deviation ellipses for eight American English vowels plotted on F1/F2 plane (normalized data)

demonstrating rather nativelike performance, in particular E, J, T, and U) and that these NR talkers, as a group, are influenced by their native phonology.

4.2.2 Euclidean distance – individual vowels

Represented below are Euclidean distances (Fox and Jacewicz 2009) between the vowels of NR (Table 4.5) and NE (Table 4.6) talkers and the NE mean. Recall that these values reflect the distance between individual talkers' vowels and the NE mean value for those vowels (see §3.4.3 for details on calculation). These values may be interpreted as a measure of how similar each individual's production was to the NE mean, with low EDV representing nativelike production and high EDV representing production very different from the NE mean.

Table 4.5 Euclidean distance – vowel, NR talkers compared to NE mean

	i	**ɪ**	**ɛ**	**æ**	**u**	**ʊ**	**ʌ**	**ɑ**	**mean**	**min**	**max**	**SD**
A	255	327	194	349	672	536	311	149	349	149	672	175.0
B	205	425	113	370	798	867	367	365	439	113	867	264.1
C	274	196	261	336	284	419	288	219	285	196	419	69.2
D	208	479	369	372	875	1025	501	243	509	208	1025	293.1
E	79	218	269	214	212	271	251	174	**211**	79	271	62.5
F	63	446	176	302	524	549	518	141	340	64	549	194.3
G	223	396	165	400	876	955	684	440	517	165	955	290.9
H	85	276	350	344	558	519	468	120	340	85	558	174.6
J	91	72	75	95	310	218	100	194	**144**	72	310	86.8
K	131	154	176	122	318	387	280	158	**216**	122	387	99.1
L	213	309	144	258	399	407	279	144	269	144	407	101.5
M	195	446	106	231	845	919	405	169	415	106	919	311.3
N	96	562	111	342	783	705	640	401	455	96	783	261.7
P	193	307	163	229	569	412	147	132	269	132	569	153.3
Q	226	465	122	288	974	1022	705	387	524	122	1022	340.4
R	176	271	246	204	619	561	322	91	311	91	619	185.8
S	222	207	535	245	529	438	404	207	349	207	535	144.0
T	187	121	220	207	207	237	197	253	**204**	121	253	39.9
U	102	137	154	201	321	211	280	248	**207**	102	321	74.5
V	69	567	96	338	792	860	347	114	398	69	860	312.5
NR mean	165	319	202	272	573	576	375	218				
NR min	63	72	75	95	207	211	100	91				
NR max	274	567	535	400	974	1025	705	440				
NR SD	67.7	148.7	112.5	85.0	246.3	276.6	168.4	103.6				
Normal Distrib	**No <.0001****						**No .0027***	**No <.0001****				

Table 4.6 Euclidean distance – vowel, NE talkers compared to NE mean

	i	ɪ	ε	æ	u	ʊ	ʌ	ɑ	mean	min	max	SD
1	122	75	69	122	135	322	182	82	139	69	322	83.1
2	282	212	151	215	338	221	165	233	**227**	151	338	60.4
3	113	106	82	75	148	161	105	96	111	75	161	30.1
4	132	113	96	105	238	178	150	354	171	96	354	87.5
5	405	186	151	142	181	149	95	205	189	95	405	93.3
6	87	100	121	100	299	132	88	164	136	87	299	70.5
7	106	165	151	287	133	95	144	181	158	95	287	59.5
8	56	74	128	185	147	232	188	48	132	48	232	68.1
9	42	83	108	86	198	182	80	82	**108**	42	198	54.0
NE mean	149	124	117	146	202	186	133	160				
NE min	42	74	69	75	133	95	80	48				
NE max	405	212	151	287	338	322	188	354				
NE SD	117.9	51.1	30.8	70.1	75.0	66.3	41.7	96.6				
Normal Distrib	**No <.0001****		**No .0053***			**No <.0001****						

As a group, NR talkers performed best (lowest mean EDV, representing the most nativelike production, closest to the NE mean) on /i/, followed in descending order by /ε/, /ɑ/, /æ/, /ɪ/, /ʌ/ and worst (indicated by highest mean EDV) on /u/ and /ʊ/ ("NR mean" row of Table 4.5). Interestingly, EDV values for NE talkers as a group ("NE mean" row of Table 4.6) were also greatest for the high back vowels, indicating that for both groups, variation among individuals is greatest in production of these two phonemes. EDV was calculated based on a sex-matched NE mean, so I do not interpret this to be a reflection of the variation in the centralization of the high back pair between the sexes the case of NE talkers, (recall that NE male talkers centralized the high back vowels more than did the female talkers; see Figure 4.8); however, within-sex variation on the high back vowels may be a contributing to the high EDV in the NE group.

NR talkers displayed more variation (indicated by high standard deviations – 2nd from bottom row of each table) than did NE talkers for all vowels but /i/, where the NR SD was nearly twice that of the NR talkers, indicating that in production of /i/, NE talkers vary more than do NR talkers, whose production of /i/ was relatively uniform. On two contrasts (/æ/ and /ɑ/), the NR and NE SD values were similar, with the NR value only slightly higher. On the other five tested vowels

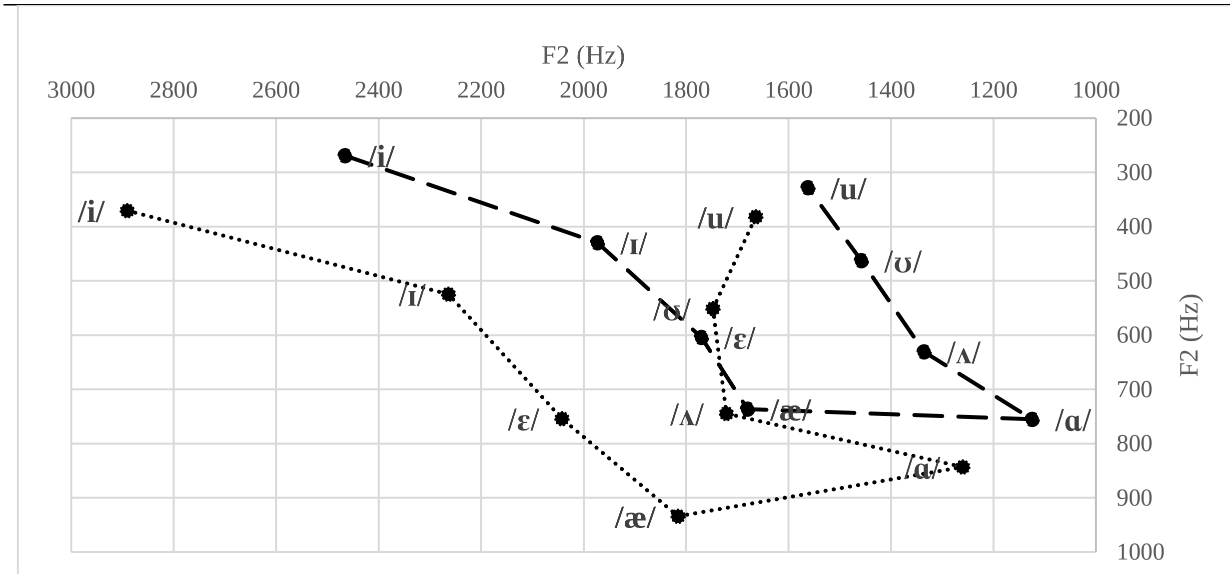


Figure 4.8 Mean F1/F2 (Hz) of eight American English vowels produced by NE males (4 talkers; dashes) and females (5 talkers; dots).

(/ɪ ɛ u ʊ ʌ/) the NR SD value was much higher than the NE (just over 4x higher in the most extreme case - /ʌ/), indicating that NR performance for these contrasts is highly talker-dependent.

Individual talkers' EDVs averaged across all eight tested vowels reveal some rather small EDVs for some NR talkers, indicating more nativelike overall performance for those talkers. In particular, Participants E, J, K, T, and U have the lowest EDV averaged across all eight vowels of the NR group (bold and shaded gray in Table 4.9) – J's is lowest and well within the range of NE SDs from the NE mean, which range (as a mean of all eight vowels) from 108 to 227 (bold and shaded gray in Table 4.10). Individual talker performance on individual vowels, too, reveals some cases of particularly nativelike performance: an EDV value of less than 100 Hz from the NE mean was observed for Participants E, F, H, J, N, and V on /i/; for J only on /ɪ/; for J and V on /ɛ/; for J on /æ/, and for R on /ɑ/. Note, however, that with the exception of J's /i - ɪ/ and /ɛ - æ/, these particularly nativelike values are restricted to only one vowel in the tested pairs; that is, when, in the production of any particular NR talker, one vowel in a given pair (e.g. /i/) is nativelike, as for E, F, H, N, and V, the other vowel in the pair (in this case, /ɪ/) is not – note the high EDV values for /ɪ/ for E, F, H, N, and V. Figures 4.6 and 4.7 shed light on why this might be the case: for NR talkers as a group, the vowels of all four tested pairs tend to be clustered together with considerable overlap, and in some instances, this clustering happens to be near the NE mean for one of those vowels. It is surprising that the EDV for /ɛ/ was small for only two NE talkers, as the NR and NE /ɛ/ mean are very close to one another in Figure 4.6, suggesting more nativelike performance by NR talkers on this vowel; however, the rather large ellipse for performance on /ɛ/ in Figure 4.7 and moderately high SD value for /ɛ/ in Table 4.9 suggest that this may be due to variation between individual talkers.

Because none of the EDV data were normally distributed for any combination of vowel and L1 background, a Wilcoxon test rather than an ANOVA was used to determine if differences in EDV between the NR and NE groups were significant. To test whether L1 predicts EDV, EDV for all available data points were used as the dependent variable, and L1 – Russian or English – was used as the independent variable. For all eight vowels tested, L1 predicted EDV (Table 4.7) AOA positively correlated with the EDV of six out of eight tested vowels (all but /i/ and /ɛ/); that is, those NR talkers who arrived to the U.S. at a younger age produced these six vowels closer to the NE mean than did later arrivals (Table 4.8). LOR in the U.S. positively correlated with the EDV of /ɪ/, /æ/, and /ʊ/; that is, surprisingly, longer residence in the U.S. correlated with less nativelike production on these three contrasts. LOR in GA negatively correlated with the EDV of /ɑ/; that is, longer residence in GA correlated with more nativelike performance on /ɑ/. The AOA and LOR in GA results are in line with what would be expected based on the literature dealing with effects of AOA and LOR; however, the three positive correlations of LOR in GA with Euclidean distance from the NE mean is unexpected.

Table 4.7 Comparison of EDV between NR (N=20) and NE (N=9) talkers on American English eight vowels. Statistically significant differences are shaded gray.

Vowel tested	Significance
/i/	S = 14979, Z = -2.31905, **p = .0204***
/ɪ/	S = 10054, Z = -9.84123, **p = <.0001****
/ɛ /	S = 13912, Z = -5.01395, **p = <.0001****
/æ/	S = 11488, Z = -8.34262, **p = <.0001****
/u/	S = 9898, Z = -10.2266, **p = <.0001****
/ʊ/	S = 9387, Z = -10.8458, **p = <.0001****
/ɑ/	S = 15113, Z = -4.15619, **p = <.0001****
/ʌ/	S = 9018, Z = -11.1583, **p = <.0001****

Table 4.8 Correlation of EDV with AOA, LOR in GA, and LOR in USA on eight vowels for (N=20) NR talkers. Statistically significant correlations are shaded gray.

	AOA	LOR in USA	LOR in GA
/i /	r = -0.03899 p = 0.5589	r = 0.029185 p = 0.6618	r = 0.062726 p = 0.3468
/ɪ/	r = 0.452879 **p = <.0001****	r = 0.16088 **p = 0.0131***	r = 0.024347 p = 0.7092
/ɛ/	r = -0.06827 p = 0.3079	r = -0.02183 p = 0.7447	r = 0.008894 p = 0.8945
/æ/	r = 0.250484 **p = 0.0001****	r = 0.194367 **p = 0.0027***	r = 0.092703 p = 0.1557
/u/	r = 0.312336 **p = <.0001****	r = 0.01331 p = 0.8385	r = -0.08553 p = 0.1895
/ʊ/	r = 0.384991 **p = <.0001****	r = 0.154276 **p = 0.0172***	r = 0.046286 p = 0.4773
/ʌ/	r = 0.436891 **p = <.0001****	r = 0.043321 p = 0.5105	r = -0.05578 p = 0.3967
/ɑ/	r = 0.135248 **p = 0.0375***	r = -0.05127 p = 0.4320	r = -0.13082 **p = 0.0442***

4.2.3 Euclidean distance – vowel pairs

Represented in Table 4.9 and Figure 4.9 are Euclidean distances (Fox and Jacewicz 2009) between the members of each of the four tested vowel pairs – /i - ɪ/, /ɛ - æ/, /u - ʊ/, and /ɑ - ʌ/ – among NR and NE talkers (see Figure 4.10 for a visual example of EDP and §3.4.3 for details on how EDP was calculated). NR talkers as a group made the greatest spectral distinction between the /i - ɪ/ pair (mean EDP of 261.09 Hz) and their smallest spectral distinction between the /u - ʊ/

pair (mean EDP of 125.74 Hz). The pairs /ɑ - ʌ/ and /æ - ɛ/ had similar, intermediate spectral distinctions (mean EDPs of 159.37 Hz and 153.00 Hz, respectively). Compared to native values, the spectral distinction between the /u - ʊ/ pair was most nativelike, followed by the /ɑ - ʌ/ and /i - ɪ/ pairs (nearly tied), and finally the /i - ɪ/ pair (see Table 4.12 for nativelikeness).

NR talkers displayed more variation (indicated by high SD) within the /i - ɪ/ and /ɑ - ʌ/ contrasts while NE talkers displayed more variation on /ɛ - æ/ and /u - ʊ/. The difference was greatest for /i - ɪ/, where NR SD was over twice that of the NE values for the same pair. Some NR

Table 4.9 Euclidean distance between members of four vowel pairs (EDP), NR and NE talkers

Talker	i - ɪ	æ - ɛ	u - ʊ	ɑ - ʌ
A	122.4	35.6	164.5	53.3
B	429.9	173.4	36.3	227.7
C	258.2	367.2	159.5	459.9
D	11.1	51.2	46.9	63.4
E	495.0	270.0	171.6	299.8
F	182.9	8.9	115.3	42.3
G	56.1	99.4	26.5	229.7
H	346.5	236.9	136.3	101.1
J	730.1	302.5	290.8	327.9
K	621.4	172.8	68.6	125.0
L	12.8	101.4	100.5	20.3
M	45.1	91.4	30.2	27.7
N	50.7	48.1	198.0	130.4
P	147.5	179.0	99.0	97.9
Q	41.3	66.6	45.1	84.3
R	94.2	72.6	36.9	116.8
S	411.8	188.3	212.2	69.7
T	691.4	396.4	298.6	496.1
U	456.1	173.0	245.8	162.1
V	17.1	25.1	32.2	52.2
Mean	261.09	153.00	125.74	159.37
Min	11.14	8.86	26.48	20.26
Max	730.2	396.4	298.6	496.1
SD	242.49	113.99	88.89	139.27
Norm Distrib	**No 0.0136***	Yes 0.0947	**No 0.0464***	**No 0.0030***

Talker	i - ɪ	æ - ɛ	u - ʊ	ɑ - ʌ
1	722.0	720.4	302.3	310.5
2	743.7	733.2	231.7	294.4
3	516.3	504.6	112.2	116.0
4	652.1	673.9	322.4	279.5
5	401.7	413.9	183.7	158.2
6	541.2	511.9	164.9	243.7
7	710.7	728.6	425.0	398.6
8	489.2	493.6	223.8	207.0
9	530.4	534.1	156.0	98.2
Mean	589.7	590.5	235.8	234.0
Min	401.7	413.9	112.2	98.2
Max	743.7	733.2	425.0	398.6
SD	120.6	122.8	98.3	98.5
Norm Distrib	Yes 0.3737	Yes 0.1000	Yes 0.6114	Yes 0.8596

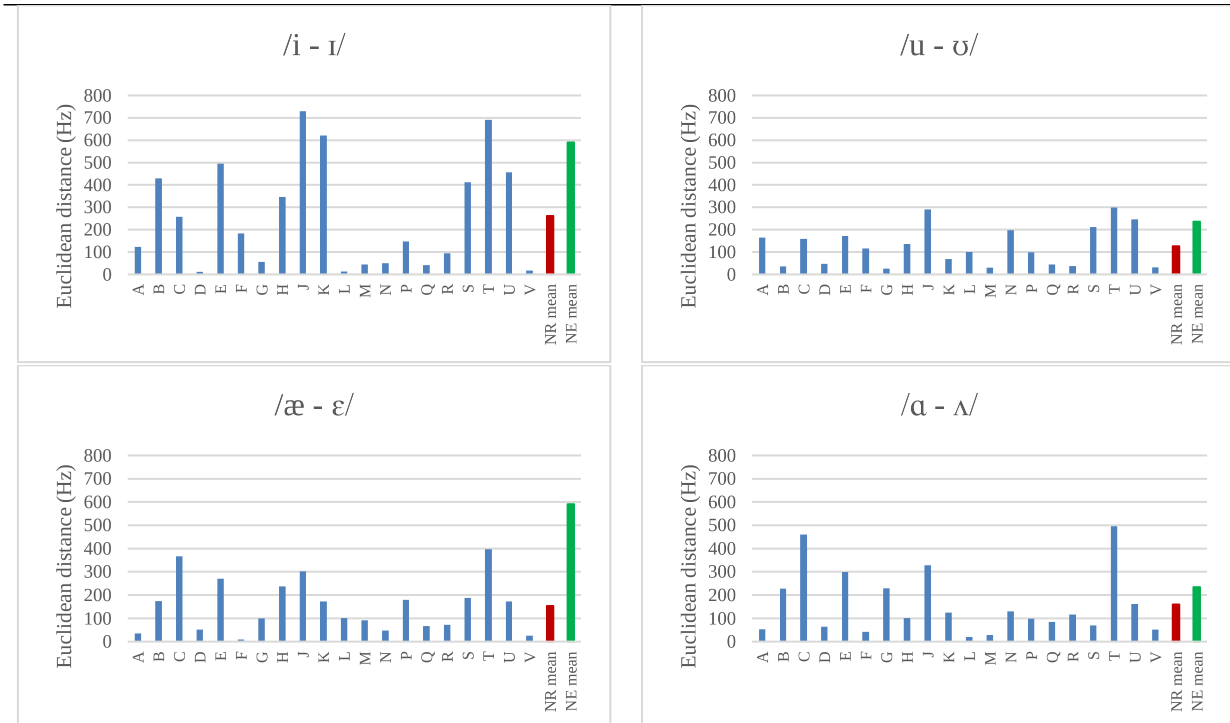
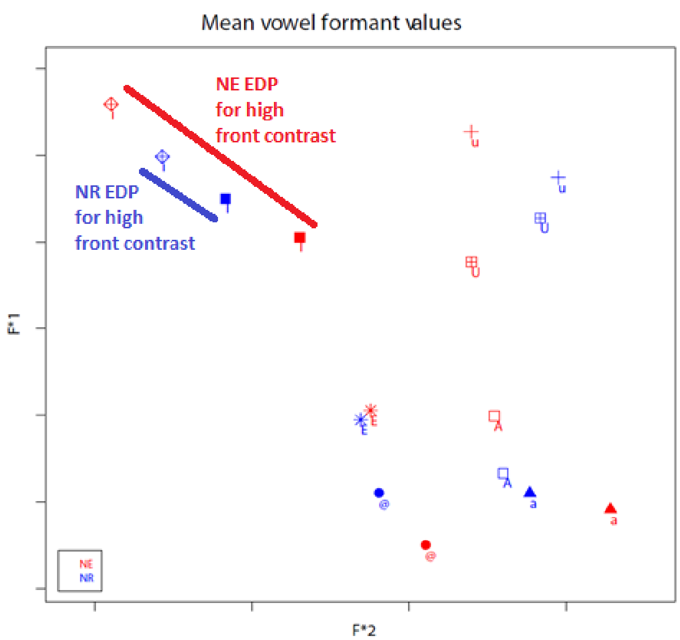


Figure 4.9 Euclidean distance between members of four vowel pairs (EDP), NR talkers and NR and NE means



Symbol	Phoneme	Represented
i	i *as in*	'bead'
I	ɪ	'bid'
E	ɛ	'bed'
@	æ	'bad'
u	u	'booed'
U	ʊ	'good'
a	ɑ	'pod'
A	ʌ	'bud'

Figure 4.10 Visual example of EDP: /i - ɪ/

talkers made a very small distinction between members of the four tested vowel pairs, indicated by low minimum EDP values (e.g., only 8.86 Hz between /u/ and /ʊ/ for Participant F), suggesting that the pair is not well differentiated by those speakers in terms of spectral differences. NE minimum values for all four contrasts were much higher than the NR minimums (the lowest NE EDP being a difference of 98.2 Hz between Participant 9's /ɑ/ and /ʌ/), indicating that as a group, the NE talkers make a greater spectral distinction (greatest for /ɛ - æ/ and smallest for /u - ʊ/) between the members of all four tested vowel pairs on the F1/F2 plane than do the NR talkers.

A somewhat smaller difference was observed in maximum values between the NR and NE groups. Maximum values for EDP between /i - ɪ/ were 730.2 Hz and 743.7 Hz for NR and NE talkers, respectively; 396.4 Hz and 733.2 Hz for /ɛ - æ/; 298.6 Hz and 425.0 Hz for /u- ʊ/; and 496.1 Hz and 398.6 Hz for /ɑ - ʌ/. In the case of /i - ɪ/, I interpret this similarity in otherwise quite distinct groups to be a consequence of interaction with other vowels in the vowel space; perhaps these maximum values represent the maximum distance by which the vowels in question *can* be separated, given the organization of the vowel space (that is, the maximum by which they may be separated before /ɪ/ begins to overlap with /e/, /ɛ/ etc.). The extraordinarily high maximum EDP observed in the NR group for the /ɑ - ʌ/ pair can be attributed to Participant T's unusually high and fronted /ʌ / paired with a more nativelike but low /ɑ/.

The data were normally distributed for all four NE contrasts and for /ɛ - æ/ in the NR group, but not normally distributed for the NR /i - ɪ/, /u - ʊ/, and /ɑ - ʌ/ contrasts. Because a portion of the data were not normally distributed, a Wilcoxon test rather than an ANOVA was used to determine if differences in EDP between /i - ɪ/, /ɛ - æ/, /u - ʊ/, and /ɑ - ʌ/ between the NR and NE groups were significant. To test whether L1 predicts EDP between the problematic vowel contrasts, each talker's mean EDP for each tested vowel pair was used as the dependent variable, and L1 – Russian or English – was used as the independent variable. These results are reported in Table 4.10. Differences between NR and NE groups were significant for 3 contrasts – /i - ɪ/, /ɛ - æ/, and /ɑ - ʌ/, and approached significance for the /u - ʊ/ contrast (p = 0.0695). Since the group sizes were small and standard deviations were large, I interpret this to mean that the NR and NE groups performed most similarly on this contrast in terms of EDP; indeed, mean values for this contrast were more similar between the NR and NE groups than on other contrasts. These results are in line with the conclusions drawn from the analysis of raw and normalized frequencies in 4.2.1 above, which suggested that NR takers make a smaller spectral distinction between all four contrasts.

A bivariate fit model was used to test for correlations between the NR EDP for each vowel pair and talker AOA, LOR in GA, and LOR in the USA (Table 4.11). The only statistically significant correlations found were for AOA, which correlated with NR talkers' Euclidian distance between /æ - ɛ/, /u - ʊ/, and /ɑ - ʌ/ and came close to correlation on /i - ɪ/. The correlation of AOA

Table 4.10 Comparison of EDP differences among NR and NE talkers on four vowel contrasts. Statistically significant differences are shaded gray.

Contrast tested	Significance
/i - ɪ/	S = 201, Z = 3.08770, **p = .0020***
/ɛ - æ/	S = 199, Z = 2.99342, **p = .0028***
/u - ʊ/	S = 174, Z = 1.81491, ***p = .0695***
/ɑ - ʌ/	S = 188, Z = 2.47487, **p = .0133***

Table 4.11 Correlation of EDP with AOA, LOR in GA, and LOR in USA on four vowel contrasts, NR talkers. Statistically significant correlations are shaded gray.

	AOA	LOR in USA	LOR in GA
/i - ɪ/	r = -0.42734	r = -0.13238	r = -0.03517
	p = 0.0602	p = 0.5780	p = 0.8830
/æ - ɛ/	r = -0.64669	r = -0.0631	r = 0.03056
	p = 0.0021*	p = 0.7916	p = 0.8982
/u - ʊ/	r = -0.55979	r = -0.18292	r = -0.05272
	p = 0.0103*	p = 0.4402	p = 0.8253
/ɑ - ʌ/	r = -0.51162	r = 0.086453	r = 0.051487
	p = 0.0211*	p = 0.7170	p = 0.8293

with distance between members of problematic L2 vowel pairs on the F1/F2 plane is not surprising, as many sources report influence of AOA on the ability of L2 learners to acquire novel contrasts (see §2.3 for a discussion of the link between AOA and L2 phonological acquisition). LOR is also frequently implicated in the acquisition of L2 contrasts; however, the role of LOR diminishes after the first few years in adult learners, and is thought by some scholars to lose its predictive power for L2 foreign accent after only one year of residence, particularly when not controlling for factors such as amount and quality of input (Krashen, Long & Scarcella, 1979; Long, 1990; Flege et al., 1992; Flege, Munro, & Fox, 1994; Munro, 1993). Thus, it is not surprising that LOR was not found to be significant for this group of participants, all of whom had lived in the US for at least five years (LOR ranged from 5 to 27 years; mean = 16.0 years; SD = 4.5 years) and in GA for at least five years (LOR in GA ranged from 5 to 27 years; mean = 14.4 years; SD = 5.7 years).

Next, I tested for correlations between nativelikeness of EDP (EDPN – see §3.4.3 for calculation details), and the same biographical variables that were previously tested for correlation with EDP: 1) AOA, 2) LOR in the U.S., and 3) LOR in GA. Recall that for EDPN, small values indicate performance that is more nativelike, while large values indicate a greater divergence from the NE mean (i.e. less nativelike performance). These values (EDPN), along with actual corresponding NR EDP, are shown in Table 4.12. The five values in each column that are closest to the NE mean (most nativelike) are shaded in gray. These five most nativelike EDP values vary in how close they fall to the NE mean on the four contrasts: for the /i - ɪ/ contrast, they range from 5.4% to 23.8% difference from the NE mean; for /æ - ɛ/, they range from 32.9% to 59.9% difference from the NE mean (that is, even the most nativelike EDP on this contrast is quite different from the NE mean); for /u - ʊ/, they range from 4.2% to 26.6% difference from the NE mean; and for /ɑ - ʌ/, they range from 1.8% to 40.1%.

A bivariate fit model was used to test for correlations between these EDPN values for each NR and talker AOA, LOR in GA, and LOR in the USA for each vowel pair (Table 4.13). Two

Table 4.12 EDP and EDPN for four vowel contrasts or NR talkers. The five most nativelike values in each column – representing the top 25% of performance – are shaded gray.

	/i - ɪ/		/æ - ɛ/		/u - ʊ/		/ɑ - ʌ/	
Talker	**EDP**	**EDPN**	**EDP**	**EDPN**	**EDP**	**EDPN**	**EDP**	**EDPN**
A	122.4	0.792	35.6	0.940	164.5	0.302	53.3	0.772
B	429.9	0.271	173.4	0.706	36.3	0.846	227.7	0.027
C	258.2	0.562	367.2	0.378	159.5	0.324	459.9	0.965
D	11.1	0.981	51.2	0.913	46.9	0.801	63.4	0.729
E	495.0	0.161	270.0	0.543	171.6	0.272	299.8	0.281
F	182.9	0.690	8.9	0.985	115.3	0.511	42.3	0.819
G	56.1	0.905	99.4	0.832	26.5	0.888	229.7	0.018
H	346.5	0.412	236.9	0.599	136.3	0.422	101.1	0.568
J	730.1	0.238	302.5	0.488	290.8	0.233	327.9	0.401
K	621.4	0.054	172.8	0.707	68.6	0.709	125.0	0.466
L	12.8	0.978	101.4	0.828	100.5	0.574	20.3	0.913
M	45.1	0.924	91.4	0.845	30.2	0.872	27.7	0.882
N	50.7	0.914	48.1	0.919	198.0	0.160	130.4	0.443
P	147.5	0.750	179.0	0.697	99.0	0.580	97.9	0.582
Q	41.3	0.930	66.6	0.887	45.1	0.809	84.3	0.640
R	94.2	0.840	72.6	0.877	36.9	0.844	116.8	0.501
S	411.8	0.302	188.3	0.681	212.2	0.100	69.7	0.702
T	691.4	0.172	396.4	0.329	298.6	0.266	496.1	1.120
U	456.1	0.227	173.0	0.707	245.8	0.042	162.1	0.307
V	17.1	0.971	25.1	0.957	32.2	0.863	52.2	0.777

Table 4.13 Correlation of EDPN with AOA, LOR in GA, and LOR in USA on four vowel contrasts, NR talkers. Statistically significant correlations are shaded gray.

	AOA	LOR in USA	LOR in GA
/i - ɪ/	r = 0.315545 p = 0.1753	r = 0.068542 p = 0.7740	r = 0.009875 p = 0.9670
/æ - ɛ/	r = 0.64637 **p = 0.0021***	r = 0.062326 p = 0.7941	r = -0.03136 p = 0.8956
/u - ʊ/	r = 0.375329 p = 0.1030	r = 0.12961 p = 0.5860	r = 0.03865 p = 0.8715
/ɑ - ʌ/	r = -0.21305 p = 0.3671	r = 0.276601 p = 0.2378	r = 0.473102 **p = 0.0351***

statistically significant correlations emerged. AOA positively correlated with NR talkers' EDPN for /æ - ɛ/, and LOR in GA positively correlated with NR talkers' DEPN for /ɑ - ʌ/. Recall that lower EDPN values represent more nativelike EDPs between the associated vowel pairs; thus, low AOA (which might be expected based on the literature) and low LOR (which is surprising) were correlated with nativelike performance. Note, however, the considerable range, and in some cases quite high minimum values, for EDPN (see discussion on previous page). For the /æ - ɛ/ contrast, as well as for /ɑ - ʌ/, even the most nativelike EDPNs were quite far from the NE mean.

To summarize the spectral results of Experiment 1 before moving on to duration, NR talkers made a smaller distinction between all four tested contrasts and had greater standard deviations for all tested vowels than did NE talkers. Lower AOA was associated with increased nativelikeness on six out of eight tested vowels (all but /i/ and /ɛ/); surprisingly, longer residence in the U.S. was associated with less nativelike production on /ɪ/, /æ/, and /ʊ/; and longer residence in GA was associated with more nativelike performance on /ɑ/. In terms of Euclidean distance between members of the tested vowel pairs, lower AOA was associated with more nativelike performance on /ɛ - æ/, and, also surprisingly, shorter residence in GA was associated with more nativelike performance on /ɑ - ʌ/.

4.2.4 Duration

This section presents duration data obtained from Experiment 1. Mean durations and descriptive statistics for all eight tested vowels are presented first for NR (Table 4.14), then NE (Table 4.15) talkers, then presented visually in Figure 4.11. Mean and maximum durations were consistently longer in the NR group (as were minimum durations in three cases), and standard deviations for the group were consistently greater. It is not clear whether these longer vowel

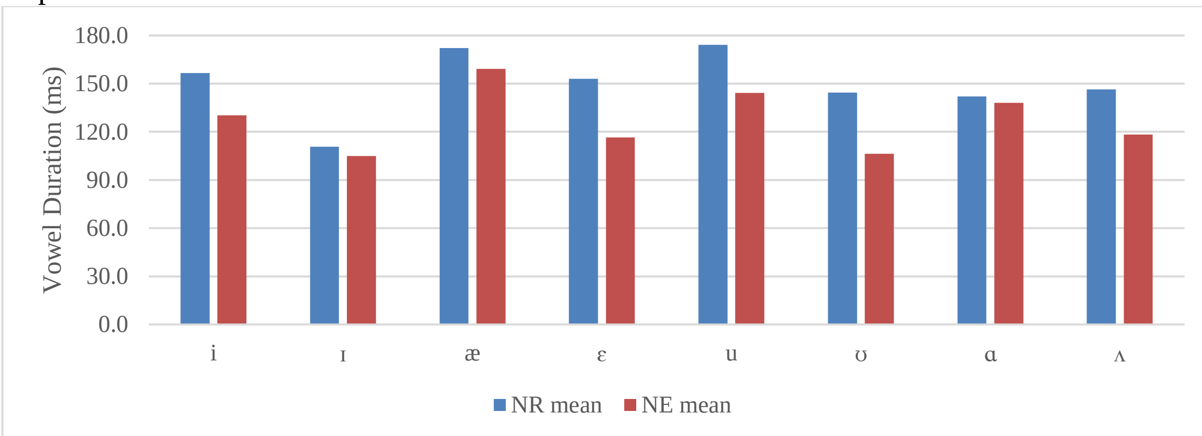
Table 4.14 Mean duration (ms) standard deviations (in parentheses) of eight American English vowels produced by N=20 native Russian talkers

	i	ɪ	ɛ	æ	u	ʊ	ɑ	ʌ
A	75 (80)	83 (21)	98 (19)	102 (15)	100 (25)	79 (22)	101 (10)	110 (17)
B	143 (50)	79 (36)	96 (23)	126 (47)	222 (123)	169 (95)	105 (29)	110 (27)
C	277 (92)	102 (42)	159 (45)	237 (63)	219 (66)	188 (66)	137 (19)	119 (34)
D	135 (28)	121 (28)	172 (30)	176 (34)	169 (48)	142 (26)	168 (27)	191 (24)
E	149 (29)	117 (31)	149 (32)	172 (26)	163 (70)	134 (38)	158 (24)	144 (42)
F	188 (62)	138 (42)	198 (32)	170 (36)	187 (64)	183 (21)	149 (37)	172 (43)
G	86 (12)	81 (12)	111 (19)	106 (12)	166 (105)	106 (59)	104 (17)	99 (14)
H	189 (33)	83 (19)	150 (51)	181 (40)	182 (44)	137 (51)	131 (14)	126 (13)
J	172 (49)	131 (26)	178 (55)	216 (55)	223 (76)	179 (53)	182 (48)	154 (36)
K	166 (39)	73 (25)	126 (26)	169 (48)	145 (66)	155 (35)	83 (20)	101 (24)
L	172 (26)	122 (37)	162 (28)	130 (20)	184 (71)	169 (55)	135 (21)	144 (39)
M	149 (25)	100 (33)	147 (34)	191 (42)	157 (67)	127 (27)	117 (31)	163 (32)
N	113 (31)	105 (24)	166 (34)	176 (36)	155 (50)	116 (42)	149 (32)	165 (33)
P	135 (40)	105 (24)	151 (34)	172 (32)	152 (80)	124 (52)	166 (65)	143 (32)
Q	246 (45)	216 (62)	262 (44)	277 (40)	269 (118)	240 (43)	237 (44)	261 (66)
R	106 (29)	75 (25)	102 (39)	105 (30)	136 (77)	96 (43)	94 (19)	103 (24)
S	146 (27)	147 (37)	199 (60)	191 (47)	143 (42)	130 (41)	174 (42)	170 (57)
T	136 (22)	130 (18)	133 (24)	175 (35)	145 (47)	133 (28)	149 (27)	147 (20)
U	229 (39)	129 (35)	160 (39)	236 (66)	212 (62)	144 (60)	166 (46)	141 (34)
V	121 (38)	99 (24)	146 (31)	134 (13)	155 (67)	136 (44)	138 (35)	165 (52)
Mean	156.6	110.8	153.1	172.2	174.2	144.4	142.1	146.4
Min	75	73	96	102	100	79	83	99
Max	277	216	262	277	269	240	237	261
SD	51	34	39	47	39	36	36	38

Table 4.15 Mean duration (ms) and standard deviations (in parentheses) of eight American English vowels produced by nine native English talkers

	i	ɪ	ɛ	æ	u	ʊ	ɑ	ʌ
1	124 (21)	110 (32)	118 (30)	152 (31)	133 (36)	96 (31)	127 (29)	120 (38)
2	152 (21)	142 (24)	138 (30)	187 (29)	168 (30)	150 (28)	164 (28)	151 (30)
3	120 (18)	99 (17)	114 (30)	146 (24)	150 (39)	98 (21)	134 (22)	120 (29)
4	90 (10)	72 (12)	87 (20)	132 (21)	102 (20)	79 (14)	106 (12)	87 (14)
5	110 (16)	83 (18)	104 (26)	131 (25)	116 (24)	83 (16)	138 (29)	91 (20)
6	114 (18)	83 (16)	93 (12)	165 (18)	116 (24)	91 (25)	145 (18)	103 (13)
7	129 (19)	89 (18)	104 (13)	158 (34)	142 (28)	98 (17)	137 (19)	104 (22)
8	197 (64)	153 (49)	149 (47)	207 (34)	228 (116)	150 (68)	153 (56)	149 (44)
9	136 (27)	115 (36)	139 (36)	154 (27)	142 (38)	115 (43)	140 (40)	138 (53)
Mean	130.4	105.0	116.4	159.2	144.2	106.4	138.0	118.2
Min	90	72	87	131	102	79	106	87
Max	197	153	149	207	228	15	164	151
SD	30	28	22	25	37	27	16	24

Top: Mean Durations of Individual Vowels as a Function of L1[37]



Bottom: Mean Durations of Individual Vowels Organized by Typically Long-Short Pairs

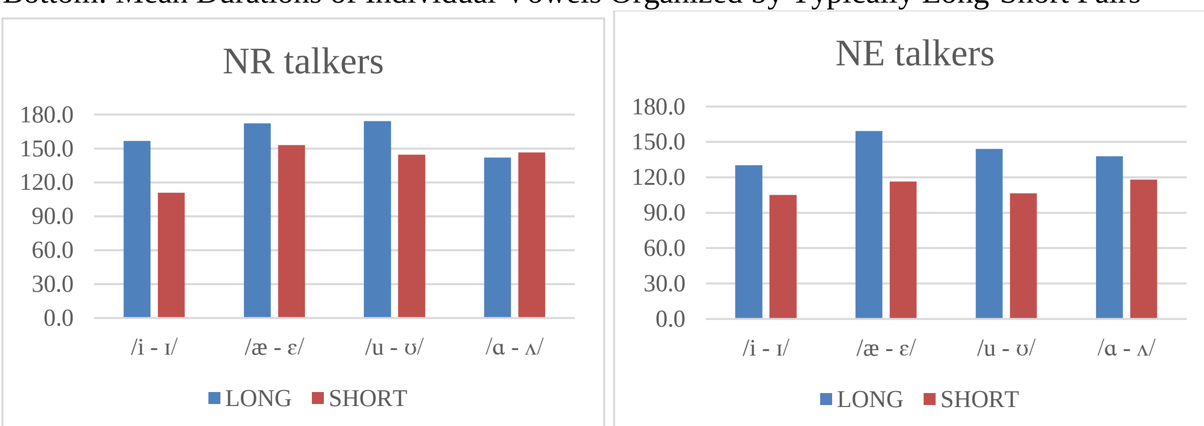


Figure 4.11 Individual Durations of Eight American English Vowels

durations indicate an overall slower speaking rate, as vowel duration was not normalized to word duration. I chose not to normalize vowel duration to word duration for three reasons having to do with variation in the duration of onset and coda consonants: 1) pre-voicing of onset plosives p/b varied greatly both between and within L1 backgrounds (especially for NR), 2) some NR talkers affricated their coda t/d, extending word duration, and 3) some NE talkers did not release those final consonants, shortening word duration.

[37] Note that these durations are not normalized as a function of word duration, as onset and coda consonants displayed great variation between language groups, as well as some individuals (see §3.4.3 Experiment 1 Analysis).

Duration ratios (**DRs**) between the long and short members of the four tested vowel pairs were calculated by dividing the mean duration (in ms) of the longer vowel by that of the shorter vowel in each pair for each speaker (see §3.4.3). These results are given in Table 4.16 and represented visually in Figure 4.12. As a group, the NE talkers made a greater duration distinction (represented by a greater ratio of long to short vowel) between three of the four tested vowel pairs than did the NR group; only on /i - ɪ/ did the NR group make a greater duration distinction. In native production of the four vowel contrasts tested, /i/ was on average 1.26 times as long as /ɪ/,/æ/ 1.39 times as long as /ɛ/, /u/ 1.36 times as long as /ʊ/, and /ɑ/ 1.19 times as long as /ʌ/. The NR

Table 4.16 Duration Ratios (DRs) in between Four Vowel Contrasts

	DRs in NR Speech						**DRs in NE Speech**				
	i:ɪ	**æ:ɛ**	**u:ʊ**	**ɑ:ʌ**	**Mean**	**ID**	**i:ɪ**	**æ:ɛ**	**u:ʊ**	**ɑ:ʌ**	**Mean**
A	0.895	1.043	1.262	0.922	1.031	**1**	1.129	1.286	1.386	1.053	1.214
B	1.806	1.317	1.310	0.955	1.347	**2**	1.072	1.356	1.121	1.084	1.158
C	2.721	1.496	1.166	1.151	1.634	**3**	1.202	1.284	1.541	1.111	1.285
D	1.116	1.022	1.189	0.883	1.053	**4**	1.255	1.513	1.295	1.209	1.318
E	1.267	1.155	1.217	1.098	1.184	**5**	1.356	1.256	1.404	1.514	1.382
F	1.366	0.861	1.026	0.867	1.030	**6**	1.380	1.768	1.279	1.400	1.457
G	1.066	0.958	1.563	1.048	1.159	**7**	1.444	1.520	1.455	1.313	1.433
H	2.274	1.207	1.327	1.047	1.464	**8**	1.284	1.387	1.518	1.030	1.305
J	1.312	1.215	1.244	1.179	1.237	**9**	1.185	1.108	1.242	1.014	1.137
K	2.270	1.338	0.933	0.812	1.338	**Mean**	**1.256**	**1.386**	**1.360**	**1.192**	
L	1.407	0.803	1.086	0.939	1.059	**Min**	1.072	1.108	1.121	1.014	
M	1.485	1.300	1.236	0.716	1.184	**Max**	1.444	1.768	1.541	1.514	
N	1.081	1.066	1.340	0.905	1.098	**SD**	0.122	0.192	0.138	0.179	
P	1.283	1.137	1.224	1.161	1.201	**Norm Distr**	Yes 0.9604	Yes 0.6633	Yes 0.8513	Yes 0.1839	
Q	1.143	1.056	1.119	0.907	1.056						
R	1.419	1.034	1.414	0.915	1.195						
S	0.989	0.958	1.097	1.023	1.017						
T	1.042	1.318	1.092	1.010	1.115						
U	1.768	1.482	1.471	1.179	1.475						
V	1.556	0.923	1.138	0.835	1.113						
Mean	**1.463**	**1.134**	**1.223**	**0.978**							
Min	0.895	0.803	0.933	0.716							
Max	2.721	1.496	1.563	1.179							
SD	0.485	0.196	0.153	0.132							
Norm Distr	No **0.0117***	Yes 0.6844	Yes 0.9678	Yes 0.4842							

group produced /i/ 1.46 times as long as /ɪ/, /æ/ 1.13 times as long as /ɛ/, /u/ 1.22 times as long as /ʊ/, and /ɑ/ 0.98 times as long as /ʌ/.

Minimum values were consistently lowest in the NR group; that is, for all four contrasts, some NR talkers made a smaller difference in duration between the members of that vowel pair

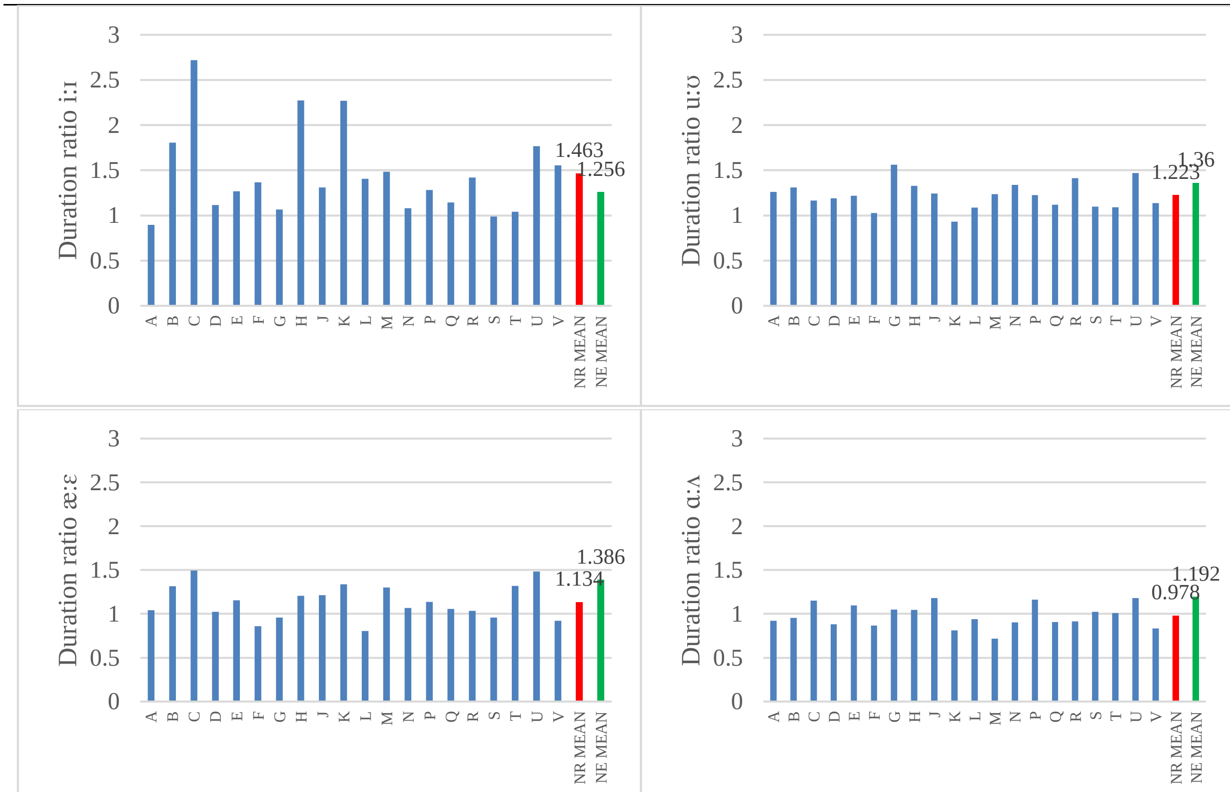


Figure 4.12 Comparison of NR talkers' and NE mean DR between long and short vowels in each pair. Top: /i - ɪ/ (left), /u - ʊ/ (right); bottom /ɛ - æ/ (top), /ɑ - ʌ/ (bottom)

than any NE talker. Maximum values were greatest in the NR group for /i - ɪ/ (/i/ 2.721 as long as /ɪ/), greatest in the NE group for /ɛ - æ/ and /ɑ - ʌ/ (/æ/ 1.768 times as long as /ɛ/ and /ɑ/ 1.514 times as long as /ʌ/), and nearly equal between the NR and NE groups for /u - ʊ/ (/u/ 1.563 as long as /ʊ/ for the NR group and 1.541 as long as /ʊ/ for the NE group).

Among the four vowel contrasts tested, native DRs were greatest for /æ/ and /ɛ/ and least for /ɑ/ and /ʌ/. For the NR group, DRs were greatest for /i/ and /ɪ/ and least for /ɑ/ and /ʌ/. That is, the NE and NR groups differed in which contrast they produced with the greatest DRs, but both groups had their smallest DRs for /ɑ/ and /ʌ/. The NR group has a very high standard deviation on the /i - ɪ/ contrast, caused by the performance of a few individuals (in particular, C, H, and K) who made a very large duration distinction between /i/ and /ɪ/, possibly indicating different approaches

Table 4.17 Difference in DRs in four vowel pairs, NR and NE talkers. Statistically significant differences are shaded gray.

Contrast tested	Significance
/i - ɪ/	S = 119, Z = -0.73068, p = .4650
/æ - ɛ/	S = 113, Z = 2.71058, **p = .0067***
/u - ʊ/	S = 181, Z = 2.14489, **p = .0320***
/ɑ - ʌ/	S = 116, Z = 2.85200, **p = .0040***

to this vowel contrast. NR performance was overall more uniform on the other three contrasts, indicated by lower standard deviations and means more similar to the NE mean.

This variation in NR performance on the /i - ɪ/ contrast is reflected in statistical analysis of these data: the DRs reported in Table 4.16 were normally distributed for both the NR and NE groups with the exception of NR /i - ɪ/. Because a portion of the DRs data were not normally distributed, a Wilcoxon test was used to determine if differences in DRs between the members of the pairs /i - ɪ/, /ɛ - æ/, /u - ʊ/, and /ɑ - ʌ/ between the NR and NE groups were significant. To test whether L1 predicts the size of the DRs between long and short vowels in the tested pairs, each talker's ratio of long vowel to short vowel of each tested pair was used as the dependent variable, and L1 – Russian or English – was used as the independent variable. These results are presented in Table 4.17. Differences between NR and NE groups were significant for 3 contrasts – /ɛ - æ/, /u - ʊ/, and /ɑ -ʌ/. Only the contrast /i - ɪ/ did not reveal a statistically significant difference between the two groups, likely due to the variation in NR production on this contrast, discussed above.

Next, I tested for correlations between two different measures of duration in NR production: 1) DRs between members of the four tested vowel pairs (as above), and 2) *nativelikeness* of NR talkers' DRs (DRN – see 3.4.3 for details on calculation), and the same biographical variables that were previously tested for correlation with distance between vowels: 1) AOA, 2) LOR in the U.S., and 3) LOR in GA. Recall that for this measure, small values indicate performance that is more nativelike, while large values indicate a greater divergence from the NE mean (i.e. less nativelike performance). These values (DRN), along with actual NR duration ratios (DR) between long and short vowels, are shown in Table 4.18. The five values in each column that are closest to the NE mean (most nativelike) are shaded in gray; all shaded values fall within 9% of the NE mean, indicating relatively nativelike performance on duration for those vowel contrasts. Note that some speakers (E, F, J, P, R, U) fall within this narrow range of relatively nativelike values by producing a slightly greater duration difference in the indicated vowel contrasts than do NE talkers (all values higher than the NE mean are bold in Table 4.18), while others fall just below the NE mean.

A bivariate fit model was used to test first for correlations between DRs for the four tested vowel pairs and each talker's AOA, LOR in the USA, and LOR in GA for each vowel pair, then for correlations between DRN for the four tested vowel pairs and each talker's AOA, LOR in the USA, and LOR in GA for each vowel pair (Table 4.19). AOA was negatively correlated with DR and positively correlated with DRN (recall that ***larger*** DR values indicate a larger duration ratio between the two vowels, while ***smaller*** DRN values indicate more nativelike performance) for /ɑ/ and /ʌ/ ($p = 0.0085^*$); that is, those NR talkers who had arrived to the US at a younger age made a significantly larger and more nativelike duration distinction between the vowels /ɑ/ and /ʌ/ than did later arrivals. LOR in GA was positively correlated with DRN for /u - ʊ/ ($p = 0.0385^*$), indicating that those talkers who had resided in GA the least produced a more nativelike duration

Table 4.18 NR duration ratios between four vowel contrasts and adjusted distance of those ratios from the NE mean. The five most nativelike values in each column – representing the top 25% of performance – are shaded gray; all values exceeding NE mean are in bold.

	/i - ɪ/		/æ - ɛ/		/u - ʊ/		/ɑ - ʌ/	
ID	**DR**	**DRN**	**DR**	**DRN**	**DR**	**DRN**	**DR**	**DRN**
A	0.895	0.288	1.043	0.247	1.262	0.072	0.922	0.227
B	**1.806**	0.438	1.317	0.050	1.310	0.037	0.955	0.199
C	**2.721**	1.167	**1.496**	0.080	1.166	0.142	1.151	0.035
D	1.116	0.111	1.022	0.262	1.189	0.126	0.883	0.259
E	**1.267**	0.009	1.155	0.167	1.217	0.105	1.098	0.079
F	**1.366**	0.087	0.861	0.379	1.026	0.246	0.867	0.273
G	1.066	0.151	0.958	0.309	**1.563**	0.149	1.048	0.121
H	**2.274**	0.810	1.207	0.129	1.327	0.024	1.047	0.122
J	**1.312**	0.044	1.215	0.123	1.244	0.085	1.179	0.011
K	**2.270**	0.807	1.338	0.035	0.933	0.314	0.812	0.319
L	**1.407**	0.120	0.803	0.421	1.086	0.201	0.939	0.212
M	**1.485**	0.182	1.300	0.062	1.236	0.091	0.716	0.400
N	1.081	0.139	1.066	0.231	1.340	0.015	0.905	0.241
P	**1.283**	0.022	1.137	0.180	1.224	0.100	1.161	0.026
Q	1.143	0.090	1.056	0.238	1.119	0.177	0.907	0.239
R	**1.419**	0.130	1.034	0.254	**1.414**	0.040	0.915	0.233
S	0.989	0.212	0.958	0.309	1.097	0.193	1.023	0.141
T	1.042	0.171	1.318	0.049	1.092	0.197	1.010	0.153
U	**1.768**	0.407	**1.482**	0.070	**1.471**	0.081	1.179	0.011
V	**1.556**	0.238	0.923	0.334	1.138	0.163	0.835	0.300

ratio between /u/ and /ʊ/. It is not immediately clear why this should be the case, as greater, not smaller, LOR would be expected to yield more nativelike results, especially compared to a localvariety. Two other correlations approached statistical significance: the correlation of LOR in GA and DR for /u - ʊ/ (p = 0.0575) – that is, those NR talkers who resided in GA longest produced the greatest duration ratios between /u/ and /ʊ/; and the correlation of LOR in the U.S. with DRN for the /i - ɪ/ contrast (p = 0.0627) – that is, those NR talkers who resided in the U.S. longest produced least nativelike duration ratios between /i/ and /ɪ/. As with the correlation of LOR in GA

Table 4.19 Correlation of duration with AOA, LOR in GA, and LOR in USA on four vowel contrasts for NR talkers. Correlations with DR on top, with DRN on bottom. Statistically significant correlations are shaded gray.

	Correlation with DR		
	AOA	LOR in USA	LOR in GA
/i - ɪ/	r = 0.152353	r = 343426	r = 0.28466
	p = 0.5214	p = 0.1382	p = 0.2238
/æ - ɛ/	r = -0.33555	r = 0.096213	r = 0.104229
	p = 0.1481	p = 0.6866	p = 0.6619
/u - ʊ/	r = -0.11269	r = -0.2053	r = -0.4314
	p = 0.6362	p = 0.3852	***p = 0.0575***
/ɑ - ʌ/	r = -0.57117	r = -0.32178	r = -0.21874
	p = 0.0085*	p = 0.1665	p = 0.3542
	Correlation with DRN		
	AOA	LOR in USA	LOR in GA
/i - ɪ/	r = 0.18438	r = 0.423643	r = 0.368943
	p = 0.4365	***p = 0.0627***	p = 0.1094
/æ - ɛ/	r = 0.353743	r = 0.024963	r = 0.022588
	p = 0.1260	p = 0.9168	p = 0.9247
/u - ʊ/	r = 0.281808	r = 0.210723	r = 0.465698
	p = 0.2287	p = 0.3725	**p = 0.0385***
/ɑ - ʌ/	r = 0.571166	r = 0.32178	r = 0.218744
	p = 0.0085*	p = 0.1665	p = 0.3542

and DRN for /u - ʊ/, it is not clear why this should be the case, as longer, not shorter, LOR would be expected to correlate with more nativelike performance.

To summarize the duration results of Experiment 1 before moving on to the results of Experiments 2 and 3, NR duration performance was significantly different from NE duration performance on all contrasts but /i - ɪ/, where the NR group showed considerable variation and data were not normally distributed, with several individuals making a very large duration distinction while others performed more like natives. The NR talkers' duration ratios (DR) between /ɑ/ and /ʌ/ correlated negatively with AOA, while DRN, calculated by comparing NR individual performance to the NE mean, positively correlated with AOA on the /ɑ - ʌ/ contrast and with LOR in GA on the /u - ʊ/ contrast; that is, NR talkers who arrived to the US at a younger age made a greater and more nativelike duration distinction between /ɑ/ and /ʌ/, and those who had resided in GA the least produced a more nativelike duration ratio between /u/ and /ʊ/.The next section presents results of Experiment 2, dealing with the perception of NE speech by NR listeners.

4.3 Experiment 2 – NR Listeners' Perception of a NE talker

Results from Experiment 2, which investigated NR perception of vowels spoken by a single NE talker, are summarized in Table 4.20, and individual performance on all eight vowels (grouped into pairs to facilitate within-contrast comparison) is presented in Figure 4.13. As a group, NR listeners most accurately identified /ɑ/ (83.3% accuracy) and least accurately identified /u/ (53.3% accuracy) in NE speech. Accuracy of identification of the eight vowels tested, in order form highest to lowest, was: ɑ > æ > ʊ > ɪ > ʌ > i > ɛ > u.

All stimulus vowels except /ʊ/ had at least one score of 100% accuracy of perception by a NR listener in Experiment 2. That is, with the exception of /ʊ/, each vowel was correctly perceived by at least one listener every time it was heard by that listener, while /ʊ/ had a maximum perception score of only 93.30%, potentially indicating a ceiling effect for the most proficient listeners on all

Table 4.20 Identification of eight American English vowels in NE speech by 20 NR listeners

		Auditory Stimulus							
		/i/	**/ɪ/**	**/ɛ/**	**/æ/**	**/u/**	**/ʊ/**	**/ʌ/**	**/ɑ/**
Response	**/i/**	65.8%	25.8%						
	/ɪ/	34.2%	74.2%						
	/ɛ/			65.4%	23.3%				
	/æ/			34.6%	76.7%				
	/u/					53.3%	25.7%		
	/ʊ/					46.7%	74.3%		
	/ʌ/							71.1%	16.7%
	/ɑ/							28.9%	83.3%
	Min	25.00%	33.30%	25.00%	41.70%	0.00%	46.70%	33.30%	55.60%
	Max	100.00%	100.00%	100.00%	100.00%	100.00%	93.30%	100.00%	100.00%
	SD	0.234	0.218	0.165	0.152	0.332	0.141	0.174	0.137

Figure 4.13 Identification of eight American English vowels, grouped by adjacent pairs, in NE speech by NR listeners. See Appendix D for actual values.

vowels but /ʊ/. Minimum values for NR perception of these eight vowels are perhaps even more revealing: while minimum perception scores for the other seven vowels fell between 25% and 55.6%, the minimum perception score for /u/ was zero. That is, while some NR listeners identified /u/ with 100% accuracy, three NR listeners never correctly identified /u/ when they heard it as theauditory stimulus in Experiment 2, identifying it instead as /ʊ/; interestingly, these three participants (D, G, and R) identified /ʊ/ with 60%, 93.3%, and 80% accuracy, respectively, when it was the auditory stimulus. Four other listeners (C, S, U, and V) also scored quite low on perception of /u/ (33.3%) but relatively high on perception of /ʊ/ (73.3-93.3%). No NR listeners showed the opposite trend – identifying /u/ disproportionately better than /ʊ/; E came closest, with 100% accuracy on /u/ and 73.3% accuracy on /ʊ/, and only five other NR listeners performed better on /u/ than /ʊ/, all with a smaller margin than participant E. It is not clear why this should be the case; however, two possible explanations stand out:

1. These listeners' category for English /u/ is influenced by the Russian /u/, which is considerably more back than English /u/ (see NR and NE production results compared with Russian vowels in Figures 4.2 and 4.4 above), causing these listeners to judge both /u/ and /ʊ/ - strongly centralized in southern English – to be poor exemplars of /u/ and thus, by process of elimination, better exemplars of /ʊ/.
2. These listeners are aware of the challenging nature of the /u - ʊ/ contrast (and particularly the phoneme /ʊ/ for L2 English speakers of their language background), and thus chose /ʊ/ disproportionately more often than /u/, fearing that they would not perceive /ʊ/ accurately when they heard it.

The latter seems to be a weaker standalone explanation than the former), as it is hard to imagine a NR listener identifying both /u/ and /ʊ/ as /ʊ/ only in nearly all instances that they were heard (as was the case for participant G) due entirely to self-doubt. However, it is possible that the two explanations operated in tandem to cause some listeners to much more frequently identify both /u/ and /ʊ/ stimuli as /ʊ/ but never vice versa.

In order to facilitate comparison of performance between vowel contrasts rather than between individual vowels, individual vowel results from Experiment 2 were averaged between members of each of the four tested vowel pairs; that is, one mean value was obtained for each listener's performance on /i/ and /ɪ/, one mean value for performance on /u/ and /ʊ/, etc.. This approach masks some of the details having to do with the perception of individual vowels (e.g. the idiosyncratic case of /u/ discussed above), but offers a more streamlined glimpse at how each listener performs on each tested vowel pair. These results are presented in Figure 4.14.

When vowel pair accuracy is calculated as a mean of accuracy on the two individual vowels that make up the pair, NR listeners most accurately perceived the /ʌ-ɑ/ contrast (the most accurately discriminated contrast for 10 NR listeners, with a mean accuracy of 77.2% across all

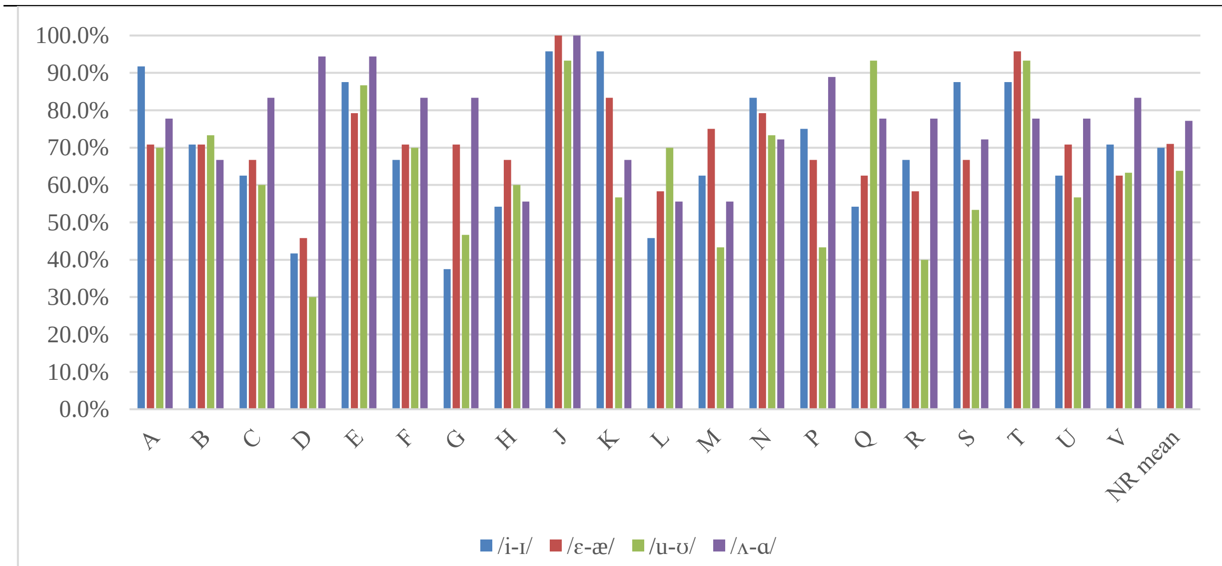


Figure 4.14 NR listener performance on all four contrasts. Calculated as mean of listening performance on individual vowels.

listeners) compared to the other three contrasts tested and least accurately perceived the /u - ʊ/ contrast for only three listeners, with a slightly higher mean accuracy (70.0%). Accuracy of contrast (the most accurately perceived contrast for only 3 listeners, with a mean accuracy of 63.8% across all listeners) in NE speech. The /i-ɪ/ contrast was also the most accurately perceived identification of the four vowel contrasts tested, in order form highest to lowest, was: /ʌ-ɑ/ (77.2%) > /ɛ - æ/ (71.0%) > /i-ɪ/ (70.0%) > 63.8% > /u - ʊ/ (63.8%). See Appendix D for a table of all vowel contrast accuracy values.

Two particularly low contrast values stand out: Participant D's performance on the /u - ʊ/ contrast (30% accuracy) and Participant G's performance on the /i-ɪ/ contrast (37.5% accuracy). As discussed above, Participant D had a 0% accuracy on identifying /u/ and a 60% accuracy on identifying /ʊ/, which resulted in her rather low accuracy on the /u - ʊ/. Participant G's low /i-ɪ/ accuracy may be attributed to low accuracy on both of the contrasted vowels : 41.7% accuracy on /i/ and 33.3% on /ɪ/. These two participants' overall low performance (see next section) may be at least partially explained by their relatively high AOA values (26 and 27, respectively), although their AOA values were not the highest the NR group, with five other NR participants having an equal or higher AOA (B – 26 years, F – 28 years, K – 32 years, L – 27 years, and N 29 years); however, it is not immediately clear why they should perform poorly on these particular vowels and contrasts. Participant G especially might be expected to perform well on the /i - ɪ/ contrast due to both academic and naturalistic exposure to Ukrainian; however, her performance on the /i - ɪ/ contrast, which might be hypothesized to be bolstered by experience with a similar contrast in Ukrainian, was relatively low (41.7% accuracy for /i/ and 33.3% - in fact, the lowest individual performance - for /ɪ/).

On the other end of the spectrum, two particularly high vowel pair perception values are associated with Participant J, who performed with 100% accuracy on the /ɛ - æ/ and /ʌ-ɑ/ contrasts. As discussed below for overall performance on all eight tested vowels, Participant J's high performance might be attributed to her low AOA and relatively lengthy LOR, particularly given her age when she arrived in the US.

Individually, with each listener's performance averaged across all eight tested vowels (Figure 4.15), some participants (E, J, and T) performed quite well (86.94%, 97.29%, and 88.61% accuracy, respectively), while others (D, G, H, L, and M) performed rather poorly (52.99%, 59.58%, 59.10%, 57.43%, and 59.10%, respectively) – barely above the probability of correctly identifying the stimulus vowel by chance.

Early AOA may explain the high performance of E, J, and T, who represent three of the four lowest AOAs in the present study (19, 14, and 13 years old, respectively; another participant arrived in the US at age 17 but has only lived in the US for five years – the lowest LOR in the group; 15 others arrived in their 20's and one in her early 30's.). The five lowest performers (D, G, H, L, and M) had AOAs ranging between 21 and 27 years (comparable to the AOAs of the 12

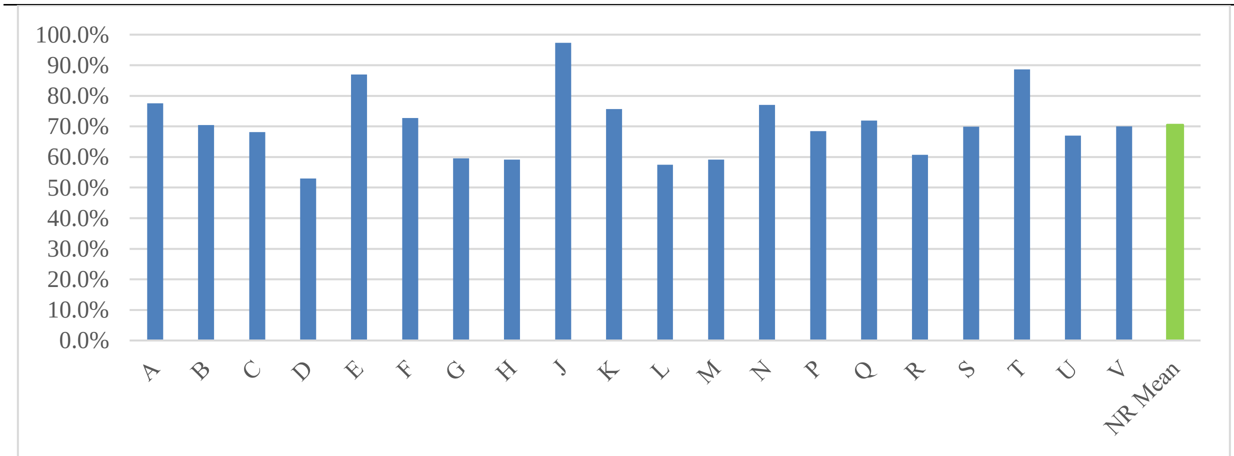


Figure 4.15 Mean accuracy of NR perception on all 8 vowels tested

NR listeners whose performance on Experiment 2 was intermediate) and spent between 13 and 17 total years in the US and between 10 and 17 years in GA (relatively comparable to the LORs of the intermediate performers on this task), and it is not immediately clear why their performance should be so low relative to other listeners with a similar AOA and LOR. The role of influence from other languages to which participants were exposed cannot be excluded, but does not offer a reliable explanation for performance (as discussed above for Participant G's /i - ɪ/).

Perception accuracy of NR listeners was normally distributed for five out of the eight vowels tested; accuracy was normally distributed on all four vowel contrasts when accuracy per contrast is calculated as a mean of accuracy on the two vowels that make up that contrast (Table 4.21). A bivariate fit model was used to test for correlations between the NR listeners' accuracy on the eight individual vowels tested as well as the vowel contrast values, calculated as a mean of the accuracy on each vowel individually, and AOA and LOR in the US and GA (Table 4.22).

Only one statistically significant correlation was identified in this part of the study: a negative correlation between AOA and perception accuracy on /ɑ/ (p = 0.0321), indicating that a lower AOA is correlated with higher performance on identification of this vowel in NE speech for NR listeners. AOA did not correlate with listening performance on any of the other vowels or

Table 4.21 Distribution of NR perception of NE speech. Vowel pair values calculated as a mean of the individual accuracy on the two vowels in the pair

/i/	normal 0.3184	**/i - ɪ/**	normal 0.3705
/ɪ/	normal 0.0813		
/ɛ/	normal 0.2593	**/æ - ɛ/**	normal 0.1994
/æ/	normal 0.3567		
/u/	**NOT normal 0.0248***	**/u - ʊ/**	normal 0.4472
/ʊ/	**NOT normal 0.0454***		
/ʌ/	normal 0.5005	**/ɑ - ʌ/**	normal 0.2932
/ɑ/	**NOT normal 0.0166***		

Table 4.22 Correlation of perception accuracy of NR listeners of NE talker with AOA and LOR on eight vowels and four vowel contrasts. Contrast accuracy value is mean of the two vowels; statistically significant correlations are shaded gray.

	AOA	LOR in USA	LOR in GA
/i/	r = -0.34023	r = -0.24481	r = -0.17405
	p = 0.1422	p = 0.2982	p = 0.4630
/ɪ/	r = -0.08511	r = 0.146388	r = 0.104136
	p = 0.7213	p = 0.5380	p = 0.6622
/ɛ/	r = -0.4295	r = 0.081451	r = -0.01368
	p = 0.0588	p = 0.7328	p = 0.9544
/æ/	r = -0.16947	r = -0.4201	r = -0.19857
	p = 0.4751	***p = 0.0652***	p = 0.4013
/u/	r = -0.26414	r = -0.08661	r = 0.007144
	p = 0.2604	p = 0.7165	p = 0.9762
/ʊ/	r = -0.16444	r = 0.342637	r = 0.190946
	p = 0.4885	p = 0.1392	p = 0.4200
/ʌ/	r = -0.1069	r = 0.04528	r = 0.135422
	p = 0.6537	p = 0.7231	p = 0.5692
/ɑ/	r = -0.48028	r = -0.34367	r = -0.23369
	p = **0.0321***	p = 0.1379	p = 0.3214
/i - ɪ/	r = -0.27518	r = -0.07134	r = -0.05062
	p = 0.2403	p = 0.7650	p = 0.8322
/æ - ɛ/	r = -0.39032	r = -0.20301	r = -0.1305
	p = 0.0889	p = 0.3907	p = 0.5834
/u - ʊ/	r = -0.29884	r = 0.052855	r = 0.079267
	p = 0.2006	p = 0.8249	p = 0.7397
/ɑ - ʌ/	r = -0.33321	r = -0.12811	r = -0.03366
	p = 0.1511	p = 0.5904	p = 0.8880

contrasts tested, nor did LOR correlate with listening performance. AOA also came close to negatively correlating with perception of /ɛ/, as did LOR in the U.S. and perception of /æ/.

To summarize the results of Experiment 2, investigating listening accuracy in Russian-English bilinguals, /ɑ/ was the most successfully identified vowel among the group of NR listeners, and /u/ was least successfully identified, although considerable individual variation was observed in listening accuracy, particularly for /u/, where individual accuracy ranged from 0% to 100% (reflected in this vowel's high standard deviation of 33.2%). Accuracy for all eight vowels tested was, in order of greatest to least: ɑ > æ > ʊ > ɪ > ʌ > i > ɛ > u. Accuracy for the four contrasts tested (calculated as a mean of accuracy on individual vowels making up each contrast) was, in order of greatest to least: /ʌ - ɑ/ > /ɛ - æ/ > /i - ɪ/ > /u - ʊ/. Early AOA was associated with increased accuracy in identification of /ɑ/ and approached statistically significant correlation on /æ/.

The next section presents results of Experiment 3, dealing with the perception of NR speech by NE listeners.

4.4 Experiment 3 – NE Listeners' Perception of NR Talkers

Results from Experiment 3, which investigated NE perception of vowels spoken by the 20 NR talkers that participated in the study, are summarized in Table 4.23, and individual performance on all eight vowels (grouped into pairs to facilitate within-contrast comparison) is presented in Figure 4.16. NE listeners most accurately identified /i/ (77.8% accuracy) and least accurately identified /ʊ/ (18.9% accuracy) in NR speech. Accuracy of identification of the eight vowels tested was, in order from highest to lowest: i > u > ɑ > ɪ > ɛ > æ > ʌ > ʊ. Many more accuracy scores of zero were noted in this part of the study than in Experiment 2 (recall that in

Table 4.23 Identification of eight American English vowels of NR talkers by NE listeners. N=20 NR talkers, mean of performance of N=9 NE listeners.

		Auditory stimulus							
		/i/	**/ɪ/**	**/ɛ/**	**/æ/**	**/u/**	**/ʊ/**	**/ʌ/**	**/ɑ/**
Response	**/i/**	77.8%	22.2%						
	/ɪ/	22.2%	65.6%						
	/ɛ/			63.3%	46.1%				
	/æ/			36.7%	53.9%				
	/u/					76.1%	81.1%		
	/ʊ/					23.9%	18.9%		
	/ʌ/							53.3%	32.2%
	/ɑ/							46.7%	67.8%
	Min	22.20%	11.10%	0.00%	0.00%	33.30%	0.00%	0.00%	11.10%
	Max	100%	100%	100%	100%	100%	66.70%	100%	100%
	SD	0.231	0.310	0.353	0.425	0.181	0.184	0.332	0.297

Figure 4.16 Identification of eight American English vowels, grouped by adjacent pairs, in NR speech (N=20 talkers) by (N=9) NE listeners. See Appendix E for individual values.

Experiment 2, only three accuracy scores of zero were recorded, all for /u/, from Participants D,G, and R). On Experiment 3, a total of 12 accuracy scores of 0% were observed: two on /ɛ/ (H, U), four on /æ/ (A, G, L, V), five on /ʊ/ (E, F, J, K, M), and one on /ʌ/ (H). This is not entirely unexpected, as the NR listeners are accustomed to hearing NE speech in their daily lives and heard the production of only one speaker, which may have allowed them to improve performance over the duration of the experiment, while the NE listeners were not prepared for hearing Russian-accented speech and heard the production of 20 different talkers, minimizing the possibility of attuning to any particular listener.

In order to facilitate comparison of performance between vowel contrasts rather than between individual vowels, as in Experiment 2, individual vowel results from Experiment 3 were averaged between members of each of the four tested vowel pairs; that is, one mean value was obtained for each listener's performance on /i/ and /ɪ/, one mean value for performance on /u/ and /ʊ/, etc. This approach masks some of the details having to do with the perception of individual vowels, but offers a more streamlined glimpse at how each listener performs on each tested vowel pair. These results are presented in Figure 4.17.

When vowel pair accuracy is calculated as a mean of the individual accuracy on the two vowels that make up the pair, NE listeners most accurately perceived the /i - ɪ/ contrast (mean accuracy of 71.7% across all listeners) compared to the other three contrasts tested and least accurately perceived the /u - ʊ/ contrast (with a mean accuracy of 47.5% across all listeners) in NR speech. Accuracy of identification of the four vowel contrasts tested, in order form highest to lowest, was: /i - ɪ/ > /ʌ - ɑ/ > /ɛ - æ/ > /u - ʊ/. See Appendix E for a table of all vowel contrast accuracy values. Note that the /u - ʊ/ contrast was also the least accurately identified contrast by NR listeners, but the NR listeners performed better on /ʊ/ (74.3% accuracy) than /u/ (53.3%), while

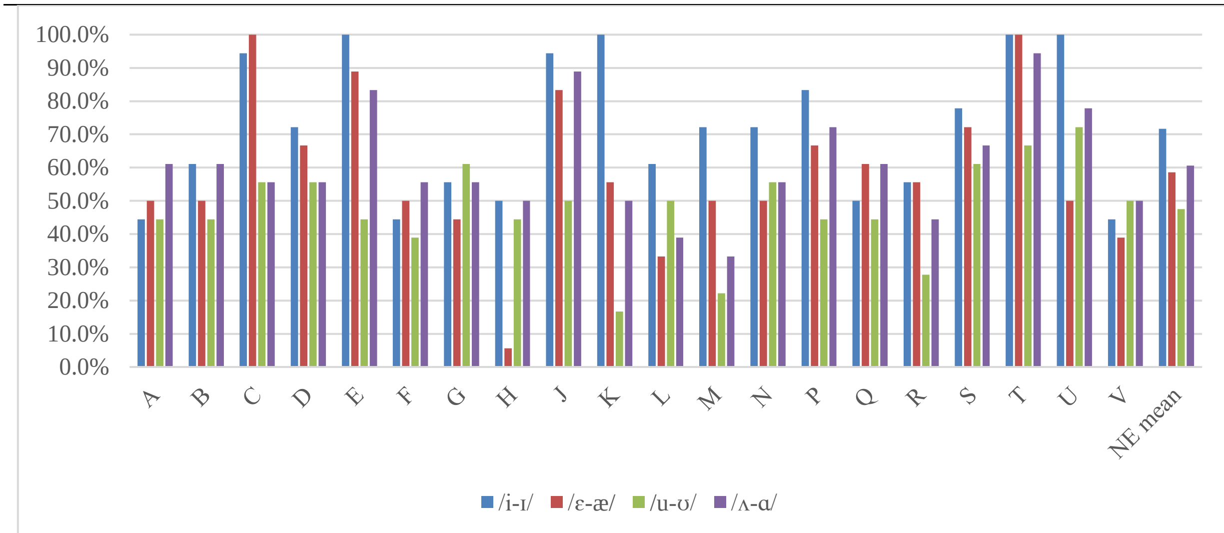


Figure 4.17 NE listener performance on all four contrasts. Calculated as mean of listening performance on individual vowels.

the NE listeners performed much better on /u/ (76.1%) than /ʊ/ (18.9%). Additionally, these results were highly speaker dependent, underscored by the large standard deviations on most contrasts: on three contrasts, NE listening performance ranged from 0% to 100%, and from 11.1% to 100% on two others.

To explain why NR listeners frequently confuse /u/ for /ʊ/ but not /ʊ/ for /u/, I posited that NR listeners are influenced by the Russian /u/, which is high and much more back than the /u/ of the NE talkers in this study, and that this is causing the NR listeners to judge both /u/ and /ʊ/ - strongly centralized in southern English – to be poor exemplars of /u/ and thus, by process of elimination, better exemplars of /ʊ/. As a secondary explanation, I suggested that NR listeners may be disproportionately choosing /ʊ/ due to their awareness of it as a particularly challenging sound for NR speakers to perceive and produce.

In the case of NE listeners, recall that /u/ and /ʊ/ in native production are differentiated more via height rather than backing, and that NR talkers do not make as great a distinction in F1 between /u/ and /ʊ/, producing both vowels quite high and more back than the NE talkers in the present study but generally in line with NE talkers in other parts of the country. Thus, it seems likely that native listeners are likely to identify the high, back vowels of NR talkers as /u/ rather than /ʊ/, which they expect to be considerably lower regardless of dialect and possibly more central (if approximating southern values) (see a comparison of NR and NE production and Russian vowels in Figures 4.2 and 4.4). Additionally, NE listeners more often misidentified NR's /ɪ/ as /i/ than vice versa (recall that both high front vowels in NR production were positioned closer to NE /i/ in the vowel space), and more often misidentified NR's /æ/ as /ɛ/ than vice versa (recall that both /ɛ/ and /æ/ in NR production were positioned closer to NE /ɛ/ in the vowel space), lending support to the idea that NE listeners are assimilating the audio stimuli to the closest NE vowel in the part of the vowel space.

As pointed out above, very low performance on individual vowels was observed for the NE perception of several NR talkers – accuracy scores of 0% on /ɛ/ (H, U), /æ/ (A, G, L, V), /ʊ/ (E, F, J, K, M), and /ʌ/ (H); considerably more than for the NR listeners. On the other hand, very high performance on individual vowels was also more frequently observed in the NE listener results – 30 (vs. 20 among the NR listeners) individual vowel accuracy scores of 100%. This underscores the greater overall amount of variability present in the results of Experiment 3 (also supported by greater standard deviations in Experiment 3).

Individually, with NE perception of each NR talker averaged across all eight tested vowels (Figure 4.18), some participants (C, E, J, T, and U) were understood relatively well (76.4%, 79.2%, 79.2%, 90.3%, and 75% accuracy, respectively), while others (F, H, L, M, R, and V) were understood rather poorly (47.2%, 37.5%, 45.8%, 44.5%, 45.8%, and 45.8% accuracy, respectively) – below the probability of the vowels in their production being correctly identified by chance. The highest accuracy over all tested vowels in Experiment 3 was associated with Participant T, and the

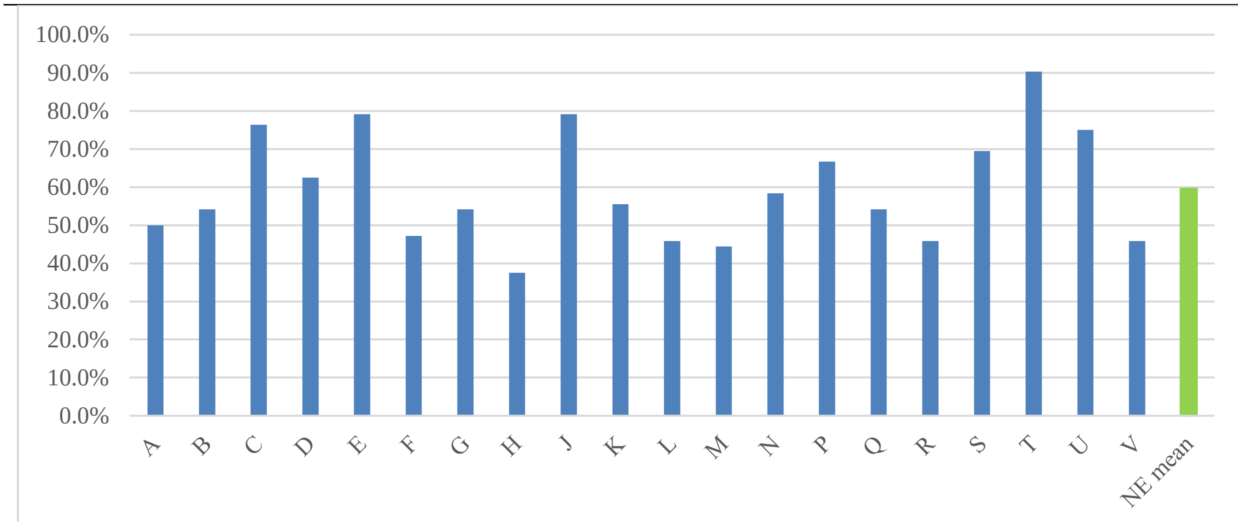


Figure 4.18 Mean accuracy of NE perception on all 8 vowels tested

lowest, with Participant H. This is slightly different from the results of Experiment 2, where Participant D had the lowest accuracy scores and Participant J had the highest; however, Participant H was also in the lower range on Experiment 2 (among the bottom four performers), while Participant T had the 2nd highest listening accuracy in Experiment 2.

Accuracy of NE listeners' perception of NR talkers was largely not normally distributed: only /ʌ/ and the pairs (calculated as a mean of each vowel's individual accuracy) /æ - ɛ/, /u - ʊ/, and /ɑ - ʌ/ had a normal distribution (Table 4.24). A bivariate fit model was used to test for correlations between accuracy of NE listeners' perception of NR talkers (both on individual vowels and on vowel contrasts, calculated as a means of the individual vowels' accuracy) and those talkers' AOA and LOR (Table 4.25). Two negative correlations were significant: AOA with NE perception accuracy on the /æ - ɛ/ ($p = 0.00189$) and /ɑ - ʌ/ ($p = 0.0017$) contrasts; NR talkers with lower AOA were more accurately perceived by NE listeners on these contrasts.

Next, a bivariate fit model was used to test for correlations between the accuracy of NE listeners' perception of NR talkers (both on individual vowels and on vowel contrasts, calculated as a mean of performance on the individual vowels making up the contrasts) and 1) those NR talkers' DRs between the tested vowel pairs, 2) those talkers' DRN – as compared to the NE mean

Table 4.24 Distribution of NE perception of NR speech. Vowel pair values are calculated as a mean of the individual values for the two vowels in that pair.

/i/	**NOT normal 0.0008***	**/i - ɪ/**	**NOT normal 0.0244***
/ɪ/	**NOT normal 0.0153***		
/ɛ/	**NOT normal 0.0054***	**/æ - ɛ/**	normal 0.2262
/æ/	**NOT normal 0.0011***		
/u/	**NOT normal 0.0098***	**/u - ʊ/**	normal 0.3190
/ʊ/	**NOT normal 0.0058***		
/ʌ/	normal 0.0957	**/ɑ - ʌ/**	normal 0.3273
/ɑ/	**NOT normal 0.0151***		

Table 4.25 Correlation of NE listeners' perception accuracy of NR talkers with the talkers' AOA and LOR on four vowel contrasts. Statistically significant correlations are shaded gray.

	AOA	LOR in USA	LOR in GA
/i/	r = -0.17657	r = -0.0448	r = -0.01324
	p = 0.4565	p = 0.8512	p = 0.9558
/ɪ/	r = -0.39687	r = -0.14639	r = 0.005586
	p = 0.0832	p = 0.5380	p = 0.9814
/ɛ/	r = -0.28086	r = 0.090339	r = 0.098265
	p = 0.2303	p = 0.7049	p = 0.6802
/æ/	r = -0.32093	r = -0.05818	r = 0.043206
	p = 0.1677	p = 0.8075	p = 0.8565
/u/	r = -0.22422	r = -0.05099	r = 0.056471
	p = 0.3419	p = 0.8309	p = 0.8131
/ʊ/	r = -0.20698	r = 0.12422	r = 0.111143
	p = 0.3813	p = 0.6018	p = 0.6409
/ʌ/	r = -0.31163	r = -0.06315	r = -0.02334
	p = 0.1811	p = 0.7914	p = 0.9222
/ɑ/	r = -0.36076	r = -0.23018	r = -0.1121
	p = 0.1181	p = 0.3289	p = 0.6380
/i - ɪ/	r = -0.39101	r = -0.13332	r = -0.00367
	p = 0.0882	p = 0.5752	p = 0.9877
/æ - ɛ/	r = -0.51958	r = 0.015749	r = 0.116873
	p = 0.00189*	p = 0.9475	p = 0.6236
/u - ʊ/	r = -0.28263	r = 0.049528	r = 0.110706
	p = 0.2273	p = 0.8357	p = 0.6422
/ɑ - ʌ/	r = -0.65659	r = -0.27882	r = -0.12793
	p = 0.0017*	p = 0.2339	p = 0.5909

– on the tested vowel pairs, 3) those talkers' EDV on the tested vowels[38], 4) those NR talkers' EDP on the four tested vowel contrasts, and 5) EDPN for each vowel pair (Table 4.26). Accuracy of NE perception negatively correlated with nativelikeness of the NR talkers' duration ratios (DRN) on the tested pairs for /æ/ (p = 0.0073), /ɑ/ (p = 0.0141), and the /ɑ - ʌ/ contrast (p = 0.0011); that is,

Table 4.26 Correlation of NE listeners' perception accuracy of NR talkers with two measures of duration and three measures of Euclidean distance.
(1) those NR talkers' DR on the associated vowel contrasts,
(2) those talkers' DRN on the same vowel contrasts,
(3) Euclidean distance between the talkers' vowels and the NE mean
(4) Euclidean distance between members of vowel pairs
(5) difference between NR talkers' Euclidean distance and NE mean
(Statistically significant correlations are shaded gray.)

	NE perception vs. NR DR	NE perception vs. NR DRN	NE perception vs. NR EDV	NE perception vs. NR EDP	NE perception vs. EDPN
/i/	r = 0.059334 p = 0.8038	r = -0.12068 p = 0.6123	r = 0.06547 p = 0.7839	r = 0.35249 p = 0.1274	r = -0.32805 p = 0.1579
/ɪ/	r = 0.255933 p = 0.2761	r = 0.335644 p = 0.1480	r = -0.68021 **p = 0.0010***	r = 0.671463 **p = 0.0012***	r = -0.66168 **p = 0.0015***
/ɛ/	r = 0.075121 p = 0.7529	r = 0.286624 p = 0.2205	r = 0.017445 p = 0.9418	r = 0.117476 p = 0.6218	r = -0.11748 p = 0.6218
/æ/	r = 0.006463 p = 0.9784	r = -0.58072 **p = 0.0073***	r = -0.53909 **p = 0.0142***	r = 0.54248 **p = 0.0135***	r = -0.54248 **p = 0.0135***
/u/	r = 0.072336 p = 0.7618	r = 0.21422 p = 0.3644	r = -0.08623 p = 0.7177	r = 0.458299 **p = 0.0421***	r = -0.5092 **p = 0.0218***
/ʊ/	r = -0.0555 p = 0.8162	r = -0.02494 p = 0.9169	r = -0.13219 r = 0.5785	r = 0.370091 p = 0.1082	r = -0.293 p = 0.2099
/ʌ/	r =0.125267 p = 0.5987	r = -0.16859 p = 0.4774	r = -0.46859 **p = 0.0372***	r = 0.262746 p = 0.2631	r = -0.09647 p = 0.6858
/ɑ/	r = -0.10448 p = 0.6611	r = -0.5396 **p = 0.0141***	r = 0.11712 p = 0.6229	r = 0.390007 p = 0.0891	r = -0.04725 p = 0.8432
/i - ɪ/	r = 0.221952 p = 0.3470	r = 0.181768 p = 0.4431	r = -0.66799 **p = 0.0013***	r = 0.691361 **p = 0.0007***	r = -0.67074 **p = 0.0012***
/æ - ɛ/	r = 0.064111 p = 0.7883	r = -0.32161 p = 0.1667	r = -0.10399 p = 0.6626	r = 0.600092 **p = 0.0052***	r = -0.60009 **p = 0.0052***
/u - ʊ/	r = -0.16836 p = 0.4780	r = -0.12599 p = 0.5966	r = -0.15669 p = 0.5094	r = 0.54413 **p = 0.0131***	r = -0.52606 **p = 0.0172***
/ɑ - ʌ/	r = 0.03306 p = 0.8900	r = -0.67558 **p = 0.0011***	r = -0.2253 p = 0.3396	r = 0.633869 **p = 0.0027***	r = -0.14394 p = 0.5449

[38] EDV for the two vowels comprising each pair was averaged for the purposes of this calculation only, to compare with averaged perception accuracy in the bottom four rows.

NE listeners were significantly more accurate in identifying the vowels /æ/ and /ɑ/ and the /ɑ - ʌ/ contrast (when contrast accuracy is calculated as a mean of accuracy on each vowel) when NR talkers had a more nativelike duration ratio on the associated vowel pair.

Accuracy of NE perception negatively correlated with EDV on /ɪ/ (p = 0.0010), /æ/ (p = 0.0142), and /ʌ/ (p = 0.0372), and the /i - ɪ/ contrast (p = 0.0013); that is, NE listeners were significantly more accurate in identifying the vowels /ɪ/, /æ/, and /ʌ/, and the /i - ɪ/ contrast (when contrast accuracy is calculated as a mean of accuracy on each vowel) when NR talkers produced more spectrally nativelike vowels.

Accuracy of NE perception positively correlated with EDP on /ɪ/ (p = 0.0012), /æ/ (p = 0.0135), /u/ (p = 0.0421), and all four contrasts; that is, NE listeners were significantly more accurate in identifying the vowels /ɪ/, /æ/, and /u/, and all four vowel contrasts (when contrast accuracy is calculated as a mean of accuracy on each vowel) when NR talkers made a greater spectral difference between the paired vowels.

Accuracy of NE perception negatively correlated with EDPN on /ɪ/, /æ/, and /u/ and the /i - ɪ/, /æ - ɛ/, and /u - ʊ/ contrasts; that is, NE listeners were significantly more accurate in identifying the vowels /ɪ/, /æ/, and /u/, and the /i - ɪ/, /æ - ɛ/, and /u - ʊ/ contrasts (when contrast accuracy is calculated as a mean of accuracy on each vowel) when NR talkers had a more nativelike Euclidean distance between the paired vowels (recall that smaller EDPN values, closest to zero, indicate higher rates of nativelikeness).

To summarize the results of Experiment 3, NE listeners most accurately identified /i/ and least accurately identified /ʊ/ in NR speech, with accuracy of identification of the eight vowels tested falling, in order from highest to lowest: i > u > ɑ > ɪ > ɛ > æ > ʌ > ʊ. Accuracy of identification of the four vowel contrasts tested, in order form highest to lowest, was: /i - ɪ/ > /ʌ -

ɑ/ > /ɛ - æ/ > /u - ʊ/. NR talkers with lower AOA were significantly more accurately perceived by NE listeners on the /æ - ɛ/ and /ɑ - ʌ/ contrasts. NE listeners were significantly more accurate in identifying /æ/ and /ɑ/ and the /ɑ - ʌ/ contrast when NR talkers had a more nativelike duration ratio on the associated vowel pair, significantly more accurate in identifying /ɪ/, /æ/, and /ʌ/, and the /i - ɪ/ contrast when NR talkers produced more spectrally nativelike vowels, significantly more accurate in identifying /ɪ/, /æ/, and /u/, and all four vowel contrasts when NR talkers made a greater spectral difference between the paired vowels, and significantly more accurate in identifying /ɪ/, /æ/, and /u/, and the /i - ɪ/, /æ - ɛ/, and /u - ʊ/ contrasts when NR talkers had a more nativelike Euclidean distance between the paired vowels.

Compared to Experiment 2, higher individual variation was observed in Experiment 3 than Experiment 2 (NE listener accuracy varied more than did NR listening accuracy, suggesting that listener variables also play an important role in communication), and a different pattern emerged in which vowel of three of the four tested pairs had better accuracy: in Experiment 2, NR listeners more accurately identified /ɪ/, /æ/, and /ʊ/, while NE listeners more accurately identified /i/, /ɛ/, and /u/ in the speech of the same NR participants. In Experiment 2, early AOA was associated with increased accuracy in identification of /ɑ/ and approached statistically significant correlation on /æ/; in Experiment, 3, NR talkers with early AOA were significantly more accurately perceived by NE listeners on the /æ - ɛ/ and /ɑ - ʌ/ contrasts.

Finally, a bivariate fit model was used to test for correlations between NR listening accuracy in Experiment 2 and the accuracy of NE listeners listening to those same NR talkers in Experiment 3. Three statistically significant correlations and two more correlations approaching statistical significance emerged (Table 4.27). Those NR participants who more accurately identified /ɪ/and /ʊ/ and more accurately discerned the /ɑ - ʌ/ contrast in NE speech were also more accurately perceived by NE listeners on those vowels and contrast.

Table 4.27 Table 4.27 Correlations between NR listening accuracy and NE listening accuracy for the same talkers' production. Statistically significant correlations are shaded gray.

/i/	r = 0.090404 p = 0.7047	**/i - ɪ/**	r = 0.431335 ***p = 0.0576***
/ɪ/	r = 0.561595 **p = 0.0100***		
/ɛ/	r = 0.084201 p = 0.7241	**/æ - ɛ/**	r = 0.395298 p = 0.0845
/æ/	r = 0.313749 p = 0.1780		
/u/	r = 0.090672 p = 0.7038	**/u - ʊ/**	r = 0.182619 p = 0.4409
/ʊ/	r = 0.480141 **p = 0.0321***		
/ʌ/	r = 0.141902 p = 0.5507	**/ɑ - ʌ/**	r = 0.605051 **p = 0.0047***
/ɑ/	r = 0.0430244 ***p = 0.0583***		

4.5 Conclusion

This chapter has presented the data obtained from Experiments 1, 2, and 3, on the perception and production of four American English vowel contrasts by NR and NE participants.

Experiment 1 revealed smaller distinctions between all four tested contrasts and greater standard deviations for all tested vowels for NR talkers than NE talkers. Lower AOA was associated with increased nativelikeness on six out of eight tested vowels (all but /i/ and /ɛ/); surprisingly, longer residence in the U.S. was associated with less nativelike production on /ɪ/, /æ/, and /ʊ/; and longer residence in GA was associated with more nativelike performance on /ɑ/. In terms of Euclidean distance between members of the tested vowel pairs, lower AOA was associated with more nativelike performance on /ɛ - æ/, and, also surprisingly, shorter residence in GA was associated with more nativelike performance on /ɑ - ʌ/. NR duration performance was significantly different from NE duration performance on all contrasts but /i - ɪ/, where the NR group showed considerable variation and data were not normally distributed, with several individuals making a very large duration distinction while others performed more like natives. NR talkers who arrived to the US at a younger age made a greater and more nativelike duration distinction between /ɑ/ and /ʌ/, and those who had resided in GA the least produced a more nativelike duration ratio between /u/ and /ʊ/.

In Experiment 2, the order of successful identification of English vowels by NR listeners was, from greatest to least: ɑ > æ > ʊ > ɪ > ʌ > i > ɛ > u. Early AOA was associated with increased accuracy in identification of /ɑ/ and approached statistically significant correlation on /æ/.

In Experiment 3, the order of successful identification by NE listeners of English vowels produced by NR talkers was, in order from greatest to least: i > u > ɑ > ɪ > ɛ > æ > ʌ > ʊ. NR talkers with lower AOA were significantly more accurately perceived by NE listeners on the /æ - ɛ/ and /ɑ - ʌ/ contrasts. NE listeners were significantly more accurate in identifying /æ/ and /ɑ/ and the /ɑ - ʌ/ contrast when NR talkers had a more nativelike duration ratio on the associated vowel pair, significantly more accurate in identifying /ɪ/, /æ/, and /ʌ/, and the /i - ɪ/ contrast when NR talkers produced more spectrally nativelike vowels, significantly more accurate in identifying /ɪ/, /æ/, and /u/, and all four vowel contrasts when NR talkers made a greater spectral difference between the paired vowels, and significantly more accurate in identifying /ɪ/, /æ/, and /u/, and the /i - ɪ/, /æ - ɛ/, and /u - ʊ/ contrasts when NR talkers had a more nativelike Euclidean distance between the paired vowels.

Comparing results from Experiment 2 with those from Experiment 3, those NR participants who more accurately identified /ɪ/and /ʊ/ and more accurately discerned the /ɑ - ʌ/ contrast in NE speech were also more accurately perceived by NE listeners on those vowels and contrast.

The following chapter will discuss how these results inform the research questions posed in Chapter 2 and fit in with the findings of existing literature.

CHAPTER 5

DISCUSSION OF RESULTS AND THEORETICAL IMPLICATIONS

5.1 Introduction

Chapter 5 discusses the results of Experiments 1, 2, and 3, presented above, as they relate to the literature reviewed and the research questions posed in Chapter 2, before addressing additional considerations related to and theoretical implications of the present works' findings as well as the limitations of the current work.

5.2 Question 1: Duration Distinction

This question asked whether Russian-English bilinguals would make a different duration ratio between the typically longer and typically shorter vowels of each tested pair (/i - ɪ/, /ɛ - æ/, /u - ʊ/, and /ɑ - ʌ/) than do monolingual native speakers of American English. Data relevant to this question comes from Experiment 1, dealing with production, and the specific measure that informs this research question comes from §4.2.4, dealing with the duration ratios (DR) between the typically longer and typically shorter vowels of each tested vowel pair.

Data from Experiment 1 revealed that NE talkers consistently produced the vowel expected to be the longer vowel in each tested pair longer than the vowel anticipated to be shorter; in each case, the difference in vowel duration was great enough to be expected to be discernible in normal speech (see Figure 4.11). NR talkers produced /i/ perceptibly longer than /ɪ/ and /u/ perceptibly longer than /ʊ/, but /æ/ was only slightly longer than /ɛ/ and /ʌ/ just slightly longer than /ɑ/, reversing the anticipated duration relationship. However, for the latter two contrasts – /ɑ - ʌ/ and /ɛ - æ/ – the duration distinction made between the vowels of the tested pairs was not great enough to be expected to be discernible in normal speech (see Figure 4.12). These results suggest that the NR talkers as a group made an intentional and perceptible duration distinction between only two of the four tested vowel contrasts: the high pairs /i - ɪ/ and /u - ʊ/. This result is unexpected if duration is more salient than spectral differences (Bohn 1995) and because NR listeners have been shown by previous studies to either over-rely or develop nativelike reliance (rather than under-rely or fail to rely at all) on duration in perception of at least three of these contrasts (Makarova 2010; Kondaurova & Francis 2008). This contradiction is addressed in more detail further in this chapter.

Differences between the NR and NE groups were significant for all contrasts but the high front /i - ɪ/ (see Table 4.17), likely attributable to the variation on performance on this contrast for the NR group (see Figure 4.12): performance was normally distributed for all contrasts but /i - ɪ/ (see Table 4.16). A handful of NR talkers, namely C, H, and K, produced an unusually large duration distinction between /i/ and /ɪ/, with /i/ more than twice as long as /ɪ/, while five other NR talkers had relatively nativelike performance, within one standard deviation of the NE mean, and six others did not appear to make a detectable duration distinction between these two vowels at all. I do not believe that C, H, and K produced a large duration distinction between /i/ and /ɪ/ because they fall in the earlier stages of acquisition first proposed for native-Spanish learners of English by Escudero (2000) and subsequently noted by Makarova (2010) among her NR listeners. C, H, and K have relatively high LOR (24, 13, and 17 years in the U.S., respectively) and make a fairly large spectral distinction between the high front pair (258, 347, and 621Hz, respectively) compared to NE talkers (mean of 590Hz) and other NR talkers (mean of 261 and min of 11Hz, with a total of 8 NR talkers producing a spectral distinction of 100Hz or less between the high front pair), indicating that they are most likely not in the early stages of acquisition of these vowel contrasts. Instead, I argue that this is indicative of an idiosyncratic pattern of acquisition displayed by these three NR talkers: they make an unusually large duration distinction between /i/ and /ɪ/ in addition to a large (in the case of K, quite nativelike) spectral distinction. These three talkers may be argued to indeed over-rely on duration in production of the /i - ɪ/ contrast, K in conjunction with a nativelike spectral distinction (so, over-reliance on duration and nativelike reliance on spectral cues), and C and H in conjunction with a spectral distinction that is high for the NR group but about half that of the NE mean (so over-reliance on duration and under-reliance on spectral cues).

A final noteworthy point about duration differences in long-short pairs, the NE group in the present study made their greatest duration distinctions between /æ - ɛ/ and /u - ʊ/, followed by /i - ɪ/, and lastly by /ɑ - ʌ/. Interestingly, this duration relationship differs somewhat from values reported for northern English, as values for the /ɑ - ʌ/ and /æ - ɛ/ pairs appear to be reversed (Hillenbrand 1995), which will be addressed in more detail in §5.7. The NR group in the present study made their greatest duration distinction between /i - ɪ/, followed by /u - ʊ/, then /æ - ɛ/, and finally by /ɑ - ʌ/. These values are compared in Table 5.1.

When the three extraordinarily high duration ratios of C, H, and K for the high front pair are excluded, the mean NR duration ratio between /i/ and /ɪ/ falls to 1.294:1, close to the NE mean

Table 5.1 Duration ratios of four tested vowel pairs for NE (9) and NR (20) talkers

	NE south	NR south
highest duration ratio	/æ - ɛ/ 1.39:1	/i - ɪ/ 1.46:1 (1.294:1)
	/u - ʊ/ 1.36:1	/u - ʊ/ 1.22:1
	/i - ɪ/ 1.26:1	/æ - ɛ/ 1.13:1
lowest duration ratio	/ɑ - ʌ/ 1.19:1	/ɑ - ʌ/ 0.98:1

of 1.256:1, and the NR values become normally distributed (p=0.5657), although there are still no statisticallly significant differences on this contrast between the NR and NE groups; this modification is reflected in the values in parentheses in the "NR south" column for /i - ɪ/. This suggests that the NR group, as a whole, if the outliers C, H, and K which appear to be implementing a different strategy to producing the /i - ɪ/ contrast are excluded, has acquired the /i - ɪ/ duration relationship. The NR duration ratio for the /u - ʊ/ contrast is most likely detectable in natural speech but falls some 14 percentage points below the NE ratio, suggesting that the NR talkers are under-using duration in producing the /u - ʊ/ contrast. The NR duration ratio for the /æ - ɛ/ contrast falls some 26 percentage points below the NE ratio (the contrast with the greatest difference between NE and NR groups) and may not be detectable in natural speech; it may indicate a partially acquired but under-used duration distinction, as for /u - ʊ/, or a lack of a duration distinction, as displayed by this group for the /ɑ - ʌ/, where a nativelike duration distinction has certainly not been acquired by the NR talkers as a group (and only four NR talkers even approach the NE mean duration ratio, with the rest of the group falling well below). The present study's findings are overall in line with the duration results of Romano et al. (1998), whose four NR talkers produced an average /i - ɪ/ duration ratio slightly higher than but rather close to the NE mean for their geographic region and a much lower duration ratio than native speakers on the other three vowel contrasts.

Interestingly, compared to Makarova's (2010) results (the /u - ʊ/ contrast, which had the smallest duration difference in local native speech, was the most accurately acquired and least over-relied on in duration by the NR listeners, while the /æ - ɛ/ contrast, which had the greatest duration difference in local native speech, was the least accurately acquired and most over-relied on in duration), the contrast with the greatest duration difference in the present study's NE speech

(/æ - ɛ/ with a 1.39:1 duration ratio) was also the least accurately acquired in production (recall that the NR duration ratio for this contrast fell 26 percentage points below the NE ratio). However, rather than over-relying on duration (as did Makarova's 2010 listeners), the present study's NR talkers either produced the pair with a much smaller duration difference than did NE talkers or did not utilize a duration contrast for the pair at all. As in Makarova's (2010) study, the contrast that had the smallest duration ratio in local NE speech (/ɑ - ʌ/ in the present study) was also the contrast in the production of which NR talkers least relied on duration, falling some 21 percentage points below the NE mean, without a detectable duration difference. It is possible that the frequently noted (Hillenbrand et al. 1995; Hagiwara 1997; Labov 1998; Thomas 2001; Clopper et al. 2005) spectral and duration differences between northern (Makarova 2010) and southern (present study) English vowels lead non-native speakers to adopt different strategies and follow different trajectories in acquiring them. These questions are addressed in more depth in §5.7.

The present study's production findings regarding duration do not align with the perception findings of Makarova (2010) and Kondaurova and Francis (2008), who found over-reliance on duration in perception of /i - ɪ/, /æ - ɛ/, and /u - ʊ/, and in perception of /i - ɪ/, respectively; however, if perception and production (in SLA or otherwise) are at least partially controlled by different mechanisms, it is possible that these incongruent findings have more to do with testing methods than inconsistent learning trajectories. Makarova tested only perception and asked listeners to identify manipulated stimuli created from one speaker's production and varying along five-point spectral and duration continua; the present study, however, analyzed production data, as well as asked listeners to identify un-manipulated stimuli from multiple speakers (addressed below). However, since perception results from Makarova (2010) generally align with production results from Hillenbrand et al. (1995) and since it would be unintuitive to posit that native speakers rely on different cues in perception and production of specific contrasts, the explanation dealing with dialectal variation appears to be the stronger one.

5.3 Question 2: Spectral Distinction

This question asked whether Russian-English bilinguals would make a different spectral distinction between the /i - ɪ/, /ɛ - æ/, /u - ʊ/, and /ɑ - ʌ/ contrasts than do monolingual native speakers of American English. Data relevant to this question comes from Experiment 1, dealing with production, and the specific measure that informs this research question comes from §4.2.3, dealing with the Euclidean Distance between the members of the four tested vowel pairs (EDP).

Data from Experiment 1 revealed that while NE talkers produced a large Euclidean distance (around 600Hz) between the vowels of the two front pairs (/i - ɪ/, /ɛ - æ/) and a smaller Euclidean distance (around 250Hz) between the other two contrasts tested (/u - ʊ/, /ɑ - ʌ/), NR talkers, on average, made their highest spectral distinction, measured as Euclidean distance between the members of each pair, between the /i - ɪ/ pair (around 260Hz), followed by the /ɑ - ʌ/ pair (around 160Hz), then the /ɛ - æ/ pair (153Hz), and finally the /u - ʊ/ pair (125Hz) (see Table 4.9 and Figure 4.9). These mean NR spectral distances are considerably lower than the NE means – for the high vowel pairs, the NR EDP is about half that of the corresponding NE EDP; for /ɑ - ʌ/, the distinction made by the NR group is about 64% that made by the NE group; for the /ɛ - æ/ pair, the mean NR EDP is about ¼ of the corresponding NE EDP. As a group, the NR talkers make a smaller spectral distinction between all four contrasts tested, especially for the /ɛ - æ/ pair.

Of the NR results, only the /ɛ - æ/ EDP results were normally distributed, indicating considerable individual variation in performance on the other three contrasts; indeed, Figure 4.9 illustrates some obvious outliers (re-summarized below in Table 5.2). For the high front contrast, talkers J, K, and T produced a large Euclidean distance, greater in fact than the NE mean, while D, L, M, Q, and V produced almost no spectral distinction on this pair (defined for the purposes of this discussion as less than 50Hz). For the high back contrast, J, T, and U produced a larger EDP than the NE mean, while B, D, M, Q, R, and V produced almost no spectral distinction. For the /ɑ - ʌ/ pair, C, E, J, and T produced a spectral distinction greater than the NE mean, while D, L, and M made almost no distinction. Only on the /ɛ - æ/ contrast did no NR talker produce an EDP higher than the NE mean for that contrast (EDP of C and T were highest), and A, D, F, N, Q, R, and V produced an EDP near zero.

Table 5.2 Mean AOA and LOR of Highest and Lowest EDP Performers

	highest EDP	AOA	LOR US	lowest EDP	AOA	LOR US
/i - ɪ/	J, K, T	19.7	14.3	D, L, M, Q, V	24.2	18.4
/ɛ - æ/	*C, T*	17	20	F, N, V	27.3	18.4
/u - ʊ/	J, T, U	14.7	14.3	B, D, M, Q, R, V	23.3	17.8
/ɑ - ʌ/	C, E, J, T	16.8	16.8	D, L, M	24.7	16.7

Some individual consistency is immediately visible in the groups of highest and lowest performers: T appears in the list of NR talkers with highest EDP on all four contrasts, J on three contrasts, and C on two; of the NR talkers who make a spectral distinction of 50Hz or less on the tested contrasts, D, M, and V each do so on three contrasts, and L and Q each do so on two. Mean AOA values for these highest and lowest performing groups, differing by 5-10 years, with the higher value always observed in the higher performing group, suggest that AOA plays a strong role in their acquisition. indeed, AOA negatively correlated with EDP performance for all contrasts but the high front pair, which approached statistical significance (see Table 4.11). Differences between the NR and NE groups were statistically significant for all contrasts but /u - ʊ/, which approached statistical significance (see Table 4.10).

These data are consistent with the literature on the effect of AOA on acquisition (more nativelike performance in early arrivals) and with data reported by Romano et al. (1998 – recall that all four talkers in this pilot study made a smaller spectral distinction between the vowels of the four vowel pairs tested by the present study), as well as with casual observation of the speech of accented Russian-English bilinguals – that these speakers may struggle to make a sufficiently large distinction between the vowel pairs /i - ɪ/, /ɛ - æ/, /u - ʊ/, and /ɑ - ʌ/ on the F1/F2 plane. It is curious that not a single NR talker – including talkers who performed the other three contrasts with quite large EDP values – performed like the native group on the /ɛ - æ/ pair in terms of spectral distinction between the vowels. While the mean NE EDP for /ɛ - æ/ was 590Hz (min: 414Hz; max: 733Hz), the largest distance between /ɛ/ and /æ/ produced by any Russian-English bilingual was 396Hz – slightly less than the minimum NE value. It is not immediately clear why this contrast should be the only one to have no performance even approaching nativelike, much less – as is the case for the other three contrasts – a few examples of performance exceeding nativelike in terms of distance between the paired vowels. It should be noted that Makarova (2010) found least nativelike NR perception performance on the /ɛ - æ/ contrast of the three contrasts tested in her study; however, this may be coincidence, as the two studies assessed very different aspects of non-native performance, using radically different methodologies, and tested speakers in different parts of the country, and it is not immediately clear how these low results might be correlated between the two studies; that is, what single, related factor or process might cause lowest performance on this particular contrast in both studies.

5.4 Question 3: Overall Accuracy

This question asked whether Russian-English bilinguals more accurate in perceiving or more nativelike in producing certain English vowels than others. Data relevant to this question comes from all three experiments. From Experiment 1 (which deals with production), acoustic data are compared between the NR and NE groups. From Experiment 2 (which deals with NR perception of NE speech), NR perception accuracy is analyzed and compared within the group and between vowels and contrasts. From Experiment 3 (which deals with NE perception of NR speech), NE perception accuracy – that is, the comprehensibility to a native listener of NR speech – is treated as a secondary measure of NR production accuracy. The specific measures that inform this research question include Euclidean distance between NR vowels and the NE mean for that vowel (EDV) from §4.2.2, Euclidean Distance between the members of the four tested vowel pairs (EDP) from §4.2.3, and nativelikeness of duration ratios (DRN) from §4.2.4, as well as NR and NE perception results from §4.3 and 4.4, respectively.

5.4.1 Production accuracy (acoustic analysis)

Speaking generally about the acoustic properties of vowels produced by the NE and NR groups in the present study and the relationships between tested vowel pairs on the F1/F2 plane, Figure 5.1 demonstrates visually that NR means for all tested vowel pairs are closer together than in NE production; that is, the NR group makes a smaller distinction between members of these vowel pairs than do NE talkers. Difference in performance (indicating lowest rate of nativelikeness for the NR group) appears to be greatest on the /i - ɪ/ and /ɑ - ʌ/ contrasts, where the NE group makes a rather large difference between vowels on the F1/F2 plane while the NR group makes a much smaller one, followed by /u - ʊ/, and finally /ɛ - æ/, where there is a relatively smaller

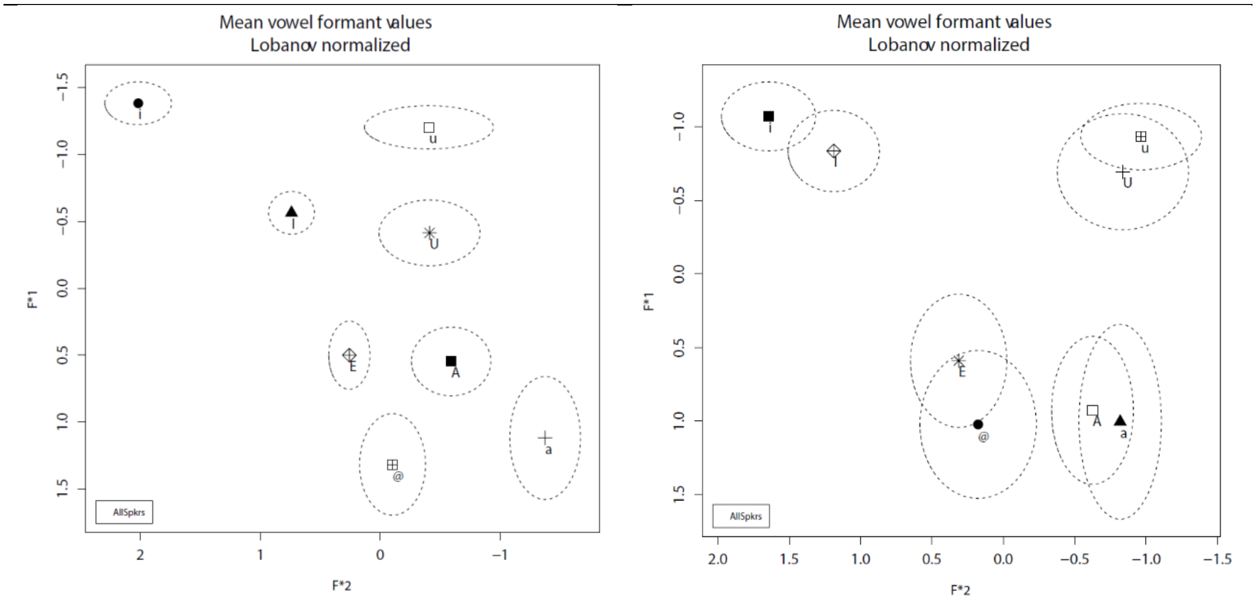


Figure 5.1 Means and 1SD Ellipses of 8 American English vowels by NE (left) and NR (right) speakers on F1/F2 plane (Lobanov normalized values)

difference between two groups. Ellipses at 1 standard deviation demonstrate that NR talkers vary considerably more than NE within their production of each vowel, and the parts of the vowel space dedicated to members of the tested vowel pairs overlap much more in NR than in NE production. No overlap is observed in the NE group at 1SD, while the NR group overlaps all adjacent vowels.

Since spectral and duration distinctions made between the four tested vowel contrasts by NR talkers were addressed in detail in §5.2 and §5.3 above, here, I will focus on overall accuracy on individual vowels, as measured by Euclidean distance between each NR talker's vowels compared to the NE mean (EDV), and only briefly address spectral and duration accuracy on vowel contrasts.

Mean NR EDV (§4.2.2) ranged from 165Hz (/i/) to 576Hz (/ʊ/); that is, NR talkers as a group produced /i/ closest its mean NE value and /ʊ/ furthest from its mean NE value. Accuracy, as measured by EDV - distance of each vowel in NR production from its NE mean - was as follows:

best for /i/, followed in descending order by /ɛ/, /ɑ/, /æ/, /ɪ/, /ʌ/, /u/, and finally /ʊ/ (represented in Figure 5.2). There was a large gap between the EDV for the first six vowels listed above and the latter two – the third to largest EDV was 375 (/ʌ/) after which the 2nd largest rose to 573 (/u/). The exceptionally poor performance noted for the high back pair may be attributable to the fronting observed for these phonemes among native speakers in the southern variety; that is, an issue of

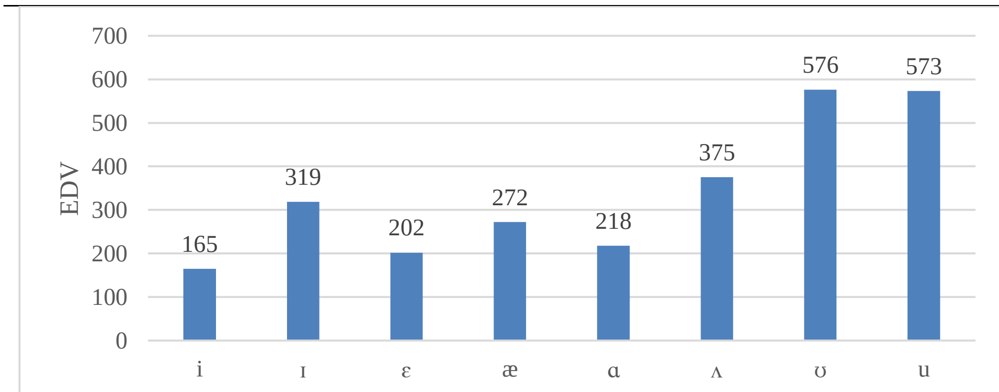


Figure 5.2 Mean NR EDV for 8 American English Vowels

dialect and, importantly for Question 6 below, an indication that the Russian-English bilinguals in this study are not particularly successful at replicating this aspect of southern English.

Considered in terms of EDP between the four tested vowel pairs, NR talkers produced a much smaller spectral distinction in the front vowel pairs than did NE talkers (Figure 5.3). Least nativelike performance, when viewed proportionally to NE performance on the same contrast, was on the /ɛ - æ/ pair, where NE talkers made a spectral distinction of 591Hz between /ɛ/ and /æ/ while NR talkers made a distinction of only 153Hz; the NR spectral distinction on this pair was only 25.9% that of the NE talkers. Most nativelike performance was observed on the /ɑ - ʌ/ pair, where NE talkers made their smallest spectral distinction of 234Hz and NR talkers made a distinction of

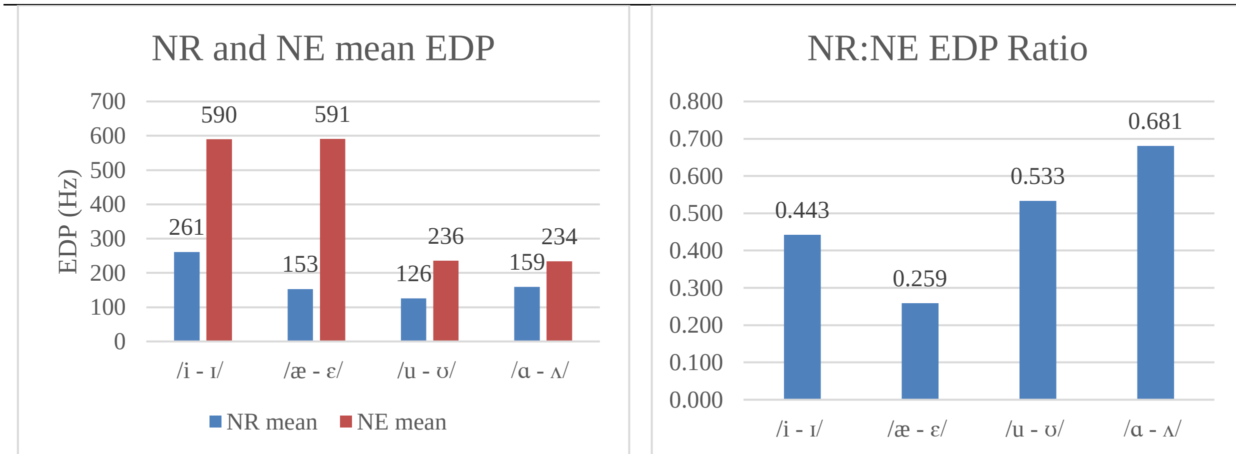


Figure 5.3 Left: NR and NE mean EDP for 4 tested vowel pairs; right: ratios of NR to NE EDP

159Hz (68.1% of the NE value). Viewed this way, it might be said that the NR talkers were most nativelike on the contrasts where NE talkers made their smallest spectral distinction and progressively less nativelike in the other three contrasts as native spectral distinction increased. Of course, individual variation was observed in production accuracy as measured by EDP. The following two figures (Figures 5.4 and 5.5) offer four examples each of particularly well distinguished and poorly distinguished vowels of NR talkers on an F1/F2 plane. These individuals' vowel plots were selected due to being the best visual examples of most and least distinguished vowels. Statistical analysis of EDP revealed that AOA is negatively correlated with performance, and indeed, AOA was low for talkers E, J, T, and U and high for A, D, Q, V, a difference of 7.25 years between their respective means (Table 5.3; see §5.8 for discussion of roles of AOA, LOR). As for EDP, EDV also correlated, for most vowels, with AOA; unlike EDP, EDV also correlated positively in a handful of cases with LOR in the U.S and in one case negatively with LOR in GA (to be discussed further in §5.8).

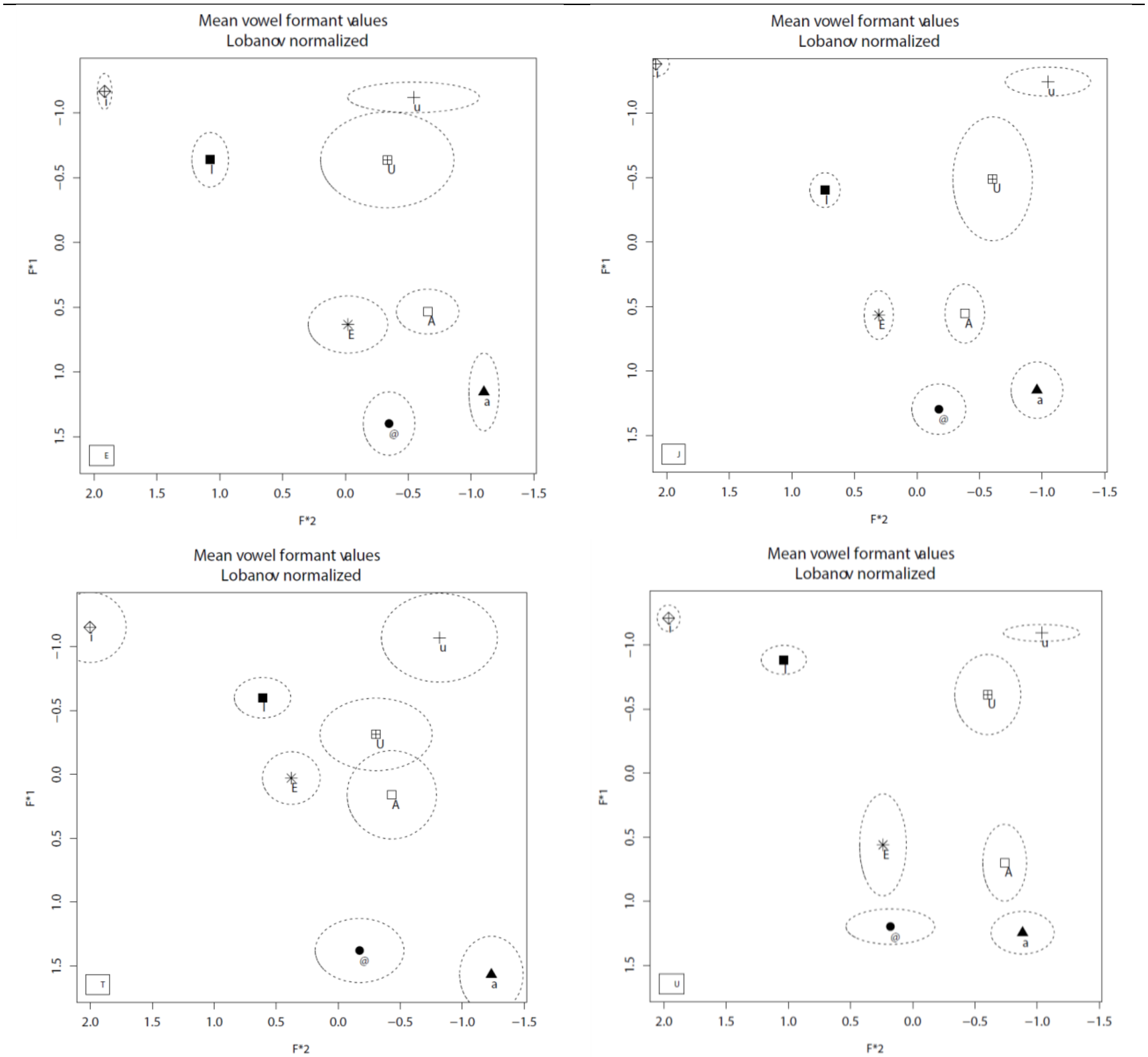


Figure 5.4 Means and 1SD Ellipses of 8 American English vowels by speakers E, J, T, U (low spectral overlap; Lobanov normalized values)

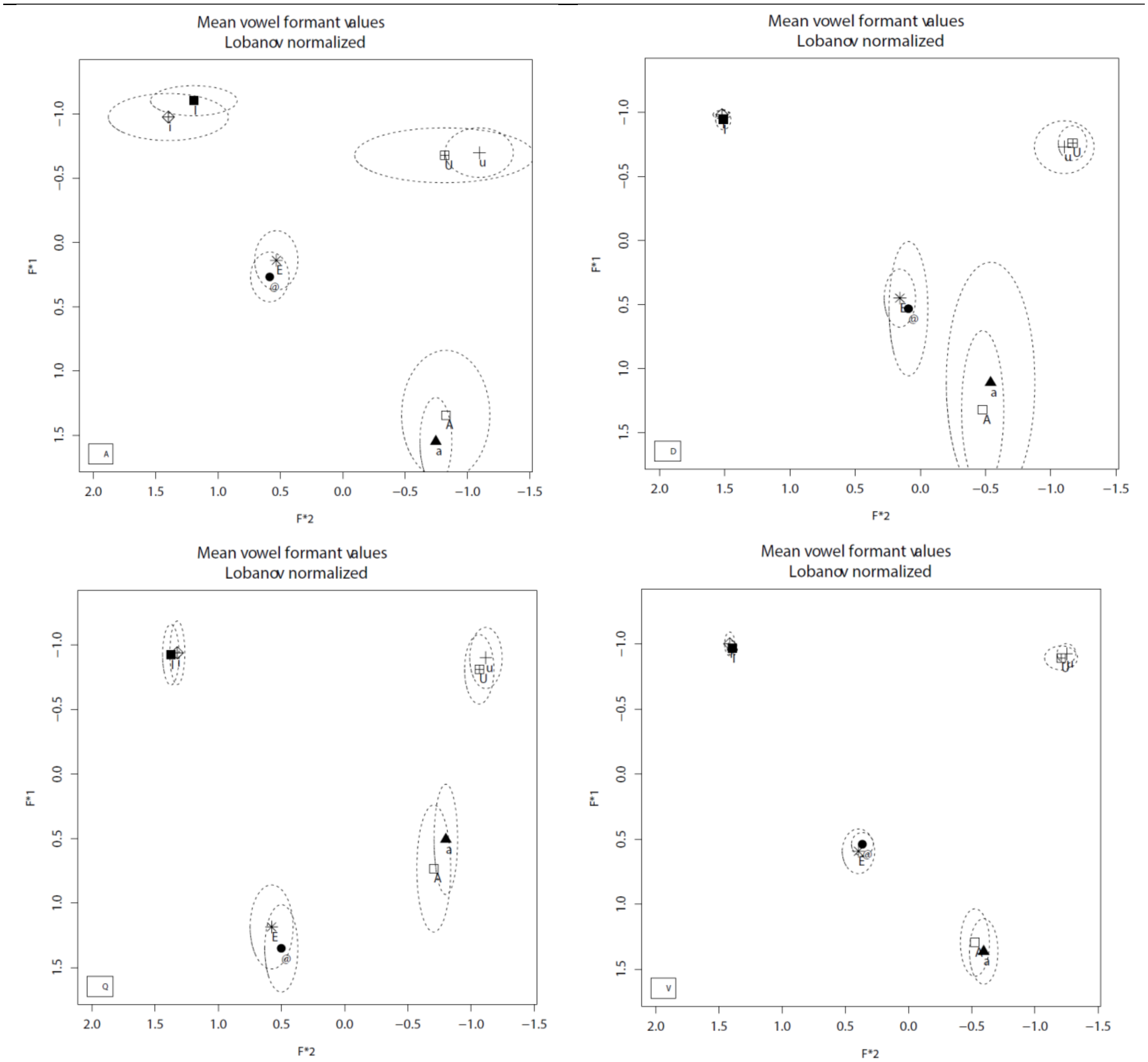


Figure 5.5 Means and 1SD Ellipses of 8 American English vowels by speakers A, D, Q, V (high spectral overlap; Lobanov normalized values)

Table 5.3 AOA for talkers represented by Figures 5.4 and 5.5

Talker	AOA	Talker	AOA
E	19	A	23
J	14	D	26
T	13	Q	22
U	21	V	25
mean	16.75	mean	24

The last measure of production nativelikeness considered here is DRN – nativelikeness of duration ratios between members of the tested vowel contrasts for NR talkers (Figure 5.6). Recall that for this measure, smaller values represent higher rates of nativelikeness (see §3.4.3 for details on this calculation). The most nativelike DRN performance across all NR talkers was observed for /u - ʊ/, followed by /ʌ - ɑ/, then /ɛ - æ/, and, finally, least nativelike performance was observed for the /i - ɪ/ pair.

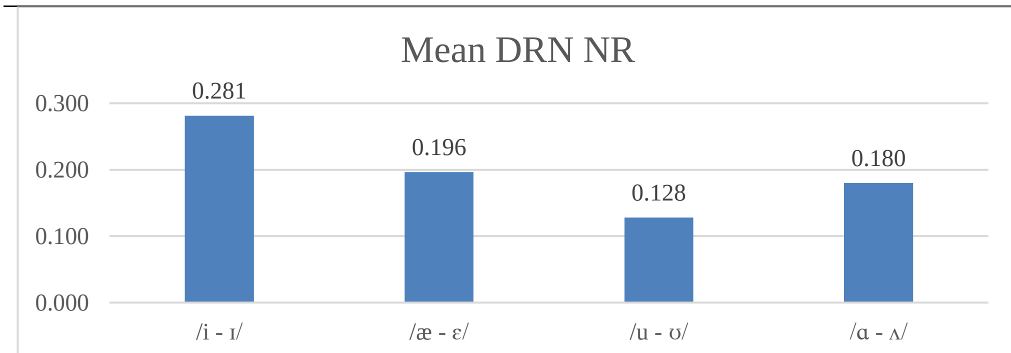


Figure 5.6 Mean DRN, all NR talkers pooled

5.4.2 Production accuracy as measured by NE listener peceptions

Another, albeit secondary, measure of production accuracy for NR talkers is the accuracy with which NE listeners identify vowels in the production of these NR talkers (see §4.4) in Experiment 3.

NE listeners most accurately identified /i/ and least accurately identified /ʊ/ in NR speech overall. Accuracy on the eight vowels tested ranged from 77.8% to 18.9%: i (77.8%) > u (76.1%) > ɑ (67.8%) > ɪ (65.6%) > ɛ (63.3%) > æ (53.9%) > ʌ (53.3%) > ʊ (18.9%). It is worth noting that there was considerable individual variation within each tested vowel. Accuracy rates of native listener judgment of NR speech fell as low as 0% for at least one speaker for four (/ɛ/, /æ/, /ʊ/, /ʌ/)

of the eight vowels tested (a total of twelve scores of 0%), as low as 11.1% for two vowels (/ɪ/, /ɑ/), and 22.2% and 33.3% for the remaining two vowels (/i/ and /u/, respectively). All tested vowels but /ʊ/ had at least three and as many as six NR talkers whose production of that vowel was identified correctly by natives 100% of the time (a total of 29 scores of 100%); the highest score achieved on /ʊ/ was 66.7%, for only one speaker, and the next highest scores for /ʊ/ were 44.4%. /æ/ had the greatest standard deviation (SD 0.425) and /u/, the smallest (SD 0.181).

This individual variation may indicate different learning trajectories or present levels of proficiency among individuals. Indeed, some individuals' scores cluster at high or low ranges for comprehensibility by NE listeners; for example, participant T appears to be generally well understood, with a mean comprehensibility score of 90.3% and five vowels out of eight understood with 100% accuracy. In other cases, however, such as that of participant E and U, relatively high overall comprehensibility and several individual vowel accuracy scores of 100% (four for E and five for U) nonetheless combine with a comprehensibility score of zero for another vowel (/ʊ/ in the case of E and /ɛ/ in the case of U). In these cases, it appears that some individuals have acquired some contrasts with very high accuracy but fail to perform nearly as well on other contrasts.

When accuracy for vowel contrasts is calculated as a mean of the accuracy on the two vowels in each contrast, NE listeners most accurately identified the /i-ɪ/ contrast and least accurately identified the /u - ʊ/ contrast in NR speech. Accuracy on the four contrasts tested ranged from 71.7% to 47.5%: /i - ɪ/ (71.7%) > /ʌ - ɑ/ (60.6%) > /ɛ - æ/ (58.6%) > /u - ʊ/ (47.5%). Individual variation was apparent, although not as great in degree as on the vowels individually due to the averaging of values for each pair: low scores for the four vowel contrasts dropped to 44.4% for /i - ɪ/, 5.6% for /ɛ - æ/, 16.7% for /u - ʊ/, and 33.3% for /ʌ - a/. The /i - ɪ/ and /ɛ - æ/ contrasts had some scores of 100% (four and two talkers, respectively), but the maximum scores for /u - ʊ/ and

/ʌ - ɑ/ were only 72.2% and 94.4%, respectively. The /ɛ - æ/ pair had the greatest standard deviation (SD 0.227) and /u - ʊ/, the smallest (SD 0.139).

The pattern of NE listener accuracy in identifying vowels produced by NR talkers represented by these results (i > u > ɑ > ɪ > ɛ > æ > ʌ > ʊ for individual vowels and /i - ɪ/ > /ʌ - ɑ/ > /ɛ - æ/ > /u - ʊ/ for pair averages) is fairly consistent with results reported by Bell-Berti et al. (1998) for worst performance on individual vowels (/ɪ ɛ ɔ ʊ ʌ/) but not as consistent with the adjacent pairs reported as causing the most confusion: /i/-/ɪ/ and /ɛ/-/æ/. Importantly, Bell-Berti et al.'s NE listeners chose from all ten tested vowels for each audio stimuli, while NE listeners in the present study had only two options for each audio stimulus.

An additional perspective underscoring the importance of spectral cues comes from a comparison of relatively well understood and poorly understood NR talkers in Figures 5.4 and 5.5 above. The four plots in Figure 5.4 were selected as examples with well-dispersed means and little to no overlap between members of adjacent vowel pairs, which is predicted to make them easy to discriminate. Indeed, the four tested contrasts in these four participants' production are relatively well identified (left of Table 5.4). The four plots in Figure 5.5 were selected as examples with poorly-dispersed (in some cases overlapping) means and particularly great overlap between members of adjacent vowel pairs. Since these vowels overlap considerably in spectral information, they are predicted to be difficult to discriminate. Indeed, the four tested contrasts in these four participants' production are relatively poorly identified (right of Table 5.4).

Table 5.4 NE identification accuracy for talkers represented by Figures 5.4 and 5.5

	/i - ɪ/	/ɛ - æ/	/u - ʊ/	/ɑ - ʌ/		/i - ɪ/	/ɛ - æ/	/u - ʊ/	/ɑ - ʌ/
E	100.00%	88.90%	44.40%	83.30%	A	44.40%	50.00%	44.40%	61.10%
J	94.40%	83.30%	50.00%	88.90%	D	72.20%	66.70%	55.60%	55.60%
T	100.00%	100.00%	66.70%	94.40%	Q	50.00%	61.10%	44.40%	61.10%
U	100.00%	50.00%	72.20%	77.80%	V	44.40%	38.90%	50.00%	50.00%

5.4.3 Perception accuracy

A final measure of overall accuracy of the Russian-English bilinguals' performance is their perception of the eight tested vowels in NE speech (see §4.3) in Experiment 2.

As a group, NR listeners most accurately identified /ɑ/ (83.3% accuracy) and least accurately identified /u/ (53.3% accuracy) in NE speech although considerable individual variation was observed in listening accuracy, particularly for /u/, where individual accuracy ranged from 0% to 100% (reflected in this vowel's high standard deviation of 33.2%). Accuracy of identification of the eight vowels tested, in order form highest to lowest, was: ɑ > æ > ʊ > ɪ > ʌ > i > ɛ > u. Accuracy for the four contrasts tested (calculated as a mean of accuracy on individual vowels making up each contrast) was, in order of greatest to least: /ʌ - ɑ/ > /ɛ - æ/ > /i - ɪ/ > /u - ʊ/. Early AOA was associated with increased accuracy in identification of /ɑ/ and approached statistically significant correlation on /æ/.

To briefly summarize accuracy findings,

- Accuracy of NR perception of NE speech was:
  ɑ > æ > ʊ > ɪ > ʌ > i > ɛ > u
  /ʌ - ɑ/ > /ɛ - æ/ > /i - ɪ/ > /u - ʊ/ (pair values mean of individual vowel accuracy)

- Accuracy of NE judgments of NR production was:
  i > u > ɑ > ɪ > ɛ > æ > ʌ > ʊ
  /i - ɪ/ > /ʌ - ɑ/ > /ɛ - æ/ > /u - ʊ/ (pair values mean of individual vowel accuracy)

- Spectral accuracy of NR production (EDV) was:
  /i/ > /ɛ/ > /ɑ/ > /æ/ > /ɪ/ > /ʌ/ > /u/ > /ʊ/

- Spectral accuracy of NR production (EDP) was
  /ɑ - ʌ/ > /u - ʊ/ > /i - ɪ/ > /ɛ - æ/

- Duration ratio accuracy (DRN) of NR production was:
  /u - ʊ/ > /ʌ - ɑ/ > /ɛ - æ/ > /i - ɪ/

As discussed above, in NE production, the front vowel pairs enjoy a much larger spectral difference than the /u - ʊ/ and /ɑ - ʌ/ contrasts (about 600Hz vs. about 250Hz; see §5.3) and the /æ

- ɛ/ and /u - ʊ/ pairs have a greater duration distinction than do /i - ɪ/ and /ɑ - ʌ/ (1.39:1 and 1.36:1 vs. 1.26:1 and 1.19:1, respectively; see §5.2) in the local variant examined in the present work. Thus, we might expect that the /æ - ɛ/ contrast, having a comparatively larger difference between its vowels both in terms of spectrum and duration, might be the easiest to discriminate and acquire, while the /ɑ - ʌ/ contrast, having the smallest difference, the most difficult. In fact, this expectation was not met entirely in any measure of accuracy and was reversed complete in spectral accuracy of NR production (EDP), where the /ɑ - ʌ/ contrast had the highest accuracy and /æ - ɛ/ the lowest.

5.5 Question 4: Spectrum vs Duration

This question asked whether native-English listeners more accurately identify English vowels produced by Russian-English bilingual talkers when they match native values more in ***spectral*** or ***duration*** cues; that is, which cues contribute more to overall comprehensibility? It is worth noting that the present study did not focus on which cues NE listeners rely in general (as did Makarova 2010 and Kondaurova and Francis 2008), but rather on which cues (?) NE listeners relied when they accurately perceived the intended vowel: a measure of which cue contributes to accurate identification.

To answer this question, NE listener perception accuracy was compared with measures of duration and Euclidean distance in the NR talkers' production. Table 4.26 presents all data. NE listener accuracy was correlated (negatively in the case of measures of 'nativelikeness' and EDV – recall that for these measures, low values represent more nativelike performance; positively in the case of EDP, where higher values usually represent more nativelike performance) with only three measures of duration (DRN) but with many more measures of Euclidean distance (four statistically significant correlations of vowels or contrasts with EDV, seven with EDP, and six with EDPN). This indicates that native listeners rely more on spectral than duration cues to correctly identify vowels spoken by NR talkers.

On the one hand, this is in line with what is expected based on existing literature: recall that native, monolingual speakers of English typically rely more on spectral cues and only marginally on duration cues to distinguish the English vowel contrasts characterized by both spectral and duration differences, such as the tense and lax vowel pairs in the same part of the vowel space (Peterson and Barney 1952; Hillenbrand et al. 1995; Hillenbrand et al. 2000). On the other hand, the anticipated reliance on spectral cues is based on native spectrum and duration values, and the NR talkers' production does not necessarily match those values; thus, for example, NE listeners relying primarily on spectral cues may fail to correctly identify the vowels of NR talkers who rely primarily on duration cues to produce those vowels.

Interestingly, no measure of duration or spectrum correlated at a statistically significant level with NE listener accuracy of identifying /i/ or /ʊ/ in NR speech. Recall that /i/ was the vowel most accurately identified by NE listeners (and associated with the highest spectral accuracy for NR talkers) while /ʊ/ was the vowel least accurately identified by NE listeners (and associated with the lowest spectral accuracy for NR talkers), while in terms of duration accuracy (DRN) this relationship was reversed (duration ratio accuracy highest for /u - ʊ/ and lowest for /i - ɪ/). NE listener accuracy on the pair vowels of /i/ and /ʊ/ - /ɪ/ and /u/ - was significantly correlated with all three and two, respectively, measures of Euclidean distance. It is not immediately clear why this should be the case.

It is not surprising that NE listeners' accuracy on vowel pairs (calculated as a mean of accuracy on individual vowel accuracy) correlated on each pair with EDP and for three pairs (all but /ɑ - ʌ/ with EDPN, as these are measures of the size and nativelikeness of Euclidean distance between these paired vowels. It is surprising, however, that EDV – a measure of spectral nativelikeness – did not always correlate (in fact, only correlated for three vowels - /ɪ/, /æ/, and /ʌ/) with NE listeners' accuracy.

5.6 Question 5: Perception-Production Link

This question asked whether Russian-English bilingual listeners who more accurately discriminate certain vowels also perform those contrasts better. Data relevant to this question comes from Experiment 2 (which deals with NR perception of NE speech) and from Experiment 3 (which deals with NE perception of NR speech), as well as acoustic measures of NR production from Experiment 1.

I will first compare NR perception accuracy to NR production accuracy as measured by the ability of NE listeners to identify vowels spoken by the NR talkers. Admittedly, this is not a direct measure of perception and production: NR listeners heard the production of only one NE talker, which may have allowed them to become more accustomed to the talker over the course of the experiment than the NE listeners, who heard the speech of 20 different NR talkers (see Chapter 3 for experiment design). Additionally, the NR participants, due to living in an English-speaking community, are accustomed to hearing native English speech in their daily lives, while only one of the NE participants reported any exposure to native speakers of Russian via a university course in the Russian language; thus, the NR listeners in this study have had more exposure to the accent they heard in the listening part of the experiment than did the NE listeners. These two factors may be artificially inflating NR perception accuracy as compared to 'production accuracy' as measured by NE listeners' ability to accurately identify the tested vowels in NR speech. Nonetheless, I compare these two sets of results because they are similar (0-100%) measures of accuracy. In addition to this comparison, I compare (using statistical analysis) NR perception accuracy with NR production accuracy as measured by EDV, EDP, EDPN, DR, and DRN, which are more accurate measures of production accuracy but cannot be directly compared with percent accuracy.

On the surface, the Perception-Production link appears to have been moderately supported by the finding that as a group, and across all tested contrasts, NR participants discriminated the tested vowel contrasts better (71% accuracy across all participants and vowels) than they performed them (as measured by NE perception of NR speech – 61% accuracy). When all participants and contrasts are pooled together, the ratio of NR *perception* accuracy (accuracy of NR listeners identifying vowels in NE speech) to NR *production* accuracy (measured in this case as accuracy of NE listeners identifying vowels in NE speech) is 1.22:1; that is, NR participants, as a group, produced all tested contrasts (as measured via NE perceptions of NR speech) about 22% better than they discriminated the same contrasts in NE speech. Again, it is important to remember that if the differences in testing methods between the NR and NE listener groups inflated the NR perception score more than 22%, then this result cannot be interpreted as higher perception than production accuracy supporting the Perception-Production link.

NR participants were more accurate in perception than in production of three of the four vowel contrasts tested, with reversed but very close perception and production values in the fourth (/i - ɪ/ the only contrast where production accuracy exceeded perception) (Table 5.5).

Table 5.5 Perception vs. production (as measured by NE listener judgment) accuracy

	/i - ɪ/	/ɛ - æ/	/u - ʊ/	/ɑ - ʌ/
NR Perception	70.00%	71.00%	63.80%	77.20%
(NE Perception of) NR Production	71.70%	58.60%	47.50%	60.60%
Ratio Perception:Production	0.98	1.21	1.34	1.27

For all contrasts and participants pooled together, perception accuracy was higher that production. The contrast with the greatest difference between perception and production means across participants was /u - ʊ/ (1.34:1 ratio of perception to production), followed by /ɑ - ʌ/ (1.27:1), and finally /ɛ - æ/ (1.21:1).

Four participants (C, D, T, and U) performed slightly better in production (as measured by NE listener judgment) than in perception when performance was pooled across all four contrasts (0.89:1, 0.85:1, 0.98:1, and 0.89:1 perception-production ratios, respectively); that is, they were better understood by the NE listeners than they themselves identified vowels in NE speech. No notable pattern emerges from these exceptions: C's perception:production ratio was reduced by her high production of the front vowel contrasts; D produced better than discriminated all contrasts but /a-ʌ/, T's values were closest together of this group with perception exceeding production only for /u - ʊ/, and U performed better on production than perception of the high vowel contrasts.

There was more variation in production (lowest accuracy 48% on /u - ʊ/ and highest 72% on /i - ɪ/) than in perception (lowest accuracy 64% on /u - ʊ/ and highest 77% on /a-ʌ/). This may be related to the above-mentioned problem of comparing NR perception (of only one talker by listeners accustomed to native English speech) to NE perception (of 20 talkers by listeners unaccustomed to a Russian accent) as a measure of 'production'. The /i - ɪ/ contrast displayed the least change between perception and production accuracy (difference of only 2%) than the other contrasts (/ɛ - æ/ had a difference of 12% between perception and production accuracy, and /u - ʊ/ and /a-ʌ/ had a difference of 16%). A bivariate fit model revealed that NR perception accuracy per contrast (as the *factor* variable, calculated as a mean of accuracy on the vowels in the pair) predicts NE perception accuracy for that contrast (as a secondary measure of NR production) as *response*

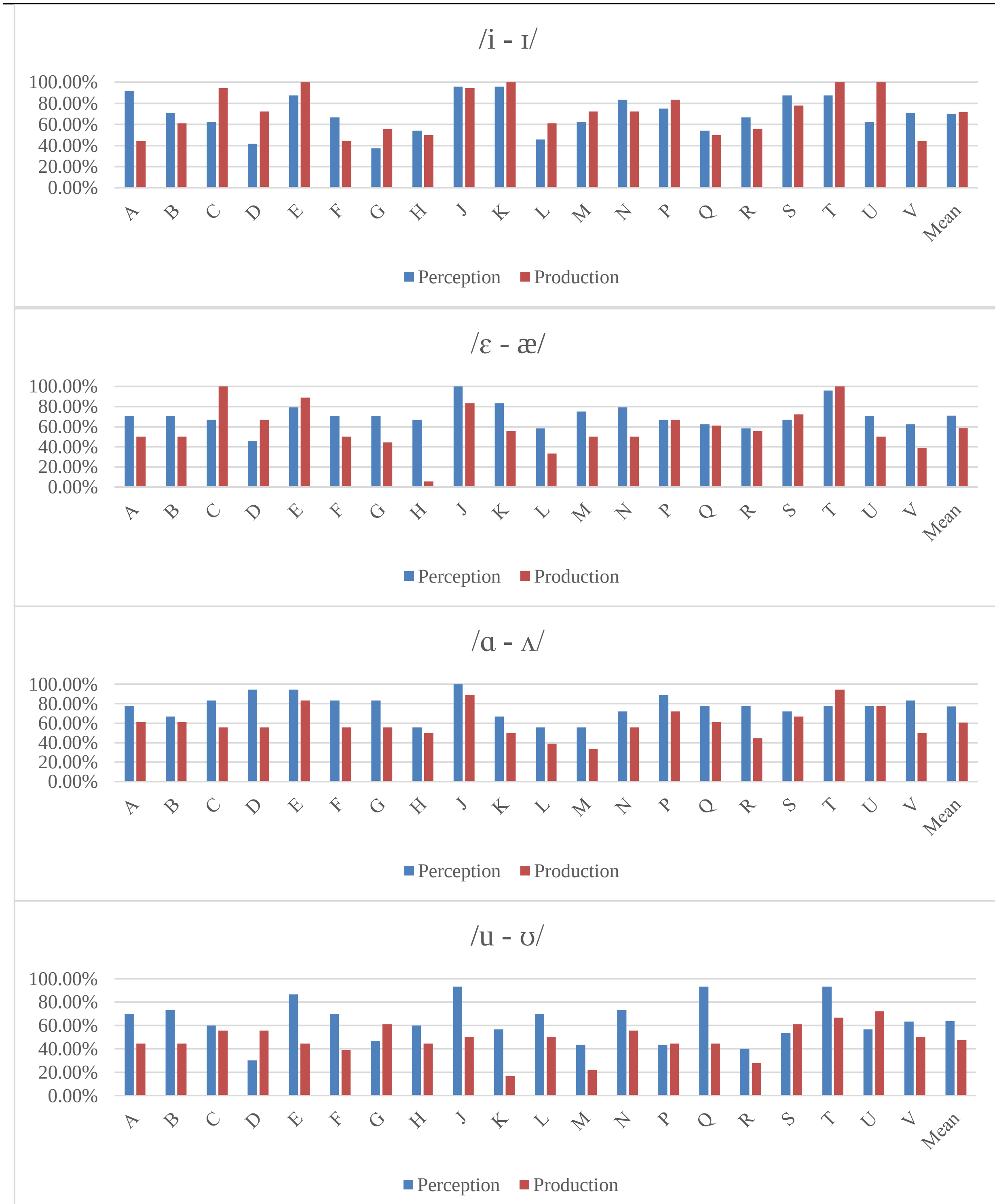


Figure 5.7 NR Perception vs. NR Production. NR Perception accuracy calculated as mean of NR perception accuracy on each vowel in pair; NR Production accuracy calculated via NE listener accuracy in identifying vowels produced by that talker (pair value is calculated as mean of individual vowels; see Appendices D and E for actual values).

(Y) on only one contrast: /ɑ - ʌ/ r = 0.605051; p = 0.0047*. These perception and production results, as measured by NR perception of NE speech vs. NE perception of NR speech, generally support the idea that perception precedes production.

Turning now to a comparison of NR perception with NR production as measured by acoustic analysis, no correlations between perception accuracy and EDV per vowel were statistically significant; however, a bivariate fit model revealed that NR perception accuracy (as the *factor variable)* predicts NE perception accuracy (as the *response* variable) on three vowels:

/ɪ/ r = 0.561532; p = 0.0100*
/ʊ/ r = 0.480036; p = 0.0322*
/a/ r = 0.43019; p = 0.0583 (near correlation)

A bivariate fit model also revealed that NR perception accuracy per vowel pair (as the *factor variable)* predicts EDP (for /i - ɪ/ and /æ - ɛ/), EDPN (for /i - ɪ/ and /æ - ɛ/), DR (for /æ - ɛ/), and DRN (for /ɑ - ʌ/ and /æ - ɛ/).

Overall, these results indicate that NR perception and production ability is correlated on at least some vowels and contrasts, and that perception ability appears to generally exceed production ability; however, due to incongruent testing methods (e.g. NE listeners heard 20 different talkers while NR listeners heard only one NE talker) and the difficulty of comparing perception accuracy with Euclidean distance- and duration ratio- based measures of production accuracy, it is not possible to conclude that these data offer strong support for the Perception-Production link.

Table 5.6 NR perception accuracy per vowel pair predicts EDP, EDPN, DR, and DRN

EDP	**DR**
/i - ɪ/ r = 0.633088; p = 0.0027* /æ - ɛ/ r = 0.745002; p = 0.0002*	/æ - ɛ/ r = 0.444739; p = 0.0494*
EDPN	**DRN**
/i - ɪ/ r = -0.59114; p = 0.0061* /æ - ɛ/ r = -0.59516; p = 0.0056*	/æ - ɛ/ r = -0.52186; p = 0.0183* /ɑ - ʌ/ r = -0.44799; p = 0.0476*

5.7 Question 6: Local Variety

5.7.1 Spectrum

This question asked whether Russian-English bilinguals in the present study are acquiring a local variety of American English or a generalized standard. Data relevant to this question come from Experiment 1 (which provided acoustic measurements for NR and NE speech) and the findings of previous publications on varieties of American English, particularly Peterson and Barney (1952) which serves as the General American standard, Hillenbrand et al.'s (1995) northern data, Hagiwara's (1997) California data, and Clopper et al.'s (2005) and Stanley et al.'s (2018) southern data. I will begin by addressing whether the NR talkers in the present study appear to be acquiring a local variety based on the formant values of vowels in their production, then turn to duration ratios between the long/short vowel pairs.

In terms of formant values, the NR group in the present study do not appear to be acquiring a local variety.

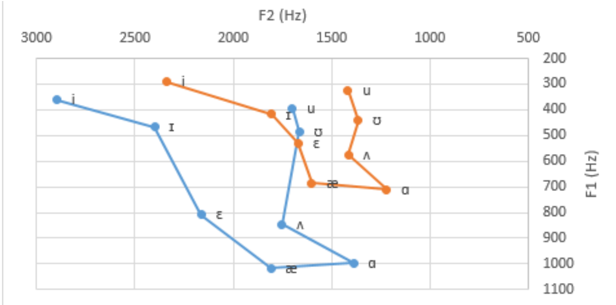
As addressed in Chapter 2, the Peterson and Barney (1952: 1052) values (top right of Figure 5.8) represent a generalized standard for American English. A handful of the speakers in this sample were not native speakers of English or spoke other languages before or in addition to English in childhood; however, most were native. The female participants in this study were largely from the Mid-Atlantic area, while the male participants were more representative of a 'General American' dialect. This figure represents a fairly typical organization of the eight tested vowels on the F1/F2 plane for a General American dialect. The other five plots in Figure 5.8 represent either monolingual English or Russian-English bilingual (bottom two plots) production collected from participants living in a particular part of the U.S. Romano et al.'s (1998) Russian-English bilinguals are included here as another example of native-Russian L2 speakers of English, to serve as a comparison with native speakers in the north as well as with Russian-English bilinguals in the south; admittedly, this comparison is not particularly robust, as the data come from a pilot study and represent the production of only two male and two female talkers, but it is the only available source of formant data for the English vowels of Russian-English bilinguals. I turn next to describing notable features of these different regional dialects spoken in the United States, before discussing whether the NR talkers in the present study appear to be acquiring the southern variety versus a generalized standard.

Typical features of California English, observed in Hagiwara (1997) (top left of Figure 5.8), include fronting of /u/ and /ʊ/ (these vowels are unrounded in Californian speech, which raises their F2 compared to other dialects) and fronting of /ʌ/. Compare this plot with the Western figures from Clopper et al. (2005), provided in Figure 5.9.

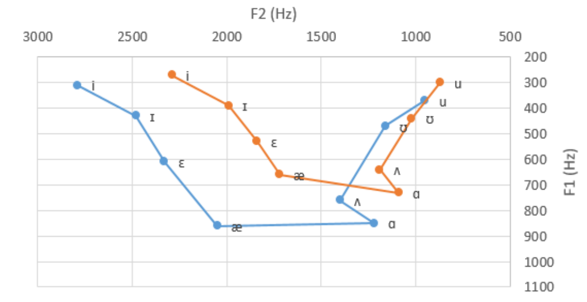
The Northern Cities Chain Shift (Labov 1998) is characterized by raising and fronting of /æ/, backing of /ɛ/ and /ʌ/, lowering of /ɔ/ (not included in this analysis), and lowering and fronting of /ɑ/; all are evident in the production of northern speakers in Hillenbrand et al. (1995) (middle left of Figure 5.8). Compare this plot with the Northern figures from Clopper et al. (2005), provided in Figure 5.10.

The Southern Vowel Shift (Labov 1998) is characterized by fronting of /u/ and /o/ (not included in the analysis of the present study), raising and fronting of /ɪ/ and /ɛ/, and lowering and backing of /i/ and /e/ (not included in this analysis). NE talkers in the present study (middle right of Figure 5.8) exhibit evidence of some of the features of the SVS – namely, fronting of /u/ and moderate raising of /ɪ/ (there is no evidence in this sample of fronting of /ɪ/, raising and fronting of /ɛ/, or lowering and backing of /i/, and /o/ and /e/ were not tested). Compare this with the Southern figures from Clopper et al. (2005) and Stanley et al. (2018) in Figures 5.11 and 5.12.

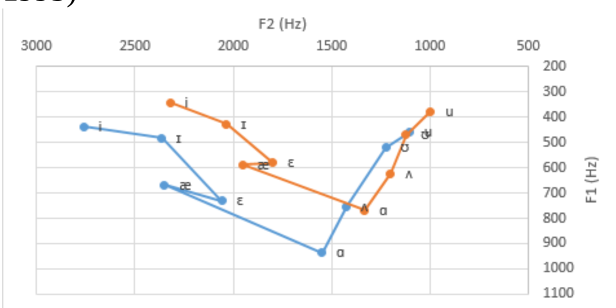
Top left: English: CA (Hagiwara 1997)



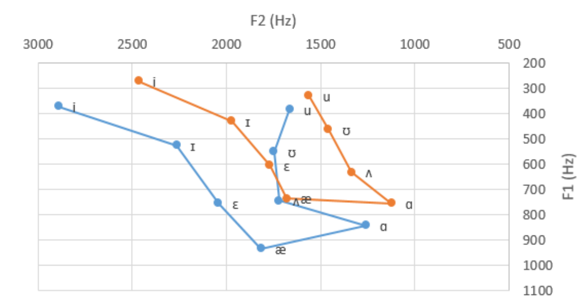
Top right: General American (Peterson & Barney 1952)



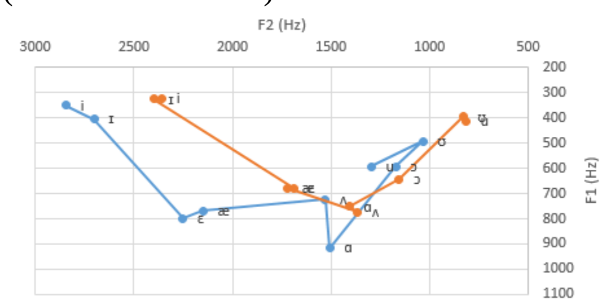
Middle left: English: north (Hillenbrand et al. 1995)



Middle right: English: south (present study)



Bottom left: Russian-English bilinguals: north (Romano et al. 1998)



Bottom right: Russian-English bilinguals: south (present study)

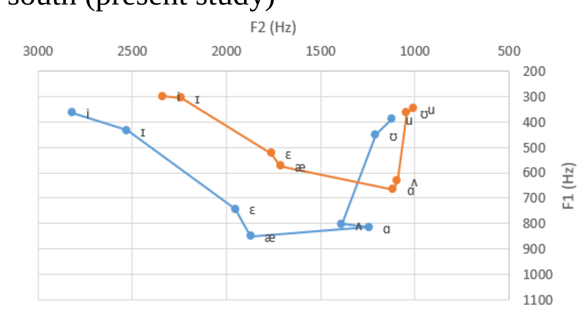


Figure 5.8 Four American English vowel charts compared with two studies of NR production. Blue: female speakers; orange: male speakers.

Compared to these three varieties of English, the NR talkers in the present study did not display evidence of the Northern Cities Chain Shift typical of northern dialects or /ʌ/ fronting typical of western speech. There appears to be some fronting of /u/, which is consistent with a Southern Vowel Shift (see Figures 5.11 and 5.12) but is also observed in the Western dialect and

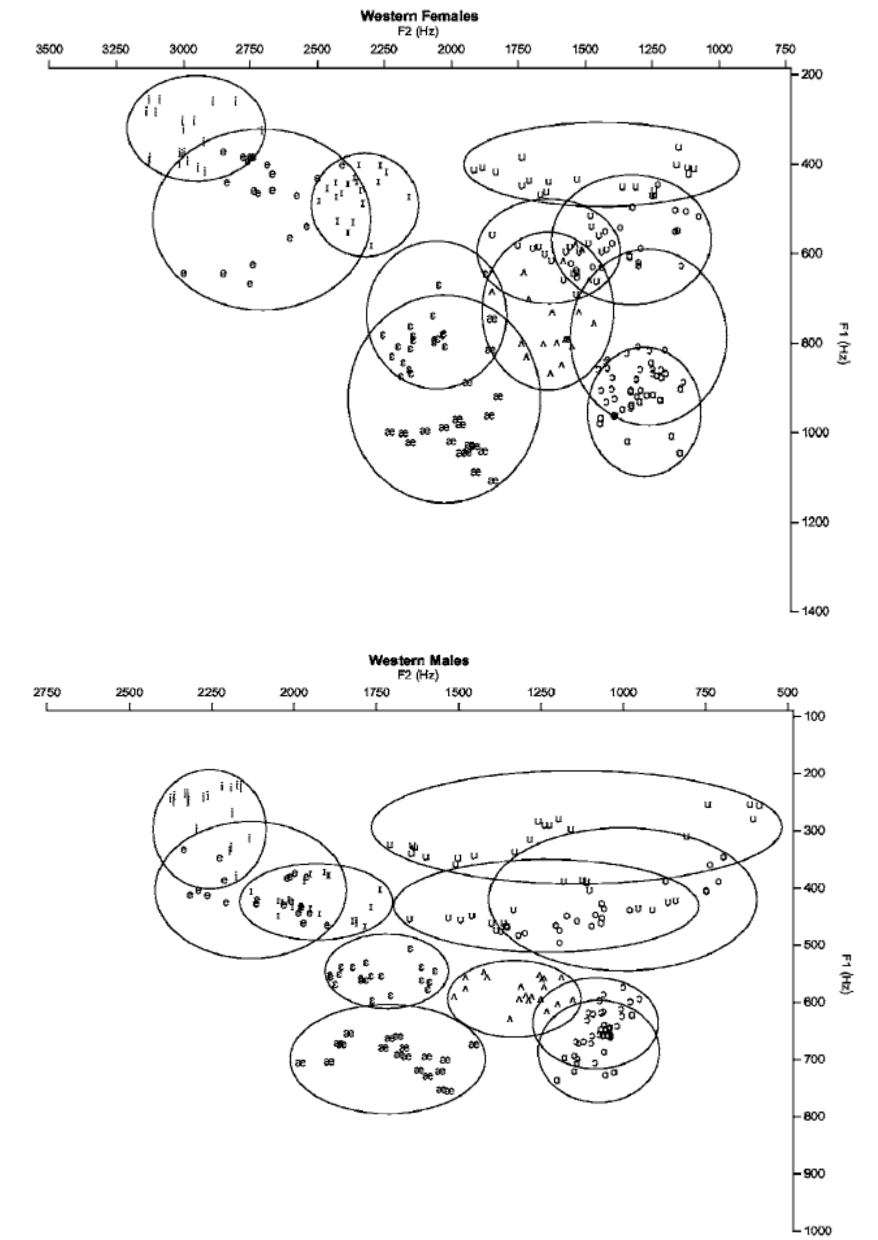


Figure 5.9 Western US speech (Clopper et al. 2005)

could also simply be influence from the L1 (failure to subdivide the part of the vowel space occupied by the native /u/ to make a nativelike distinction between L2 /u/ and /ʊ/ - see Figures 4.2 and 4.4 for a comparison of NR and NE English vowels and Russian vowels). /ɪ/ is raised and fronted in both male and female NR speakers in the present study, but, as with /u/, this could be

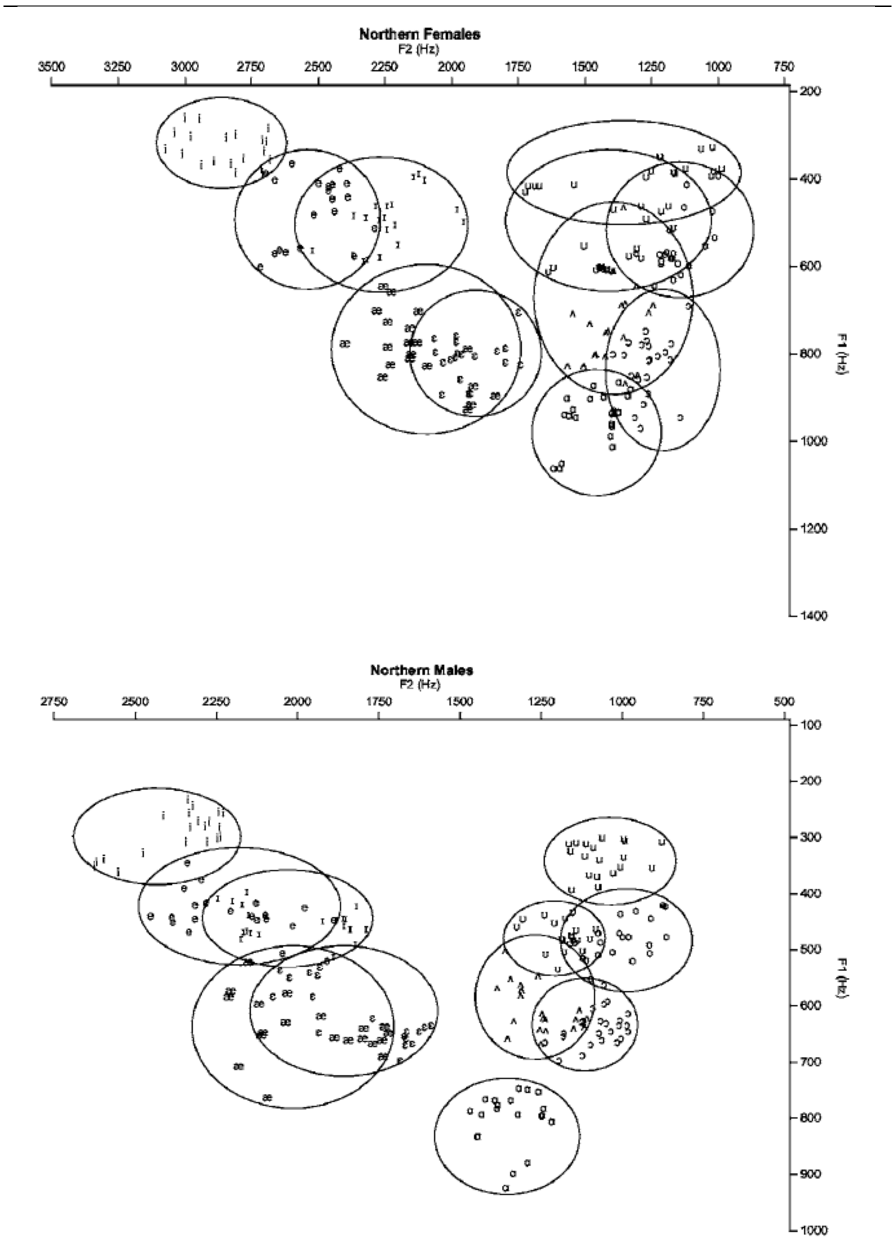


Figure 5.10 Northern US speech (Clopper et al. 2005)

interpreted *either* as a feature of the Southern Vowel Shift (in which case it would lend credibility to the possibility that the Russian-English bilinguals living in the south have at least partially acquired the southern variety) *or* influence from the L1, as this configuration of the vowel space puts the NR /i/ and /ɪ/ quite close to one another on the perceptual plane, in the perceptual space

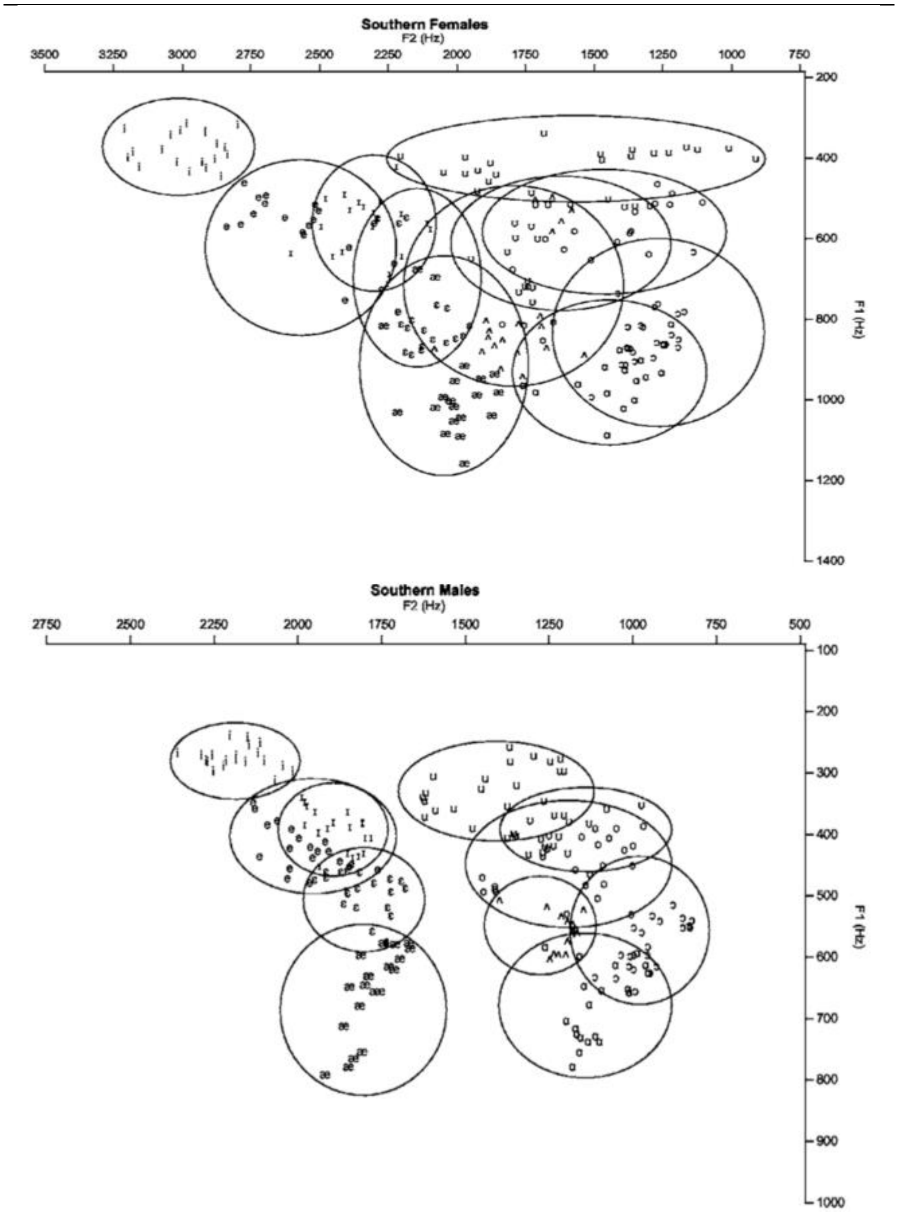


Figure 5.11 Southern US speech (Clopper et al. 2005)

of these talkers' native (Russian) /i/. Overall, then, based on formant analysis, it is more certain that these NR talkers living in the south are ***not*** acquiring a northern or western variety of American English (which is not surprising), but evidence for their acquiring a southern variety, illustrated by the NE talkers in the present study, rather than a generalized standard (Peterson & Barney 1952) is much less convincing. While it is possible to posit these Russian-English bilinguals are at least

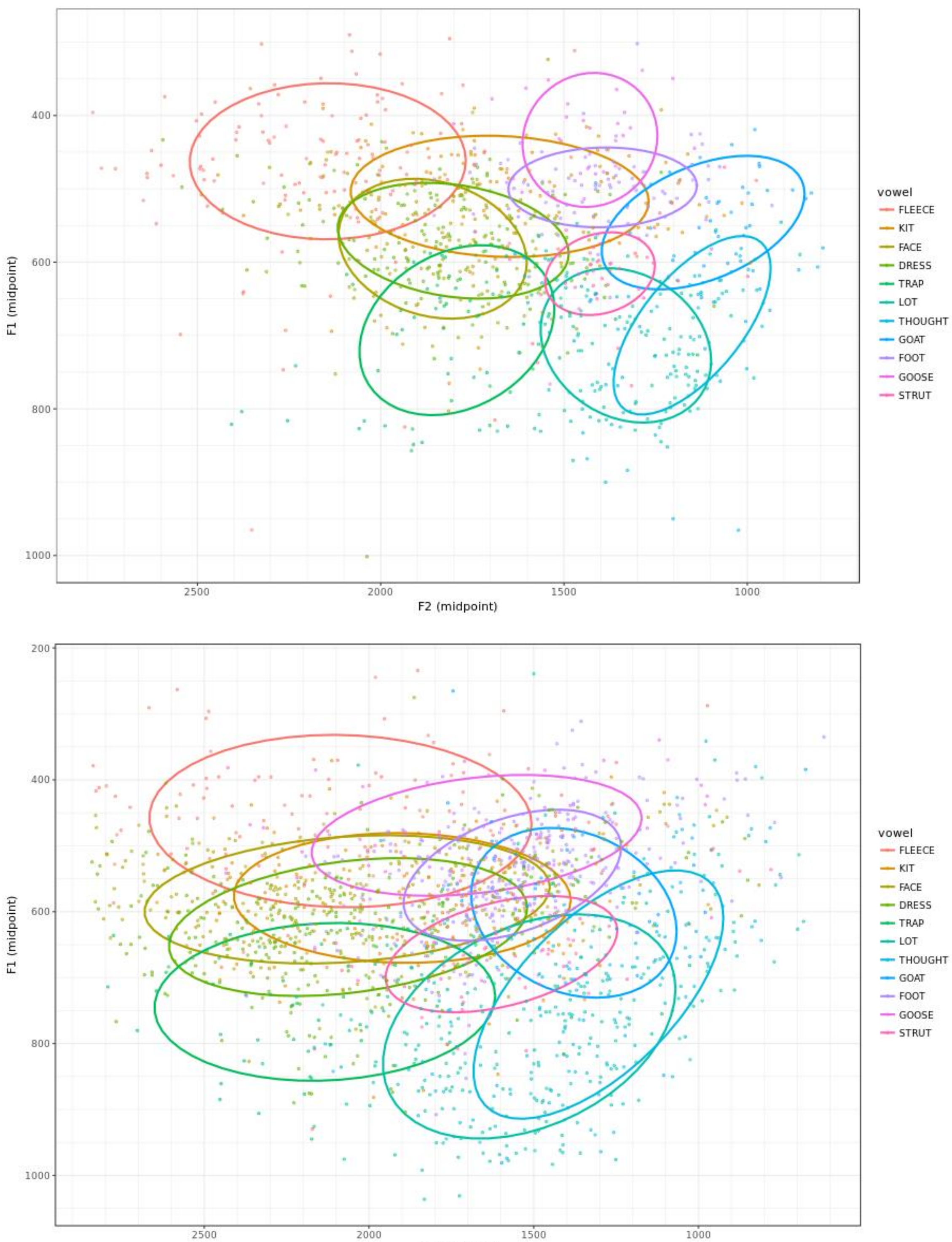


Figure 5.12 11 Southern English vowels - GA (Stanley et al. 2018)[39]

[39] Vowels /i ɪ e ɛ æ ɑ ɔ ʌ o ʊ u/ in Vt and Vd contexts; produced by three male (top) and three female (bottom) talkers; all tokens, ellipses at 67% of data.

partially acquiring the southern variety, as evidenced by a fronted /u/ and raised and fronted /ɪ/, I think, particularly given the organization of the rest of the vowel space for these talkers, it is more likely that they are acquiring a standard variety, and that those features which look reminiscent of a Southern Vowel Shift are simply influence from the L1.

The discussion so far has dealt with the production of all 20 Russian-English bilinguals, separated by sex. This necessarily ignores individual production; however, it may be that any evidence for the acquisition of a regional variety would be seen in individuals – especially those who arrived at a younger age and have lived in the south for a long time. Of the 20 Russian-English bilinguals who participated in the study, 14 had spent all of their time in the US in the state of Georgia, while 6 others had resided in other states, including NY, KS, TN, KY, FL, TX, NJ, WA, MD, and CA before moving to GA and spending at least five years in this southern state. Of those 14 who had lived only in GA, all but 3 had arrived in their 20's. Of those three who had arrived to GA in their teenage years and had lived only in GA, one arrived at the age of 17 and lived in GA for 5 years, one had arrived at age 14 and lived in GA for 10 years, and one had arrived at age 13 and lived in GA for 16 years. It is these latter two individuals, participants J and T, both females who had arrived to the US (and GA) in their early teen years and spent a decade or more living in English-speaking communities (including receiving college degrees from American universities) that seem the most likely to me to have acquired a local language variety. Plots for their production are provided in Figure 5.13. Indeed, the organization of the vowel space for these two Russian-English bilinguals has more in common with the southern variety represented by the present study's NE female talkers, suggesting that perhaps a combination of factors, including low AOA, high LOR, and cultural factors, such as education in the L2, are necessary for L2 learners to display evidence of acquiring a local L2 variety rather than a generalized standard.

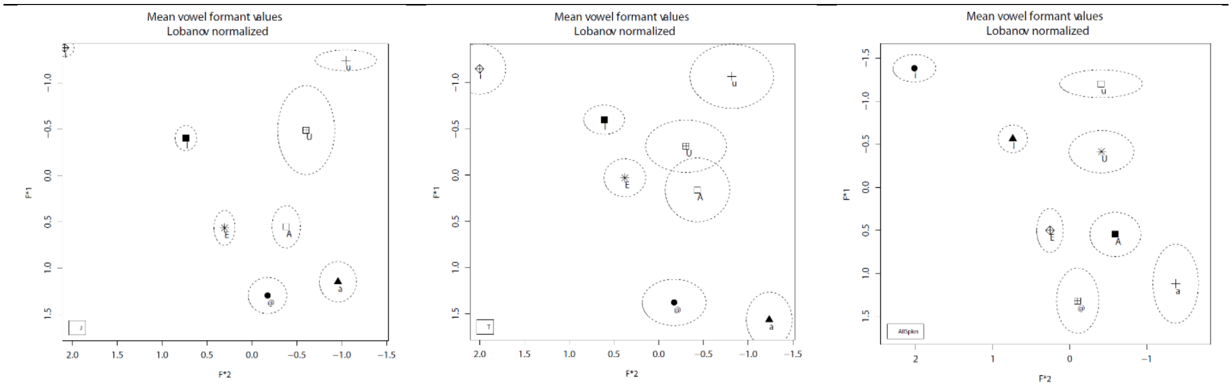


Figure 5.13 Means and 1SD Ellipses of 8 American English vowels by talkers J (left) and T (center) and all NE talkers (right) on F1/F2 plane (Lobanov normalized values)

5.7.2 Duration

Turning now to duration ratios between typically long and short paired vowels, fewer studies report these data than formant frequencies or vowel plots; for example, Clopper et al. (2005) do not provide raw vowel duration or between-vowel duration ratios, but they do state that in their study, "...Southerners did not produce generally longer vowels or have an overall slower speaking rate... but that the vowel duration differences based on dialect were selective in nature and were due to longer lax vowels for Southern talkers than for the other dialect groups. That is, the durational distinction between lax and tense vowels was reduced for the Southern talkers" (2005: 1665). However, it may be possible to draw some conclusions about the acquisition of local duration norms by L2 learners by comparing NE and NR duration data from the present study with duration values for native (Hillenbrand 1995) and Russian-English bilingual (Romano et al. 1998) northern talkers. Table 5.7 and Figure 5.14 present the duration ratios between the four long:short vowel pairs tested in the present study for monolingual and Russian-English bilingual talkers in the north (Hillenbrand 1995; Romano et al. 1998) and south (present study).

Table 5.7 Duration ratios of four vowel pairs: northern and southern native, Russian-English bilingual talkers. Ratios too small to represent an intended duration distinction shaded gray.[40]

duration ratio	NE north	NE south	NR south	NR north
highest	/ɑ - ʌ/ 1.39:1	/æ - ɛ/ 1.39:1	/i - ɪ/ 1.46:1 (1.294:1)	/i - ɪ/ 1.27:1
	/ɛ - æ/ 1.37:1	/u - ʊ/ 1.36:1	/u - ʊ/ 1.22:1	/ɑ - ʌ/ 1.08:1
	/i - ɪ/ 1.25:1	/i - ɪ/ 1.26:1	/æ - ɛ/ 1.13:1	/u - ʊ/ 1.05:1
lowest	/u - ʊ/ 1.19:1	/ɑ - ʌ/ 1.19:1	/ɑ - ʌ/ 0.98:1	/æ - ɛ/ 0.96:1

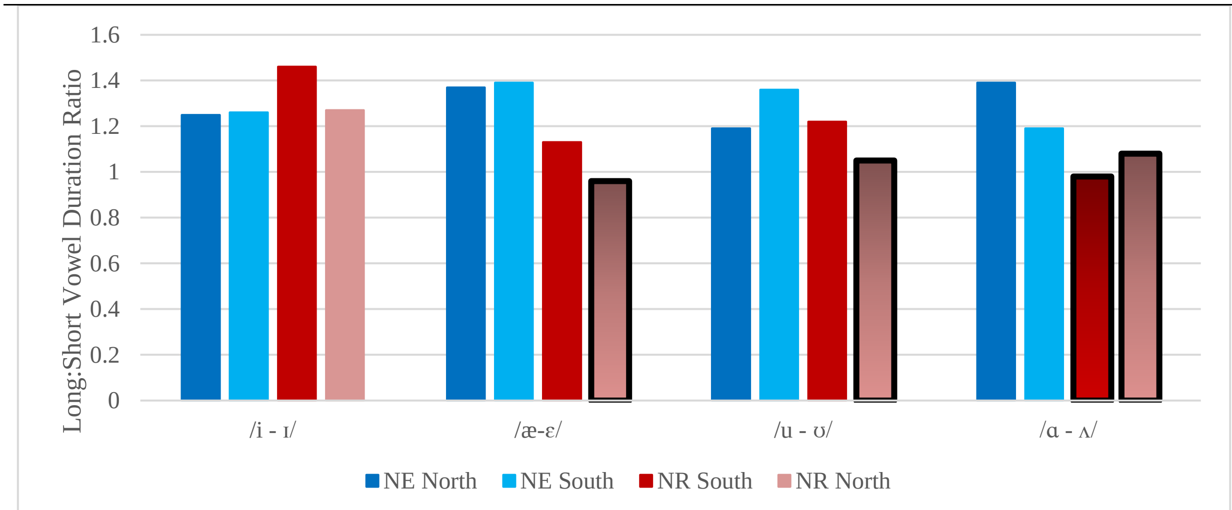


Figure 5.14 Duration ratios in production of four American English vowel contrasts in two variants of American English and two groups of native Russian speakers [41]. Ratios too small to represent an intended duration distinction have a dark border and shading.

I first want to establish what appear to be differences in native duration use between the four vowel contrasts under investigation here. Compare the first (NE north) and second (NE south) from the left data columns in Table 5.7 and the dark blue (NE north) and light blue (NE south) bars in Figure 5.14. Native English speakers in the north and south appear to have similar duration ratios between /i - ɪ/ and between /æ - ɛ/; however, their performance on the other two vowel pairs is reversed. For northern speakers, the duration ratio between /u - ʊ/ is 1.19:1 (/u/ 19% longer than

[40] From left to right: Northern English (Hillenbrand 1995), northern Russian-English bilinguals (Romano et al. 1998), southern English (present study), southern Russian-English bilinguals (present study).
[41] NE North - Hillenbrand et al. (1995); NE South, NR South - present study; NR North - Romano et al. (1998).

/ʊ/ - this group's smallest duration ratio) and the ratio between /ɑ - ʌ/ is 1.39:1 (/ɑ/ 39% longer than /ʌ/ - this group's largest duration ratio); that is, northern speakers make a greater distinction in duration between /ɑ - ʌ/ than between /u - ʊ/. For southern speakers, the duration ratio between /u - ʊ/ is 1.36:1 (/u/ 36% longer than /ʊ/ - their second largest duration ratio, exceeded only by the slightly greater duration ratio between /æ - ɛ/) and the ratio between /ɑ - ʌ/ is 1.19:1 (/ɑ/ 19% longer than /ʌ/ - their smallest duration ratio); that is, southern speakers make a greater duration distinction between /u - ʊ/ than /ɑ - ʊ/.

Next, I will compare the Russian-English bilingual data to native values for the corresponding region. The Russian-English bilinguals in the present study (3rd data column from left in Table 5.7 and dark red bars in Figure 5.14) make a large duration distinction between /i/ and /ɪ/ (likely led by the performance of three individuals; see §4.2.4 and §5.2). If this value were adjusted to exclude these three outliers, the /i - ɪ/ duration ratio would still remain the largest of the four vowel pairs for NR talkers in the present study, just slightly higher than but now rather similar to the northern and southern native ratios. The Russian-English bilinguals in the north (rightmost column in Table 5.7 and salmon bars in Figure 5.14) perform similarly to the other three groups on the /i - ɪ/ contrast. Since the native groups perform similarly on this contrast, it is not possible for the bilingual talkers to demonstrate acquisition of a local norm on this contrast; however, it does appear that this duration contrast has been acquired well by both bilingual groups.

For the Russian-English bilinguals in the north, the /i - ɪ/ contrast appears to be the only contrast on which duration ratio was acquired in a nativelike fashion. For the other three contrasts, this group produces the long:short vowel pairs with a duration difference so small it is unlikely to be detected in normal speech and likely does not represent an intended duration distinction. This is also the case for one contrast – /ɑ - ʌ/ – for the Russian-English bilingual group in the present study. For the remaining two contrasts, the present study's bilingual group produced a detectable duration distinction; however, the NR ratio for /æ - ɛ/ is lower than either of the native groups, who performed similarly on this contrast, suggesting that as a group the southern NR talkers have not fully acquired this contrast, and the southern NR ratio for /u - ʊ/ is more similar to that of northern natives than the southern native group in the present study, whose duration ratio for /u - ʊ/ is higher. Of course, individual performance on duration ratios varies (see Tables 4.16 and 4.18 and Figure 4.12), and the single most nativelike duration performance was observed in participant J, one of the two speakers with the lowest AOA and most nativelike spectral performance. The second most nativelike performance across all contrasts was observed in the talker with the 3rd lowest AOA (P), and the third most nativelike, in the talker with the 4th lowest AOA (E).

As with the discussion of formant values above, the NR group in the present study does not appear to have acquired a local variety in terms of duration ratios in long:short vowel pairs; however, it is possible that some individuals have indeed acquired these nativelike duration relationships. Taken together, these spectral and duration data provide little support for the idea that non-native speakers acquire a local variety rather than a generalized standard. Individually, however, some speakers in the present study, particularly those with low AOA, perform quite similarly to native locals in terms of both formant values and duration ratios (see §5.8 below for correlation of AOA with seven specific measures of nativelikeness). Perhaps a study that examined Russian-English bilinguals with a lower average AOA and higher LOR may have detected acquisition of local norms across the entire study group rather than in just a few individuals.

5.8 Question 7: AOA and LOR

This question asked whether age of arrival to and length of residence in an English-speaking community correlate with performance. Data relevant to this question come from all three experiments, as production and perception (both of NR and NE listeners) are tested for correlation with the NR participants involved (either as talkers or as listeners).

AOA was correlated positively with the following three measures:

- EDV on *all vowels* except /i/ and /ɛ/ (low AOA; low/nativelike EDV)
- EDPN on /æ - ɛ/ (low AOA; low/nativelike EDPN)
- DRN on /ɑ - ʌ/ (low AOA; low/nativelike DRN)

and correlated negatively with the following four measures:

- EDP on *all contrasts* except /i - ɪ/ (low AOA; high/nativelike EDP)
- DR on /ɑ - ʌ/ (low AOA; high DR – in this case more nativelike)
- NR listener accuracy on /ɑ/ (low AOA; high accuracy)
- NE listener accuracy on /ɛ - æ/ and /ɑ - ʌ/ (low AOA; high accuracy)

To briefly review these findings in prose, low AOA was associated with low EDV (that is, more nativelike performance in spectral values for individual vowels) for the vowels /ɪ/, /æ/, /ʌ/, /ɑ/, /ʊ/, and /u/. Low AOA was also associated with low EDPN (that is, more nativelike performance on distance between the vowels of tested contrasts) on /æ - ɛ/. Further, low AOA was associated with low DRN (that is, more nativelike performance on the duration ratios between typically longer and shorter vowels) on /ɑ - ʌ/. Further, low AOA was associated with high EDP (that is, more nativelike Euclidean distances between the vowels of the tested pairs) on the /ɛ - æ/, /u - ʊ/, and /ɑ - ʌ/ contrasts. Low AOA was associated with high DR (that is a higher, in this case, more nativelike, duration ratio) between /ɑ/ and /ʌ/. Finally, low AOA was also associated with high accuracy for NR listeners on /ɑ/ and for NE listeners on the /ɛ - æ/ and /ɑ - ʌ/ contrasts. These correlations were consistent with what might be expected based on existing literature on AOA and accuracy of L2 performance; one puzzling element, however, is why in some cases more nativelike performance was not associated with low AOA. It is possible that the sample in the present study was not sufficiently large or the testing methods not sufficiently targeted to elicit these correlations.

LOR in the US was correlated positively with EDV on /ɪ/, /æ/, and /ʊ/; that is, low LOR (short period of residency in the U.S.) was associated with low – nativelike – Euclidean distance between NR talkers' /ɪ/, /æ/, and /ʊ/ and the NE means for those vowels). This was an unexpected correlation based on previous literature, as, typically, *longer* and not shorter residence in an L2 speaking community would be expected to correlate with more nativelike production, and it is not clear why this should be the case. LOR in GA was correlated negatively with EDV on /ɑ/; that is, high LOR (longer period of residency in GA) was associated with low – nativelike – Euclidean distance between NR talkers' /ɑ/ and the NE means for the same vowel. This result is in line with existing literature, which holds that longer residence in an L2 speaking community should correlate with more nativelike production. Two additional correlations with LOR in GA are not in line with existing literature: that of a positive correlation between LOR in GA and EDPN on /ɑ - ʌ/ and ERN on /u - ʊ/ (recall that for both EDPN and DRN lower values represent more nativelike production). In this case, a shorter period of residence in GA is associated with lower – that is, more nativelike – scores. It is also not clear why this should be case.

5.9 Additional Considerations

Additional questions raised by the present work and not represented in the research questions included whether the present research participants' acquisition of the American English vowel system is best described in terms of acoustic measures or features, and whether the data better support an argument for the redeployment of L1 features or for the acquisition of new L2 features. These additional considerations are addressed below.

I will first address the question of whether the acquisition of American English vowels by the present study's Russian-English bilinguals is best described in terms of acoustic measures or features. In §2.2.8, I posited that the phonological features distinguishing the Russian monophthongs are [high], [back], and [round]. This feature inventory presupposes the possibility of a [+high], [+back], [-round] vowel which may fall in the range of English /ʊ/, particularly (since lack of rounding would raise the F2, centralizing the vowel - for dialects where /ʊ/ is centralized. There was no evidence, however, in the present study for Russian-English bilinguals redeploying their L1 features in this way to distinguish the high back vowels more effectively than they distinguish the high front, /ɛ - æ/, or /ʌ - ɑ/ pairs.

If we assume another set of features, such as [high], [low], and [back], as is assumed by Barrios et al. (2016) for Spanish, which has a five-vowel system similar to Russian, we presuppose the possibility of a [-high], [+low], [-back] vowel similar to English /æ/. Barrios et al. (2016) hypothesized that if native-Spanish speakers of L2 English can redeploy L1 features to acquire novel L2 contrasts, they should be able to do so to distinguish /ɑ/ and /æ/ but would not have sufficient L1 features in order to do the same for /i/ and /ɪ/. A listening task revealed no empirical support for this idea, as their Spanish-English bilinguals did not differ in their performance on the two contrasts. The present study did not test Russian-English bilinguals' ability to distinguish English /ɑ/ from /æ/ in perception; rather, I focused on the /ɑ - ʌ/ and /æ - ɛ/ distinctions. The production task, however, provided an opportunity for Russian-English bilinguals to 'confuse' /ɑ/ with /æ/ by producing them in the same perceptual space. Unsurprisingly given existing literature on this topic, Russian-English bilinguals failed to sufficiently dissimilate /æ/ from /ɛ/ rather than /ɑ/. I do not believe this supports a featural approach to the question of acquisition of American English vowels by native speakers of Russian.

Furthermore, the asymmetry in accuracy of both NR and NE listeners' perception (e.g. NE listeners identified /u/ quite well but /ʊ/ rather poorly in NR speech, while NR listeners identified the /ʊ/ of native talkers well but /u/ poorly, with some individuals never choosing /u/ on this task) suggests to me that most of the Russian-English bilinguals that participated in the present study utilized a range of strategies and attuned (sometimes in rather idiosyncratic ways, as in the case of the three talkers who produced an extraordinarily large duration distinction between their /i/ and /ɪ/) to various acoustic cues (e.g. spectrum, duration) rather than relying on redeployed L1 features to perceive and produce English vowel contrasts.

Several Russian-English bilinguals may indeed have acquired the novel L2 features responsible for the problematic contrasts. J and T in particular, whose vowels look quite nativelike on an F1/F2 plot and were exceptionally well identified by native listeners with the notable exception of the /u - ʊ/ contrast (which they perceive well), appear to have acquired these contrasts with almost nativelike proficiency. However, a feature-based approach could not explain why these two talkers (as well as K) would perform poorly on /u - ʊ/ but quite well on /i - ɪ/, two contrasts which rely on the same feature, [tense], or why C would produce a large spectral distinction between the high front vowels, have a great deal of spectral overlap between the high back pair, and consistently gravitate toward the lax variant in perception of both contrasts. These data offer little empirical support for a feature-based approach but underscore the importance of individual differences in the treatment of acoustic cues in the acquisition process.

5.10 Limitations of the Present Study

It is important to acknowledge several limitations of the present study. First, no measure of English proficiency was utilized, and thus it was not possible to compare pronunciation or listening proficiency with grammatical or lexical competence. Second, several important ecological factors were not addressed in the survey administered to participants, for example education, occupation, or percentage language use in the case of Russian-English bilinguals. Therefore, the study cannot examine the basis for and/or explain the recorded individual variation with regard to such factors. Most if not all monolingual English participants at least had high school degrees (most were University of Georgia undergraduates). Several Russian-English bilinguals mentioned having advanced degrees, but this information was not collected from all participants as part of the research process.

5.11 Conclusion

To summarize the present chapter, three NR talkers appear to be treating the duration relationship between /i/ and /ɪ/ differently than native speakers or the rest of the NR group; the rest (note that distribution is normalized by removing these three outliers) appear to have acquired the /i – ɪ/ duration contrast, producing a duration ratio very close to the NE ratio for this contrast. The NR group under-uses duration to produce the other three contrasts, falling 14 percentage points below the NE ratio for the /u - ʊ/ contrast, 26 percentage points below the NE ratio for the /æ - ɛ/ contrast (where the NR distinction may or may not be detectable), and 21 percentage points (consequently not making a duration distinction at all) for /ɑ - ʌ/. These findings are in line with the duration results of Romano et al. (1998), if considering regional variation, whose four NR talkers produced an average /i - ɪ/ duration ratio slightly higher than but rather close to the NE mean for their geographic region and a much lower duration ratio than local native speakers on the other three vowel contrasts.

In terms of spectral distance between the tested vowel contrasts, as a group, the NR talkers make a smaller spectral distinction between all four contrasts tested: about 64% the EDP of the NE group for /ɑ - ʌ/, about half the NE value for the high vowels, and about ¼ of the NE value for the /ɛ - æ/ pair. Only the /ɛ - æ/ EDP results were normally distributed, indicating considerable individual variation in performance on the other three contrasts; indeed, extreme highs and lows were observed, with a handful of participants responsible for some of the highest or lowest performance on several (in some cases all four) contrasts; the /ɛ - æ/ contrast lacked the distinct 'highs' observed in the other three contrasts – no NR talker came close to the NE mean EDP for this contrast, and the highest NR EDP for /ɛ - æ/ fell just below the minimum NE value for the same contrast.

Production accuracy on individual vowels, as measured by EDV, was as follows: best for /i/, followed in descending order by /ɛ/, /ɑ/, /æ/, /ɪ/, /ʌ/, /u/, and finally /ʊ/. In terms of vowel pairs, most nativelike performance, as measured by EDP, was observed for the /ɑ - ʌ/ pair, decreased for /u - ʊ/, decreased further for /i - ɪ/, and was least nativelike for the /ɛ - æ/ pair.

The most nativelike DRN performance across all NR talkers was observed for /u - ʊ/, followed by /ʌ - ɑ/, then /ɛ - æ/, and, finally, least nativelike performance was observed for the /i - ɪ/ pair.

NE listener accuracy in identifying vowels produced by NR talkers was, in decreasing order: i > u > ɑ > ɪ > ɛ > æ > ʌ > ʊ for individual vowels and /i - ɪ/ > /ʌ - ɑ/ > /ɛ - æ/ > /u - ʊ/ for pair averages. NE listener accuracy was correlated with only three measures of duration (DRN) but with many more measures of Euclidean distance (four statistically significant correlations of vowels or contrasts with EDV, seven with EDP, and six with EDPN), indicating that native listeners rely more on spectral than duration cues to correctly identify vowels spoken by NR talkers.

Regarding the Perception-Production link, NR perception and production ability in the present study appears to be correlated on at least some vowels and contrasts. Additionally, perception ability appears to generally exceed production ability; however, due to incongruent testing methods and the difficulty of comparing perception accuracy with Euclidean distance- and duration ratio- based measures of production accuracy, it is not possible to conclude that these data offer strong support for the Perception-Production link.

The Russian-English bilinguals in the present study do not appear as a group to be acquiring a local variety based on spectral or durational analysis of their production, although it is possible that some talkers – those with low AOA, relatively high LOR, and extensive cultural experience (including education) in Georgia – may have acquired a local rather than generalized standard.

AOA was correlated with multiple measures of Euclidean distance and duration; all correlations with AOA were in line with existing literature in that low AOA was correlated with more nativelike performance. LOR in the U.S. was correlated with only one measure (that of EDV on /ɪ/, /æ/, and /ʊ/); however, the direction of the correlation was unexpected in that a short period of residency in the U.S. was associated with more nativelike performance on those vowels. LOR in GA was correlated in a predictable way with EDV on /ɑ/ but in a way that is not in agreement with existing literature on two additional correlations (EDPN on /ɑ - ʌ/ and ERN on /u - ʊ/).

CHAPTER 6

CONCLUSION

The present work sought to fill research gaps in existing literature on cue weighting (specifically the treatment of spectral and duration cues in the acquisition of English vowels by native speakers of Russian, an under-researched L1-L2 pairing), individual- and vowel-based differences in learning trajectories, and the acquisition of local language norms, and to contribute evidence for the perception-production link and for the roles of AOA and LOR in the acquisition of nonnative vowels. In doing so, this research has added empirical data to the existing body of knowledge on the acquisition of English vowel contrasts by native speakers of Russian and contributed important insights into the differential treatment of acoustic cues in the acquisition of a more complex vowel system in adulthood.

In particular, this work has underscored the importance of learner differences as part of the acquisition process, including asymmetrical performance on contrasts predicted to be equally challenging from a featural perspective. It has also raised some unexpected and tantalizing questions about the impact of regional dialect on native use of duration in certain vowel contrasts.

Perhaps most importantly, the present work has highlighted the need for further inquiry into the roles of individual variability in both perception and production, as well as the role of local dialect in second language acquisition, and identified potential directions for further inquiry. More precise research is needed to determine what role dialect plays in cue weighting in the acquisition of non-native vowel contrasts and connections between perception and production accuracy. Gaps, which could be addressed with more targeted methodology, remain in our understanding of each of these specific areas – for example, how regional variety affects cue weighting and how individual learning trajectories differ in second language acquisition. Future research should carefully control for factors such as AOA, LOR, and dialect, and design experimental tasks that isolate the roles of the variables under investigation. In particular, congruent tasks are needed to compare perception and production accuracy, as well as performance on contrasts which might illuminate the role of L1 and L2 features in the acquisition of L2 segments.

REFERENCES

Abrahamsson, Niclas & Kenneth Hyltenstam. 2009. Age of Onset and Nativelikeness in a Second Language: Listener Perception Versus Linguistic Scrutiny. *Language Learning* 59(2). 249-306.

Adamson, H. Douglas & Vera M. Regan. 1991. The acquisition of community speech norms by Asian immigrants learning English as a second language. *Studies in Second Language Acquisition* 13. 1-22.

Altenberg, E. & Vago R. 1987. Theoretical implications of an error analysis of second language phonology production. In G. Ioup, & S. H. Weinberger (Eds.), *Interlanguage phonology. The acquisition of a second language sound system. Series on issues in second language research*, 427–447. Cambridge: Newbury House Publishers.

Anisman, Paul H. 1975. Some aspects of code switching in New York Puerto Rican English. *Bilingual Review* 2. 56-85.

Aoyama, Katsura, James Emil Flege, Susan G Guion; Reiko Akahane-Yamada & Tsuneo Yamada. 2004. Perceived phonetic dissimilarity and L2 speech learning: The case of Japanese /r/ and English /l/ and /r/. *Journal of Phonetics* 32. 233-250.

Archibald, John. 2005. Second language phonology as reployment of L1 phonological knowledge. *Canadian Journal of Linguistics* 50. 285-314

Asher, James T. & Ramiro Garcia. 1969. The optimal age to learn a foreign language. *The Modern Language Journal* 53(5). 334-341.

Avanesov, Ruben Ivanovič. 1972. *Russkoe Literaturnoe Proiznoshenie*. Prosvezchenie, Moskva.

Baker, Wendy & Pavel Trofimovich. 2006. Perceptual paths to accurate production of L2 vowels: The role of individual differences. *International Review of Applied Linguistics* 44. 231-250.

Barrios, Shannon; Nan Jiang & William J. Idsardi. 2016. Similarity in L2 Phonology: Evidence from L1 Spanish late-learners' perception and lexical representation of English vowel contrasts. *Second Language Research* 32(3). 367-395.

Bell-Berti, Fredericka; Lisa Jayne Romano & Eugenia Lorin. Comparison of American English vowel production and identification by native speakers of Russian. *Proceedings of the 16th International Congress on Acoustics,* Seattle, WA, 1998. 2997-2998.

Benson, Morton. 1959. English Loanwords in Russian. *The Slavic and East European Journal*. 3(3). 248-267.

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.

Best, Catherine T & Michael D. Tyler. 2007. Nonnative and second-language speech perception. In O.-S. Bohn and M.J. Munro (Eds.), *Language experience in second language speech learning*, 13–34. Amsterdam: John Benjamins.

Bialystok, Ellen. 1997. The structure of age: In search of barriers to second language acquisition. *Second Language Research* 13(2). 116-137.

Birdsong, David. 2006. Age and second language acquisition and processing: A selective overview. *Language Learning* 56. 9–49.

Boersma, Paul & D. Weenink. 2017. *Praat: doing phonetics by computer [Computer program]. Version, 6.0.27*, retrieved 17 March 2017 from http://www.praat.org/

Bohn, Ocke-Schwen. 1995. Cross-language speech perception in adults: First language transfer doesn't tell it all. In W. Strange (Ed.), *Speech perception and linguistic experience: Issues in cross-language research*, 273– 304. Timmonium, MD: York Press.

Bohn, Ocke-Schwen & James Emil Flege. 1990. Interlingual identification and the role of foreign language experience in L2 vowel perception. *Applied Psycholinguistics* 11. 303-328.

Bohn, Ocke-Schwen & Murray J. Munro. (Eds.) 2007. *Language Experience in Second Language Speech Learning: In honor of James Emil Flege.* Amsterdam / Philadelphia: John Benjamins.

Bondarko, Liya V. 1998. *Fonetika sovremennogo russkogo jazyka.* St. Petersburg.

Bongearts, Theo. 1999. Ultimate attainment in L2 pronunciation: The case of very advanced late learners. In D. Birdsong (Ed.), *Second language acquisition and the critical period hypothesis*, 133–159. Mahwah, NJ: Erlbaum.

Bongearts, Theo, Chantal van Summeren, Brigitte Planket, & Schils Erik. 1997. Age and ultimate attainment in the pronunciation of a foreign language. *Studies in Second Language Acquisition* 19. 447-465.

Bradlow, Ann R. 2008. Training non-native language sound patterns: Lessons from training Japanese adults on the English /r/-/l/ contrast. In J. G. Hansen Edwards, & M. L. Zampini (Eds.), *Phonology and Second Language Acquisition*, 287–308. John Benjamins Publishing Company.

Bradlow, Ann R., David B. Pisoni, Reiko Akahane-Yamada & Yoh'ichi Tohkura. 1997. Training Japanese listeners to identify English /r/ and /l/: IV. Some effects of perceptual learning on speech production. *The Journal of the Acoustical Society of America*. 101(4). 2299-2310.

Brannen, Kathleen. 2002. The role of perception in differential substitution. *Canadian Journal of Linguistics* 47. 1-46.

Brown, Cynthia A. 1998. The role of the L1 grammar in the L2 acquisition of segmental structure. *Second Language Research* 14. 136-93.

Brown, Cynthia A. 2000. The interrelation between speech perception and phonological acquisition from infant to adult. In: Archibald J (eds) *Second language acquisition and linguistic theory*, 4–63. Oxford: Blackwell.

Busch, Deborah. 1982. Introversion-extraversion and the EFL proficiency of Japanese students. *Language Learning* 32. 109-132.

Canepari, Luciano & Daniele Vitali. 2018. *Russian Pronunciation & Accents*. LINCOM GmbH

Cardoso, Walcir. 2007. The variable development of English word-final stops by Brazilian Portuguese speakers: A stochastic optimality theoretic account. *Language Variation and Change* 19. 1-30.

Cardoso, Walcir. 2011. The development of coda perception in second language phonology: A variationist perspective. *Second Language Research* 27. 433-465.

Cebrian, Juli. 2006. Experience and the use of non-native duration in L2 vowel categorization. *Journal of Phonetics* 34. 372-387.

Chen, Matthew. 1970. Vowel Length Variation as a Function of the Voicing of the Consonant Environment. *Phonetica* 22(3). 129-159.

Chew, P. A. 2003. A Computational Phonology of Russian. Dissertation. Dissertation.com.

Chomsky, Noam & Morris Halle. 1968. *The Sound Pattern of English*. The MIT Press, Cambridge, Massachusetts, London, England.

Clements, G.N. 2001. Representational economy in constraint-based phonology. In Hall T. Alan. (Ed.) Distinctive Feature Theory, 71–146. Mouton de Gruyter, Berlin.

Clopper, Cynthia G, David B. Pisoni & Kenneth de Jong. 2005. Acoustic characteristics of the vowel systems of six regional varieties of American English. *The Journal of the Acoustical Society of America*. 118(3). 1661-1676.

Crosswhite, Katherine Margaret. 2000. Vowel Reduction in Russian: A Unified Account of Standard, Dialectal, and "Dissimilative" Patterns. University of Rochester Working Papers in the Language Sciences, Vol. 1(1): 107-172. Katherine M. Crosswhite and Joyce McDonough (eds.)

Crowther, Court S & Virginia Mann. 1992. Native language factors affecting use of vocalic cues to signal consonant voicing in English. *The Journal of the Acoustical Society of America* 92. 711–722. doi: 10.1121/ 1.403996

Cubberley, Paul. 2002. *Russian: A Linguistic Introduction*, Cambridge University Press.

Cutler, Anne; Andrea Weber, Roel Smits & Nicole Cooper. 2004. Patterns of English phoneme confusions by native and non-native listeners. *The Journal of the Acoustical Society of America* 116(6). 3668-3678

Darcy, Isabelle & Franziska Krüger F. 2012. Vowel perception and production in Turkish children acquiring L2 German. *Journal of Phonetics* 40. 568-581.

DeKeyser, Robert M. 2000. The robustness of critical period effects in second language acquisition. *Studies in Second Language Acquisition* 22. 499-533.

Dupoux, Emmanuel & Sharon Peperkamp. 2002. Fossil markers of language development: Phonological “deafnesses” in adult speech processing. In J. Durand & b. Laks (Eds.), *Phonetics, phonology, and cognition*, 168–190. Oxford: Oxford University Press.

Elliott, A. Raymond. 1995. Field independence/dependence, hemispheric specialization, and attitude in relation to pronunciation accuracy in Spanish as a foreign language. *Modern Language Journal* 79. 356-371.

Escudero, Paola. 2000. Developmental patterns in the adult L2 acquisition of new contrasts: The acoustic cue weighting in the perception of Scottish tense/lax vowels by Spanish speakers. Unpublished Master’s thesis, University of Edinburgh, Scotland, UK.

Escudero, Paola. 2007. Second-language phonology: the role of perception. In: Pennington M.C. (eds) Phonology in Context. Palgrave Advances in Linguistics. Palgrave Macmillan, London.

Escudero, Paola & Paul Boersma. 2004. Bridging the gap between L2 speech perception research and phonological theory. *Studies in Second Language Acquisition* 26. 551-585.

Escudero, Paola, Titia Benders, & Silvia C. Lipski. 2009. Native, non-native and L2 perceptual cue weighting for Dutch vowels: The case of Dutch, German, and Spanish listeners. *Journal of Phonetics* 37(4). 452-465.

Flege, James Emil. 1987. The production of ‘new’ and ‘similar’ phones in a foreign language: evidence for the effect of equivalence classification. *Journal of Phonetics* 15. 47-65.

Flege, James Emil. 1988a. Factors affecting degree of perceived foreign accent in English sentences. *The Journal of the Acoustical Society of America* 84. 70-79.

Flege, James Emil. 1988b. The production and perception of foreign language speech sounds. In Winitz H., (Ed.), *Human communication and its disorders: a review*, 224-401. Norwood, NJ: Ablex Publishing.

Flege, James Emil. 1991. The interlingual identification of Spanish and English vowels: Orthographic evidence. *Quarterly Journal of Experimental Psychology* 34A. 701-731.

Flege, James Emil. 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 Emil; Bohn, Ocke-Schwen & Sunyoung Jang. 1997. Effects of experience on non-native speaker production and perception of English vowels. *Journal of Phonetics* 25. 437-70.

Flege, James Emil & Kathryn L. Fletcher. 1992. Talker and listener effects on degree of perceived foreign accent. *The Journal of the Acoustical Society of America* 91(1). 370-389

Flege, James Emil & Ian R. A. MacKay. 2004. Perceiving vowels in a second language. *Studies in Second Language Acquisition* 26. 1-34.

Flege, James Emil; MacKay Ian R. A. & Diane Meador. 1999. Native Italian speakers' perception and production of English vowels. *The Journal of the Acoustical Society of America* 106. 2973-2987.

Flege, James Emil & Murray J. Munro. 1994. Auditory and categorical effects on cross-language vowel perception. *The Journal of the Acoustical Society of America* 95. 3623-3641.

Flege, James Emil; Munro, Murray J. & MacKay IRA. 1995. Effects of age of second-language learning on the production of English consonants. *Speech Communication* 16. 1-26.

Flege, James Emil; Munro, Murray J. & Robert Allen Fox. 1994. Auditory and Categorical effects on cross-language vowel perception. *The Journal of the Acoustical Society of America* 95(6). 3623-3641.

Flege, James Emil & Robert Port. 1981. Cross-language phonetic interference: Arabic to English. *Language & Speech* 24. 25-146.

Flege, James Emil; Schirru Carlo & Ian R. A. MacKay. 2003. Interaction between the native and second language phonetic subsystems. *Speech Communication* 40. 467-491.

Flege, James Emil; Yeni-Komshian, Grace H & Serena Liu. 1999b. Age constraints on second-language acquisition. *Journal of Memory and Language* 41. 78-104.

Flemming, E & S. Johnson. 2007. Rosa's roses: reduced vowels in American English. *Journal of the International Phonetic Association* 37(1). 83-96.

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.

Fox, Robert Allen & Ewa Jacewicz. 2009. Cross-dialectal variation in formant dynamics of American English vowels. *The Journal of the Acoustical Society of America* 126(5). 2603–2618.

Gass, S. M. & L. Selinker. 2008. *Second Language Acquisition*. (3rd ed.). New York: Routledge.

Giegerich, Heinz J. 1992. *English phonology*. Cambridge University Press.

Gilichinskaya, Yana D & Winifred Strange. 2010. Perceptual assimilation of American English vowels by inexperienced Russian listeners. *The Journal of the Acoustical Society of America* 128 (2). EL80-EL85.

Grimaldi, Mirko, Bianca Sisinni, Barbara Gili Fivela, Sara Invitto, Donatella Resta, Paavo Alku & Elvira Brattico. 2014. Assimilation of L2 vowels to L1 phonemes governs L2 learning in adulthood: a behavioral and ERP study. *Frontiers in Human Neuroscience* 8(279). 1-14.

Guion, Susan G., James E. Flege & Jonathan D. Loftin. 2000. The effect of L1 use on pronunciation in Quichua-Spanish bilinguals. *Journal of Phonetics* 28. 27-42.

Hagiwara, Robert. 1997. Dialect variation and formant frequency: The American English vowels revisited. *The Journal of the Acoustical Society of America* 102. 655–658.

Hall, T. Alan. 2007. Segmental features. In Paul de Lacy (ed.) *The Cambridge Handbook of Phonology,* 311-333. Cambridge University Press.

Halle, Morris. 1971. *The Sound Pattern of Russian*. Mouton, The Hague, Paris.

Hancin-Bhatt, Barbara. 1994. Segment transfer: a consequence of a dynamic system. Second Language Research 10(3). 241-269.

Harley, Birgit & Doug Hart. 1997. Language aptitude and second language proficiency in classroom learners of different starting ages. *Studies in Second Language Acquisition* 19: 379-400.

Hillenbrand, James M. 2003. American English: Southern Michigan. *Journal of the International Phonetic Association* 33(1). 121–126

Hillenbrand, James M., Michael J. Clark & Robert A. Houde. 2000. Some effect of duration on vowel recognition. *The Journal of the Acoustical Society of America* 108. 3013-3022.

Hillenbrand, James M., Laura A.Getty, Michael J. Clark & Kimberlee Wheeler. 1995. Acoustic characteristics of American English vowels. *The Journal of the Acoustical Society of America* 97(5 Pt 1). 3099-3111.

Holt, Lori L. & Andrew J. Lotto. 2006. Cue weighting in auditory categorization: Implications for first and second language acquisition. *The Journal of the Acoustical Society of America* 119. 3059 –71.

Hyltenstam, Kenneth. 1992. Non-native features of near-native speakers: On the ultimate attainment of childhood L2 learners. In R. J. Harris (Ed.), *Cognitive processing bilinguals*, 351-368. Amsterdam: Elsevier Science.

Hyltenstam, Kenneth & Niclas Abrahamsson. 2000. Who can become native-like in a second language? All, some, or none? On the maturational constraints controversy in second language acquisition. *Studia Linguistica* 54. 150-166.

Hyltenstam, Kenneth & Niclas Abrahamsson. 2003. Maturational constraints in SLA. In C. J. Doughty & M. H. Long (Eds.), *The handbook of second language acquisition*, 539-588. Malden, MA: Blackwell.

Ioup, Georgette. 2005. Age and second language development. In E. Hinkel (ed.), *Handbook of research in second language teaching and learning*, 419-35. Mahwah, JN: Lawrence Erlbaum.

Ioup, Georgette, Elizabeth Boustagui, Manal El Tigi & Martha Moselle. 1994. Reexamining the critical period hypothesis: A case study of successful adult SLA in a naturalistic environment. Studies in Second Language Acquisition 16. 73-98.

Roach, P.J. 1989. Report on the 1989 Kiel Convention: INTERNATIONAL PHONETIC ASSOCIATION. *Journal of the International Phonetic Association*. 19(2). 67-80.

Isbell, Daniel Richard. 2016. The perception-production link in L2 phonology. MSU Working Papers in SLS, Vol 7. 57-67.

Iverson, Paul & Patricia K. Kuhl. 1995. "Mapping the perceptual magnet effect for speech using signal detection theory and multidimensional scaling". *The Journal of the Acoustical Society of America* 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. B47– B57. doi: 10.1016/S0010-0277(02)00198-1

Jacewicz, Ewa, Robert A. Fox & Joseph Salmons. 2007. Vowel duration in three American English dialects. *American Speech* 82(4). 367-385.

Janulienė, Aušra & Justina Andriulaitytė J. 2016. On English IT field borrowings in Modern Russian. *Sustainable Multilingualism* 8. 81-100.

Jarvis, Scott & Aneta Pavlenko. 2008. *Crosslinguistic Influence in Language and Cognition*. New York & London: Routledge.

Johnson, Jacqueline S. & Elissa L. Newport. 1989. Critical period effects in second language learning: the influence of maturational state on the acquisition of English as a second language. *Cognitive Psychology* 21. 60-99.

Jones, Daniel & Dennis Ward. 1969. *The Phonetics of Russian*. Cambridge University Press.

Kim, Donghyun, Meghan Clayards & Heather Goad. 2017. A longitudinal study of individual differences in the acquisition of new vowel contrasts. *Journal of Phonetics* 67. 1-20.

Kinney, John. 2013. The L2 English production of [ð] in word-initial onset and intervocalic onset position: A pilot study. George Mason University Working Papers in Linguistics Vol. 9.

Klatt, Dennis H. 1976. Linguistic uses of segmental duration in English: Acoustic and perceptual evidence. *The Journal of the Acoustical Society of America* 59(5). 1208–21.

Kondaurova, Maria V. & Alexander L. Francis. 2004. Perception of the English tense/lax vowel contrast by native speakers of Russian. *The Journal of the Acoustical Society of America* 116. 2572.

Kondaurova, Maria V. & Alexander L. Francis. 2006. Russian and Spanish listener's perception of the English tense/lax vowel contrast: Contributions of native language allophony and individual experience. *The Journal of the Acoustical Society of America* 120(5). 3293-3293.

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. *The Journal of the Acoustical Society of America* 124(6). 3959-3971.

Krashen, Stephen D., Michael A. Long & Robin C. Scarcella. 1979. Accounting for child-adult differences in second language rate and attainment. *TESOL Quarterly* 13. 573-82.

Kretzschmar, William A., Paulina Bounds, Jacqueline Hettel, Lee Pederson, Ilkka Juuso, Lisa Lena Opas-Hänninen & Tapio Seppänen (2013). "The Digital Archive of Southern Speech (DASS)." *Southern Journal of Linguistics*, 27(2). 17–38.

Kuhl, Patricia K. 1991. Human adults and human infants show a 'perceptual magnet effect' for the prototypes of speech categories, monkeys do not. *Perception and Psychophysics* 50. 93-107.

Kuhl, Patricia K. 2000. A new view of language acquisition. *Proceedings of the National Academy of Sciences* 97. 11850-11857.

Kuhl, Patricia K & Paul Iverson. 1995. Linguistic experience and the "perceptual magnet effect". In W. Strange (Ed.), *Speech Perception and Linguistic Experience*, 121–154. Timonium, MD: York Press.

Kuhl, Patricia K., Karen A.Williams, Francisco Lacerda, Kenneth N. Stevens & Björn Lindblom. 1992. Linguistic experience alters phonetic perception in infants by 6 months of age. *Science* 255. 606-608.

Kulikov, Vladimir. 2012. Voicing and voice assimilation in Russian stops (Ph.D. dissertation), University of Iowa.

Kulikov, Vladimir. 2013. Voicing contrast in consonant clusters: evidence against sonorant transparency to voice assimilation in Russian. *Phonology* 30.3 (Dec 2013): 423-452.

Labov, William. 1994. *Principles of Linguistic Change: Internal Factors*. Blackwell, Malden, MA.

Labov, William. 1998. The three dialects of English. In *Handbook of Dialects and Language Variation*, ed. by MD Linn. Academic Press, San Diego, pp. 39-81.

Labov, William, Sharon Ash & Charles Boberg. 2006. *The Atlas of North American English*. Berlin/New York: 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. *Journal of the International Phonetic Association*.

Ladefoged, Peter. 2001. *A Course in Phonetics* (4th edn.) Boston, MA: Heinle & Heinle.

Ladefoged, Peter. 2005. *Vowels and Consonants: An Introduction to the Sounds of Languages* (2nd edn.). Malden, MA: Blackwell.

Lado, Robert. 1957. *Linguistics across cultures: Applied linguistics for language teachers*. University of Michigan Press: Ann Arbor.

Lehiste, Ilse. 1970. *Suprasegmentals*. Cambridge, MA–London: The MIT Press.

Lipski, Silvia C., Paola Escudero & Titia Benders. 2012. Language experience modulates weighting of acoustic cues for vowel perception: An event-related potential study. *Psychophysiology* 49. 638-650.

Lobanov, Boris M. 1971. Classification of Russian vowels spoken by different speakers. *The Journal of the Acoustical Society of America* 49(2B). 606–608.

Lombardi, Linda. 2003. Second language data and constraints on manner: explaining substitutions for the English interdentals. Second Language Research 19. 225-50.

Long M. 1990. Maturational constraints on language development. *Studies in second language acquisition* 12. 251-285.

MacKay, Ian R.A., Diane Meador D & James Emil Flege. 2001. The identification of English consonants by native speakers of Italian. *Phonetica* 58. 103-125.

Makarova, Aleksandra Olegovna. 2010. Acquisition of Three Vowel Contrasts by Russian Speakers of American English. PhD Dissertation. Harvard University.

Marinova-Todd, S. H. 2003. Native, near-native or non-native: Comprehensive analysis of ultimate attainment in adult second language acquisition. Unpublished doctoral dissertation, Harvard University.

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. *Journal of Phonetics* 30. 229-258

Miller, George A. & Patricia E. Nicely. 1955. An analysis of perceptual confusions among some English consonants. *The Journal of the Acoustical Society of America* 27(2). 338-352.

Fox, Michelle Minnick & Kazuaki Maeda. 1999. Categorization of American English vowels by Japanese speakers. In J. J. Ohala, Y. Hasegawa, M. Ohala, D. Granville, & A. Baily (Eds.), *Proceedings of the 14th international congress of phonetic sciences* (pp. 1437–1440). Berkeley, CA: University of California.

Moreton, Elliott & Joe Pater. 2012. Structure and Substance in Artificial-phonology Learning, Part I: Structure. *Language and Linguistics Compass* 6/11. 686-701.

Morrison, J. S. 2006. LI & L2 Production and Perception of English and Spanish Vowels: a Statistical Model. Doctoral dissertation, University of Alberta.

Moyer, Alene. 1999. Ultimate attainment in L2 phonology: the critical factors of age, motivation, and instruction. *Studies in Second Language Acquisition* 21. 81-108.

Muñoz, Carmen & David Singleton. 2007. Foreign accent in advanced learners. Two successful profiles. *Eurosla Yearbook* 7. 171-190

Munro, Murray J. 1993. Production of English vowels by native speakers of Arabic: Acoustic measurements and accentedness ratings. *Language and Speech* 36. 39-66.

Munro, Murray J, James Emil Flege, & Ian R.A. MacKay. 1996. The effects of age of second language learning on the production of English vowels. *Applied Psycholinguistics* 17. 313-334.

NORM source http://lingtools.uoregon.edu/norm/norm1.php, Language Variation and Computation (LVC) Lab, University of Oregon

Oyama, Susan. 1976. A sensitive period for the acquisition of a nonnative phonological system. *Journal of Psycholinguistic Research* 5. 261-283.

Padgett, Jaye. 2001. Contrast Dispersion and Russian Palatalization, in Hume E, Johnson K, *The role of speech perception in phonology*, Academic Press, p. 187-218.

Padgett, Jaye. 2003. Contrast and post-velar fronting in Russian. *Natural Language & Linguistic Theory* 21. 39-87.

Peterson, Gordon E., & Lehiste, Ilse. 1960. Duration of syllable nuclei in English. *Journal of the Acoustical Society of America,* 32(6). 693-703.

Piske, Thorsten, Ian R.A. MacKay IRA & James E. Flege. 2001. Factors affecting degree of foreign accent in an L2: a review. *Journal of Phonetics* 29. 191-215.

Plonsky, Luke. 2012. Replication, meta-analysis, and generalizability. In G. Porte (Ed.), *Replication research in applied linguistics*, 116–132. New York: Cambridge University Press.

Porte, G. (ed.): *Replication Research in Applied Linguistics*. Cambridge University Press, 2012.

Purcell, Edward T & Richard W. Suter. 1980. Predictors of pronunciation accuracy: A reexamination. *Language Learning* 30. 271-287.

Repp, Bruno H. 1982. Phonetic trading relations and context effects: New evidence for a phonetic mode of perception. *Psychological Bulletin*, 92. 81–110. doi: 10.1037/0033-2909.92.1.81

Romano, Lisa Jayne, Fredericka Bell-Berti F & Eugenia Lorin. 1998. Acoustic Analysis of American English vowels by native speakers of Russian. *The Journal of the Acoustical Society of America* 103(5). 3092-4.

Ščerba, Lev Vladimirovic. 1912. *Russkie glasnye v kačestvennom i količestvennom otnošenii*, St. Petersburg.

Seliger, Herbert W. 1978. Implications of a multiple critical periods hypothesis for second language learning. In W. Ritchie (Ed.), *Second language acquisition research*, 11–19. New York: Academic Press.

Seliger, Herbert W. 1975. Maturational constraints in the acquisition of a second language accent. *Language Sciences* 36. 20-22.

Singleton, David Michael. 1989. *Language acquisition. The age factor*. Clevendon: Multilingual Matters.

Stanley, Joseph A., Margaret E. L. Renwick, William A. Kretzschmar Jr., Rachel M. Olsen, & Michael Olsen. 2018. "The Gazetteer of Southern Vowels." The American Dialect Society Annual Meeting. Salt Lake City, UT.

Suter, Richard W. 1976. Predictors of pronunciation accuracy in second language learning. *Language Learning* 26. 233-253.

Tahta, Sonia, Margaret Wood, & Kate Loewenthal. 1981. Foreign accents: factors relating to transfer of accent from the first language to a second language. *Language and Speech* 24(3). 265-272.

Thomas, Erik R. 2001. *An Acoustic Analysis of Vowel Variation in New World English*. Durham, NC: Duke University Press; 2001

Thomas, Erik R. & Tyler Kendall. 2007. NORM: The vowel normalization and plotting suite. [Online Resource: http://ncslaap.lib.ncsu.edu/tools/norm/]

Thompson, Irene. 1991. Foreign accents revisited: the English pronunciation of Russian immigrants. *Language Learning* 41(2). 177-204.

Thompson, Roger M. 1976. Mexican-American English: Social correlates of regional pronunciation. *American Speech* 50. 18-24.

Timberlake, Alan. 2004. *A Reference Grammar of Russian*. Cambridge University Press.

Tsukada, Kimiko, David Birdsong, Ellen Bialystok, Molly Mack, Hyekyung Sung, & James Flege. 2005. A developmental study of English vowel production and perception by native Korean adults and children. *Journal of Phonetics* 33. 263-290.

Vokic, Gabriela. 2010. L1 allophones in L2 speech production: the case of English learners of Spanish. *Hispania* 93(3). 430-452.

Wolfram, Walt, Philip Carter, & Beckie Moriello. 2004. Emerging Hispanic English: New dialect formation in the American South. *Journal of Sociolinguistics* 8. 339-358.

Yanushevskaya, Irena, & Daniel Bunčić D. 2015. Illustrations of the IPA: Russian. *Journal of the International Phonetic association* 45(2). 221-228.

Ylinen, Sari, Maria Uther, Antti Latvala, Sara Vepsäläinen, Paul Iverson, Reiko Akahane-Yamada, & Risto Näätänen. 2010. Training the brain to weight speech cues differently: A study of Finnish second-language users of English. *Journal of Cognitive Neuroscience* 22. 1319–32.

APPENDIX A: Production tokens excluded from analysis

Table A.1 Production words excluded from analysis.

Speaker	Word #	Target	Reason excluded
2	#57	beat /i/	Misread as "bet"
3	#17	bead /i/	Misread as "bed"
3	#29	bead /i/	Misread as "bed"
4	#17	bead /i/	Misread as "bed"
6	#17	bead /i/	Misread as "bed"
6	#97	hid /ɪ/	Misread as "bid"
7	#29	bead /i/	Misread as "bed"
A	17	bead /i/	Misread as "bed"
A	42	bead /i/	Misread as "bed"
C	17	bead /i/	Misread as "bed"
C	29	bead /i/	Misread as "bed"
C	42	bead /i/	Misread as "bed"
C	3	boot /u/	Participant error (skipped)
E	8	head /ɛ/	Misread as "heed"
E	21	head /ɛ/	Misread as "heed"
E	91	head /ɛ/	Misread as "heed"
F	79	bid /ɪ/	Background noise
F	86	hood /ʊ/	Background noise
F	80	pat /æ/	Background noise
F	83	pet /ɛ/	Background noise
F	39	bed /ɛ/	Misread as "bead"
F	44	bet /ɛ/	Misread as "beat"
F	91	head /ɛ/	Misread as "heed"
G	4	bet /ɛ/	Misread as "beat"
G	14	hoot /u/	Misread as "haught"
G	91	head /ɛ/	Misread as "heed"
G	45	heat /i/	Misread as "hett"
G	95	heat /i/	Misread as "hett"
G	7	spud /ʌ/	Misread as "spood"
G	59	spud /ʌ/	Misread as "sprat"
G	13	spud /ʌ	Misread as "sprat"
H	81	bad /æ/	Participant error (skipped)
H	79	bid /ɪ/	Participant error (skipped)

H	82	foot /ʊ/	Participant error (skipped)
H	78	hit /ɪ/	Participant error (skipped)
H	80	pat /æ/	Participant error (skipped)
J	7	spud /ʌ/	Misread (hesitation)
L	17	bead /i/	Misread as "bed"
L	29	bead /i/	Misread as "bed"
L	42	bead /i/	Misread as "bed"
L	69	heed /i/	Misread as "need"
M	91	head /ɛ/	Misread as "heed"
M	80	pat /æ/	Participant error (skipped)
N	58	hut /ʌ/	Misread as "hoot"
N	62	hut /ʌ/	Misread as "hoot"
N	77	hut /ʌ/	Misread as "hoot"
N	84	pot /a/	Misread as "pot/"
P	17	bead /i/	Misread as "bed"
P	10	booed /u/	Misread (hesitation)
Q	8	head /ɛ/	Misread as "heed"
Q	21	head /ɛ/	Misread as "heed"
R	11	pot /a/	Background noise
R	21	head /ɛ/	Misread as "heed"
R	91	head /ɛ/	Misread as "heed"
S	65	pot /a/	Misread as "spot"
T	91	head /ɛ/	Misread as "heed"
U	26	beat /i/	Background noise

APPENDIX B: Normalized F1,F2, NR talkers

Table B.1 Lobanov normalized F1,F2 values for eight American English vowels, means for individual NR talkers

	i		ɪ		ɛ		æ	
Speaker	**F*1**	**F*2**	**F*1**	**F*2**	**F*1**	**F*2**	**F*1**	**F*2**
A	-0.98	1.40	-1.10	1.20	0.14	0.53	0.27	0.59
B	-1.02	1.62	-1.03	1.05	0.58	0.46	0.71	0.24
C	-1.21	1.52	-0.52	1.16	0.15	0.35	1.52	0.19
D	-0.98	1.52	-0.94	1.51	0.45	0.16	0.53	0.09
E	-1.17	1.92	-0.64	1.08	0.63	-0.02	1.40	-0.35
F	-0.99	1.57	-1.00	1.30	0.59	0.32	0.64	0.32
G	-0.90	1.42	-0.91	1.35	0.78	0.54	0.81	0.41
H	-1.13	1.85	-0.82	1.32	0.91	0.10	1.18	-0.26
J	-1.38	2.09	-0.40	0.74	0.57	0.31	1.30	-0.17
K	-1.20	1.99	-0.70	0.76	0.76	0.19	1.39	-0.01
L	-1.17	1.46	-1.15	1.44	0.32	0.54	0.57	0.36
M	-1.07	1.41	-0.89	1.38	0.74	0.38	1.16	0.36
N	-0.73	1.39	-0.75	1.32	0.98	0.39	1.27	0.40
P	-1.08	1.57	-0.95	1.30	0.57	0.39	1.15	0.10
Q	-0.94	1.33	-0.93	1.38	1.19	0.58	1.35	0.50
R	-1.15	1.48	-1.04	1.31	0.45	0.22	0.77	0.30
S	-1.09	1.83	-0.63	1.17	0.71	-0.19	1.12	0.07
T	-1.15	2.00	-0.60	0.61	0.03	0.38	1.38	-0.17
U	-1.21	1.96	-0.89	1.04	0.56	0.24	1.20	0.18
V	-1.00	1.42	-0.97	1.40	0.59	0.40	0.54	0.37
Mean	-1.073	1.638	-0.838	1.191	0.595	0.314	1.024	0.176
Min	-1.38	1.33	-1.15	0.61	0.03	-0.19	0.27	-0.35
Max	-0.64	2.09	-0.4	1.51	1.19	0.58	1.52	0.59
SD	0.154	0.249	0.207	0.247	0.302	0.197	0.380	0.261
	u		**ʊ**		**ɑ**		**ʌ**	
Speaker	**F*1**	**F*2**	**F*1**	**F*2**	**F*1**	**F*2**	**F*1**	**F*2**
A	-0.70	-1.09	-0.68	-0.82	1.55	-0.75	1.34	-0.83
B	-0.76	-1.17	-0.68	-1.13	0.83	-0.66	1.36	-0.42
C	-0.99	-0.65	-0.83	-0.95	0.57	-1.08	0.93	-0.21
D	-0.73	-1.10	-0.76	-1.17	1.11	-0.54	1.32	-0.48
E	-1.12	-0.54	-0.64	-0.33	1.16	-1.10	0.53	-0.65
F	-0.93	-0.92	-0.57	-0.79	1.29	-0.80	1.09	-0.83
G	-0.84	-1.05	-0.75	-1.02	0.44	-0.87	1.72	-0.72
H	-1.03	-0.86	-0.83	-0.66	0.79	-0.76	0.93	-0.61
J	-1.25	-1.05	-0.49	-0.60	1.15	-0.96	0.56	-0.38
K	-0.99	-0.64	-0.70	-0.67	0.81	-0.93	0.63	-0.69
L	-0.87	-0.67	-0.43	-0.81	1.24	-0.92	1.11	-0.92
M	-0.91	-1.23	-0.94	-1.19	0.97	-0.52	1.09	-0.53
N	-1.01	-1.15	-0.86	-0.90	1.01	-0.89	0.23	-0.85
P	-0.94	-1.02	-0.91	-0.84	1.23	-0.79	0.76	-0.66
Q	-0.90	-1.12	-0.81	-1.06	0.51	-0.80	0.73	-0.70
R	-0.68	-0.98	-0.65	-1.04	1.31	-0.55	1.17	-0.76
S	-1.03	-0.96	-0.64	-0.65	0.88	-0.71	0.76	-0.61
T	-1.07	-0.82	-0.31	-0.31	1.57	-1.24	0.16	-0.43
U	-1.10	-1.04	-0.61	-0.60	1.25	-0.88	0.70	-0.74

V	-0.93	-1.25	-0.89	-1.21	1.36	-0.59	1.30	-0.53
Mean	-0.935	-0.966	-0.695	-0.838	1.003	-0.817	0.929	-0.628
Min	-1.25	-1.25	-0.94	-1.21	0.04	-1.24	0.16	-0.92
Max	-0.68	-0.54	-0.31	-0.31	1.57	-0.52	1.72	-0.21
SD	0.147	0.207	0.162	0.266	0.394	0.194	0.387	0.181

Table B.2 Lobanov normalized F1,F2 values of eight American English vowels, means of all NR talkers combined

Speaker	**Vowel**	**N**	**F*1**	**F*2**
AllSpkrs	i	227	-1.076	1.645
AllSpkrs	ɪ	237	-0.841	1.188
AllSpkrs	ɛ	225	0.580	0.313
AllSpkrs	æ	236	1.012	0.177
AllSpkrs	u	237	-0.938	-0.965
AllSpkrs	ʊ	238	-0.700	-0.837
AllSpkrs	ɑ	237	1.050	-0.818
AllSpkrs	ʌ	233	0.921	-0.624

Table B.3 Lobanov normalized F1,F2 values for eight American English vowels, means for individual NE talkers

	i		ɪ		ɛ		æ	
ID	**F*1**	**F*2**	**F*1**	**F*2**	**F*1**	**F*2**	**F*1**	**F*2**
1	-1.33	2.04	-0.61	0.78	0.63	0.39	1.30	-0.10
2	-1.39	2.09	-0.36	0.70	0.59	0.16	1.41	-0.16
3	-1.58	2.08	-0.55	0.68	0.51	0.24	1.25	0.19
4	-1.24	2.05	-0.59	0.76	0.53	0.28	1.52	-0.19
5	-1.19	1.82	-0.65	0.76	0.56	0.26	1.32	-0.06
6	-1.53	1.97	-0.48	0.84	0.25	0.48	0.81	0.20
7	-1.38	2.22	-0.59	0.64	0.49	0.19	1.54	-0.61
8	-1.42	1.95	-0.66	0.88	0.59	0.15	1.42	-0.19
9	-1.41	1.95	-0.59	0.62	0.36	0.17	1.29	-0.01
Mean	-1.386	2.019	-0.564	0.740	0.501	0.258	1.318	-0.103
Min	-1.58	1.82	-0.66	0.62	0.25	0.15	0.81	-0.61
Max	-1.19	2.22	-0.36	0.88	0.63	0.48	1.54	0.20
SD	0.124	0.113	0.093	0.088	0.122	0.112	0.216	0.241
	u		**ʊ**		**ʌ**		**ɑ**	
ID	**F*1**	**F*2**	**F*1**	**F*2**	**F*1**	**F*2**	**F*1**	**F*2**
1	-1.25	-0.56	-0.47	-0.74	0.70	-0.54	1.02	-1.26
2	-1.48	-0.22	-0.25	-0.34	0.55	-0.44	0.82	-1.62
3	-1.19	-0.14	-0.43	-0.46	0.56	-0.81	1.18	-1.42
4	-1.24	-0.94	-0.54	-0.23	0.23	-0.31	1.23	-1.25
5	-1.06	-0.75	-0.49	-0.16	0.59	-0.37	0.92	-1.49
6	-1.22	-0.85	-0.16	-0.52	0.58	-0.63	1.59	-1.24
7	-1.24	-0.27	-0.40	-0.23	0.53	-0.39	0.93	-1.37
8	-1.13	-0.05	-0.41	-0.67	0.57	-0.85	1.05	-1.21
9	-1.02	0.11	-0.58	-0.36	0.61	-0.98	1.33	-1.50
Mean	-1.203	-0.408	-0.414	-0.412	0.547	-0.591	1.119	-1.373
Min	-1.48	-0.94	-0.58	-0.74	0.23	-0.98	0.82	-1.62
Max	-1.02	0.11	-0.16	-0.16	0.70	-0.31	1.59	-1.21
SD	0.133	0.378	0.134	0.202	0.128	0.240	0.240	0.144

Table B.4 Lobanov normalized F1,F2 values of eight American English vowels, means of all NE talkers combined

Speaker	**Vowel**	**N**	**F*1**	**F*2**
AllSpkrs	i	102	-1.383	2.015
AllSpkrs	ɪ	107	-0.564	0.738
AllSpkrs	ɛ	108	0.5	0.256
AllSpkrs	æ	108	1.317	-0.104
AllSpkrs	u	108	-1.203	-0.407
AllSpkrs	ʊ	108	-0.414	-0.412
AllSpkrs	ɑ	108	1.119	-1.375
AllSpkrs	ʌ	108	0.547	-0.592

APPENDIX C: Individual vowel plots, NR talkers

Figure C.1 Lobanov normalized values of NR speaker means for eight American English vowels, speakers A, B, C, D.

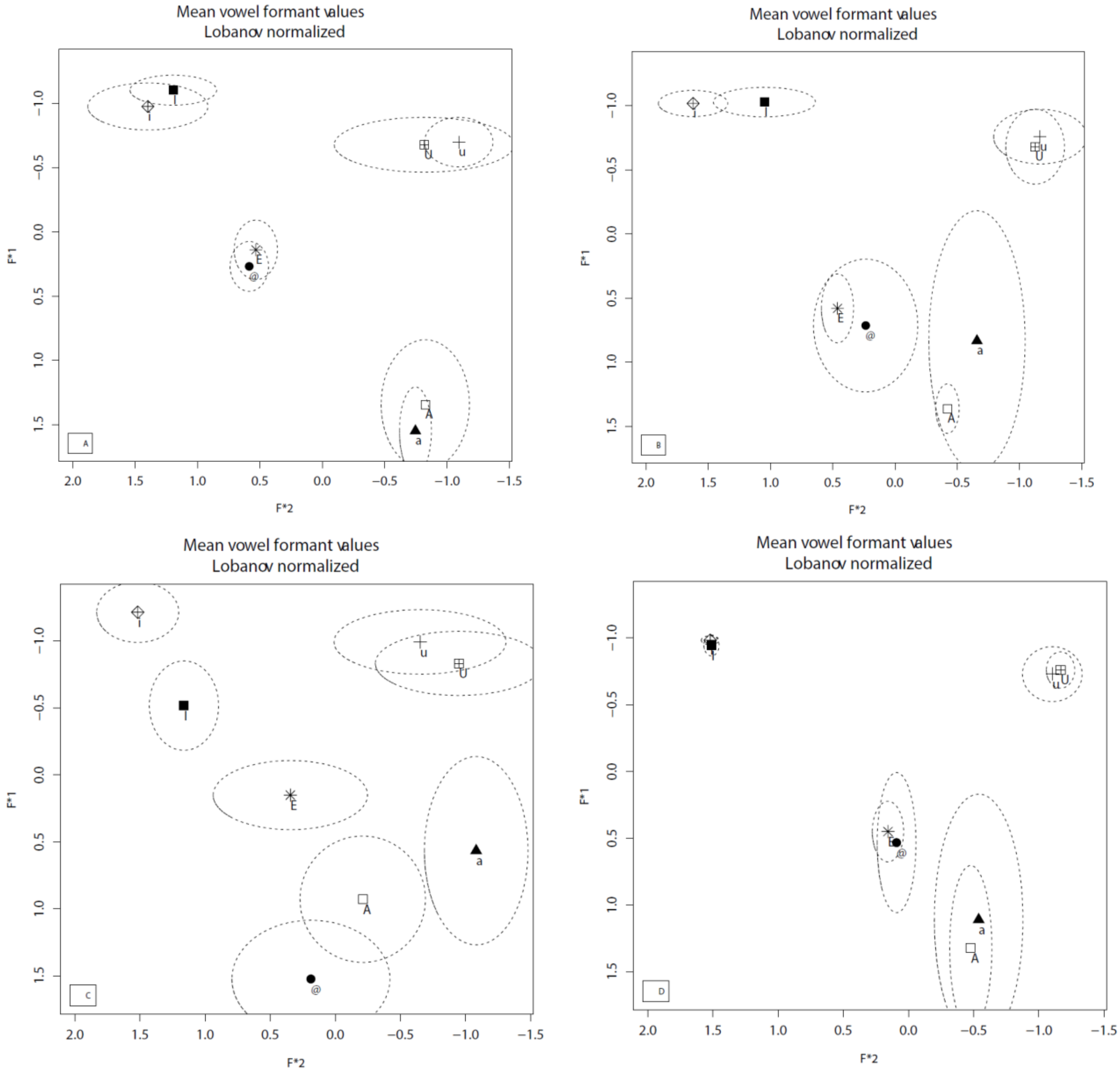


Figure C.2 Lobanov normalized values of NR speaker means for eight American English vowels, speakers E, F, G, H.

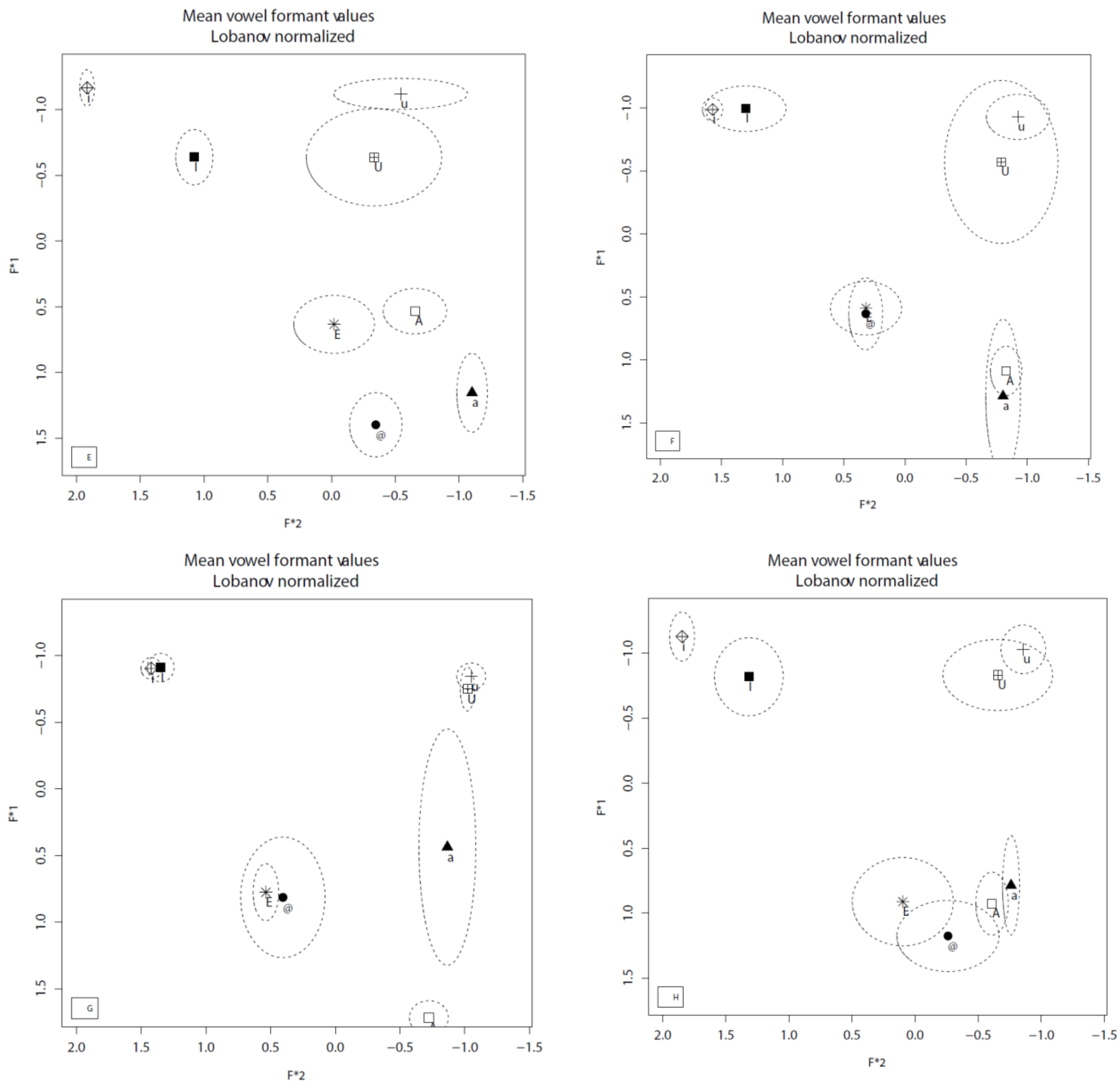


Figure C.3 Lobanov normalized values of NR speaker means for eight American English vowels, speakers J, K, L, M.

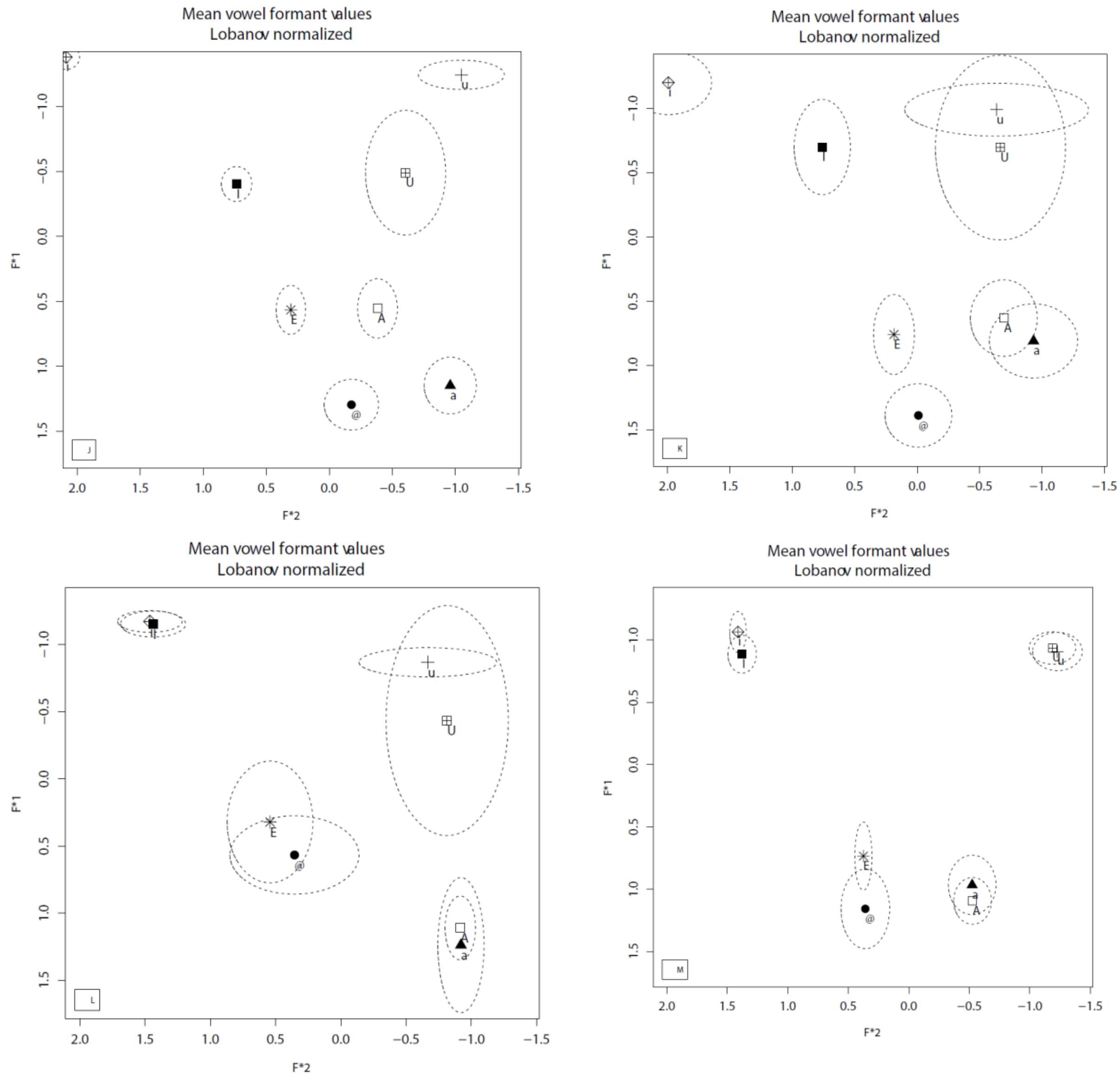


Figure C.4 Lobanov normalized values of NR speaker means for eight American English vowels, speakers N, P, Q, R.

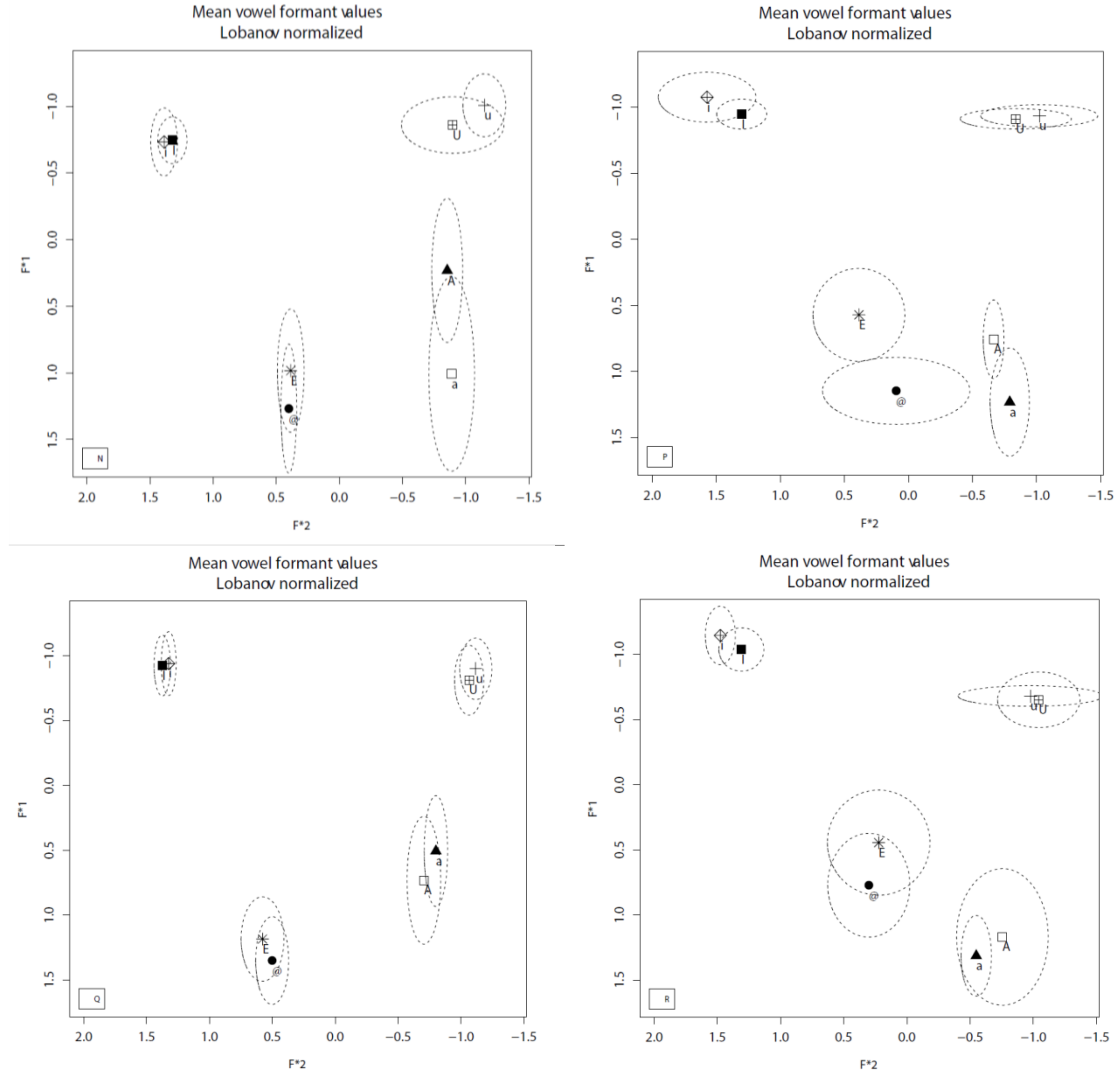
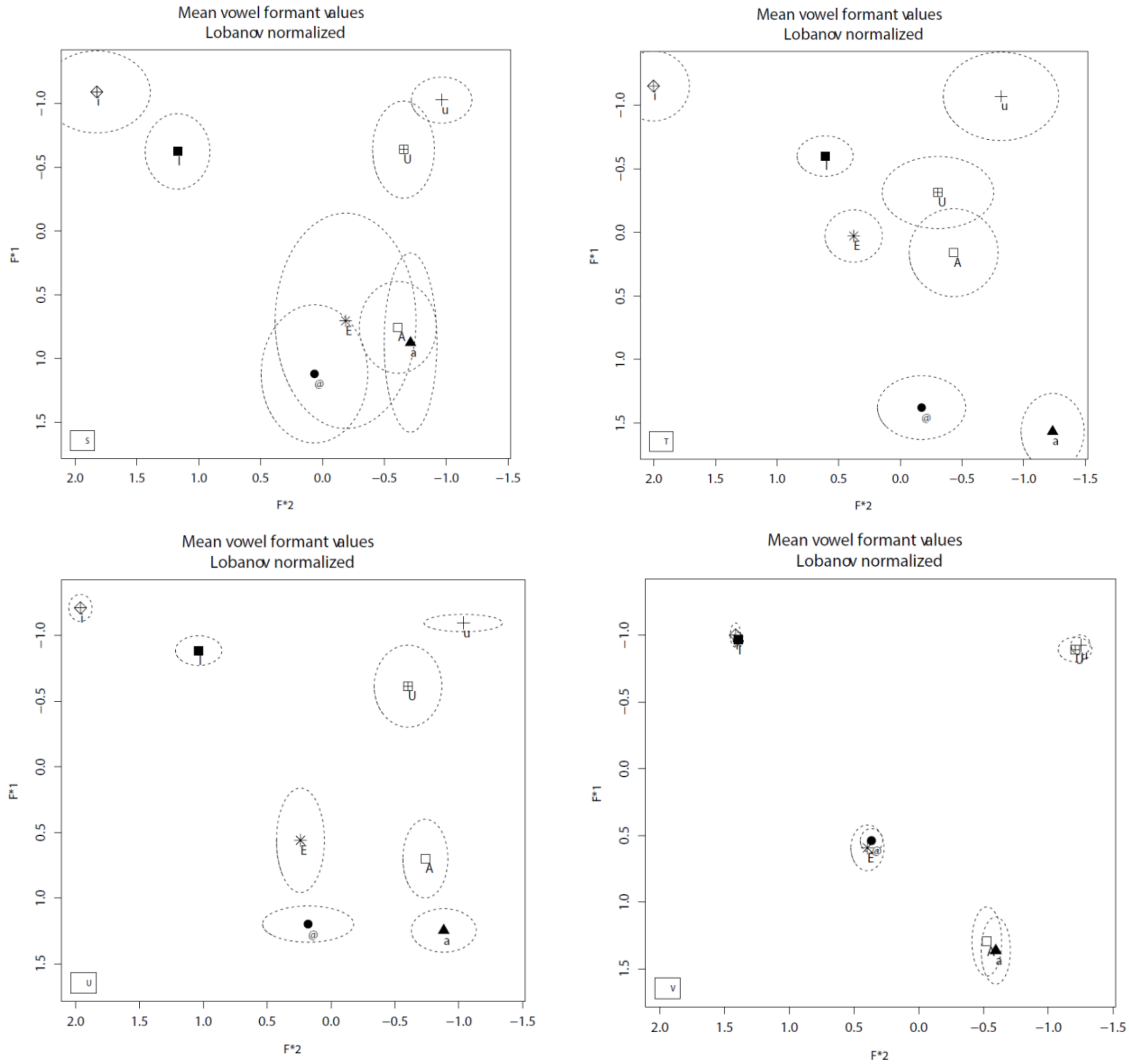


Figure C.5 Lobanov normalized values of NR speaker means for eight American English vowels, speakers S, T, U, V.



APPENDIX D: NR listener accuracy

Table D.1 Accuracy of NR listeners' identification of eight American English vowels in the production of one NE talker

	i	**ɪ**	**ɛ**	**æ**	**u**	**ʊ**	**ʌ**	**ɑ**	**Mean***
A	83.3%	100.0%	58.3%	83.3%	66.7%	73.3%	66.7%	88.9%	77.57%
B	58.3%	83.3%	75.0%	66.7%	66.7%	80.0%	66.7%	66.7%	70.42%
C	25.0%	100.0%	75.0%	58.3%	33.3%	86.7%	77.8%	88.9%	68.13%
D	41.7%	41.7%	25.0%	66.7%	0.0%	60.0%	88.9%	100.0%	52.99%
E	100.0%	75.0%	75.0%	83.3%	100.0%	73.3%	100.0%	88.9%	86.94%
F	50.0%	83.3%	50.0%	91.7%	66.7%	73.3%	88.9%	77.8%	72.71%
G	41.7%	33.3%	66.7%	75.0%	0.0%	93.3%	66.7%	100.0%	59.58%
H	50.0%	58.3%	50.0%	83.3%	66.7%	53.3%	33.3%	77.8%	59.10%
J	100.0%	91.7%	100.0%	100.0%	100.0%	86.7%	100.0%	100.0%	97.29%
K	91.7%	100.0%	75.0%	91.7%	66.7%	46.7%	77.8%	55.6%	75.63%
L	50.0%	41.7%	50.0%	66.7%	66.7%	73.3%	55.6%	55.6%	57.43%
M	66.7%	58.3%	66.7%	83.3%	33.3%	53.3%	44.4%	66.7%	59.10%
N	66.7%	100.0%	66.7%	91.7%	66.7%	80.0%	55.6%	88.9%	77.01%
P	83.3%	66.7%	50.0%	83.3%	33.3%	53.3%	77.8%	100.0%	68.47%
Q	58.3%	50.0%	58.3%	66.7%	100.0%	86.7%	66.7%	88.9%	71.94%
R	66.7%	66.7%	75.0%	41.7%	0.0%	80.0%	77.8%	77.8%	60.69%
S	91.7%	83.3%	58.3%	75.0%	33.3%	73.3%	55.6%	88.9%	69.93%
T	91.7%	83.3%	91.7%	100.0%	100.0%	86.7%	66.7%	88.9%	88.61%
U	25.0%	100.0%	75.0%	66.7%	33.3%	80.0%	66.7%	88.9%	66.94%
V	75.0%	66.7%	66.7%	58.3%	33.3%	93.3%	88.9%	77.8%	70.00%
Mean**	65.83%	74.17%	65.42%	76.67%	53.33%	74.33%	71.11%	83.33%	
Min	25.00%	33.30%	25.00%	41.70%	0.00%	46.70%	33.30%	55.60%	
Max	100.00%	100.00%	100.00%	100.00%	100.00%	93.30%	100.00%	100.00%	
SD	0.234	0.218	0.165	0.152	0.332	0.141	0.174	0.137	

Mean* indicates mean performance across all vowels tested by individual speaker
Mean** indicates mean performance for an individual vowel across all speakers

Table D.2 Accuracy of NR listeners' identification of eight American English vowels in NE speech organized by vowel pairs (pair value is calculated as mean performance on each member of the pair individually – see Table A.8)

	/i-ɪ/	**/ɛ - æ/**	**/u - ʊ/**	**/ʌ-ɑ/**
A	91.7%	70.8%	70.0%	77.8%
B	70.8%	70.8%	73.3%	66.7%
C	62.5%	66.7%	60.0%	83.3%
D	41.7%	45.8%	30.0%	94.4%
E	87.5%	79.2%	86.7%	94.4%
F	66.7%	70.8%	70.0%	83.3%
G	37.5%	70.8%	46.7%	83.3%
H	54.2%	66.7%	60.0%	55.6%
J	95.8%	100.0%	93.3%	100.0%
K	95.8%	83.3%	56.7%	66.7%
L	45.8%	58.3%	70.0%	55.6%
M	62.5%	75.0%	43.3%	55.6%
N	83.3%	79.2%	73.3%	72.2%
P	75.0%	66.7%	43.3%	88.9%
Q	54.2%	62.5%	93.3%	77.8%
R	66.7%	58.3%	40.0%	77.8%
S	87.5%	66.7%	53.3%	72.2%
T	87.5%	95.8%	93.3%	77.8%
U	62.5%	70.8%	56.7%	77.8%
V	70.8%	62.5%	63.3%	83.3%
Mean**	70.0%	71.0%	63.8%	77.2%
Min	37.50%	45.80%	30.00%	55.60%
Max	95.80%	100.00%	93.30%	100.00%
SD	0.179	0.124	0.185	0.127

APPENDIX E: NE listener accuracy

Table E.1 Accuracy of NE listeners' identification of eight American English vowels in the production of NR talkers

	i	**ɪ**	**ɛ**	**æ**	**u**	**ʊ**	**ʌ**	**ɑ**	**Mean***
A	22.2%	66.7%	100.0%	0.0%	77.8%	11.1%	100.0%	22.2%	50.0%
B	88.9%	33.3%	88.9%	11.1%	66.7%	22.2%	55.6%	66.7%	54.2%
C	88.9%	100.0%	100.0%	100.0%	88.9%	22.2%	22.2%	88.9%	76.4%
D	88.9%	55.6%	55.6%	77.8%	100.0%	11.1%	22.2%	88.9%	62.5%
E	100.0%	100.0%	77.8%	100.0%	88.9%	0.0%	66.7%	100.0%	79.2%
F	55.6%	33.3%	88.9%	11.1%	77.8%	0.0%	11.1%	100.0%	47.2%
G	66.7%	44.4%	88.9%	0.0%	77.8%	44.4%	22.2%	88.9%	54.2%
H	22.2%	77.8%	0.0%	11.1%	77.8%	11.1%	0.0%	100.0%	37.5%
J	88.9%	100.0%	66.7%	100.0%	100.0%	0.0%	88.9%	88.9%	79.2%
K	100.0%	100.0%	11.1%	100.0%	33.3%	0.0%	77.8%	22.2%	55.6%
L	88.9%	33.3%	66.7%	0.0%	77.8%	22.2%	11.1%	66.7%	45.8%
M	66.7%	77.8%	22.2%	77.8%	44.4%	0.0%	55.6%	11.1%	44.4%
N	66.7%	77.8%	11.1%	88.9%	88.9%	22.2%	88.9%	22.2%	58.3%
P	88.9%	77.8%	100.0%	33.3%	77.8%	11.1%	66.7%	77.8%	66.7%
Q	88.9%	11.1%	55.6%	66.7%	77.8%	11.1%	22.2%	100.0%	54.2%
R	88.9%	22.2%	88.9%	22.2%	44.4%	11.1%	33.3%	55.6%	45.8%
S	77.8%	77.8%	66.7%	77.8%	77.8%	44.4%	66.7%	66.7%	69.4%
T	100.0%	100.0%	100.0%	100.0%	66.7%	66.7%	100.0%	88.9%	90.3%
U	100.0%	100.0%	0.0%	100.0%	100.0%	44.4%	100.0%	55.6%	75.0%
V	66.7%	22.2%	77.8%	0.0%	77.8%	22.2%	55.6%	44.4%	45.8%
Mean**	77.8%	65.6%	63.3%	53.9%	76.1%	18.9%	53.3%	67.8%	
Min	22.20%	11.10%	0.00%	0.00%	33.30%	0.00%	0.00%	11.10%	
Max	100%	100%	100%	100%	100%	66.70%	100%	100%	
SD	0.231	0.310	0.353	0.425	0.181	0.184	0.332	0.297	
Dist.	**0.0008***	**0.0153***	**0.0054***	**0.0011***	**0.0098***	**0.0058***	0.0957	**0.0151***	

Mean* indicates mean performance across all vowels tested by individual speaker
Mean** indicates mean performance for an individual vowel across all speakers

Table E.2 Accuracy of NE listeners' identification of eight American English vowels in NR speech organized by vowel pairs (pair value is calculated as mean performance on each member of the pair individually – see Table A.10)

	/i-ɪ/	**/ɛ - æ/**	**/u - ʊ/**	**/ʌ-ɑ/**
A	44.4%	50.0%	44.4%	61.1%
B	61.1%	50.0%	44.4%	61.1%
C	94.4%	100.0%	55.6%	55.6%
D	72.2%	66.7%	55.6%	55.6%
E	100.0%	88.9%	44.4%	83.3%
F	44.4%	50.0%	38.9%	55.6%
G	55.6%	44.4%	61.1%	55.6%
H	50.0%	5.6%	44.4%	50.0%
J	94.4%	83.3%	50.0%	88.9%
K	100.0%	55.6%	16.7%	50.0%
L	61.1%	33.3%	50.0%	38.9%
M	72.2%	50.0%	22.2%	33.3%
N	72.2%	50.0%	55.6%	55.6%
P	83.3%	66.7%	44.4%	72.2%
Q	50.0%	61.1%	44.4%	61.1%
R	55.6%	55.6%	27.8%	44.4%
S	77.8%	72.2%	61.1%	66.7%
T	100.0%	100.0%	66.7%	94.4%
U	100.0%	50.0%	72.2%	77.8%
V	44.4%	38.9%	50.0%	50.0%
Mean**	71.7%	58.6%	47.5%	60.6%
Min	44.40%	5.60%	16.70%	33.30%
Max	100%	100%	72.20%	94.40%
SD	0.209	0.227	0.139	0.160