THE ROLE OF SECONDARY-STRESSED AND UNSTRESSED-UNREDUCED
SYLLABLES IN WORD RECOGNITION: ACOUSTIC AND PERCEPTUAL STUDIES
WITH RUSSIAN LEARNERS OF ENGLISH

Elina Banzina

A Dissertation

Submitted to the Graduate College of Bowling Green
State University in partial fulfillment of
the requirements for the degree of

DOCTOR OF PHILOSOPHY

August 2012

Committee:

Laura Dilley, Advisor

Lynne Hewitt, Chair

Sheri Wells-Jensen,
Graduate Faculty Representative

Alexander Goberman

John Folkins

© 2012

Elina Banzina

All Rights Reserved

ABSTRACT

Laura Dilley, Advisor

Identifying those phonological factors that native listeners rely on most when perceiving non-native speech is critical for setting priorities in pronunciation instruction. The importance of accurate lexical stress production, particularly primary stress, has been explored. However, little is known about the role of Secondary-stressed (SS) syllables and Unstressed-unreduced (UU) syllables, and the importance of their accuracy for speech perception. These questions are of relevance for Russian learners of English, who often reduce English SS and UU vowels—a phenomenon which is arguably due to the fact that only one stressed syllable per word is allowed in Russian phonology. Moreover, second language research has not addressed the issue of vowel over-reduction, which is a pattern typical of Russian learners. Low-accuracy productions of SS and UU syllables are generally not expected to lead to unintelligibility; however, they might interfere with the ease and accuracy with which speech is perceived. An acoustic study first compared realization of SS and UU syllables in words produced in isolation by six Russian learners of English and six native English speakers. Words were selected to contain low vowels and specific UU and SS syllable positions to optimally reflect vowel reduction by Russian speakers. Acoustic analyses revealed significant vowel quality and duration reductions in Russian-spoken SS and UU vowels, which were only half the duration of native English productions and significantly centralized. A subsequent psycholinguistic perceptual study investigated the degree of interference that inaccurate productions of SS and UU syllables have on native listeners' speech processing. A cross-modal phonological priming technique combined

with a lexical decision task assessed speech processing of 28 native English speakers as they listened to (1) native English speech, (2) unmodified Russian speech, and (3) modified Russian speech with SS and UU syllables altered to match native productions. Unmodified UU vowels led to significant inhibition of lexical access, while unmodified SS vowels revealed less of such interference. Acoustically “improving” vowel quality and duration in UU and SS syllables greatly facilitated word recognition only for UU-syllable-containing words. A recommendation is made that UU syllables are incorporated into pronunciation instruction for Russian learners of English.

ACKNOWLEDGEMENTS

I would like to thank my advisor Dr. Laura Dilley for her valuable guidance throughout my dissertation writing process, and chair and co-advisor Dr. Lynne Hewitt for all the professional support I received. I am also grateful to my committee members, Dr. Sheri Wells-Jensen, Dr. Alexander Goberman and Dr. John Folkins, whose help and constructive input helped shape this research.

TABLE OF CONTENTS

	Page
INTRODUCTION	1
Significance of the Study	1
Dissertation Outline	7
CHAPTER 1. LITERATURE REVIEW	10
Theoretical Accounts of Speech Rhythm and Lexical Stress Degrees	11
Rhythm Classes and Rhythmic Classification of English and Russian	11
Lexical Stress in English Phonology	16
Lexical Stress in Russian Phonology	23
Research on Lexical Stress in Native-speaker Perception/Production	30
Production of Lexical Stress Degrees by Native Speakers of English	30
Perception of Lexical Stress Degrees by Native Speakers of English	35
Second Language Phonological Acquisition	40
Theories of L2 Phonological Acquisition	41
Perceptual Assimilation Model	41
Speech Learning Model and Feature Hypothesis	44
Markedness Differential Hypothesis	47
Application of Theoretical Models to the Current Study	49
Research on Non-native Production and Perception of Suprasegmentals	55
Stress Correlates	56
Vowel Reduction	58
Lexical Stress Degrees	59

Suprasegmentals and Perception of Non-native Speech.....	60
Dimensions of Non-native Speech Processing	61
Suprasegmentals and Non-native Speech Processing.....	65
CHAPTER 2. ACOUSTIC STUDY.....	72
Methods	72
Design	73
Participants	73
Materials	74
Procedure	77
Acoustic Analysis	79
Results and Discussion	86
Duration	86
Raw Duration	86
Rate-normalized Duration	90
Intensity	92
Raw Intensity	93
Normalized Intensity	95
Fundamental Frequency	98
Raw Fundamental Frequency	98
Normalized Fundamental Frequency	100
F1 and F2 Frequencies	102
First Formant (F1)	103
Second Formant (F1)	106

General Discussion and Conclusions	109
CHAPTER 3. PERCEPTUAL STUDY	116
Methods	117
Design	117
Participants	119
Stimulus Materials	120
Procedure	126
Results and Discussion	127
Secondary-stressed Syllables	128
Unstressed-unreduced Syllables	131
General Discussion and Conclusions.....	134
CHAPTER 4. SUMMARY AND CONCLUSIONS	140
REFERENCES	152
APPENDIX A. PRIMING STUDY STIMULI	165
APPENDIX B. RUSSIAN SPEAKER DEMOGRAPHIC FORM	167
APPENDIX C. NATIVE ENGLISH SPEAKER DEMOGRAPHIC FORM	168
APPENDIX D. PERCEPTUAL STUDY DEMOGRAPHIC FORM	169
APPENDIX E. HSRB APPROVAL AND CONSENT FORMS	170

LIST OF FIGURES/TABLES

Figure/Table	Page
1 Degrees of prominence in English (Ladefoged)	19
2 English language stress system (Beckman & Edwards)	21
3 Russian vowel inventory in stressed and pre-stress syllables (Crosswhite)	24
4 Vowel space shrinkage in Russian unstressed syllables (Padgett & Tabain)	25
5 Russian stress pattern within a lexical word	26
6 Russian participant demographics	74
7 Experimental stimuli for the acoustic analysis	76
8 Spectrogram illustrating vowel boundaries	80
9 Spectrogram illustrating voiced stop consonant segmentation from vowels	81
10 Spectrogram illustrating aspirated voiceless stop and vowel boundaries (1)	82
11 Spectrogram illustrating aspirated voiceless stop and vowel boundaries (2)	82
12 Spectrogram illustrating approximant and vowel boundaries	83
13 SS, UU and PS syllable raw durations for English and Russian speakers	88
14 Spectrogram of Russian-spoken and English-spoken SS syllables	88
15 Spectrogram of Russian-spoken and English-spoken UU syllables	89
16 SS, UU and PS syllable raw intensity for English and Russian speakers	94
17 SS, UU and PS syllable normalized intensity	96
18 SS, UU and PS syllable raw F_0 for English and Russian speakers	99
19 SS, UU and PS syllable normalized F_0 for English and Russian speakers	101
20 Low front vowel /æ/ F_1 and F_2 values for English and Russian speakers	103

21	Low back vowel /ɑ/ F1 and F2 values for English and Russian speakers	103
22	Low vowel /ɑ/ and /æ/ F1 for Russian and English speakers	105
23	F1 in SS and UU syllables for Russian and English speakers	105
24	F2 for low back vowel /ɑ/ for Russian and English speakers	107
25	F2 for low front vowel /æ/ for Russian and English speakers	108
26	Priming effect for words with Secondary-stressed syllables	128
27	Priming effect for words with Unstressed-unreduced syllables	132

INTRODUCTION

Significance of the Study

Native English speakers in the United States are exposed to a large number of non-native English speakers and various foreign accents on a daily basis. It is relatively unclear how such accented speech and various speech errors affect native speaker comprehension, as well as precisely which phonological elements affect intelligibility most. A clear understanding of those phonological features that are most critical for successful speech perception by native English listeners is imperative for ensuring effective and informed pronunciation instruction to non-native speakers of English. Current pronunciation research and pedagogy is based on the “intelligibility principle” (Levis, 2005), which holds that different phonological features have different effects on native listener speech comprehension, and some matter more for overall intelligibility. As a result, “instruction should focus on those features that are most helpful for understanding and should deemphasize those that are relatively unhelpful” (Levis, 2005, p. 370) in order for such instruction to be effective, time-efficient and based on real-life needs. In order to set priorities in teaching pronunciation in accordance with the intelligibility principle, the first step is identifying those phonological factors that native listeners rely on most when perceiving non-native speech, and those that have the most adverse effects on on-line speech processing and intelligibility (Morley, 1991). Derwing and Munro (2005) call for considerably more empirical, replicable research to understand the nature of foreign accents and the way they affect communication. They point out that “teachers are often left to rely on their own intuitions with little direction” (p. 379) and that “reliance on anecdotal evidence and personal impressions in

language pedagogy (...) cannot resolve many of the critical questions that face classroom instructors” (p. 380)—a sentiment shared by Levis (2005) as well.

The importance of accurate production of lexical prominence such as main word stress has been well documented in second language research literature (Field, 2005; Small, Simon, & Goldberg, 1988). However, there are underexplored aspects of lexical prominence whose role for speaker intelligibility is unknown. Secondary-stressed (SS) and unstressed unreduced (UU) syllables (e.g., the first syllable in the word “optimistic” and the second syllable in “robot”, respectively) would be one such unexplored realm of second language phonology. Along with primary-stressed (PS) and reduced vowels, SS and UU syllables collectively form the four levels of lexical prominence in English, creating the typical English language rhythm with strong and weak syllable alternations (Beckman, 1986; Beckman & Edwards, 1994; Fear, Cutler, & Butterfield, 1995; Ladefoged, 2005). Compared to PS and reduced vowels, the role and functions of SS and UU syllable types have been underexplored, which is surprising, given the fact that 41% of all English words contain at least one SS syllable (Mattys, 2000). Contributing to the general understanding regarding the relative importance of SS and UU syllables for word recognition is thus the goal of this study, the findings of which will be informative for pronunciation instruction for certain non-native speaker groups.

Additionally, while there have been a number of studies in second language literature on insufficient vowel reduction by non-native speakers of English (Braun, Lemhöfer, & Mani, 2011; Flege & Bohn, 1989), there are practically no widely known studies on just the opposite phenomenon—vowel *over-reduction*. Vowel over-reduction is not representative of those speaker groups that have been most widely researched in the second language literature; however, such patterns do exist, the most notable example being Russian learners of English.

This second language learner group has been under-researched despite the great numbers of these speakers in the United States. Russian speakers have been often observed to excessively reduce vowels in English, specifically, vowels in SS and UU syllables. Thus, exploration of two phenomena can be done simultaneously: first, the unique pattern of vowel reduction exhibited by Russian learners of English that has not been researched before, and second, the consequences of inappropriate vowel reduction in UU and SS syllables, which would then be indicative of the importance of these syllable types in native speaker perception.

Non-native English speech production often differs from native-like performance both in terms of segmental and suprasegmental aspects. Trubetzkoy (1939/1969) was the first to suggest an existence of a “phonological filter” that makes the second language learner perceive non-native sounds and features according to native language categories, and ever since this view has been widely explored and accepted in second language research. The first language “filter”, or the interference that one’s native language exerts on another language acquired later, could account for the accented Russian production of lexical stress and rhythm in English, and more specifically, the reduction of full vowels in SS and UU syllable positions. The Russian language, unlike the English language, permits only one stressed syllable per lexical word—all the other syllables undergo various degrees of vowel reduction. As a result, SS syllables do not exist in Russian (long compound words and certain loan words are the only exception), and neither do syllables with UU vowels (Avanesov, 1956). The absence of these syllables in the Russian phonological system may have a hindering effect on the acquisition of English SS and UU syllable categories, typically resulting in their temporal and vowel quality reduction by Russian speakers of English. Such inappropriate reductions may be a considerable contributor to the perceived accentedness of Russian speakers. However, what is more serious is that SS and UU

syllable reductions might interfere with the ease and accuracy with which a message is perceived by placing additional demand on listeners' cognitive processing capacity. It is conceivable that low accuracy of production of SS and UU syllables could even lead to utter unintelligibility of speakers' utterances where word recognition is completely hindered. This could be especially true regarding SS syllables that are prosodically prominent due to being stressed, and thus potentially more critical for listeners. Second language research studies have suggested that suprasegmental properties—stress, rhythm, timing, intonation, phrasing—generally are even more critical for intelligibility than segmental properties, since suprasegmental deviations from native-like production have more detrimental effects on native comprehension than segmental deviations (Anderson-Hsieh, Johnson, & Koehler, 1992; Anderson-Hsieh & Koehler, 1988; Derwing, Munro, & Wiebe, 1998), even more so in less optimal listening conditions (Mattys, Brooks, & Cooke, 2009; Wilson & Spaulding, 2010). Whether that means that UU syllable accuracy is less critical than SS syllable accuracy is unclear. It is also not clear whether low accuracy of SS and UU syllables in speech would seriously undermine the comprehension of native listeners, or perhaps on-line speech processing simply gets more effortful yet still accurate. Alternatively, mispronunciation of SS and UU syllables might not affect native perception at all. Overall, the individual contribution of UU and SS syllables towards intelligibility so far has been seriously underexplored.

Another gap in second language acquisition research concerns the consequences of improper SS and UU *vowel reduction*. Most acoustic studies on rhythm acquisition have focused on speakers that struggle with exactly the opposite—not sufficiently reducing unstressed syllables, but none has explored the acoustics or perceptual effects of excessive vowel reductions. A recent study by Braun, Lemhöfer and Mani (2011) was the only one that

investigated the on-line processing of words with insufficiently reduced vowels, which is a measure of speech processing used also in the current study.

As earlier pointed out by Derwing and Munro (2005), instructors working with Russian learners of English as a second language (L2) so far have had to rely on their own intuitions of the importance of addressing SS and UU vowels in pronunciation instruction. Moreover, many Russian learners of English are proficient users of English language dictionaries that typically mark SS syllables and possibly aware of the existence of SS in multisyllabic words; however, if that is the case, such knowledge is not always reflected in their speech, most likely due to no clear understanding of how important SS and UU vowels are. The goal of this study is therefore to clarify the relevance of these syllables for speech processing and lexical access, so that instructors can make informed decisions about whether to include them in the pronunciation instruction or not, and whether they are worth attending to.

It should be noted, however, that obtaining answers to these questions is not going to be equally critical for all pronunciation instructors, as reduction of certain full vowels is a feature that might characterize the speech of the Russian language speakers, and, possibly, speakers of other Slavic languages. The Russian speaker population in the USA is large, and more importantly, the overall number of international students pursuing higher education in the U.S. is growing (data from IIE Annual Report, 2011), many of them from Russia. The findings of this study might be especially relevant for teaching oral proficiency courses to Russian-speaking International Teaching Assistants (ITAs). Having proper speech patterns might be especially critical working with undergraduate native speakers, who often have had no exposure to nonnative speech and who struggle with the cognitively demanding context in academia as it is. In fact, universities have to follow state legislation that requires teaching assistants to meet

certain oral proficiency standards in English before they are allowed to teach. The legislation mandates that they attend oral proficiency classes specifically created for ITAs. In accordance with the Revised Code 3345.281 of the state of Ohio, where the current study was conducted, a university board of trustees “shall establish a program to assess the oral English language proficiency of all teaching assistants providing classroom instruction to students and shall ensure that teaching assistants who are not orally proficient in the English language attain such proficiency prior to providing classroom instruction to students.” What areas exactly the oral proficiency standards and pronunciation instruction should include has yet to be determined; the task is not easy because pronunciation goals will be different for various speaker groups. Specifically for Russian speakers, it would be important to determine whether SS and UU syllable inclusion among pronunciation teaching goals is critical, so that instructors of ITA oral proficiency courses are in touch with the specific needs of Russian speaker population and can increase Russian L2 learners’ intelligibility if the findings of this study confirm the adverse effects of improperly produced SS and UU syllables on speech processing.

The findings of this study are not limited to second language teaching and research alone. Rather, they will simultaneously contribute to the body of psycholinguistic research regarding the status and functions of UU and SS syllables for word recognition in English, and the structure and functions of stressed syllables in general. Surprisingly, despite the substantial body of research on suprasegmentals and their role in word recognition, the focus of research has traditionally been on PS and reduced syllables. The status and role of SS and UU syllables has been explored in an extremely limited number of research studies (Fear, Cutler, & Butterfield, 1995; Mattys, 2000). Stressed syllables are considered to have important prosodic functions and considered to be critical for word recognition—but does this apply only to PS syllables, or could

it be extended to SS syllables as well? Are stressed syllables really more critical for word recognition than unstressed ones? For example, do deviations in the vowel /æ/ that occurs in an unstressed position (e.g., second syllable in “abstract”) have a less detrimental effect on perception than deviations in a vowel /æ/ that is in a stressed position (first syllable in “graduation”)? These findings will provide deeper insights into the intricate functions of lexical stress levels for speech processing in English, and will help facilitate a better understanding of the stressed syllable category as such.

Dissertation Outline

The study has two main goals. The first goal is to identify acoustic differences between native English speaker and Russian learner of English productions of SS and UU syllables, and provide evidence for or against the existence of vowel over-reductions in Russian L2 learner population. Contingent upon the findings in the acoustic study, the second goal will be to examine whether improper pronunciation of SS and UU vowels can affect native listener’s real-time speech processing and word recognition performance. This will carry implications for second language pronunciation teaching and learning, and will allow making certain conclusions regarding the overall status and importance of SS and UU vowels for speech perception from a psycholinguistic perspective.

More specifically, the study sought to answer the following research questions:

- (i) What are the acoustic differences between English and Russian productions of (1) SS syllables, and (2) UU syllables in English?
- (ii) Does improper SS and UU vowel use in English by Russian speakers of English affect the ease of speech processing by native speakers of English, and to what degree?

(iii) Do deviations in vowels that are considered to have important prosodic functions (SS), and those that do not (UU), have different effects on speech processing by native listeners?

This dissertation consists of an acoustic analysis of native English and Russian-accented speech, and a perceptual experiment, which explores native speaker perception of non-native speech with UU and SS syllables properly and improperly produced. The acoustic study will assess vowel quality, duration, intensity and F_0 differences between SS, UU as well as PS syllables as produced by native and non-native speakers, and these acoustic differences will further be used for creating stimuli for the subsequent perceptual experiments. The perceptual experiment uses a cross-modal phonological priming method and a lexical decision task to assess speech processing speed when perceiving native and non-native productions with a goal to determine which syllables are most critical for American listeners when perceiving Russian-accented speech.

The dissertation is divided into five chapters. Chapter 1 reviews theoretical accounts and acoustic evidence of English and Russian language rhythmic systems and lexical prominence degrees in these languages. It also provides an overview of second language phonological acquisition theories with respect to lexical stress acquisition in particular, and examines the available body of research on the effects of lexical prominence degrees on intelligibility and speech processing. Chapter 2 describes methods used to obtain acoustic data of native English and native Russian speaker production of SS and UU syllables in English, and discusses findings of the acoustic study. It also prepares the reader for the subsequent priming study that is based on the findings and speech stimuli of the acoustic study. Chapter 3 describes the perceptual study design, stimuli creation and procedure, and discusses the results of the phonological priming experiment. Finally, Chapter 4 summarizes the findings of both studies and discusses them with

respect to second language and psychological research, draws implications for second language pronunciation teaching and pronunciation goal selection, recognizes limitations of the study, and identifies further avenues for research.

CHAPTER 1. LITERATURE REVIEW

A significant body of research has shown that adult learners of a second language perceive and produce second language phonological features differently than native speakers of a language (e.g., Flege & Bohn, 1989; Flege, MacKay, & Meador, 1999; Kondaurova & Francis, 2008; McAllister, Flege, & Piske, 2002). The differences in perception and production apply not only to segmental aspects of languages but also to suprasegmental aspects, which are defined as those phonological features that are not confined to any one segment but extend over syllables, words and phrases (Baker, Baese-Berk, Bonnasse-Gahot, Kim, Van Engen, & Bradlow, 2011; Lehiste & Fox, 1992; Tajima, Port, & Dalby, 1997; Trofimovich & Baker, 2006; Yu & Andruski, 2010; Zhang, Nissen, & Francis, 2008).

This study specifically examines differences in production of linguistic stress and rhythm between native and non-native speakers of English. Stress has been defined as a suprasegmental feature of utterances that applies to whole syllables which makes them more prominent in the flow of speech (Ladefoged, 2005). From a physiological standpoint, a syllable is stressed when it is produced with greater effort of muscles involved in respiration, relative to other syllables (Lehiste, 1970). Stress is part of rhythm, which is characterized as “patterns of timing and accentuation that characterize the flow of syllables in sentences” (Patel, 2008, p. 97).

Understanding native and non-native differences in stress realization entails knowledge of the structure and workings of the native and target language suprasegmental systems, and rhythm in particular. The first part of Chapter 1 discusses rhythmic patterns of the English and Russian languages with regard to rhythmic categories that have been long established in phonetic research. Next, the types of lexical stress degrees that contribute to the rhythm of each language

are described from an acoustic and perceptual standpoint, setting the stage for acoustic studies on Russian and English productions and further literature on cross-linguistic issues. The second part of Chapter 1 discusses the theoretical framework and empirical research of second language phonological acquisition with a focus on suprasegmentals to account for differences observed in the English speech of Russian language learners. Finally, research on the importance of suprasegmental accuracy for native speaker perception and word recognition is reviewed, making specific predictions for the outcomes of the priming experiment.

Theoretical Accounts of Speech Rhythm and Lexical Stress Degrees

Rhythm Classes and Rhythmic Classification of English and Russian. Every language has different patterns of timing, accentuation, and grouping, collectively called the rhythm (Patel, 2008). Linguists have been long trying to capture differences between rhythmic patterns of languages and categorize them according to their rhythmic properties, giving rise to various accounts and theories. Originally, Pike (1945) and Abercrombie (1967) suggested a basis for a typological classification of language rhythm and proposed that all languages can be classified as either *stress*-timed or *syllable*-timed. The cornerstone of this typological dichotomy was the assumed temporal equidistance of interstress and intersyllable intervals, respectively. Syllable-timed languages were hypothesized to exhibit acoustic isochrony at the syllable level, while stress-timed languages were hypothesized to exhibit isochrony at foot level, that is, have equidistant intervals between strong beats, or stressed syllables (Pike, 1945; Abercrombie, 1967). The prototypical stress-timed languages were considered to be English, German, Dutch and Russian; in contrast, Spanish, Japanese, and French were assumed to fall in the syllable-timed language category.

However, empirical studies that applied instrumental measures failed to show that such isochrony exists. Successive syllables in syllable-timed languages varied greatly in duration, and interstress interval durations in both stress-timed and syllable-timed languages were comparable (Dauer, 1983; Roach, 1982). Following the unfruitful attempts to provide convincing acoustic evidence for isochrony, Lehiste (1977) later suggested that isochrony could be a purely perceptual phenomenon that has psychological importance. Support for the perception-based theory of rhythmic isochrony comes from perceptual studies with infants and adults which demonstrated that listeners do make a distinction between the two proposed rhythm classes and perceptually lump together languages that belong to the same class (Nazzi & Ramus, 2003; Ramus, Nespor, & Mehler, 1999). More recently, however, White, Mattys and Wiget (in press) showed that adults could perceptually distinguish between languages that belong to the same rhythm class based on solely temporal information, and proposed that it is durational cues that make it possible to distinguish between languages, not rhythm class.

A different view was represented by Dauer (1983), who offered a phonological account of differences between the posited linguistic rhythm classes. She proposed that the impression of different rhythmic patterns stems not from some intrinsic timing factors, but rather from a complex interaction of phonological features of a language such as syllable structure, vowel reduction and phonetic realization of stress. The syllable structure of stress-timed languages like English and Russian includes considerable variation in syllable length due to the permissible syllable types: complex consonant clusters with several consecutive consonants allowed both in syllable onsets and codas. Syllable-timed languages, on the other hand, feature syllables of mostly simple CV structure, where open syllables are the predominant syllable type. The existence of vowel reduction in stress-timed languages makes stressed syllables maximally

prominent relative to the short unstressed syllables, while in syllable-timed languages there are no reduced variants of vowels, and prominence differences between consecutive vowels are not as great. Finally, stress-timed languages have lexical stress that is realized by a complex interplay of vowel duration, loudness, pitch and vowel quality, while in syllable-timed languages, syllables are often marked with merely a pitch-accent (Dauer, 1983). Overall, this means that some syllables are more prominent than others in stress-timed languages, and all syllables are of more or less the same prominence in syllable-timed languages. Thus, stress-timed languages like English and Russian are considered to have greater differences between strong and weak elements than syllable-timed languages. Moreover, given the fact that not all languages can be classified as easily as the “prototypical” ones (e.g., English or Spanish) and quite many are in fact “mixed” cases (e.g., Polish, Estonian), it has been suggested that a continuum between the strictly “stress-timed” and “syllable-timed” language polarities would be a better representation of the variability of linguistic rhythm rather than a strict typological dichotomy. Thus, a language can be considered to be, for example, “more stress-timed” or “less stress-timed” (Dauer, 1983).

The phonological account has been the basis for devising various rhythm metrics that have attempted to categorize language rhythms based on vocalic and consonantal variability (Dellwo & Wagner, 2003; Grabe & Low, 2002; Ramus, Nespors, & Mehler, 1999). For example, the rhythm metric created by Ramus et al. (1999) showed that proportion of vocalic intervals, consonantal interval durations and variability in vowel durations (measured as standard deviations of vocalic intervals) are those features that can distinguish between language rhythms like English, and Japanese or Spanish. The hypothesis that English belongs to a rhythmic category that is distinct from that of syllable-timed languages was confirmed by using another

rhythm metric—Grabe and Low’s (2002) Normalized Pairwise Variability Index (nPVI). This metric measures variability in *successive* vocalic and consonantal intervals, which allows it to capture the alternating nature of rhythmic patterns by averaging the durational difference between consecutive vocalic or consonantal intervals (Barry et al., 2003). The focus is on the actual distribution of strong and weak intervals, not just their presence. The study found that *sequential* durational variability is greater in stress-based languages than in syllable-based languages, thus emphasizing the alternating nature of strong and weak element contrasts in stress-based languages.

While the rhythm of English has been examined extensively, Russian has received less attention. Abercrombie (1967) was the first to classify the rhythm of Russian, and came to a conclusion that Russian, just like English, can be classified as a stress-based language. Thus English and Russian, while distinct in various linguistic aspects, rhythm-wise have been categorized as very similar. Roach (1982) tested Abercrombie’s discredited isochrony assumptions by applying inter-stress and inter-syllable measures to the rhythm of English and Russian among other languages. Roach concluded that merely measuring time intervals is not sufficient to distinguish linguistic rhythm types, yet the values that he obtained indicated that Russian and English might indeed exhibit features that make them different from syllable-based languages.

A study by Stockmal, Markus and Bond (2005) might have been the first one to use rhythm metric measurements on Russian and Russian productions in Latvian, which is a language previously unclassified according to stress-timed/ syllable-timed rhythm typology. However, the rhythm values obtained for Russian speakers cannot be directly compared to findings of other studies as the different tasks, contexts and materials used in each study can be

strong confounding factors. A similar study with second language learners was conducted by Ordin and Setter (2008), who applied Grabe's PVI rhythm metric to Russian, English, and English spoken by Russian learners of English who were retelling a fable in English and Russian. The applied rhythm metric would indicate placement of English and Russian—relative to each other—on the assumed rhythmic continuum. The obtained PVI values showed that both Russian and English were stress-timed languages, as expected; however, English appeared to be “more stress-based” than Russian and have greater variability in successive vocalic interval durations. This last fact might be indicative of certain within-class rhythmic differences between English and Russian, and Grabe's PVI might be one of the rhythm metrics that could capture these subtleties. Specifically, one main difference between the rhythms of these two languages lies in the *successive* variability between vocalic intervals, or the *distribution* of these intervals. English might have greater successive variability in vocalic interval durations than Russian due to a strong preference for systematically alternating rhythm pattern. Russian phonology, on the other hand, allows long sequences of unstressed syllables and gradually increasing successive syllable durations, features which could acoustically manifest themselves as less dramatic contrasts in successive vocalic interval durations.

Ramus et al. (1999) advanced the hypothesis that rhythm contrasts are accounted for by differences in the variety of syllable structures; however, this could arguably not be the case for distinguishing languages like English and Russian which have rather similar syllable structures with rich consonant clusters and vowel reduction. Also, measuring solely the vocalic and consonantal interval durations might fail to capture the essence of differences between language rhythms of Russian and English, as strong and weak elements in these languages are realized in a similar manner. However, the “consecutive element” dimension to capturing rhythm, made

possible by nPVI metric, might be potentially important for distinguishing certain aspects of language rhythms like English and Russian, where differences arise precisely in the distribution of stressed and unstressed syllables. The phonological reality of these potential differences will be addressed in the following sections, along with a discussion of their effects on second language rhythmic acquisition.

Lexical Stress in English Phonology. English language rhythmic structure is based on the opposition of more prominent and less prominent elements (Dauer, 1983; Hayes, 1984; Ladefoged, 2005; Kager, 2007). The more prominent or salient elements typically are syllables that bear *stress*. From a physiological perspective, stress can be defined as greater muscular effort and respiratory energy relative to neighboring unstressed elements (Ladefoged, 2005) and, indirectly, greater subglottal pressure (Lehiste, 1970). Phonetically and acoustically, there is no one single parameter that would correspond to stress—English employs a combination of various features that allow a certain syllable to achieve prominence. Stressed syllables demonstrate greater intensity, fundamental frequency, and duration relative to surrounding syllables (Flege & Bohn, 1989; Fry, 1955; Lehiste, 1970), which is perceptually manifested as increased loudness and duration, and higher (or sometimes lower) pitch, thus definition of stress differs depending on whether it is viewed from the point of view of the speaker or the listener. Finally, in order to achieve the alternating pattern of English rhythm, some elements have to be of lesser prominence. These are syllables that feature shorter duration, lower intensity, typically lower fundamental frequency, and, frequently, vowel quality reduction relative to the more prominent elements. While stressed syllables always contain a full vowel, unstressed vowels can contain

either full or reduced vowels, thus *vowel quality* is yet another important dimension of lexical prominence.

Metrical phonology, a theory of stress and rhythm, holds that English strives towards a rhythmic alternation of strong and weak syllables, avoiding dysrhythmic situations, known as ‘stress lapse’ and ‘stress clash’. Stress lapse is defined as a situation where several unstressed elements occur right next to each other without a single stressed element in between them, whereas stress clash refers to two stressed elements occurring right next to each other without an intervening unstressed element (Hayes, 1984; Kager, 2007). This means that English language rhythm is based on systematic strong and weak syllable alternations, where multisyllabic words will often receive more than one strong element in order to prevent long sequences of unstressed syllables (Dauer, 1983; Ladefoged, 2005).

As mentioned before, syllable types that create the alternating pattern of English rhythm vary in *prominence*. Prominence is achieved by a combination of stress and vowel quality, thus “prominence” is a more inclusive term than “stress” (Lehiste & Fox, 1992). Generally, four degrees of prominence can be distinguished in English: primary-stressed syllables (e.g., the first syllable in the word “badger”), secondary-stressed syllables (e.g., the second syllable in the word “imagination”), unstressed unreduced syllables (e.g., the second syllable in the word “robot”) and unstressed reduced syllables that contain a “schwa” (e.g., the second syllable in the word “husband”). It is not entirely clear which degrees of prominence form the strong and weak elements, because strong and weak syllables can mean either (i) stressed and unstressed syllables, or (ii) syllables that contain full vowels and reduced vowels. The complex interplay of suprasegmental (duration, fundamental frequency, intensity) and segmental (vowel quality) features in English phonology has generated various definitions of what exactly constitutes a

strong and weak syllable (Fear, Cutler, & Butterfield, 1995). A number of linguists view vowel quality as the main factor; thus, all full vowels are viewed as strong (Beckman & Edwards, 1994; Bolinger, 1981; Norris, McQueen, & Cutler, 1995). Others believe it is the presence of *stress* that makes the difference, thus, according to this definition, all stressed syllables are strong, and all unstressed syllables, regardless of the vowel quality, are weak (Hayes, 1984; Kager, 2007). For the purposes of this study, the latter, stress-based definition was adopted, where the presence of stress makes a syllable strong, and its absence makes it weak.

Traditionally, English is considered to have syllables of four levels of stress: (1) primary-stressed (abbreviated as PS further in the text); (2) secondary-stressed (SS); (3) unstressed unreduced (UU), and (4) unstressed reduced. In general, one syllable in each content word will have one PS syllable, which is the most prominent or salient syllable in a word; it is also the syllable that has a potential to receive an additional phrase or sentence stress, typically in the form of a pitch accent. PS always features a full vowel. In longer multisyllabic words, another stressed syllable will appear in addition to the PS, namely, SS, which is the weaker of the two possible stress degrees. SS typically does not receive phrase stress, but it will always contain a full vowel, just like the PS syllable.

As for unstressed syllables, these are considered to be syllables that never bear stress and can be further divided into reduced and unreduced in terms of vowel quality. Thus lack of stress makes UU vowels different from stressed vowels, and vowel quality make them different from reduced vowels. Unlike in UU syllables, vowels in unstressed *reduced* syllables undergo significant reduction in their quality and become centralized as a result.

As an alternative to the traditional view, Ladefoged (2005) supports a binary account of stress degrees. He proposed that there are more than two degrees of prominence; however, these

variations in prominence between the syllable types “are not all associated with what we want to call stress” (p. 113) but other factors as well, for example, intonation peaks or vowel quality, thus he clearly separates the notion of degrees of stress from that of degrees of prominence.

According to the binary account of stress, English syllables are either stressed or unstressed, but no further distinctions are phonologically justified. The difference in prominence between PS and SS syllables in reality is due to superimposition of the intonation pattern on one of the stressed syllables, or the so-called “tonic stress”. In contexts where there are no intonation effects and no tonic accent needs to be assigned, such as in phrase non-final positions, these two types of syllables are of equal prominence. Thus there is no PS or SS in a traditional sense, but rather just a single stress degree—corresponding to a stressed syllable—either with a tonic accent, or without it.

Furthermore, Ladefoged states that UU vowels and reduced vowels cannot be considered to be two distinct “degrees of stress” either, because vowel quality creates the impression of their being different, not stress. He refers to differences in these syllable types as degrees of prominence, but collectively they represent only one stress level—that of unstressed syllables.

Figure 1 shows the degrees of prominence according to Ladefoged:

(a)

(b)

-	-	+	-	tonic accent
+	-	+	-	stress
+	-	+	-	full vowel

explanation

Figure 1. From Ladefoged, 2005: (a) a schematic representation of degrees of prominence as exemplified by the word “explanation”; in (b), it is converted to a multilevel stress representation.

A somewhat similar view was expressed by Beckman and Edwards (1994), who approached the issue from the perspective of the theory of intonation and prosodic hierarchy. They proposed an account of English stress system hierarchy with four distinct levels, where at the lower level, the contrast is one of vowel quality, and at the higher level, the contrast lies in the intonation pattern. They adopted a vowel-based approach to strong-weak syllable distinction, where all syllables that contain a full vowel are considered to be strong. Thus they posit that at the bottom of the hierarchy there is a weak syllable with a reduced vowel, because any syllable that contains a reduced vowel will be considered less prominent than a full vowel. Full vowels, on the other hand, will always occupy a higher position in the prosodic hierarchy than reduced, which is why UU syllables form the next level of the hierarchy. Phonetically, the difference between reduced and UU vowels might lie not only in the difference of vowel quality, but also accompanying differences in duration and loudness. UU vowels, however, do not receive pitch accents, which in turn makes them lower in hierarchy than the so-called SS vowels. Any syllable that receives pitch accent gives it “absolute prominence” over syllables that have no associated pitch accents, which is why UU vowels and SS vowels represent separate categories. Finally, if the word receives tonic stress (or “nuclear stress”, which is the term that Beckman & Edwards

use), then the syllable that bears it will be on top of the hierarchy. If a word has more than one full syllable, only one is marked in the lexicon as a potential “association site” for a sentence stress. The idea that the so-called primary and secondary stressed syllables are identical in prominence unless the PS one bears a tonic accent resonates with Ladefoged’s (2005) view. Beckman and Edwards’s four-level stress system is illustrated in Figure 2.

Figure 2. English stress system according to Beckman and Edwards (1994).

Lehiste (1970) suggests that word stress is more of an abstract phenomenon—a *potential* for being stressed, or “a capacity of a syllable within a word to receive sentence stress when the word is realized as part of the sentence” (p. 150). Thus word stress only gets realized phonetically when it happens to be stressed at sentence level; the evidence for that would be the fact that not always syllables that are perceived as stressed have an increase in subglottal pressure. According to Lehiste, the knowledge that we possess as language speakers and users tells us which syllables are the ones that have the potential to be stressed even if the stress is not physically present.

Fear, Cutler and Butterfield (1995), however, provided acoustic evidence for the hypothesized levels of prominence and showed that they actually correspond to four distinct degrees of stress. Importantly, this study analyzed words that occurred in phrase non-final

positions and that intonation peaks did not apply to, for example, “*it may be inspiring, but authentic artwork is expensive*” or “*it may be tedious, but authorization is required*”, where words “authentic” in the first sentence and “authorization” in the second sentence were analyzed for UU and SS syllables, respectively. They found convincing acoustic evidence that UU syllables form a stress category on their own, which is an intermediate position between stressed and reduced vowels. The status of PS and SS when there was no tonic stress involved was less clear: they did not differ on two of the stress measures—intensity and fundamental frequency, yet were significantly different in terms of duration. Since both of them contain full vowels, the difference in duration cannot be attributed to purely segmental effects of vowel quality but rather treated as a stress correlate. This last finding thus could in a way support the tonic stress effect that makes one syllable more prominent when the tonic stress is present, but there are no differences in stressed syllable prominence when it is absent, expressed by Ladefoged (2004) and Beckman and Edwards (1994).

The current study focuses on SS and UU syllables. Based on the previously reviewed accounts of English stress system, SS and PS syllables in intonationally unaccented positions are expected to be of the same prominence, except for greater durations for PS syllables. Thus while they might not show great differences phonetically in such positions from a phonological perspective, SS syllables might not be viewed as critical due to their lack of potential to attract tonic stress.

The status of UU vowels has to be viewed relative to SS syllables, as their specifications and functions overlap and vary considerably in various accounts. Ladefoged (2005) claimed that UU vowels have no strong increase in respiratory activity, and in terms of stress, they can be considered just as unstressed as reduced syllables. Beckman and Edwards (1994), on the other

hand, consider UU vowels to share some of the stress correlates with the full and stressed SS and PS vowels, specifically, duration and intensity; thus, PS, SS and UU are all treated as strong syllables. This is still different from other accounts that consider UU vowels, such as the second syllable in “neon” or “icon” to be in fact secondary-stressed (Davenport & Hannahs, 2010). Such accounts are clearly based on only vowel quality considerations, and treat all syllables that do not bear PS but whose vowels are not reduced, as secondary-stressed, ignoring other differences between these two syllable types. Moreover, having two adjacent strong beats would create a stress clash and in general would be in disharmony with the alternating structure of English rhythm.

The study by Fear, Cutler and Butterfield (1995) showed that acoustically, UU syllables form a stress category on their own, which is significantly different from the SS syllable category and also unstressed reduced vowel category. Such a finding challenges both Davenport and Hannahs’s classification of UU syllables as SS, and Ladefoged’s claim that UU vowels exhibit complete lack of stress. Overall, the issue of stress levels in English is extremely complex due to the interplay of segmental and suprasegmental factors, and also due to many phonetic studies exploring stress degrees without controlling for sentence intonation, i.e., pitch accents—a concern raised by Beckman and Edwards (1994).

Lexical Stress in Russian Phonology. English and Russian both are categorized as stress-based languages; however, beyond the surface similarity a large number of differences exist. While Russian does have great durational variability of vocalic intervals (Abercrombie, 1967), lexical stress and vowel reduction just like English, it is nevertheless quite different from English with respect to its lexical stress patterns.

Russian features three degrees of prominence: the main stress, and two levels of unstressed syllables, which represent two distinct levels of vowel reduction—moderate and extreme (Avanesov, 1956; Bondarko, 1998; Crosswhite, 2000; Padgett & Tabain, 2005). Russian exhibits a more complex and abundant pattern of phonological vowel reduction than English, which has only one level of vowel reduction. Russian typically allows only one stressed syllable per multisyllabic word, while all the surrounding syllables remain unstressed and undergo vowel reduction of some degree (Avanesov, 1956). In fact, it is safe to say that a vowel that is not stressed gets reduced to a certain extent.

The three levels of stress in Russian are realized primarily in terms of vowel quality and different degrees of vowel duration. Vowels at each stress level feature different vowel sub-inventories (Avanesov, 1956; Barnes, 2004; Crosswhite, 2000), shown in Figure 3:

Second Pre-stress syllable:	First pre-stress syllable:	Stressed syllable:
i u	i u	i u
ə	a	e o
		a

Figure 3. Russian vowel inventory in stressed and pre-stress syllables (Crosswhite, 2000).

The stressed syllable vowel inventory consists of 5 vowels — [i], [e], [a], [o], [u]¹. In the syllable immediately preceding the stressed syllable (i.e., the “pre-stress” syllable), the 1st

¹ To indicate the low back vowel, some authors use the symbol /a/ (Avanesov, 1956; Padgett & Tabain, 2005), and others use /ɑ/ (Bondarko, 1977; 1998). In the current study, the latter (/ɑ/) is used.

degree, or moderate, vowel reduction occurs and results in a vowel inventory consisting of only [i], [u], or [a]. The 2nd degree, or extreme, vowel reductions occur in all the other syllables before the main stress or after it (in the so-called “post-stress” syllables”), where [a] is further reduced to a schwa (Avanesov, 1956; Bondarko, 1998; Crosswhite, 2000); therefore the most dramatic reduction in vowel quality is seen for the lower vowels /a/ and /o/. Mid vowels do not appear in unstressed syllables (Barnes, 2004). The following dataset (Table 4, from Padgett & Tabain, 2005) demonstrates how vowels change depending on the syllable they occur in; the word root in each case is the same. Since low vowels rise, Padgett and Tabain (2005) describe this process as vowel space floor rising in unstressed syllables:

Table 4.

Illustration of Vowel Space Shrinkage in Unstressed Syllables (Padgett & Tabain, 2005).

[‘d <u>i</u> m] “smoke”	[d <u>i</u> mΛ’voj] (adj.)
[‘s <u>u</u> dnə] “ship”	[s <u>u</u> dΛ’voj] (adj.)
[‘tse <u>h</u>] “(factory) shop”	[tsi <u>h</u> Λ’voj] (adj.)
[‘g <u>o</u> t] “year”	[gə <u>d</u> Λ’voj] “annual”
[‘pr <u>a</u> f] “law”	[prə <u>v</u> Λ’voj] “legal”

Note: In (a), the vowel is located in a stressed syllable; in (b), the underlined vowel occurs in the second pre-stress syllable and illustrates low vowel raising.

An important exception here occurs with vowels before palatalized consonants: the previously mentioned three-way vowel contrast reduces to a two-way contrast [i] and [u] before palatalized consonants in unstressed syllables, such that [a] neutralizes to higher vowels (Avanesov, 1956; Jones & Ward, 1969; Padgett & Tabain, 2005). Another exception is unstressed vowels located in absolute word beginnings or endings, because the lack of a

consonant before and after a vowel in word-initial and word-final positions, respectively, makes vowel reductions less dramatic (Barnes, 2004; Bondarko, 1998).

Besides changes in vowel quality, equally critical vowel reductions take place also in the temporal domain. According to Bondarko (1977, 1998), duration is the main difference between stressed and unstressed vowels in Russian, thus the three levels of stress are realized primarily in terms of different degrees of vowel duration. Specifically, Bondarko claims that the low vowels undergo three levels of temporal reduction, which can be expressed as a ratio of 1:2:3, where 3 is the duration of the stressed syllable (Figure 5). Higher vowels, on the other hand, show only one level of reduction, which is approximately the same in both first and second pre-stress syllables—only a half of the stressed vowel duration (Bondarko, 1977).

Figure 5. Russian stress pattern within a 4-syllable lexical word “водо’качка” (Eng. “waterpump”), illustrating degrees of vowel quality and duration reductions in low vowels, depending on syllable position, where 3—full vowel (stressed), 2—slightly reduced vowel, and 1—greatly reduced vowel.

The observation that vowels undergo reduction in unstressed syllables was supported by an acoustic study by Padgett and Tabain (2005), which found that Russian stressed vowels are indeed significantly longer than unstressed vowels, and that there are significant durational distinctions between low vowels in the first and second pre-stress syllables after non-palatalized

consonants. In fact, most speakers in their study maintained general durational distinctions between vowels in first and second pre-stress syllables as long as they occurred after non-palatalized consonants; after palatalized consonants, no significant durational differences were found. This finding could be explained with vowel height: only high vowels can appear after palatalized consonants, and those vowels are already intrinsically shorter, providing less opportunity for various degrees of reduction (Bondarko, 1998; Padgett & Tabain, 2005).

The degree of reduction in post-stress syllables slightly differs from reduction in pre-stress syllables. Reduction is greater in non-final post-stress syllables, commensurate to duration of second pre-stress syllables, but less pronounced in final syllables, especially if they are open (Bondarko, 1998). According to Avanesov (1956), regardless of this variability, all post-stress syllables undergo a second degree of reduction and feature a three-way vowel contrast: [i], [u] and [ə].

The importance of duration over other stress cues is supported not only by production studies, but also by possibly the only perceptual study available on Russian speaker perception of stress, which was conducted by Eek (1987) on Estonian and Russian. He systematically modified duration, F_0 and intensity, and found that Russian speakers relied on temporal information most when making stress judgments. F_0 was perceived as a weak indicator of lexical prominence—only when two syllables were of equal duration did Russian listeners consider F_0 as a cue to stress.

The lexical stress pattern of Russian suggests that, unlike in English, there are no SS syllables that would create an alternating rhythm pattern in longer words; long sequences of unstressed syllables are allowed in Russian phonology. SS in Russian occurs only in compounds and a few loanwords, but typically multisyllabic words in Russian contain only one stressed

syllable. The absence of SS syllables in Russian might contribute to the Russian language-specific pattern that makes Russian rhythm different from English, and could potentially contribute to the Russian accent. Russian allows only one stressed syllable per word, and all the vowels before and after that stressed syllable undergo reduction (Avanesov, 1956). Therefore, Russian learners of English might reduce SS syllables in English as well, producing long sequences of unstressed syllables in longer words. Thus, if Russian speakers do transfer the Russian rhythm pattern to English, the result will most likely be SS syllable reductions: English SS location within a word coincides with the location of most extreme vowel reductions in Russian. Presumably, Russian speakers might be unable to pronounce words like “confirmation” or “abbreviation” in an English-like manner, which would be secondary-stressing the first syllable in “confirmation”, and the second syllable in “abbrevation”. Instead, the Russian pattern would feature only one prominent syllable, with the surrounding elements more or less reduced, thus being quite far from the alternating rhythm patterns of English.

Upon transferring Russian rhythm patterns to English, Russian learners of English could apply the same pattern of reduction also to another phonological prominence degree in English—unstressed yet full vowels that occur right next to PS syllables, e.g., the second syllable in “robot” or “mascot”. Unlike in English, the idea of an unstressed vowel retaining its full vowel quality clashes with the basic Russian stress rules, and therefore might be violated when producing English UU vowels. In compliance with Russian rhythm rules, “robot” would be produced with the second vowel temporally shorter and more neutralized in terms of vowel quality than in native English productions, which would have the vowel produced with its full vowel quality. All in all, UU syllables present the same type of problem to the Russian speaker as SS syllables, the only difference between the two being the fact that SS involves attending to

phonetic stress correlates such as pitch, loudness and duration, rather than just vowel quality.

These two types of vowels—SS and UU—are the main object of investigation in order to determine their degree of interference with intelligibility of Russian speakers.

Finally, the differences between vowel reduction in English and Russian could also lie in their exact nature. Generally, a distinction is made between two types of vowel reduction—phonological and phonetic. The phonetic one is a gradient articulatory undershoot due to fast speech rates or register; the phonological reduction, on the other hand, refers to neutralization of vowel phoneme contrasts and is categorical (Padgett & Tabain, 2005). While vowel reduction in English is phonological in nature (i.e., it does not depend on speech rates, thus even carefully articulated speech would preserve the reduced “schwa”), in Russian the situation might not be as straightforward.

Typically, the two degrees of vowel reduction in unstressed syllables are viewed as phonological in nature. However, alternative views exist: Barnes (2004) has proposed that the most “extreme” vowel reduction in the second pre-stress syllable is a result of a phonetic undershoot, thus being phonetically conditioned in addition to phonologically. He provided experimental data where all speakers showed a highly significant correlation between vowel duration and vowel height in the second pre-stress syllable, indicating that reduction to “schwa” in second degree reduction environments has a clear gradient character. The conclusion he drew about Russian phonology is that two types of vowel reduction exist in Russian, “but only one is accomplished in the phonology. The other is a gradient process accomplished by the phonetics. /ə/ is not a phonological category in Russian” (Barnes, 2004, p. 9). Since phonology has no role in reduction of unstressed /a, o/ to schwa, according to Barnes, this means that the “schwa” that low vowels approach in their quality in the second pre-stress syllable does not appear in carefully

produced speech, but rather only at faster speech rates. For the purposes of this study, this also suggests that if transfer of native language phonological norms to second language phonology indeed takes place, it might not be possible to observe extreme vowel reductions in secondary-stressed and unstressed-unreduced syllables in carefully and slowly articulated speech of Russian learners of English.

Research on Lexical Stress in Native-speaker Perception/Production

Production of Lexical Stress Degrees by Native Speakers of English. A considerable number of studies have explored the acoustics of stressed versus unstressed syllable production in English by native speakers, without delving into finer lexical stress distinctions like SS syllables or UU syllables (Flege & Bohn, 1989; Fry, 1955; Lieberman, 1960; Zhang, Nissen, & Francis, 2008). These studies have consistently found that greater fundamental frequency, duration and intensity are associated with PS syllables, compared to unstressed. Unstressed reduced vowels, on the other hand, are produced by native speakers with significantly lower frequency, less intensity and shorter duration than PS vowels.

Fry (1955) examined disyllabic words that were noun/verb pairs and differed only in stress placement, e.g., “object” or “permit”. He concluded that both duration and intensity varied depending on whether the syllable was stressed or unstressed; fundamental frequency was not addressed in this study. Lieberman (1960) conducted acoustic measurements on similar noun/verb pairs, and concluded that fundamental frequency was the most relevant stress correlate, followed by amplitude and duration. These findings were supported by Zhang et al. (2008), whose study examined the same type of words. They found that F_0 for stressed syllables was significantly higher than for unstressed (164 Hz and 145 Hz, respectively). Stressed

syllables were also of significantly longer duration than unstressed (329 ms and 250 ms, respectively) and of significantly greater intensity than unstressed (5 dB difference between stressed and unstressed syllables). Such differences between stressed and unstressed syllables were confirmed by Flege and Bohn (1989) and Lee, Guion and Harada (2006); the latter study found the same exact 5 dB intensity difference as Zhang et al., and discovered that native speakers produced unstressed reduced vowels with roughly half the duration of stressed vowels (ratio = .45). Additional support for duration differences comes from Parmenter and Trevino (as cited in Lehist, 1970) who established that in English an average stressed vowel is approximately 50% longer than an average unstressed vowel.

In practically all of these studies, however, word stress is confounded with phrase level prominence. That is, the material used is words produced in isolation, and those receive a tonic stress on the stressed syllable. In such cases, it is unclear whether the effects shown are the effect of accent or the effect of stress (Plag, Kunter, & Schramm, 2011; Braun, Lemhöfer, & Mani, 2011). The other limitation is studies often ignoring the fact that the term “stressed syllable” refers not only to PS but also SS syllables, and including only PS syllables to represent the whole category. Similarly, studies include only reduced vowels in the unstressed syllable category, neglecting unstressed unreduced syllables, and as a result the conclusions drawn are not fully representative of the category. Most of these studies thus focus on the acoustic differences between PS syllables and unstressed-reduced syllables. The production of finer lexical stress degrees, such as SS and UU syllables, however, has been relatively underexplored.

As noted before, there has been lack of consistency in defining such finer stress distinctions, therefore some clarification is needed regarding their status in different studies. Cutler has done extensive work on UU syllables: although in her 2007 study (Cutler, Wales,

Cooper & Janssen, 2007) she refers to the first syllables in “MUsic” and “muSEum” as primary and *secondary* stressed, respectively, these would be considered primary stressed and *unstressed unreduced* syllables according to stress category standards adopted in this paper. Moreover, in an earlier study by the same author (Fear, Cutler, & Butterfield, 1995) the syllable immediately adjacent to the PS syllable, e.g., the first vowel in “auTOmata”, was referred to as UU, and defined as “non-central”, and that “carries neither primary nor secondary stress” (p. 1893). These are some of the very few acoustic studies available on native speaker production of finer stress distinctions than stressed/ unstressed, and might be the only ones that have done acoustic measurements specifically on UU syllables.

Cutler et al. (2007) instrumentally measured UU syllables and PS syllables, which also received phrasal stress and thus an additional tonal accent. Results revealed that UU syllables were significantly shorter, quieter and lower in pitch than PS syllables. The effect size was large only for pitch measures; thus, pitch might be the most reliable distinction between the two stress degrees in an accented position. However, the scope of differences in non-tonal contexts is not known. Next, Fear et al. (1995) acoustically examined all four prominence degrees, with a special focus on UU vowels. The target words they used were strategically placed in phrase-initial positions; thus, unlike in other studies, they were not intonationally accented. The study found that UU vowels were significantly different from the rest of the stress degrees in terms of duration, intensity and vowel quality. That provides evidence that UU vowels form a separate stress category that patterns neither with stressed nor reduced vowels, occupying an intermediate position between SS and reduced vowels in all measures but fundamental frequency. Regarding the last one, UU did significantly differ from PS vowels in terms of pitch, supporting the findings by Cutler et al. even in the absence of the tonic stress on the PS syllable. As mentioned earlier,

UU vowels occupied an intermediate position between stressed and reduced vowels also in terms of vowel formant characteristics; F1 and F2 measurements showed that UU vowels differed significantly from stressed vowels by becoming more centralized, which is an interesting finding, given that all of these syllable types contain full vowels.

SS syllables recently have been explored in slightly more studies. The same study by Fear et al. (1995) showed that when PS syllables occurred in un-accented positions, SS and PS syllables tended to cluster together. There were no significant differences found between PS and SS syllables in regard to F_0 , intensity and vowel quality characteristics except duration—PS syllables were found to be significantly longer than SS syllables. These findings with regard to SS syllables were supported by Mattys' (2000) acoustic measurements of SS and PS syllables. Mattys measured the stress correlates of words produced in isolation and thus with PS subject to intonational accent, which has to be taken into consideration when comparing results to other studies. Specifically, Mattys found that the average F_0 of PS syllables was significantly greater than that of SS syllables; this is in contrast to the findings by Fear et al. (1995). However, Mattys' finding is hardly surprising given the fact that PS syllables in his study bore additional pitch accent. The other significant finding by Mattys concerned duration and also intensity—PS syllables were significantly longer than SS syllables and possibly louder, which is consistent with other studies.

A recent study by Plag, Kunter and Schramm (2011) was the first ever to compare PS and SS in accented and unaccented positions, directly addressing the long-standing concerns of uncontrolled covariation of stress and accent (Beckman & Edwards, 1994). Furthermore, the study investigated the role of SS syllable location within a word in relation to the PS syllable. SS can occur both before PS (e.g., the first syllable in “nominee”, where “no-” bears SS and “-nee”

bears PS), and after PS (e.g., the third syllable in “modernize”, where “-nize” bears SS, and “mo-” bears PS), and the location of the syllable can potentially have various effects on the use of stress correlates. Plag et al. found that in words that had a tonally-accented PS syllable as the first stressed syllable (e.g., in the word “finalize”), there were large differences in F0 and intensity between the SS and PS syllable. On the other hand, when the SS syllable was the first stressed element in the word (e.g. in the word “isolation”) and PS was intonationally accented, there were significant, yet small differences in F0 and intensity. Duration, however, did not differ between SS and PS syllables, and it was not dependent on location. In unaccented positions the difference between PS and SS in terms of F0 and intensity was significant, but quite small, and there was again no effect of vowel duration.

Since the study by Plag et al. brought in another variable, namely, stressed syllable location, in the discussion of acoustic differences between PS and SS syllables, one has to be careful when relating the findings of Plag et al. to previous studies. Other studies have typically investigated SS syllables only in one position, specifically, *before* the PS syllable (Fear et al., 1995; Mattys, 2000), where it appears to have greater prominence than in a position after PS. Also, since the location of the target syllable proved to be an important factor in acoustic analyses, the same could also apply to UU vowels. UU vowels can appear in a position *before* the PS (e.g., in the word “tattoo”) and *after* PS (e.g., in the word “python”), and the location could affect the results obtained.

In sum, these studies show that the two least explored types of phonological prominence—UU and SS syllables—have a distinct phonetic reality to them. Although containing full vowels, UU vowels were found to be generally shorter, quieter and lower in pitch than PS and SS syllables, and yet longer and louder than unstressed reduced syllables (Fear et al.,

1995; Cutler et al., 2007). This acoustically places them in a category separate from the reduced and the other stressed syllables. Acoustic findings of SS syllables, on the other hand, varied depending on whether they were compared to PS syllables in accented or unaccented position. In *accented* positions, and in positions before PS (which is of interest for the current study), SS syllables had significantly lower pitch and intensity values than PS syllables (Mattys, 2000; Plag et al., 2011). In *unaccented* positions, however, studies provide somewhat conflicting results—in Plag et al. (2011), PS syllables were higher in F_0 and intensity than SS syllables, but of the same duration as SS syllables. In Fear et al.'s (1995) study, PS syllables were only longer in duration, with F_0 and intensity being the same as in SS syllables.

If the traditional notion of “primary stress” in fact indicates a stressed syllable with a tonic accent (as opposed to a stressed syllable with no tonic accent, i.e., SS syllable), then indeed there are great acoustic differences between the SS and PS syllable types, and they form separate stress categories. Thus production of stress degrees depends on various combined effects of prosody, linguistic structure, and individual stress and vowel quality features.

Perception of Lexical Stress Degrees by Native Speakers of English. Listeners are assumed to be sensitive to stress cues that are employed in their native language. In English, the three most important phonetic correlates of stress are increased intensity, duration and F_0 , as well as vowel quality, or, from the perspective of a listener, change in loudness, length, and pitch. There is a reason to think that listeners would show mastery in employing all three in order to distinguish strong syllables from weak, and correctly identify the finer distinctions of prominence. However, are all stress correlates perceptually equally important? Can listeners tell syllable types of various stress degrees and vowel quality apart? Errors in any of stress correlates

or vowel quality could potentially interfere with correctly identifying the syllable prominence and affect the ease of speech processing. Therefore, the question that is relevant for the current second language study is: what cues are most critical for native listeners to perceive a word accurately?

Fry (1958) conducted a pioneering perceptual experiment on listeners' use of stress correlates. Using noun/verb pairs that only differed in stress patterns, he systematically varied intensity, duration and F_0 , and assessed their effects on perceiving each target word as a noun or a verb. He concluded that F_0 was the strongest stress cue, followed by duration and intensity. It should be noted, however, that Fry experimented with words produced in isolation, where tonic stress could partially account for the great effect of pitch in PS syllables. Since word stress in this study was confounded with phrase level prominence, it is unclear whether identical results would be achieved with target words used in un-accented positions.

A study by Lehiste and Fox (1992) examined listeners' perception of prominence using just two stress correlates—duration and intensity; fundamental frequency was kept constant. Listeners were presented with synthetic speech stimuli—stressed and unstressed syllable sequences of various duration and intensity values. The study found that when there was a competition between amplitude and duration cues, English speakers tended to rely on amplitude more than duration.

Vowel quality as a stress cue was explored by Zhang and Francis (2010), who designed a series of experiments to determine its role in perceiving stressed syllables in comparison to that of F_0 , intensity, and duration. Three separate experiments with re-synthesized speech stimuli of noun/ verb pairs (e.g., “desert”) tested the importance of vowel quality cues versus the three other stress dimensions. The study found that native English speakers employed all four acoustic

cues in lexical stress perception. However, vowel quality was weighted more strongly than other stress correlates and it stood out as the most reliable lexical stress dimension.

Additional evidence for vowel quality being an important perceptual cue to the distinction of stressed and unstressed syllables was provided by Braun, Lemhöfer and Mani (2011), who approached the issue from the perspective of improper (i.e., non-native) stress implementation and its impact on native listener perception. Their study showed that an improperly produced vowel, specifically, an insufficiently reduced vowel, in the presence of correctly produced suprasegmental cues resulted in seriously obstructed word recognition, and English listeners were hampered by the absence of vowel quality cues to lexical stress.

Overall, it is clear that all three stress correlates and vowel quality are actively used in speech perception to detect stressed and unstressed syllables; however, no clear pattern emerges regarding the superiority of specific stress cues, as findings from different studies so far are quite contradictory. The reviewed studies, however, addressed listeners' perception from a binary stress perspective, namely, stressed versus unstressed syllable categories, whereas perception of finer differences in syllable prominence were not within the scope of these studies.

A study by Fear et al. (1995) was the first one to investigate listeners' perception of the four syllable prominence types, specifically, unstressed reduced (UR), UU, SS and PS syllables. The perceptual experiment involved listening to words with cross-spliced syllables of four different stress types: the first syllables of words such as “authorized” (PS), “authorization” (SS), “authentic” (UU) and “authority” (UR) were spliced out of their original word and spliced into the first syllable location of the other three words. Listeners were then asked to rank the acceptability of these words. Results showed that native speakers tended to perceptually group UU, SS and PS syllables together. This finding hence indicates that listeners made a binary

strong/ weak distinction between syllables with full vowels and syllables with reduced vowels. Moreover, Fear et al. also carried out a correlation between acceptability judgments of tokens and the acoustic properties of those tokens in order to identify those stress correlates that listeners tend to rely on most. Overall, listeners tended to rely on vowel quality, intensity and duration most when making acceptability judgments; however, vowel quality characteristics again showed the strongest correlation with acceptability of a word, followed by intensity and only then duration. Interestingly, F_0 showed almost no correlation with the obtained acceptability ratings. Based on these results, Fear et al. hypothesized that the observed results were due to listeners' ability to make absolute judgments faster than relational judgments. Specifically, vowel quality offers immediate information about the absolute prominence of a syllable and its association with a certain stress category, while duration, intensity and other prosodic dimensions are relational in nature and can only be judged in relation to other syllables. Thus prosodic dimensions reflecting stress differences require comparison, which can take time, while vowel quality information is available immediately.

Finally, acceptability ratings of the same study also showed that listeners viewed reduction of UU vowels as relatively acceptable, and more so in context than in citation form. On the other hand, SS vowels that were substituted with a reduced vowel were rated as relatively unacceptable. Such findings are of great relevance for the current perceptual study. Mattys (2000), however, cautions that "acceptability results cannot be directly equated with perceptual sensitivity" (p. 256) in his response to the findings of Fear et al.'s study. The same study by Fear et al. also found that listeners tended to lump SS and PS syllables in non-accented positions together; however, this finding was obtained by acceptability ratings, and the actual perceptual sensitivity to differences between SS and PS degrees remains unknown.

Previous studies had traditionally looked at listeners' ability to distinguish stress degrees that differ in both stress and vowel quality patterns, but had not exclusively explored stress-discriminatory perceptual capacity. Mattys (2000) was the first to conduct a perceptual study in order to clarify whether listeners have the perceptual capacity to distinguish between stress degrees that do not entail change in vowel quality, namely, between PS and SS syllables. Since there is no vowel quality change that could aid as a cue, listeners could rely only on suprasegmental cues to distinguish the two stress degrees. In this study, listeners were presented with short word fragments that only differed in stress degree using two-syllable or one-syllable fragments containing a critical PS or SS syllable (e.g., /'prasi/ (PS) and /,prasi/ (SS), or only /'pra/ and /,pra/); listeners had to identify the word the fragment came from (e.g., "prosecution" or "prosecutor"). Mattys showed that native speakers actually exhibit a high degree of sensitivity to finer stress distinctions, such as SS and PS. Listeners were able to distinguish between SS and PS syllables both when they were presented as a two-syllable fragments with an unstressed "reference syllable", and also when SS and PS syllables were presented as a one-syllable fragment. The conditions tested thus were relational versus absolute stress perception. Performance, however, was higher when SS and PS syllables were presented together with an unstressed reference syllable. This finding thus indirectly supports Fear et al.'s hypothesis that duration, intensity and other prosodic dimensions are relational in nature and can be best judged in relation to other syllables.

Mattys' (2000) study thus provides additional evidence that listeners are sensitive to fine-grained stress distinctions. Listener sensitivity to PS and SS distinctions, along with reliance on vowel quality in making stress-related judgments suggests that such sensitivity could also potentially indicate native English listeners' perceptual reliance on these cues, and hindered

perceptual performance when such cues are not provided by non-native English speakers. The priming experiment with word fragments by Braun et al. (2011) provided first evidence that improper realization of lexical stress cues can interfere with successful on-line speech processing. How critical such cues are for native speakers when perceiving (i) words presented in full, not just fragments, (ii) with SS and UU syllables, and (iii) when vowel quality is improperly reduced, as opposed to improperly increased (Braun et al.), remains to be answered.

Second Language Phonological Acquisition

The observation that non-native speech perception and production differs from native has been empirically tested in numerous studies. These differences, or deviations from native speaker norms, have been the focus of L2 speech acquisition research, which has tried to explain what the differences are and where they are coming from (Archibald, 1993). The cause of these differences has been mainly attributed to the powerful effect that learner's first language exerts on second language phonological learning; however, universal linguistic phenomena are also proposed. This chapter will provide an overview of some of the most influential current theoretical frameworks: Flege's Speech Learning Model (Flege, 1995), Best's Perceptual Assimilation Model (Best, 1995), and Eckman's Markedness Differential Hypothesis (Eckman, 2004). These models vary in what they assume to be the basis of their explanations and predictions of L2 phonological learning difficulties, for example, perceptual or articulatory similarity, or typological markedness. They also vary in the scope of their focus, specifically, phonological acquisition or other linguistic domains as well, second language production or only perception, and segmental or also suprasegmental aspects of speech learning. The overview will be followed by a discussion on the applicability of these models and the specific predictions they

might make regarding Russian speakers' perception and production of English language lexical stress degrees.

Theories of L2 Phonological Acquisition.

Perceptual Assimilation Model. Best's (1995) Perceptual Assimilation Model (PAM) is one of the most prominent second language speech acquisition theoretical frameworks that postulates the powerful influence of speaker's first language phonological experience on second language acquisition process. Experience with one's native language is viewed as "an organizing perceptual framework that shapes discrimination of unfamiliar speech contrasts" (Best, McRobert, & Goodell, 2001, p. 776). The model was posited specifically for segmental acquisition, yet it can be further extrapolated to L2 suprasegmental learning.

PAM was developed to predict speech sound perception from the perspective of naïve non-native listeners, or beginner learners of a language, that are first exposed to second language speech. It makes predictions neither about sound production, nor speech sound perception at more advanced language learning stages. The model states that, contrary to the widespread view, not all non-native sound contrasts will be uniformly difficult to learn; instead, there is variability in the degree of difficulty that listeners will experience with certain second language sound contrasts. A learner's success, or lack of it, in acquiring non-native contrasts depends on the perceived distance between the L1 and L2 phonological categories. Learners will tend to assimilate new sounds, whenever possible, to the existing L1 phonological categories, and assimilation will happen through detecting commonalities in articulation. Best's speech perception model is based in articulatory phonology (Browman & Goldstein, 1992), which posits that listeners perceive incoming speech sounds directly as articulatory gestures, rather than as acoustic-phonetic information. Thus, the ease or difficulty of perceiving a non-native contrast

can be predicted based on “differences in the patterns of gestural similarities and discrepancies between various non-native contrasts and the properties of native phoneme distinction” (Best, 1995, p. 191).

PAM further proposes a detailed typology of how new L2 sounds can be interpreted by listeners relative to the existing L1 phonological categories, and explains why some non-native contrasts are easier to perceive than others. Three ways of assimilating non-native sounds are identified. The first is as a Categorized exemplar of some native phoneme, which ranges from good to poor in its goodness of fit. Second, a non-native sound can be perceived as an Uncategorized sound that fits somewhere between native phonemes. Third, a non-native sound can be perceived as a Non-assimilable non-speech sound with articulatory properties greatly discrepant from L1 sound categories (Best et al., 2001). Depending on what assimilation type the non-native sound undergoes, native language phonological influence can be facilitating, hindering, or neutral. Facilitation is strongest in cases when two non-native sounds assimilate to two separate L1 phonological categories, or when one assimilates well to an L1 category and the other is perceived as clearly different and thus not categorizable in terms of established L1 categories. The opposite—a hindering effect of one’s native language phonological system—would be observable in cases when a non-native contrast assimilates equally well to a single native category, making discrimination of the contrast extremely difficult, because listeners are generally more sensitive to between-category differences than within-category differences.

This typology, however, was developed to account for difficulties that naïve non-native listeners experience when perceiving sounds in a language they have never actively learned, while individuals that have some experience with the target language might show different perceptual performance. To account for this, Best and Tyler (2007) extended the principles of

PAM to second language speech perception in a model called PAM 2; the difference between these two models is based on a distinction between monolinguals who are not actively using L2 and are linguistically naïve to it, or absolute beginner learners (PAM), and more experienced L2 learners who are actively learning L2 to achieve functional, communicative goals (PAM 2). Best and Tyler (2007) hypothesize that the difference between these two listener groups lies in the level at which non-native speech sound categorization occurs. Specifically, “the phonological level is central to the perception of L2 speech by SL (Second Language) learners, who are developing an L2 (or interlanguage system), in a way that it cannot be for L2-naïve listeners perceiving unfamiliar non-native contrasts” (p. 23). Naïve non-native listeners “are unaware of which phonetic distinctions constitute phonological differences in the unfamiliar target language and could not possibly differentiate phonetic and phonological levels in non-native stimuli” (p. 23). Thus, according to PAM 2, it is possible that despite perceiving differences at a phonetic level, experienced learners may identify L1 and L2 sounds as functionally equivalent at a phonological level. Best and Tyler (2007) further state that “contrasts at the functional linguistic level of the L1 phonology and their relationship to phonological contrasts in the L2 are as important to perceptual learning as phonetic categories” (p. 26). In sum, second language learners who use the target language actively, perceive speech not only at a phonetic level, but also at a higher-order phonological level.

The hypothesized changes in perception of beginner learners and more advanced learners indicate that listeners continue to refine their perception of non-native contrasts throughout their life. Thus, in contrast to the well-known critical period hypothesis, PAM assumes that perceptual learning ability remains intact even for late learners.

Speech Learning Model and Feature Hypothesis. Another influential model that addresses second language speech acquisition is Flege's (1995) Speech Learning Model. The main tenets of this model resonate with Best's PAM in various important aspects. However, in contrast to PAM that focused only on L2 perception, Flege's SLM was designed to account for both L2 production and perception. Also, unlike PAM, SLM focuses on experienced second language learners rather than beginner learners, and is concerned with age-related factors and ultimate attainment in second language speech learning.

Overall, SLM consists of four postulates and seven hypotheses about second language sound acquisition. SLM is based on the premise that the capacity for speech learning remains intact throughout lifespan, such that the observed difficulties that second language learners experience stem from prior linguistic experiences with their native language rather than from neurological maturation. The more developed and elaborate the L1 system is at the onset of L2 exposure, the more influence L1 will exert on the L2 system, decreasing the learner's ability to discern phonetic differences between L1 and L2 sounds, and between L2 sounds that are non-contrastive in L1; moreover, fewer sounds in the L2 will be produced accurately as age of learning increases (Flege, 1995). Overall, the earlier L2 learning starts, the less dissimilarity between sounds—and, possibly, between various phonological features—is needed to perceive phonological differences and form new sound categories, which is a prerequisite for correct sound production, according to the model.

Flege's model assumes that L1 and L2 share the same phonological space, and learners' "phonetic systems reorganize in response to sounds encountered in an L2 through the addition of new phonetic categories, or through the modification of old ones" (p. 233). Thus new category formation depends on the perceived differences and similarities between first and second

language sounds. Discerning even some phonetic differences between an L2 sound and the closest L1 sound could lead to new phonetic category formation. Specifically, SLM proposes that the existence of one or more native categories that are phonetically similar to a non-native category may interfere with a new category formation, and the perception and production of that L2 category. Flege uses a term “equivalence classification” to refer to the process where learners tend to perceive new L2 sounds and phonetically similar L1 sounds as belonging to a single category. Thus similar sounds appear to be harder to acquire than dissimilar sounds, because speakers perceive them as equivalent to those used in their native language. Flege (1987) has described equivalence classification as “a basic cognitive mechanism which permits humans to perceive constant categories in the face of the inherent sensory variability found in the many physical exemplars which may instantiate a category” (p. 49). In a second language learning context, however, the equivalence classification mechanism blocks new category formation and, as a result, learning does not take place. On the other hand, second language sounds that appear radically different from sounds in the L1 inventory might not be assimilated as easily. The third hypothesis of SLM predicts that “the greater the perceived phonetic dissimilarity between an L2 sound and the closest L1 sound, the more likely it is that phonetic differences between two sounds will be discerned” (Flege, 1987, p. 239) and a new phonetic category will be established.

The notion of similar phenomena being more difficult to learn than dissimilar phenomena is central to the Speech Learning Model. Major (2008) has pointed out that, although it was not explicitly stated in SLM, the model implies “that transfer persists more for similar sounds than for dissimilar sounds” and further concludes that “the more similar the phenomena the more likely transfer will operate; however, what constitutes similar is not always clear-cut” (pp. 73, 74).

Generally, SLM and the previously described PAM-2 converge on the main issues of second language speech learning; however, certain differences exist. One of the central differences is the *level* of equivalence classification—is it phonetic or phonological? Best's PAM-2 disagrees with one of the central postulates of Flege's SLM model, where the latter holds that "sounds in L1 and L2 are related perceptually to one another at a position-sensitive allophonic level, rather than at a more abstract phonemic level" (p. 239). Instead, Best's PAM-2 posits that listeners can perceive native and non-native sounds as equivalent not only at a phonetic but also at a phonological level (e.g., listeners can perceive French and English /r/ as functionally equivalent and assimilate them despite the considerable phonetic and articulatory differences between these two sounds). The other point of disagreement concerns the *nature* of speech learning, where PAM-2 represents the direct realist approach with speech perceived as articulatory gestures, while SLM assumes mental representations of phonetic categories.

While not explicitly stated as part of the Flege's SLM, implied in the model was the Feature Hypothesis, which was further advanced by McAllister, Flege and Piske (2002). This hypothesis posits that "L2 features not used to signal phonological contrast in L1 will be difficult to perceive for the L2 learner and this difficulty will be reflected in the learner's production of the contrast based on this feature" (p. 230). McAllister et al. provided evidence in support of this hypothesis, showing that L2 learners' success with Swedish quantity contrast perception and production was dependent upon the phonological importance of duration feature in their L1. Thus a phonetic feature not exploited in L1 to signal phonological contrast will be difficult to acquire in L2, where it is used contrastively. It is hypothesized to be due to the L1 perceptual system being attuned to "phonologically meaningful phonetic features and, as such, the L1 system will underattend to phonetic features that are not phonologically meaningful" (Lee,

Guion, & Harada, 2006). Listeners develop an orientation of their attentional resources to those aspects of L1 that are phonologically meaningful, with the help of selective attention mechanism that makes the listener tune to only those aspects that are phonologically contrastive. Extensive linguistic experience with one's native language shapes perception and makes speakers weigh perceptual cues in a language-specific way, which often is different from cue weighting in L2. If such cues are not weighted heavily in L1, unlike in L2, or not present in L1 at all, modulation in such features might escape the attention of L2 learners. Since "noticing" is critical for learning a second language, not detecting such features prevents their further processing and ultimately learning (Lee, Guion, & Harada, 2006; Schmidt, 2001).

Markedness Differential Hypothesis. Linguistic universals are another domain that researchers have looked into to explain facts about L2 phonological learning; Eckman's (2004) Markedness Differential Hypothesis (MDH) is one of the most prominent L2 acquisition models that has been formulated using the concept of typological markedness. MDH claims that L1-L2 differences are important and necessary to explain L2 learning difficulty, but not sufficient, therefore it is necessary to incorporate universal constraints into the hypothesis. According to MDH, interference from learners' native language and linguistic universals can be reconciled (Eckman, 2004). Similar to Best's Perceptual Assimilation Model and Flege's Speech Learning Model, this model posits that not all phonological differences between L1 and L2 will cause equal difficulty for the learner. Instead, the degree of difficulty depends on the degree of *markedness* that the given phonological structure holds. The long-established concept of "markedness" is explained by assuming a binary opposition between linguistic representations, where one member of the opposition is more widely distributed, simpler, more basic and natural and thus "unmarked", versus the other member, which is considered "marked" (Eckman, 2008).

The concept of typological markedness, which attempts to state universal generalizations about world's languages and is at the heart of MDH, has been defined in the following way: "A structure X is typologically marked relative to another structure, Y, (and Y is typologically unmarked to relative to X) if every language that has X also has Y, but every language that has Y does not necessarily have X" (Gundel, Houlihan, & Sanders, 1986, as cited in Eckman, 2008). Thus the relationship between these structures is asymmetric, where the presence of one structure implies the presence of another, but not the other way round. In other words, less common structures will not appear in a language without similar, more common structures; thus, the less common structures will be marked.

The Eckman's Markedness Differential Hypothesis posits that typological markedness is a measure of relative difficulty in L2 phonological acquisition and second language acquisition in general. Marked structures are claimed to be harder for learners to acquire than the corresponding unmarked structures. Specifically, MDH hypothesizes that L2 phonological categories which differ from native language categories and that are more marked will be difficult to learn, but L2 phonological structures that are different but unmarked will be relatively easy to acquire. Finally, "the degree of difficulty involved is predicted to correspond directly to the relative degree of markedness" (Eckman, 2008, p. 98). If two speakers of different languages learn each other's language, the model predicts that they might not experience equal difficulty with the same type of structures because of the possible differing levels of markedness; the same may be true if learners of various linguistic backgrounds attempt to learn the same language, because one learner might encounter more marked structures than the other. To exemplify these points, studies reported in Eckman (2008) have shown that the acquisition of a word-final voice contrast by speakers of languages that have no voice contrast word-finally, such as German, are

more difficult, as the word-final voice contrast is more marked than voice contrasts in initial and medial positions. Similarly, Chinese L2 learners of English have been shown to experience more difficulty with English coda consonant clusters than Arabic L2 learners of English due to both L1—L2 differences and the degree of markedness.

Overall, MDH appears to have a wider focus than Best's PAM or Flege's SLM. Specifically, MDH attempts to explain and predict L2 acquisition in various linguistic domains and is not necessarily restricted to the phonological domain. By the same token, it attempts to account not only for segmental but also suprasegmental learning in the realm of L2 phonology. However, despite its wide applicability, empirical evidence in support of this hypothesis so far has been limited to acquisition of L2 syllable structure from a distributional perspective, for example, acquisition of consonant clusters and voice contrasts in various syllable positions.

Application of Theoretical Models to the Current Study. The three theories mentioned in the previous section differ in their focus and scope, and can generate different predictions regarding Russian learner of English acquisition of lexical stress degrees in English. Zhang, Nissen and Francis (2008) have pointed out that “although the SLM and PAM have traditionally been applied to production and perception of segmental phonemes, there is nothing about the models themselves that would necessarily restrict their predictions to the segmental domain, and either may be able to account for the acquisition of suprasegmental aspects of speech, such as intonation or stress” (p. 4500). Thus SLM's and PAM's predictions regarding speech learning might be true also for suprasegmental learning.

The Russian stress and rhythm system features only one stressed syllable per word, with surrounding syllables reduced to a certain degree. If Russian learners transfer their native language stress patterns to English, this would result in inappropriately reduced SS and UU

syllables, which are the only syllable types besides the main, or primary stress (PS), that contain full vowels in English.

Based on the tenets of SLM, two different possible predictions can be made regarding SS and UU syllable acquisition by Russian learners of English. SLM hypothesizes that, generally speaking, phonetic dissimilarities will be noticed but similarities ignored. If the existence of two stressed syllables within one lexical word in English is viewed by Russian learners of English as sufficiently “different” from Russian one-stress-per-word restriction, then SS syllable existence in addition to the PS might be noticed. This could lead to new phonological category formation for the additional stressed syllable, which is a prerequisite for learning. Similarly, in the case of UU syllables, it is possible that L2 learners will notice that English phonology relatively frequently allows not reducing vowels in unstressed syllables, which contrasts with L1 rhythm and stress rules. If the observed L2 phonetic differences between full and reduced vowels are sufficiently large, a new L2 category may be established for UU vowels and eventually lead to learning. On the other hand, a contrasting prediction is also possible: if learners put the L1—L2 differences and similarities in a much broader context and attend to the overall rhythm pattern, then the many existing similarities between English and Russian language rhythms, such as the presence of lexical stress, stressed syllable realization and vowel reduction, could make the L2 rhythm appear similar to L1 rhythm and block any further learning through equivalence classification. White and Mattys (2007) in their study on L2 rhythm learning observed that most L2 speakers make little rhythmic accommodation if their L1 and L2 are rhythmically similar, calling it “rhythmic inertia in the face of subtle distinctions” (p. 518), which is in agreement with the main principles of Flege’s SLM and Best’s PAM. Dutch is a language that is rhythmically similar to English, and belongs to the same “stress-based” rhythm class as Russian. The study by

White and Mattys showed that for certain rhythmic parameters, Dutch learners of English did not accommodate between the L1 and L2, preferring to realize the L2 just like their L1. Likewise, Russian learners of English might perceive English rhythm as identical to that of Russian and fail to perceive the more subtle differences between the two, such as the distinction between SS and UU syllables.

Russian learners' difficulty with perceiving English SS and UU syllable features in L2 is indirectly predicted also by the Feature Hypothesis, which was implied in Flege's Speech Learning Model (MacAllister, Flege, & Piske, 2002). This hypothesis states that perception and production difficulties can stem from different degrees of phonological relevance or weight that are given to phonetic features used in different languages, where "L2 features not used to signal phonological features in L1 will be difficult to perceive for the L2 learner and this difficulty will be reflected in the learner's production" (p. 230). A good example here would be Mandarin speakers' use of fundamental frequency to mark lexical stress in English in place of a combination of F_0 , intensity and duration. Since duration and intensity are not used contrastively in Mandarin, while pitch is, their experiences with this feature and its perceived importance or "weight" makes them transfer it to their second language, simultaneously making them perceptually immune to other cues that are actually used by native speakers (Chen et al., 2001). Similarly, Dutch speakers showed greater sensitivity to stress features like duration, intensity and F_0 than English speakers, because Dutch only operates with stress features and not vowel quality, while English uses both (Cutler, 2007).

The situation might be different with SS and UU syllables, however, since SS is not used *contrastively* in English and might not be phonologically meaningful. Jones (as cited in Lehiste, 1970) claimed that SS is an intermediate category that is not used for phonological contrasts; that

is, he claimed that word meaning cannot be distinguished based on SS category alone. There could be certain exceptions, e.g., words that end in “-ate” such as “graduate”, “alternate” or “advocate”, produced as a verb or adjective/noun feature a SS and a reduced vowel contrast in the third syllable. Phonological SS syllable contrasts in a position *before* the PS, which is the SS position used in this study, however, might not be possible.

The same refers to UU syllables as well. While there might be a few instances of minimal pairs whereby UU vowels show contrastive potential with reduced vowels, for example, in word pairs such as “radar” and “raider”, or “audition” and “addition”, such instances are relatively rare. Typically, neither SS nor UU syllables appear to be phonologically contrastive. Since the Feature Hypothesis predicts that contrastive L2 features not used to signal phonological features in L1 will be difficult to perceive for the L2 learner, one can only imagine how much harder it will be for the L2 learner to perceive L2 features that are contrastive and phonologically meaningful neither in their L1 nor L2. Thus the Feature Hypothesis indirectly indicates that SS and UU syllables might create difficulty for Russian L2 learners and might not be reflected in production.

Eckman’s Markedness Differential Hypothesis looks at phonological acquisition from the perspective of typological universals. The predictions that this hypothesis makes regarding Russian learner of English acquisition of SS and UU appears to be consistent with SLM and PAM. According to MDH, if the presence of one structure (A) implies the presence of another (B), but not the other way round, then A will be marked relative to B and supposedly harder to acquire for second language learners. With regards to SS syllables, every language that has SS syllables also has PS syllables; however, the presence of PS syllables in a language does not necessarily indicate the presence of SS syllables. This line of reasoning leads to a conclusion

that SS syllables are typologically marked relative to PS syllables and thus less common in world's languages. MDH further states that L2 phonological categories which differ from native language categories and that are more marked will be difficult to learn, which might be the case with SS syllables: first, they do not exist in Russian, and second, they are more marked than PS syllables, which Russian does have. Thus this model again supports the previous predictions that Russian learners of English will experience difficulty perceiving and producing SS.

The case with UU syllables is less clear. Russian has extensive vowel reduction in unstressed syllables, and no full vowels in unstressed syllables are allowed. Whether that means that UU syllable acquisition will be particularly hard for Russian learners of English, according to MDH, depends on the degree of universal typological markedness of the UU structure relative to unstressed reduced vowels. However, the task of establishing the typological markedness of these two structures might be extremely hard due to the varying degrees of vowel centralization existing in unstressed syllables of different languages. As reported in Fletcher (2010), Dutch, Arabic and to a certain extent also Polish employ stressed and unstressed syllables to create a specific rhythmic pattern. However, vowel reduction in the unstressed syllables can often be gradient, preventing a clear classification of reduced and unreduced vowel types. For example, unstressed vowels in Dutch or Arabic are not as short or centralized as unstressed vowels in English, so it is not clear whether they represent reduced or unreduced vowel category as there are no specific criteria for these two categories. If strictly looking at those cases where there is a clear distinction between reduced and unreduced vowels in unstressed syllable positions and applying the formula of typological markedness, we find that there are languages where reduced vowels exist without the presence of UU vowels, Russian being a noteworthy example. Simultaneously, there are languages where UU vowels exist without the presence of reduced

vowels, such as Polish (Jassem, 1959, in Fletcher, 2010). This leads to a conclusion that universal markedness relationships between these UU and reduced vowels cannot be established easily and cannot provide definite answers regarding second learner success with these phonological structures.

Finally, it is also possible that difficulties that Russian learners of English might potentially encounter with UU and SS syllable structures might not be of a perceptual nature, as predicted by SLM and PAM, where phonologically not meaningful structures are underattended, but rather be a more or less intentional decision of ignoring structures that appear redundant. The decision to disregard SS and UU syllable accuracy could be made at an earlier or later stage of speech processing. For example, Flege (1995) admits that it is possible that “sensory information that has initially been processed is discarded at a later processing stage by non-native speakers as nondistinctive” (p. 241).

Kondaurova and Francis (2008) showed that Russian and Spanish speakers were able to effectively employ duration cues when it was viewed as the only contrastive feature: for lax/tense vowel distinctions, instead of using vowel quality information, listeners resorted to duration contrasts. Duration is not a phonologically contrastive feature in Russian, in the sense that it can alone change lexical meaning; however, it is an acoustic correlate of lexical stress. Kondaurova and Francis’ study showed that Russian learners of English were able to operate with duration as a feature independent of lexical stress, when its use was viewed as phonologically meaningful. Failure to use duration to mark UU and SS syllables suggests that Russian speakers might indeed discard UU and SS syllables as they do not see them relevant or meaningful.

Similarly, based on the results of the study on L2 rhythm learning by speakers of the same or different rhythm classes, White and Mattys (2007) proposed that L2 learners of a rhythmically similar L1 fail to make subtle rhythmic accommodations because it may be communicatively sufficient to transfer L1 rhythmic patterns to L2 without any further adjustments. They have referred to this phenomenon as a “strategy” (p. 517), which means a rational decision by the L2 learners as to what structures are worth attending to, and which ones are not critical for success in communication.

Overall, based on the predictions of theoretical models and hypotheses, it can be concluded that Russian learners of English most likely will experience difficulty with SS and also UU syllables, and this difficulty will be evident in their production. The main reasons for predicted low accuracy with these lexical prominence types could be due to Russian learners of English perceiving them as phonologically not meaningful, or perceiving the rhythms of English and Russian as too similar to make any further adjustments specifically for SS and UU syllables. Difficulty with SS syllables is predicted also by the linguistic universals, where SS is marked and thus harder to acquire, while the markedness of UU syllables is relatively unclear, not allowing specific predictions to be made.

Research on Non-native Production and Perception of Suprasegmentals. Production and perception studies have consistently shown that non-native perception and production of stress and rhythm in English deviate from native speakers’ performance, reflecting learners’ experience with patterns and features that are employed and viewed as critical in their first language. The current section reviews the existing research on second language rhythm and stress acquisition, production being viewed from the perspective of phonetic implementation of

stress rather than the acquisition of its placement. This section looks at research on L2 suprasegmental acquisition in terms of producing and perceiving English stress correlates, learning of vowel reduction, sensitivity to lexical stress degrees, and rhythmic pattern acquisition in general.

Stress Correlates. Prior research suggests that L2 learners tend to apply phonological knowledge of their native language when perceiving stressed and unstressed syllables in second language speech. For example, Yu and Andruski (2010) conducted a cross-linguistic perceptual study with speakers of English and speakers of Mandarin Chinese in order to explore the use of stress correlates in English by non-native speakers. The study found that English and Chinese speakers relied on different acoustic cues when identifying stressed syllables in English. Chinese speakers used pitch as the most consistent cue to stress, whereas English speakers used various cues depending on the stimulus type and stress pattern (for example, trochaic vs. iambic stress).

Zhang et al. (2008) showed that Chinese speakers' special reliance on pitch was evident also in their production of PS and unstressed reduced syllables in noun/verb word pairs such as "object", "desert" and "permit". F_0 , duration and intensity were found to be significantly greater in PS syllables than unstressed syllables for both Chinese learners of English and native English speakers. This indicates that Chinese speakers were able to employ all the same stress cues in English that native English speakers used. Chinese learners of English, however, produced PS syllables with a significantly higher F_0 than native English speakers—a finding that was supported by another acoustic study on Chinese production of English sentence stress by Chen, Robb, Gilbert and Lerman (2001). Considering the fact that Chinese is a tonal language and Chinese speakers are used to operating in the tonal domain, it is not surprising that Chinese L2 learners of English rely on F_0 most when perceiving prominence distinctions. These studies,

however, also caution that one's familiarity and experience with such cues in their native language and perceptual sensitivity to them in L2 do not yet automatically translate into their accurate *production* in L2.

Lehiste and Fox (1992) provided evidence for linguistic background affecting perception of prominence for speakers of Estonian. Their cross-linguistic study used prominence judgments, where listeners heard repetitions of synthetic stimuli of various duration and amplitude values and had to indicate which syllable was more prominent. Results showed that speakers of Estonian, which is a quantity language that employs contrastive vowel length, responded to the stimuli differently than English speakers. English language speakers were perceptually most sensitive to amplitude when making prominence judgments, while speakers of Estonian, which uses duration contrastively, were most responsive to duration cues. This finding indicates that L2 learners continue to employ their native language cues also in their L2.

Lee, Guion and Harada (2006) came to the same conclusion in their study on unstressed vowel production by Korean and Japanese early and late learners of English. Both early and late Japanese learners of English maintained the same stressed/ unstressed vowel duration contrast shown by native speakers, demonstrating native-like performance with duration cues, while Korean L2 learners showed significantly less duration and intensity difference between stressed and unstressed syllables than native English speakers. These differences in performance can be explained by the fact that in Korean, only F_0 is used to realize prosodic patterns, but duration and intensity are not an important part of Korean prosodic patterns. In Japanese, on the other hand, duration is phonemic and is used to realize rhythm patterns. Thus, Lee et al. (2006) conclude that the phonological relevance of an L1 phonetic feature tunes the perceptual system and affects L2 learning.

Vowel Reduction. English unstressed syllable production, and vowel reduction in particular, has been documented as relatively hard to master for second language learners. Most studies that have investigated acquisition of vowel reduction in English have focused on speakers of languages that represent a different rhythm class, specifically, syllable-based languages, and have no reduced vowels. For example, Flege and Bohn (1989) explored Spanish speaker performance on vowel reduction in English. While Spanish speakers performed relatively well placing stress and implementing the stressed versus unstressed syllable contrast, they were generally less successful in reducing unstressed vowels. Thus, they could “unstress” weak syllables but could not reduce them quality-wise in a native-like manner. This finding again is indicative of L1 interference. In Spanish, certain suprasegmental cues but not segmental differences are used to signal relative syllable prominence. Thus, absence of vowel reductions in learners’ first language resulted in low accuracy with reduced vowels in English, while suprasegmental cues presented less of a difficulty for Spanish learners of English.

Similar difficulties reducing vowels have been reported also for Mandarin Chinese learners of English, whose native language does not feature vowel reduction. Chinese speakers demonstrated inconsistent performance in vowel reduction: unstressed vowels were either not reduced, or incorrectly reduced with a non-native sounding vowel (Zhang et al., 2008). Similar conclusions were made by Lee et al. (2006) who showed that Korean and Japanese individuals’ performance on unstressed syllables and vowel reduction in English revealed native language influences. Both groups centralized vowels in unstressed syllables. However, the Korean group, on the one hand, tended to assimilate the reduced vowel to a similar Korean high front vowel, which resulted in a non-native-like production. The Japanese group’s vowels, on the other hand, were more dispersed in the vowel space than native speakers’.

Finally, the only more widely-known study on unstressed syllable acquisition by speakers of a language that belongs to the same rhythm class as English was conducted on Dutch learners of English. Braun, Lemhöfer and Cutler (2008) provided acoustic data of treatment of English reduced vowels by Dutch learners of English. In Dutch, vowels in unstressed syllables are only slightly more centralized than in stressed vowels, instead relying more on suprasegmental information to indicate lack of lexical prominence. As a result, the degree of vowel reduction in unstressed (reduced) syllables is greater in English than in Dutch. This was evident in Dutch speakers' production in English, where Dutch speakers produced vowels that were significantly less centralized than those produced by native English speakers, which suggested that Dutch speakers were transferring their native phonology to the production of reduced vowels in English. Again, speakers' familiarity and experience with vowel reduction in their native language did not automatically translate into accurate production of reduced vowels in L2. That is, while Dutch learners of English did apply vowel reduction, it was done in L1-specific ways.

Lexical stress degrees. There are only a couple of studies that have addressed second language learner perception and production of the fine suprasegmental differences between lexical stress degrees such as PS, SS and UU syllables, where vowel quality is kept constant. A perceptual study by Cutler et al. (2007) showed that Dutch speakers significantly outperformed native English speakers in their ability to distinguish between PS and UU syllables. The tasks were an isolated PS and UU syllable fragment identification task, as well as fragment priming and gating tasks. Given the fact that both lexical stress categories—PS and UU—contain strong vowels, this finding indicates that Dutch learners of English can use suprasegmental information more effectively than English speakers. The explanation again lies in the interference of the first language phonological system: Dutch features three levels of stress just like English, but vowel

quality is not a reliable indicator of stress degree. According to Braun, Lemhöfer and Cutler (2008) vowel reduction in Dutch is phonetic, rather than phonological, as it is in English. While native English speakers rely on vowel quality information in addition to suprasegmental cues, native Dutch speakers can only rely on the latter to distinguish between stress levels. The same Dutch learners' of English mastery of suprasegmental cues was reflected also in production studies by Braun, Lemhöfer and Cutler (2008), and Braun, Lemhöfer and Mani (2011), which demonstrated that Dutch speakers produced stronger differences between UU and PS syllables than English speakers in terms of such suprasegmental features as vowel duration and spectral tilt.

The findings from these perception and production studies collectively show that native speakers appear to continue to employ phonologically contrastive L1 features also when perceiving and producing L2 stress and rhythm patterns. Such findings tend to confirm the first language phonological system interference hypothesis. From the body of research conducted in the field of second language acquisition, it is also clear that there are very few studies on Russian acquisition of suprasegmentals and relatively few studies of English rhythm acquisition of speakers of languages that are rhythmically similar to English, Dutch being the only notable exception. Moreover, there are no studies on Russian production or perception of such phonological features as SS and UU, despite the fact that the non-native-like use of SS and UU syllables might be one of the “signature” features of Russian accented speech.

Suprasegmentals and Perception of Non-native Speech

The importance of suprasegmental accuracy for L2 speech perception has been emphasized by a number of research studies (Cooper, Cutler, & Wales, 2002; Hahn, 2004;

Quené & van Delft, 2010). Suprasegmental aspects of speech have even been shown to be more critical for intelligibility than segmental aspects in some studies (Anderson-Hsieh, Johnson, & Koehler, 1992; Anderson-Hsieh & Koehler, 1988; Derwing, Munro, & Wiebe, 1998; Derwing & Rossiter, 2003). Overall, stress and rhythmic accuracy can have a great impact on listeners' ability to perceive and comprehend speech, with the degree of interference varying from incomprehensibility to slight accentedness (Anderson-Hsieh, Johnson, & Koehler, 1992; Braun, Lemhöfer, & Mani, 2011; Cutler & Clifton, 1984; Mareüil & Vieru-Dimulescu, 2006; Tajima, Port, & Dalby, 1997; Zielinski, 2008). This section reviews empirical evidence on the type and degree of difficulty posed by speech with deviant suprasegmental patterns for native listeners' speech processing and word recognition ability.

Dimensions of Non-native Speech Processing. In the L2 research literature, perception of non-native speech has been broadly characterized with respect to three dimensions: intelligibility, comprehensibility, and accentedness. There are certain inconsistencies in the field regarding the use of these terms; therefore, for the purposes of this study, definitions provided by Derwing and Munro (1997) were used. Derwing and Munro define intelligibility as an objective measure of the extent to which a message is understood by the listener (e.g., words correctly recognized). Comprehensibility is defined as listener's subjective perception of intelligibility (i.e., perceived ease and accuracy of word recognition). Finally, accentedness is defined as the degree to which speech patterns are perceived as different from native speaker pronunciation norms, but this term is neutral regarding interference with speech processing. A high degree of perceived accentedness is typically equated with unintelligibility; however, such a relationship might not be as straightforward since even highly accented speech can still be intelligible (Munro

& Derwing, 1995a). Moreover, accentedness of some degree is inevitable even in the speech of highly proficient L2 speakers (Flege, Munro, & MacKay, 1995). Thus, perceived accentedness as such is not yet a measure of the difficulty that listeners will experience when perceiving non-native speech.

Intelligibility is largely a measure of the outcome of speech perception, and provides no information regarding the actual ease or difficulty of the word recognition process. Comprehensibility scores, on the other hand, might reflect the underlying mental effort and resources the listener spends until the word is recognized. Munro and Derwing (1995b) observed that listeners would often assign low comprehensibility scores to speech because it was reported by them to be “hard to understand”, even though listeners were able to transcribe phonemes and recognize words correctly. Based on that, Munro and Derwing (1995b) suggested that it is processing difficulty that often makes listeners rate speech as incomprehensible, a difficulty which manifests itself as an increased processing time for speech. The ease or difficulty perceiving speech can thus be assessed objectively by measuring listeners’ response-time (RT) in a relevant experimental task. The most commonly used experimental task in which RT is measured is a lexical decision task in which listeners must respond as quickly as possible if a visually (or sometimes auditorily) presented target item is a word in the language or not a word in the language.

Response time is operationally defined as the lapse of time between the presentation of a stimulus and the beginning of a response to that stimulus (Crabtree & Antrim, 1988). This measure is used in experimental psycholinguistic research to reveal insights into how different types of linguistic information are used in the on-line processing of language and where

difficulties in processing occur (McDonough & Trofimovich, 2009). Longer response times reflect processing that is more difficult and requires greater effort.

Munro and Derwing (1995b) were the first ones to use response time measurements in investigations of accented speech in general. They hypothesized that processing non-native speech would generally require longer response times than processing native speech, due to segmental and suprasegmental deviations from native pronunciation norms that lead to increased recognition time and decoding effort. Results showed that Mandarin-accented speech indeed produced greater response latencies than speech by native English speakers in a sentence verification task.

Reaction time measures are also commonly used in priming research. The priming method is an experimental paradigm that is widely used in the psycholinguistic research to detect listeners' sensitivity to language forms and the degree of lexical activation by the input. Priming itself means the tendency of people to process a spoken word more quickly and more accurately when they have had recent previous exposure to related lexical material. For example, priming can be measured as a time and accuracy benefit for previously heard words as compared to unheard spoken words (McDonough & Trofimovich, 2009; McNamara, 2005). Thus, an initially heard word ("prime") facilitates the recognition of the same stimulus ("target") when it is encountered shortly afterwards.

In the speech perception literature, a clear distinction is made between an experimental situation in which the lexical material used as a prime is semantically related to the target, as opposed to phonologically related. The former case is termed semantic, or associative, priming, where the prime shares some aspect of meaning with the target. Semantic priming is based on the concept that words are organized in memory in a large semantic network; activating one word

makes that activation spread to those words that are related in meaning, (Meyer and Schvaneveldt, 1971; Norris, Cutler, McQueen, & Butterfield, 2006). Semantic priming is the most common and well-established area of priming (Odekar, Hallowell, Kruse, Moates, & Lee, 2009).

In the field of speech perception, however, phonological priming is often used, where the prime shares certain phonological features with the target. According to Norris et al. (2006), phonological priming is based on activating phonological representations, or representations based on sound, while semantic priming is based on semantic representations, or representations based on meaning. In the case of phonological priming, facilitation to a target word is expected after having heard a prime word that is phonologically related (rather than semantically related). Phonological overlap in phonological priming can be partial, or it can be total, which is the case in identity priming. Identity priming is a specific type of phonological priming paradigm in which the prime word shares total phonological overlap with the target word, i.e., the target word is simply a repetition of the prime word. In identity priming studies, reaction times are used to assess whether prior exposure to phonological forms results in faster subsequent processing, compared to the processing of forms that the listener was not exposed to. A lexical decision task is typically used, where listeners are presented with a series of letter strings and they have to make a decision whether each such string represents a real word or not. Listeners are faster to respond that the string is a real word if they have already heard the same phonological form of a word shortly before.

Overall, phonological priming is considered to be an implicit process of lexical activation, which for the most part occurs with little awareness on the part of the listener (McDonough & Trofimovich, 2009; Nicol, Swinney, Love, & Hald, 2006). Upon hearing

speech, lexical representations get activated, and the level of activation depends on the degree of match between the perceptual input and the word's stored representation. When the phonological form of the prime differs from the canonical form of that word, the priming method can show whether lexical activation still takes place or not (Norris et al., 2006). The phonological priming concept is relevant to the current study, which aims to explore whether deviations in suprasegmentals constrain lexical access in any way.

A common experimental method is cross modal priming experiments that involve two modalities: auditory (for prime presentation) and visual (for target presentation), where listeners first hear a prime word, and then see the target word spelled on computer screen. Cross-modal priming experiments have been used in several studies in order to reveal the on-line deeper lexical activation processes that might be inhibited by *suprasegmental* deviations from native pronunciation norms (Cooper, Cutler, & Whales, 2002; Cutler & Clifton, 1984; Braun, Lemhöfer, & Mani, 2011). What is measured is the extent to which the prime influences the listener's lexical decision performance on the target word. Collectively, findings of these studies suggest that non-native-like prosodic patterns can be detrimental to the speed and ease of word recognition, even if word recognition eventually takes place—something that would be hard to detect with off-line measures.

Suprasegmentals and Non-native Speech Processing. Theoretical accounts of speech perception as well as experimental studies have assigned stressed syllables a major role in speech processing. According to Lieberman (1965), stressed syllables are acoustically more reliable and easier to identify than unstressed syllables. This suggests that listeners might direct their attention to prosodically prominent syllables in speech, a proposal which has found some support

(e.g., Pitt & Samuel, 1990). Several speech perception accounts suggest that stressed syllables might provide listeners with a code that links words with their representations in the mental lexicon (Dupoux & Mehler, 1990; Grosjean & Gee, 1987). Similarly, a study on French listener's perception of stress by Peperkamp and Dupoux (2002) proposes that speakers of English might store lexical stress as part of their phonological representation of these words in the mental lexicon, because lexical stress is used contrastively in English. French, on the other hand, is a fixed-stress language; therefore, it is not necessary to "encode" such information in phonological representations.

More empirical evidence for the important role of stressed syllables comes from research on speech segmentation, which proposes that strong syllables are those elements in speech where speech segmentation is initiated (Cutler & Norris, 1988). In this study, listeners were faster to detect a specific syllable ("mint") in bisyllabic words if the word consisted of a strong and a weak syllable ("mintesh") than if it consisted of a sequence of two strong syllables ("mintayve"). Since lexical access was initiated at the strong syllable, Cutler and Norris hypothesized that no segmentation was needed for the strong-weak syllable sequence, allowing listeners to be faster, while detection of the target syllable was harder when listeners had to assemble material across segmentation boundaries when the second syllable turned out to be strong, too. It should be noted, however, that Cutler and Norris adopted a vowel-based definition of strong syllables, where every syllable that contains a full vowel is strong. Mattys (2000) pointed out that according to such a liberal definition only reduced syllables (schwas) do not trigger segmentation.

Additional evidence for the role of stressed syllables comes from studies of speech in noise. An important function of stressed syllables in speech processing might be their usefulness

for speech perception in degraded listening conditions. Studies have shown that prosodic cues, and stress among them, are resilient to high levels of noise; thus, reliance on stress significantly increases in impoverished listening conditions—in noise, when speech is faint, or when acoustic-phonetic information is not available (Smith, Cutler, Butterfield, & Nimmo-Smith, 1989; Cutler & Butterfield, 1992). According to Mattys et al. (2009), in ideal, highly contextualized listening conditions listeners largely rely on lexical-semantic knowledge when making lexical decisions. However, in less optimal listening conditions where contextual and lexical cues are not helpful, reliance on sublexical information increases. Overall, listening experiments in an idealized setting — with carefully recorded speech, and undivided attention (Mattys et al., 2009), might overlook very important functions of SS syllables that only more realistic listening conditions can unfold.

When perceiving speech, listeners can extract information of at least three possible types: acoustic-phonetic (or segmental), lexical (or pertaining to knowledge about the incoming words), and prosodic, or suprasegmental (McQueen & Cutler, 2010). Listener reliance on stress depending on the availability of other cues has been explored in a number of research studies. Listener performance has been evaluated in contexts when, for example, multiple cues to lexical prominence are readily available, when only prosodic cues are available, or when cues are either missing or misused (Bond and Small, 1983; Cooper, Cutler, & Whales, 2002; Field, 2005).

Several priming studies with word fragments as primes have attempted to clarify whether suprasegmental information is used by native English listeners to resolve lexical competition. Cutler, Wales, Cooper and Janssen (2007) conducted a cross-modal priming experiment where they tested Dutch and English listeners' ability to recognize words based on suprasegmental information contained in a single syllable. The primes were the first syllables from words such as

‘CAMpus’ with PS as the first syllable, versus ‘camPAIGN’ with UU as the first syllable; thus, the initial syllables only differed in suprasegmental information, but not vowel quality. They found that Dutch speakers were more sensitive to suprasegmental information than native English listeners. Cutler et al. concluded that since English employs both segmental and suprasegmental cues to stress (unlike Dutch), native English speakers might rely on vowel quality cues more than suprasegmental cues.

Cooper et al. (2002) used a cross-modal phonological priming experiment with monosyllabic and bisyllabic word fragments to explore how suprasegmental cues are used in word recognition. Of importance for the current study was the bisyllabic condition, where listeners were presented with a word fragment “admi-“ that could either come from a word “admiral” with an initial PS, or come from a word “admiration” with an initial SS. They found that English listeners could successfully employ suprasegmental information in recognizing words; specifically, “admi-“ with initial primary stress activated “admiral” to a greater extent than “admiration”, and “admi-“ with an initial SS primed “admiration” to a greater degree than “admiral”. Mismatch in stress patterns of these words resulted in response times comparable to those of unrelated control words. This means that the lexical competition was resolved already at an early stage of the word, and listeners could successfully employ suprasegmental information to do that.

Also, the same study by Cooper et al. (2002) proved that listeners could reliably distinguish between SS and PS syllables—a finding that is compatible with the study by Mattys (2000). He asked participants to guess the identity of a word when solely word fragments were presented. A fragment like “presi-“ could come from a word “president”, the first syllable thus bearing a PS, or from a word “presidential”, the first syllable bearing a SS. The study found that

listeners were able to utilize purely prosodic differences between PS and SS syllables for word recognition without actually hearing the PS syllable. This study again demonstrated that reliance on SS alone could lead to successful word recognition, emphasizing its role in speech processing.

On the other hand, Cooper et al. (2002) claim that, compared to speakers of other languages, for example, native Spanish speakers (Soto-Faraco, Sebastián-Gallés, & Cutler, 2001), English speakers do not fully rely on suprasegmental information. In Cooper et al. (2002), a stress mismatch did not produce inhibition of lexical access, but rather no priming or even slight facilitation. They conclude that, consistent with findings by Cutler et al. (2007), native listeners weigh segmental information more strongly in lexical activation than suprasegmental information.

In support of this hypothesis, a recent study by Braun, Lemhöfer and Mani (2011) used a cross modal phonological priming task to explore native listener on-line processing of speech by Dutch learners of English, who do not reduce the central English vowel “schwa” to the same degree as native speakers. This study used word fragments as primes, and found that Dutch-spoken first syllable of a word “absurd”, which featured lack of vowel reduction, failed to prime the word “absurd”. Thus producing an unstressed reduced vowel as part of an UU syllable seriously constrained word recognition when word stress was not implemented in a native-like manner. A study by Cutler and Clifton (1984) measured native speaker response time using a semantic decision task. The study revealed that improper implementation of lexical stress (stress shifted to the adjacent syllable) was more detrimental for speech processing and word recognition when it involved vowel reduction in addition to suprasegmental changes, rather than only suprasegmental changes.

Studies by Small, Simon and Goldberg (1988), Bond and Small (1983) and Slowiaczek (1990) found that word recognition was seriously inhibited if both segmental and suprasegmental information got altered as a result of lexical stress shift, for example, “polite” produced as “POLite”. Finally, Field (2005), using a word transcription task, demonstrated in contrast to other studies that inaccurate placement of lexical stress had a relatively minor effect on intelligibility. He also pointed out that unlike in other studies that used response time measures, participants in this study had ample time to process what they had heard, a difference which might have facilitated word recognition.

All of these studies seem to suggest that mismatching stress information does constrain lexical activation, but mismatching segmental information has more detrimental effects on lexical activation. Russian learner of English reduction of SS and UU syllables thus is expected to pose serious difficulty for native listeners in the spoken word recognition process, because such productions do not provide native listeners with the expected segmental and suprasegmental information. Studies have shown that most detrimental effect on speech processing was created by simultaneous segmental and suprasegmental inaccuracies, which is the case when SS syllables are inappropriately reduced.

On the other hand, all the reviewed studies have investigated listeners’ speech processing or sensitivity to SS syllables only when they were presented as isolated word fragments. The importance of SS for word recognition in a full word and in the presence of an accurately realized PS syllable—of great relevance for Russian learners of English—has not been examined so far. It is conceivable that the surrounding lexical material—accurately produced PS and unstressed syllables—could compensate for the inaccurately implemented SS syllables by providing enough acoustic-phonetic information that speech processing is not disrupted and

word recognition can take place. Alternatively, SS might be stored as part of the phonological representation of a word in the mental lexicon, as proposed for PS (Peperkamp & Dupoux, 2002), in which case inaccurate production of SS syllables might have a detrimental effect on speech processing and word recognition.

Additionally, in the studies reviewed, experimental words typically involve a case of lexical competition where a single phonological aspect presents a mismatch between otherwise phonologically identical lexical candidates. Whether correctly placed yet unrealized lexical prominence is still critical in words that do *not* have many lexical competitors (as in SS syllable-containing words, less UU syllable-containing words), is not clear. To answer these questions, response time measurements in the current study provided an insight into the underlying speech processing effort when lexical prominence cues were inaccurate.

CHAPTER 2. ACOUSTIC STUDY

This research study consists of an acoustic analysis of native and non-native speech, and a following perceptual experiment that was based on the acoustic study results. The acoustic study had three goals. First, it provided a detailed examination of acoustic differences between SS, UU and Primary stressed (PS) syllables as produced by native English speakers and Russian learners of English. The following perceptual study was contingent upon finding significant acoustic differences in UU and SS syllable production by the two speaker groups. Second, the same lexical items that were recorded for the purposes of the acoustic study were further used as experimental stimuli in the subsequent perceptual study. Third, the acoustic differences detected in native and non-native speech established criteria for acoustic modifications of these stimuli for the perceptual experiment.

Methods

The acoustic study aims to identify the nature and degree of both vowel quality and duration differences between SS, UU and PS vowel productions by native speakers of American English, and Russian learners of English. Do Russian speakers transfer their native language patterns to a second language, reducing English SS and UU vowels? If they do, are these reductions of temporal nature, or do they affect vowel quality as well? To answer these questions, productions of two speaker groups—speakers of American English and Russian speakers of English—were recorded and analyzed in terms of their vowel quality and suprasegmental features. The same recordings were used in the following perceptual experiment as stimuli.

Design. A 2 x 2 x 3 mixed factorial design was used, with speaker group (Russian, English) as a between-subjects independent variable and syllable type (SS, UU, PS) and vowel type (/ɑ/, /æ/) as within-subjects independent variables. Vowel duration, fundamental frequency, intensity, and F1 and F2 were separate dependent variables.

Participants. The two populations of interest for the acoustic study were, first, Russian native speakers who speak English as their second language, and second, native speakers of English. Six Russian and six native English speakers were recruited and recorded for the acoustic study. Half of the participants in each group were male, and half were female. Prior to participation in the experiments, participants were asked to complete a questionnaire on their linguistic background (see Appendix B).

The Russian participants were selected based on their length of residence in the United States. The minimum criterion for the length of residence was set at two years, and the maximum was set at five years; these cutoff points were selected to ensure that speakers had received sufficient exposure to American English and felt comfortable producing words used for this study, but also making sure that their proficiency was still far from native-like and their language retained the specific non-native language patterns of interest for this study. The average age for Russian participants was 26, and they were all graduate students at a university in Ohio (see Table 6). The length of U.S. residence for the recruited Russian speakers of English ranged from 2.5 to 4 years, with an average of 3 years spent in the U.S. The age when participants started studying English varied from 5 to 22 years of age, with an average of 13 years of age. All of the participants spoke Russian as their first language. An ESL (English as a

Second Language) specialist with considerable experience working with Russian learners of English judged the speech of all participants to be strongly marked by the Russian accent, irrespective of the language learning background.

Table 6.

Russian Participant Demographics.

Participant	Nationality	L1	Age	Years in USA	Years learning English
1	Uzbek	Russian	25	3.5	14
2	Russian/Ukrainian	Russian	24	3	11
3	Russian	Russian	26	2.5	6
4	Russian	Russian	25	3	3
5	Russian	Russian	27	4	22
6	Russian	Russian	29	4	20

The native speakers of American English, whose speech was recorded for speech samples, were all recruited from the Northwest Ohio and spoke English as their first language. All of the participants were in the 22-24 age range, with an average age of 23, and they were all graduate and undergraduate students from various departments of Bowling Green State University. Again, participants were asked to provide detailed information about their linguistic background in a demographic survey (see Appendix C).

Materials. In this study, an acoustic analysis was conducted on words produced in isolation. Each of these words contained one target syllable—an UU, SS, or PS syllable—as part of a multisyllabic word. The words were systematically selected so that the target syllables—UU and SS—occupied positions within a multisyllabic word that were most prone to reductions and could demonstrate most dramatic effects of L2 rhythm transfer (see Table 7). As discussed in

Chapter 1, the degree of vowel reduction in Russian is greatest several syllables before the PS, and immediately after the PS syllable. These positions in English correspond well to the typical locations that SS and UU vowels occupy—SS occurs several syllables before the PS (as in the word “imagination” (/ɪˌmædʒɪˈneɪʃən/)), and UU occurs immediately after the PS syllable (as in the word “abstract” (/ˈæbstrækt/)). Overall, three different syllable types were explored in this study: SS syllables in Pre-PS position (not less than two syllables before the PS syllable, e.g., “compre’hension”), UU vowels in immediately post-PS position (one syllable after the PS, e.g., “robot”), and PS syllables (e.g. the second syllable in “ro’botic”). The PS syllable duration was measured in a three-syllable word, which was the average syllable count between two-syllable UU-words and four or five syllable long SS words. Intensity and F_0 measurements for PS vowels were done in the same two-syllable word as UU-vowels and the same four or five syllable word as SS vowels. Finally, there was great variability in the words that were included, ensuring adequate sampling of words featuring the specific syllable types.

To optimally test for the hypothesized vowel reduction effects both in terms of duration and vowel quality, the study focused on a vowel that reduces to a “schwa” in these syllables in Russian, specifically, the low vowel /ɑ/ in Russian (Avanesov, 1956). High vowels /i/ and /u/ were not included, as they are expected to reduce mostly in terms of duration, much less in vowel quality. Tabain and Padgett (2005) observed that vowel space significantly contracts in unstressed syllables, and, what is most important, the bulk of this contraction comes from rising of the vowel space floor, which affects exactly the Russian lower vowels /ɑ/ and /ɔ/. The rounded low-mid back vowel /ɔ/ was excluded from this research due to occurring in a very limited number of UU vowel-containing words in English. However, English vowel /æ/, which is not part of the Russian sound inventory, could be perceived as a variation of the Russian low

vowel /ɑ/ and therefore was included in the study. Thus the two target vowels used in this study were the low back vowel /ɑ/, and the low front vowel /æ/. The stimuli with low vowels /ɑ/ and /æ/ in UU, SS and PS syllable positions are shown in Table 7. The phonological context that these vowels occurred in could skew vowel F1 and F2 measurements; however, such co-articulation effects were controlled for, since participants in both groups produced vowels in the same phonological contexts.

Table 7.

Experimental Stimuli for the Acoustic Analysis.

	Secondary-stressed (SS)	Unstressed-unreduced (UU)	Primary-stressed (PS)
/ɑ/	<u>opti</u> 'mistic <u>ope</u> 'rational <u>domi</u> 'nation <u>poli</u> 'tician <u>popu</u> 'lation <u>popu</u> 'larity <u>conden</u> 'sation <u>confis</u> 'cation <u>accomo</u> 'dation <u>hospi</u> 'tality	'robot 'radar 'neuron 'pyth <u>on</u> 'nyl <u>on</u> 'mor <u>on</u> 'i <u>con</u> 'masc <u>ot</u> 'cou <u>pon</u> 'shamr <u>ock</u>	'o <u>ptim</u> ist 'po <u>liti</u> cs 'po <u>pu</u> lar 'ho <u>spita</u> l do <u>mina</u> te ro' <u>boti</u> c i' <u>con</u> ic 'ro <u>ck</u> ery 'co <u>tt</u> ony res' <u>pon</u> dent
/æ/	<u>appli</u> 'cation <u>gratifi</u> 'cation <u>acci</u> 'dental <u>compa</u> 'tibility <u>elabo</u> 'ration <u>magnifi</u> 'cation <u>collabo</u> 'ration <u>mascu</u> 'linity <u>classifi</u> 'cation <u>fabri</u> 'cation	'synt <u>ax</u> 'abstr <u>act</u> 'contr <u>act</u> 'bor <u>ax</u> 'extr <u>act</u> 'comb <u>at</u> 'form <u>at</u> 'imp <u>act</u> 'cont <u>act</u> 'comp <u>act</u>	syn't <u>acti</u> c abs't <u>raction</u> ex't <u>ract</u> ed con't <u>raction</u> com'p <u>action</u> 'gr <u>ati</u> tude 'mag <u>neti</u> ze 'acc <u>ident</u> 'lab <u>yrin</u> th 'class <u>ify</u>

Note: SS or UU syllables are underlined in each word; PS syllables are marked with (') preceding the syllable.

Words with several alternative pronunciations were excluded to prevent ambiguity. For words like “extract”, where spelling is not informative regarding their part of speech or the location of PS, accent marks were provided on the PS syllable to indicate whether it was a noun or a verb. Unmodified loan words with SS and UU syllables of clearly French/ Hebrew/ Spanish and other origins were avoided as their pronunciations might be unknown to students, or they might produce them as in the original language (e.g., French “crouton”, “boulevard”). Finally, to ensure that all participants read the words with a falling intonation rather than rising, each word was presented with a capitalized initial letter and a period at the end to indicate that it is a complete phrase.

Procedure. Recordings were made in a sound-attenuated booth at BGSU. Prior to recordings, participants were asked to silently read through the list and report if they found words whose meaning/ pronunciation/ stress pattern they were not sure about. The meaning of unknown words was explained in English, but they were never modeled for the participant in order not to influence their production. Thus, in case the student wanted to double-check the accuracy of pronunciation of a specific word or its stress pattern, either the word transcribed in IPA was provided (Russian participants are generally familiar with IPA and their use in dictionaries) or, alternatively, the researcher merely indicated the acceptable pronunciation of the word from the several options that the participant had. All the words on the list that can serve both as nouns and verbs depending on the stress pattern had an accent mark placed on the stressed syllable. To ensure that participants understood the function of the accent mark, a short practice trial was conducted with accent-marked word pairs like “pérmit—permít”, “óbject—

objéct”, etc., where the participant read out the words and the experimenter merely indicated their acceptability, withholding direct modeling.

After the practice trial, participants proceeded to the experimental trial, where they were given a multi-page printed list of isolated multisyllabic words in English to read out loud. Participants were instructed to first read two words from the list silently and then say them out loud from memory twice, the second repetition thought to be a more natural, “non-read” production. The first 10 or 15 were filler items to give participants some practice, and ensure more casual and natural productions for the later-coming target words.

The lists that the participants read out contained 20 words with SS syllables, 20 words with UU syllables, and 20 words with PS syllables. In order to be able to compare UU and SS vowel acoustic features to PS, both speaker groups read an equal number of words with PS identical in its phonetic context to UU and SS syllables (e.g., “popular” (PS) and “popularity” (SS)), as well as syllable count—three for the PS syllables, which is the average of two syllables long UU words, and four-to-five syllables long SS words.

The word lists for recordings also contained 120 filler items, thus the total number of words on native Russian speaker list was 180. For native English speakers, the list was longer: in addition to the 60 UU, SS and PS words and 120 filler items, there were 60 words that were needed for the subsequent perceptual study, since only native speaker productions were used as unrelated auditory primes in the priming study. Thus the total number of words on native English speaker list was 240. These words were presented in a random order to mask the repetitiveness of their orthographic and phonological structure. The order of items on each list was counterbalanced by reversing the order of words on the second list. Three participants in each

speaker group were randomly assigned to the original order list, and three were assigned to the reversed order list.

Recordings were made using Shure SM58 high-quality microphone with a mouth-to-microphone distance of 20 cm. All twelve participants were recorded in the same recording booth using the exact same recording setup; the input volume on the computer was consistent across all participants. The sampling rate used was of 44.1 KHz, and digitized recordings were saved directly to computer hard disk as .wav files using Praat 5.0.32 speech analysis software (Boersma & Weenink, 2011).

Prior to data recording, all native English and non-native participants were asked to report any speech and language related problems, but none were reported. Next, a hearing screening was administered. The test was conducted in a quiet room at the Speech and Hearing Clinic of Bowling Green State University by a speech and hearing specialist. Participants were screened with the presentation of pure tones via headphones at 25 dB HL. Tones were presented at 500, 1000, 2000, 4000 and 8000 Hz in each ear separately. All participants passed the hearing test and proceeded to the recording session.

Acoustic Analysis. Acoustic measurements were obtained and further analyzed using Praat 5.0.32 acoustic analysis software (Boersma & Weenik, 2011). The acoustic analyses of native and non-native productions involved (i) duration, (ii) fundamental frequency, (iii) intensity and (iv) formant measures for SS, UU and PS vowels. Only the second repetition of each word — out of the two recorded — was used in the analyses as more natural-sounding. The first repetition was considered to serve as a “practice repetition” and was not included, unless the

second one was deemed unusable, e.g. due to the noise of shuffling papers, poor voice quality, or other situational factors.

Vowel duration measures were obtained from the linked spectrogram and waveform displays. The measures were based on multi-tiered acoustic criteria. First, the onset and offset of each vowel was marked by an onset or loss of energy in higher formants in the spectrogram (specifically, F2 and F3). Second, to help resolve spectral ambiguity, vowel boundaries were signaled by an accompanying dramatic increase and decrease in amplitude and/or onset/ offset of aperiodicity in the waveform (Figure 8).

Figure 8. Spectrogram of a word “domination”, spoken by a native English female participant, illustrates boundaries for vowels /ɑ/ and /eɪ/.

The boundary between a fricative and a vowel was indicated primarily by the onset and offset of frication energy, best seen in the waveform. Voiced and unaspirated stops were segmented from vowels right after the stop burst and frication, at a point where higher formant structure becomes visible in the spectrogram (see Figure 9).

Figure 9. Spectrogram of a word “robot”, spoken by a native English female participant, illustrates voiced stop consonant segmentation from vowels.

When measuring durations of vowels following voiceless aspirated stops /p/, /t/, /k/, if the higher formant structure became visible in the spectrogram, the period of aspiration was included as part of vowel duration starting from the formant onset point. According to Ohala (1992), aspiration that follows consonant burst and frication is typically colored by vowel resonances; more specifically, the “true” vowel is considered to start about half-way into the period of aspiration. This is in agreement with an observation by Stevens (1999) that the interval of aspiration extends into the time when formants have reached their steady state values for a vowel, and well-defined formant peaks corresponding to F2, F3 and F4 can be observed before the onset of glottal pulsing for the vowel. Moreover, “in the initial part of the vowel after voicing onset, the vocal folds are somewhat abducted, leading to an interval of breathy voicing” (p. 460), and this interval is identical to the one in the latter part of aspiration. Based on these considerations, a certain portion of consonantal aspiration was included in the total vowel

duration measure if a clear higher formant structure was observable in this interval (Figures 10 and 11).

Figure 10. The word “condensation”, spoken by a native English female participant, illustrates the boundary between an aspirated voiceless stop consonant and vowels.

Figure 11. The word “combat”, spoken by an American female participant, illustrates the boundary between an aspirated voiceless stop consonant and vowels.

The borders of vowels preceded or followed by approximants were determined in several ways. Vowels typically have lower third formant positions than /l/, which sometimes involved a well-defined change point (Peterson & Lehiste, 1960), and higher third formant positions than /r/. Vowel steady states and transitions were inspected in the spectrogram and vowel boundaries were placed at the midpoint of the formant transitions from a vowel to the approximant, and the midpoint of transition from the approximant to the vowel. In cases when the vowel had intrinsic energy different from /l/ and /r/, intensity could serve as an indicator of a boundary point (Figure 12).

Figure 12. The word “popularity”, spoken by a native English female participant, illustrates segmentation boundaries between approximants and vowels.

Finally, in case of a glottal stop preceding an initial vowel, the glottal stop was treated as a vowel onset since it marks the beginning of weak glottal pulsing (Thomas, 2011) and formants appear visible. In order to determine vowel offset, two co-occurring events were considered:

first, loss of energy in F2 and F3 that is reflected in the spectrogram as a point where formants become indistinct, and second, a dramatic change in the amplitude in the waveform.

To permit enhanced comparison of durations of UU, SS and PS vowels by native and non-native speakers, vowel durations were normalized for the speech rate of the talker, which was a potentially confounding factor. To this end, each participant's productions were divided by his or her individual average speech rate. In order to obtain an estimate of individual speech rates, a total of 20 spoken words were selected from those used in the production study for each participant. The 20 words were 10 UU and 10 SS syllable containing words, exactly the same and in the same order for all participants. The words were concatenated for each individual, and between-word pauses were removed from the concatenated material prior to calculating its total duration. Next, the time it took for each speaker to produce the 20 words was measured, and the individual average word duration in ms was obtained by dividing the total amount of time by 20. This individual average word duration was further used to divide all the syllable durations of each speaker, thus normalizing for variability in participants' speech rates.

After applying rate normalization, SS, UU and PS syllable durations of the English speaker group were compared directly to SS, UU and PS syllable durations of the Russian speaker group (Rus UU vs. Eng UU; Rus SS vs. Eng SS; Rus PS vs. Eng PS). To ensure that these measurement tools are reliable for replication, roughly 10 percent of tokens were re-measured for duration two months later. The two datasets were highly correlated ($r = 0.96$, $n = 160$, $p < 0.001$), indicating that acoustic measurements were consistent and reliable.

Next, vowel F1 and F2 measurements were taken for each vowel to assess the possible degree of reduction for Russian speakers. For the SS and UU vowel productions, measurements of the first formant, indicative of vowel height, and the second formant, indicative of tongue

advancement, were determined for vowels /ɑ / and /æ/. Measurements were made at vowels' midpoints in the spectrogram, using formant track readings provided by the Linear Predictive Coding (LPC) analysis in Praat. If the superimposed formant tracks in the spectrogram did not appear accurate upon visual inspection, or formants were not clear, a spectral slice (wideband FFT power spectrum) was taken at the same vowel midpoint to obtain formant values. The value of 12 coefficients, or five formants, was used for the LPC analysis.

Fundamental frequency and intensity of the target vowels was measured at the approximate midpoint of each vowel, using Praat acoustic software. Fundamental frequency was estimated with the help of the autocorrelation method, which displays fundamental frequency in Praat as autocorrelation pitch tracks. If the autocorrelation track was interrupted and readings were not available for the particular segment, F_0 was measured from vibrations per unit of time: glottal pulses were first identified in the wideband spectrogram, and the span of time between any two glottal pulse peaks was measured on the waveform. F_0 was further found with the help of a formula $F_0 = \text{number of pulses} / \text{span of time in s}$ (Thomas, 2011). Intensity was measured in decibels at the midpoint of each vowel from an intensity curve, using Praat speech analysis software's built-in intensity function.

Although both groups included an equal number of male and female speakers, variations in fundamental frequency and loudness of voice could be a potential confounding variable. In order to control for individual and gender-based variability in pitch and loudness that could potentially mask the real differences between groups, normalization was applied to both intensity and fundamental frequency scores in a manner identical for duration measures. This was done by dividing each speaker's intensity and F_0 scores with their mean intensity and F_0 , averaged across

the 20 concatenated words. Data were analyzed both on raw and normalized measures for more robust results and greater reliability.

Results and Discussion

Duration. A 2 x 2 x 3 mixed measures analysis of variance (ANOVA) was conducted with Group (Russian, English) as a between-subjects independent variable, and Vowel Type (/ɑ/, /æ/) and Syllable Type (SS, UU, PS) as two within-subjects independent variables, in order to determine differences in UU, SS and PS vowel durations between native English speaker and Russian learner of English productions. Analyses were conducted both on raw duration scores (in ms) and rate-normalized duration scores (expressed as a unitless ratio) to control for variability in speech rate across participants, and both will be reported.

Raw Duration. Figure 13 shows duration measures for raw syllable durations. ANOVA revealed a significant main effect of Group ($F(1, 18) = 679.95, p < 0.001, \eta_p^2 = 0.974$), with vowel durations produced by Russian speakers ($M = 90$ ms) being only about 64% of the duration of those produced by native English speakers ($M = 140$ ms). There was also a main effect of Syllable Type ($F(2, 36) = 94.37, p < 0.001, \eta_p^2 = 0.84$) and a significant interaction between Group and Syllable Type ($F(2, 36) = 53.54, p < 0.001, \eta_p^2 = 0.748$). However, there was no main effect of Vowel Type ($F(1, 18) = 2.89, p = 0.106, \eta_p^2 = 0.138$), no significant interactions between Vowel Type and Group ($F(1, 18) = 0.154, p = 0.699, \eta_p^2 = 0.008$) or Syllable Type and Vowel Type ($F(2, 36) = 0.574, p = 0.568, \eta_p^2 = 0.0031$), as well as no three-way (Syllable Type x Vowel Type x Group) interaction ($F(2, 36) = 0.565, p = 0.573, \eta_p^2 = 0.03$). This means that vowel type was not a factor in duration analyses.

Since there was a significant interaction between Syllable and Group, and Syllable was a factor that consists of three means, post hoc tests were performed to explore which means were significantly different. Three separate t-tests were conducted with group (Russian, English) as the independent variable and (1) SS duration, (2) UU duration and (3) PS duration as the dependent variables; a Bonferroni correction for three tests was applied. For SS syllables, analyses revealed that there was a statistically significant difference between Russian-spoken ($M = 58$ ms) and English-spoken syllables ($M = 116$ ms), where English speaker productions were double the duration of Russian speakers' productions ($t(38) = 15.68, p < 0.001$). For UU syllables, there was similarly a statistically significant difference between Russian-spoken ($M = 88$ ms) and English-spoken syllables ($M = 175$ ms), with the duration of English-spoken UU syllables about twice that of Russian-spoken UU syllables ($t(38) = 15.15, p < 0.001$). For PS syllables, however, there was no significant difference between Russian-spoken ($M = 129$ ms) and English-spoken ($M = 138$ ms) syllables ($t(38) = 1.55, p = 0.128$).

The results above indicate that Russian speakers were able to realize PS duration in a native-like manner, but they failed to use temporal patterns in a native-like manner when producing SS and UU vowels. The magnitude of SS and UU syllable temporal reductions by Russian learners of English is reflected in the large effect size of between-group differences ($\eta_p^2 = 0.974$). The patterns found in Russian speaker productions in English distinctly mirror stress patterns employed in Russian phonology, suggesting that Russian speakers of English might have transferred durational patterns from their native language to English. This is evidenced by correct realization of PS syllable duration but failure to do so for SS and UU syllables that have no equivalent in the Russian phonology.

Figure 13. Average raw durations of SS, UU and PS syllables in seconds, as spoken by Russian learners of English and native English speakers.

Figure 14. Acoustic differences between a native English speaker (upper spectrogram) and Russian learner of English (lower spectrogram) productions of SS vowel in the first syllable of a word “confiscation”.

Figure 15. Acoustic differences between a native English speaker (upper spectrogram) and Russian learner of English (lower spectrogram) productions of UU vowel in the second syllable of a word “robot”.

To find out which syllable types were temporally different *within* each speaker group (e.g., Russian-spoken SS vs. UU vs. PS), post-hoc paired samples t-tests with Bonferroni correction to maintain an alpha level of 0.05 were further conducted for Russian speakers and English speaker groups. Syllable type (PS, UU, SS) was the independent variable and raw vowel duration was the dependent variable.

Results for the Russian group showed that all three syllable means ($M = 58$ ms for SS; $M = 88$ ms for UU; $M = 129$ ms for PS) were significantly different from each other ($p < 0.001$). Specifically, SS syllables were significantly shorter than UU syllables ($t(19) = 5.90$, $p < 0.001$),

and both SS and UU syllables were significantly shorter than PS syllables ($t(19) = 17.60, p < 0.001$ and $t(19) = 7.77, p < 0.001$). This relationship can be best expressed as $SS < UU < PS$.

The results obtained in the English speaker group were somewhat different. Follow-up t-tests, adjusted with Bonferroni correction, revealed that all three syllable types significantly differed from each other in terms of duration for English speakers as well. However, the relative durations produced for UU, SS and PS syllables were different for the English speaker group than for Russian learners of English. SS syllables ($M = 116$ ms) were on average shorter in duration than PS syllables ($M = 138$ ms) ($t(19) = 3.77, p = 0.001$), consistent with findings from Russian speaker data. However, UU syllables ($M = 175$ ms) were the longest in duration, greater than SS ($t(19) = 9.46, p < 0.001$) and even PS syllables ($t(19) = 5.29, p < 0.001$). This relation can be expressed as follows: $SS < PS < UU$, compared to Russian $SS < UU < PS$.

Rate-normalized Duration. In order to ensure that group differences in duration were robust to variation in speech rate, rate-normalized syllable durations were subjected to ANOVA as well. Statistical analyses of rate-normalized duration revealed results identical to the ones obtained from raw durations. In particular, vowel durations produced by Russian speakers were on average shorter ($M = 90$ ms) than those produced by native speakers ($M = 140$ ms), yielding a significant main effect of Group ($F(1, 18) = 580.48, p < 0.001, \eta_p^2 = 0.97$). Moreover, there was a significant effect of Syllable Type ($F(2, 36) = 97.83, p < 0.001, \eta_p^2 = 0.85$), a significant interaction between Group and Syllable Type ($F(2, 36) = 56.14, p < 0.001, \eta_p^2 = 0.76$), but no main effect of Vowel Type ($F(1, 18) = 1.89, p = 0.185, \eta_p^2 = 0.095$), no significant interactions between Vowel Type and Group ($F(1, 18) = .10, p = 0.759, \eta_p^2 = 0.005$) or Syllable Type and Vowel Type ($F(2, 36) = .64, p = 0.531, \eta_p^2 = 0.035$), as well as no three-way (Syllable Type x Vowel Type x Group) interaction ($F(2, 36) = .60, p = 0.557, \eta_p^2 = 0.032$).

Again, three separate independent-samples *t*-tests were conducted with group (Russian, English) as a between-subjects independent variable and (1) SS duration, (2) UU duration and (3) PS duration as dependent variables, using Bonferroni corrections for multiple comparisons. Identical to the results on raw duration, analyses revealed that English-spoken SS syllables ($M = 113$) were double the duration of Russian-spoken SS syllables ($M = 60$) ($t(38) = 14.58$, $p < 0.001$). Likewise, English-spoken UU syllables ($M = 170$) were double the duration of Russian-spoken UU syllables ($M = 88$) ($t(38) = 14.92$, $p < 0.001$). Finally, PS syllable durations did not differ for Russian speaker ($M = 131$) and English speaker ($M = 135$) groups ($t(38) = 0.78$, $p = 0.441$).

Finally, the within-group analyses revealed the same pattern of relative duration among the different syllable types in each group as shown in the raw analyses. Within the Russian group, all three means—60 for SS, 88 for UU, and 131 for PS syllables—were significantly different from each other ($p < 0.001$), SS syllable duration being the shortest, followed by UU and PS syllable durations. For the English speaker group there was again a significant difference between all three syllable durations ($p \leq 0.001$). UU syllables were the longest ($M = 170$), followed by PS syllables ($M = 135$) and SS syllables ($M = 113$).

The finding that UU syllables in both groups were longer than SS syllables, and for English speakers even longer than PS syllables, is not consistent with Fear et al.'s (2005) study results. Fear et al. found that UU syllables were significantly shorter in duration than SS syllables. Such a discrepancy can be explained by the fact that Fear et al.'s study analyzed words in phrase non-final positions, while the current study used words produced in isolation. Final syllables in phrase-final words or words spoken in isolation—as in this study—typically undergo word-final lengthening to indicate phrase boundaries (Klatt, 1976; Turk & Shattuck-Hufnagel,

2007). Out of all syllable types, only UU syllables occupied absolute word-final positions in this study and were thus the most susceptible to word final lengthening effects. Word or phrase final syllables undergoing lengthening have been reported not only in the English language (Turk & Shattuck-Hufnagel, 2007) but Russian as well (Bondarko, 1998; Volskaya & Stepanova, 2004). Therefore, as the results suggest, both speaker groups may have applied it to some degree.

SS syllables were found to be shorter in duration than the phrase-stress-bearing PS syllables in both groups, which is consistent with findings for English by Mattys (2000) and Fear et al. (1995). At the same time, it should be noted that broader generalizations about SS, UU and PS syllable durations might not warranted, as syllable count within each word may have been a confounding factor and contributed to the durations observed. Research has shown that syllable duration is inversely proportional to the number of syllables within a word (Lehiste, 1972, 1973). In this study, there was some variability in the number of syllables within a word—SS syllables were measured in four-to-five-syllable long words, PS syllables were measured in three-syllable long words, and UU syllables in two-syllable long words. This was a limitation partially imposed by the focus of the study, since word-final UU syllables specifically in disyllabic words present a common context most prone to reductions by Russian speakers. The variability in syllable count within the words used was partially a limitation of the English language itself, since SS syllables only appear in long multisyllabic words, making straightforward duration comparisons with UU syllables difficult.

Intensity. A 2 x 2 x 3 mixed measures analysis of variance (ANOVA) was conducted with Group (Russian, English) as a between-subjects independent variable, and Vowel Type (/ɑ/, /æ/) and Syllable Type (SS, UU, PS) as two within-subjects independent variables in order to

determine differences in UU, SS and PS vowel intensity between native English speakers and Russian speakers of English. To ensure greater reliability of the statistical analyses, ANOVA was conducted both on raw intensity values, measured in dB, and intensity-normalized values, expressed as a ratio, and will be described in that order.

Raw Intensity. Figure 16 shows raw intensity durations for native English and Russian learner of English groups. Statistical analyses of raw intensity showed that there was no significant main effect of Group ($F(1, 18) = 1.95, p = 0.18, \eta_p^2 = .098$); thus, overall, the intensity of SS and UU syllables produced by Russian speakers was comparable to native English speaker productions. As expected, there was a significant main effect of Syllable Type ($F(2, 36) = 396.10, p < 0.001, \eta_p^2 = .957$); post-hoc examination with Bonferroni adjustment revealed that UU vowels ($M = 61.23$ dB) were of significantly lower intensity across both speaker groups ($p < 0.001$) than either PS ($M = 68.60$ dB) or SS ($M = 68.66$ dB) vowels, which did not differ in intensity ($p = 1.00$). A significant main effect was also found for Vowel Type ($F(1, 18) = 7.814, p = 0.012, \eta_p^2 = 0.303$), whereby the vowel /a/ ($M = 66.72$ dB) was overall more intense than /æ/ ($M = 65.61$ dB).

Next, two two-way interactions were found to be significant. First, there was a significant interaction between Syllable Type and Vowel Type ($F(2, 36) = 13.94, p < 0.001, \eta_p^2 = .436$). In SS and UU syllables, the vowel /a/ and the vowel /æ/ did not differ in intensity; however, in PS syllables, the vowel /a/ ($M = 70.37$ dB) was found to be of greater intensity than /æ/ ($M = 66.83$ dB). The other significant interaction was between Syllable Type and Group ($F(2, 36) = 14.80, p < 0.001, \eta_p^2 = .451$). Post-hoc analyses, adjusted for multiple comparisons with Bonferroni correction, revealed that SS syllable intensity differed significantly for Russian speaker ($M = 67.91$ dB) and English speaker groups ($M = 69.40$ dB) ($t(38) = 2.967, p = 0.005$). However,

Russian speaker UU syllable raw intensity ($M = 61.80$ dB) was not significantly different from English speaker UU syllable intensity ($M = 60.63$ dB) ($t(38) = 1.82$, $p = 0.77$); likewise, Russian speaker PS syllable intensity ($M = 69.35$ dB) was not significantly different from English speaker PS syllable intensity when adjusted with Bonferroni correction ($M = 67.84$ dB) ($t(38) = 2.19$, $p = 0.035$). This means that Russian learners of English managed to produce PS and UU syllables in a native-like manner; SS syllables, however, were not produced in a native-like way by Russian learners of English.

Figure 16. Raw intensity values (in dB) for Secondary-stressed (SS), Unstressed-unreduced (UU) and Primary-stressed (PS) syllables for two groups of speakers: Russian learners of English and native English speakers.

Despite the fact that Russian-produced SS syllables were of lower intensity than English SS syllables, their intensity was still high: post hoc paired-samples t-tests with Syllable type (PS, UU, SS) as independent variables and raw vowel intensity as the dependent variable revealed that Russian speakers' SS syllables ($M = 67.91$ dB) were actually not significantly different from their PS syllables after Bonferroni adjustment ($M = 69.35$ dB) ($t(19) = 2.255$, $p = 0.036$) and

significantly different from UU syllables ($M = 61.80$ dB) in terms of intensity ($t(19) = 10.154$, $p < 0.001$). Thus, Russian-spoken SS syllable intensity occupies a somewhat intermediate position between their own UU and PS syllables. Within the English speaker group, however, paired-samples t-tests showed that SS syllables ($M = 69.40$ dB) were significantly more intense than PS syllables ($M = 67.84$ dB) ($t(19) = 2.797$, $p = 0.011$). UU syllables ($M = 60.63$ dB) were significantly lower in intensity than SS ($M = 69.40$ dB) ($t(19) = 18.151$, $p < 0.001$) and PS syllables ($M = 67.84$ dB) ($t(19) = 11.916$, $p < 0.001$).

Normalized Intensity. Statistical analyses of intensity-normalized data revealed results identical to the ones obtained from raw intensity data. The normalized intensity analyses (Figure 17) revealed a significant main effect of Syllable Type ($F(2, 36) = 400.13$, $p < 0.001$, $\eta_p^2 = .957$). Post-hoc tests with Bonferroni adjustment showed once again that UU syllables ($M = 0.936$) were of significantly lower intensity across both speaker groups ($p < 0.001$) than either PS ($M = 1.05$) or SS ($M = 1.05$) vowels. PS and SS vowels were of the same intensity ($p = 1.00$). There was no main effect of Group ($F(1, 18) = 4.22$, $p = 0.06$, $\eta_p^2 = 0.19$), which was compatible with the raw intensity analyses that did not reveal any significant differences in the overall intensity use by Russian and English speakers. Consistent with the findings of raw duration analyses, there was a main effect of Vowel Type ($F(1, 18) = 6.94$, $p = 0.02$, $\eta_p^2 = 0.278$), where the vowel /ɑ/ ($M = 1.02$) was again higher in intensity than the vowel /æ/ ($M = 1.00$). A significant interaction was found for Syllable Type and Vowel Type ($F(2, 36) = 13.98$, $p < 0.001$, $\eta_p^2 = 0.437$). Consistent with the raw intensity analyses, in SS and UU syllables, the vowel /ɑ/ and the vowel /æ/ did not differ in intensity; however, in PS syllables, the vowel /ɑ/ ($M = 1.08$) was found to be of greater intensity than /æ/ ($M = 1.02$). The interaction

between Vowel Type and Group was not significant ($F(1, 18) = 0.256, p = 0.62, \eta_p^2 = 0.014$), indicating that both speaker groups maintained the same differences in vowel intensity.

Figure 17. Normalized intensity values (expressed as a unitless intensity ratio) for Secondary-stressed (SS), Unstressed-unreduced (UU) and Primary-stressed (PS) syllables for Russian learners of English, and native English speakers.

Next, there was a significant interaction between Syllable Type and Group ($F(2, 36) = 16.76, p < 0.001, \eta_p^2 = 0.482$), which is in agreement with the results obtained from the raw intensity analysis. To elucidate the significant interaction between Syllable Type and Group, three independent samples t-tests with Bonferroni correction were conducted with group (Russian, English) as the independent variable and (1) SS intensity, (2) UU intensity and (3) PS intensity as the dependent variables. UU intensity was the same for Russian ($M = 0.94$) and English speakers ($M = 0.93$) ($t(38) = 0.39, p = 0.699$) and so was PS syllable intensity with $M = 1.05$ for Russian speakers and $M = 1.04$ for English speakers ($t(38) = 0.884, p = 0.383$). However, the intensity of SS syllables again differed significantly for Russian speaker ($M =$

1.03) and English speaker groups ($M = 1.07$) ($t(38) = 5.204, p < 0.001$), indicating that Russian speakers did not produce SS syllables in a native-like manner.

Nevertheless, in spite of Russian SS syllables being less intense than native English speakers', they were still of high intensity. Analyses using paired samples t-tests, adjusted with Bonferroni correction, were conducted on Russian speaker productions with Syllable type (PS, UU, SS) as the independent variable and normalized vowel intensity as the dependent variable. Results revealed that Russian speakers' SS syllables ($M = 1.03$) were again not significantly different from their PS syllables ($M = 1.05$) ($t(19) = 2.38, p = 0.028$). However, Russian speakers' SS syllables ($M = 1.03$) were significantly different from their UU syllables ($M = 0.94$) in terms of intensity ($t(19) = 9.976, p < 0.001$).

Finally, paired samples t-tests with Bonferroni correction were conducted for English speakers with Syllable type (PS, UU, SS) as the independent variable and normalized intensity as the dependent variable. Results showed that within the English speaker group, UU syllables ($M = 0.93$) were significantly lower in intensity than SS syllables ($M = 1.07$) ($t(19) = 17.77, p < 0.001$) or PS syllables ($M = 1.04$) ($t(19) = 11.73, p < 0.001$). Consistent with the earlier results obtained from raw intensity analysis, English speakers' SS syllables ($M = 1.07$) were significantly higher in intensity than their own PS syllables ($M = 1.04$) ($t(19) = 3.025, p = 0.007$).

The results from the English speaker analyses are in agreement with the findings by Plag et al. (2011). Their study showed that when PS was in a phrase-stress accented position and SS preceded PS within a word (as in the current study), SS syllables were slightly more intense than PS syllables for native English language speakers. With regard to second language acquisition, intensity results show that Russian learners of English managed to produce PS and UU syllables in a native-like manner; this finding is not surprising, given that PS and unstressed syllable

realization employs the same stress cues in Russian, including intensity. SS syllable production, on the other hand, presents a challenging case. SS syllables are not part of the Russian language phonological system, and this fact could account for the slightly yet significantly less intensity in Russian speakers' productions, compared to native English ones.

Fundamental Frequency. A $2 \times 3 \times 2$ mixed measures analysis of variance (ANOVA) was conducted with Group (Russian, English) as a between-subjects independent variable, Vowel Type (/ɑ/, /æ/) and Syllable Type (PS, UU, SS), as two within-subjects independent variables, and fundamental frequency (F_0) as the dependent variable. ANOVA was conducted both on raw F_0 values, measured in Hz, and on normalized F_0 values, expressed as a unitless ratio coefficient, in order to establish the robustness of findings across individual variation in F_0 . A one-way ANOVA was conducted separately only for the English speaker group to find out the relative differences in syllable intensity *within* this group; this information was further used for designing stimuli for the following perceptual experiment.

Raw Fundamental Frequency. Figure 18 shows raw fundamental frequency data. The analysis of variance yielded two significant main effects. First, there was a main effect of Syllable Type ($F(2, 36) = 615.382, p < 0.001, \eta_p^2 = 0.972$); post-hoc pairwise comparisons showed that all three Syllable Types—SS, UU, and PS—were significantly different from each other ($p < 0.001$), with PS syllables having the highest F_0 , and UU syllables having the lowest. Second, there was a main effect of Vowel Type ($F(1, 18) = 20.837, p < 0.001, \eta_p^2 = 0.537$), where the low back vowel /ɑ/ had an overall higher frequency than the low front vowel /æ/. There was no main effect of Group, however; the overall fundamental frequency across SS, UU

and PS syllables and vowel types was maintained the same by both English and Russian speaker groups ($F(1, 18) = 0.021, p = 0.886, \eta_p^2 = 0.001$).

Results also showed a significant two-way interaction between Syllable Type and Vowel Type ($F(2, 36) = 5.115, p = 0.011, \eta_p^2 = 0.221$), but no significant interactions were found between Syllable Type and Group ($F(2, 36) = 2.872, p = 0.07, \eta_p^2 = 0.138$), or Vowel Type and Group ($F(1, 18) = 2.856, p = 0.108, \eta_p^2 = 0.137$). Finally, there was no three-way interaction between Syllable Type, Vowel Type and Group ($F(2, 36) = 1.661, p = 0.204, \eta_p^2 = 0.085$).

Figure 18. Raw fundamental frequency values (in Hz) for Secondary-stressed (SS), Unstressed-unreduced (UU) and Primary-stressed (PS) syllables for two groups of speakers: Russian learners of English, and native English speakers.

Finally, three separate paired-samples t-tests, adjusted with Bonferroni correction, were conducted separately for English speakers to find out the relative differences in syllable intensity *within* this group. Syllable type (PS, UU, SS) was the independent variable, and raw fundamental frequency was the dependent variable. Analyses revealed that within the English speaker group, all three syllable types significantly differed in F_0 ($p < 0.001$), with PS having the highest F_0 (M

= 180 Hz), followed by SS syllables ($M = 169$ Hz) and UU syllables ($M = 125$ Hz); this pattern of results is in agreement with other studies (Mattys, 2000; Plag et al., 2011).

Normalized Fundamental Frequency. Figure 19 shows fundamental frequency data after normalization. ANOVA on normalized fundamental frequency values yielded results identical to the raw F_0 analysis. Again, there was a main effect of Syllable Type ($F(2, 36) = 677.352, p < 0.001, \eta_p^2 = 0.974$), with post-hoc pairwise comparisons confirming that all three Syllable Types—SS, UU, and PS—were significantly different from each other ($p < 0.001$), with PS syllables representing the highest F_0 , and UU syllables the lowest. Consistent with raw F_0 analysis, there was a main effect of Vowel Type on fundamental frequency of the speaker ($F(1, 18) = 25.894, p < 0.001, \eta_p^2 = 0.59$), with /a/ produced with higher frequency than the low front vowel /æ/. However, again there was no main effect of Group, demonstrating that Russian learners of English were able to produce SS, UU and PS syllables with the same relative fundamental frequency that native English speakers ($F(1, 18) = 0.207, p = 0.655, \eta_p^2 = 0.011$).

The two-way interaction found between Syllable Type and Vowel Type was stronger after fundamental frequency normalization procedure ($F(2, 36) = 7.028, p = 0.003, \eta_p^2 = 0.281$). There continued to be no significant interactions between Syllable Type and Group ($F(2, 36) = 0.069, p = 0.934, \eta_p^2 = 0.004$), or Vowel Type and Group ($F(1, 18) = 2.259, p = 0.150, \eta_p^2 = 0.112$). Finally, there was no three-way interaction between Syllable Type, Vowel Type and Group ($F(2, 36) = 1.648, p = 0.207, \eta_p^2 = 0.084$).

These results are consistent with the raw fundamental frequency analysis in tentatively suggesting that Russian speakers are able to use fundamental frequency in English at a native-like level. However, these results should be treated with caution, in that the present analyses focused narrowly on the performance on certain isolated syllables, rather than the whole pitch

pattern. The native-like F_0 pattern is reflected in pitch alternations across syllables, such as higher pitch on SS syllables, lower pitch on the intervening unstressed syllables, and again higher pitch on PS syllables. Since the intervening unstressed syllable F_0 was not measured, there is a certain chance that Russian speakers applied higher pitch continuously to all syllables from SS to PS, unstressed ones included. During recordings, there were a number of instances where Russian participants were observed to use F_0 in idiosyncratic and arbitrary ways, apparently without much consideration for stressed syllables. Considering the fact that duration is the main stress cue in Russian, F_0 being less reliable (Bondarko, 1977; 1998), perhaps the anecdotally observed randomness in F_0 use by Russian learners of English is a reflection of its low importance as a stress cue. Thus the current F_0 results, although seemingly in conformity with native-like standards for F_0 use, need to be treated with caution until F_0 values of reduced syllables within a word are measured to reveal the full pattern of F_0 use.

Figure 19. Normalized fundamental frequency values for SS, UU and PS syllables for two groups of speakers: Russian learners of English, and native English speakers.

Finally, a separate one-way within-subjects ANOVA was conducted on normalized fundamental frequency values *within* native English speaker group to obtain data for stimulus creation in the following perceptual experiment. Syllable type (PS, UU, SS) was the within-subjects independent variable, and normalized fundamental frequency was the dependent variable. Analyses confirmed the finding revealed in the raw F_0 analysis that within English speaker group, all three syllable types significantly differed in F_0 : PS syllables had the highest F_0 ($M = 1.05$), followed by SS syllables ($M = 1.00$) and UU syllables ($M = 0.71$) ($p < 0.001$).

F1 and F2 Frequencies. Figures 20 and 21 show first and second formant values for Russian- and English-spoken UU and SS vowels, graphed separately for /æ/ and /ɑ/ vowels. Two separate analyses were performed on UU and SS vowel first formant (F1) and second formant (F2) values to identify vowel quality differences in English SS and UU vowel production by Russian and English speakers. A mixed measures 2 x 2 x 2 ANOVA with Group (Russian, English) as a between-subjects independent variable, and Vowel Type (/ɑ/, /æ/) and Syllable Type (SS, UU) as two within-subjects independent variables was conducted first on F1 frequency values. A separate ANOVA was conducted on F2 frequency values. Results of the vowel quality analyses will further be described separately for F1 and F2.

Figure 20. F1 and F2 values for the low front vowel /æ/ in SS and UU syllables, as produced by native Russian and native English speaker groups.

Figure 21. F1 and F2 values for the low back vowel /ɑ/ in SS and UU syllables, as produced by native Russian and native English speaker groups.

First Formant (F1). The analysis on F1 produced a significant main effect of Group: both speaker groups, native English speakers ($M = 839$ Hz) and Russian learners of English ($M =$

580 Hz), differed from each other in terms of F1 frequency across syllable and vowel types ($F(1, 18) = 629.08, p < 0.001, \eta_p^2 = 0.972$), yielding a large effect size. There was a significant main effect of Syllable Type ($F(1, 18) = 15.67, p = 0.001, \eta_p^2 = 0.465$), with the mean F1 overall higher for vowels in UU syllables ($M = 725$ Hz) than for vowels in SS syllables ($M = 693$ Hz). Finally, there was also a significant main effect of Vowel Type ($F(1, 18) = 7.38, p = 0.014, \eta_p^2 = 0.291$), where the low back vowel /ɑ/ ($M = 719$ Hz) was found to have an overall higher F1 and thus a more open jaw position than the low front vowel /æ/ ($M = 700$ Hz).

Vowel Type interacted with Group ($F(1, 18) = 5.95, p = 0.025, \eta_p^2 = 0.248$), as shown in Figure 22. Post-hoc paired samples t-tests showed that English speakers had a significantly higher F1 and thus a more open jaw position for the low back vowel /ɑ/ than for the low front vowel /æ/ ($t(19) = 3.335, p = 0.003$), while Russian speakers maintained the same F1 for both vowel types ($t(19) = 0.125, p = 0.902$).

Another significant interaction was observed between Syllable Type and Vowel Type ($F(1, 18) = 6.17, p = 0.023, \eta_p^2 = 0.255$); post-hoc analyses showed that the low front vowel /æ/ was produced with a lower F1 than /ɑ/ in SS syllables ($t(19) = 3.64, p = 0.002$), but not UU syllables ($t(19) = 0.453, p = 0.655$). No interaction was found between Syllable Type and Group ($F(1, 18) = 1.70, p = 0.209, \eta_p^2 = 0.086$), which means that both groups maintained the same relative differences between SS and UU syllables, shown in Figure 23. Finally, no three-way interaction was found between Syllable Type, Vowel Type and Group ($F(1, 18) = 0.713, p = 0.409, \eta_p^2 = 0.038$).

Figure 22. First formant values for Russian and English speaker groups for low vowels /a/ and /æ/, averaged across syllable types (SS, UU).

Figure 23. SS and UU vowel F1 values for Russian and English speaker groups, averaged across vowel types.

These results show that Russian and English speaker groups significantly differed from each other in their F1 when producing low vowels in SS and UU syllables. The average F1 measured for Russian learners of English was $M = 580$ Hz, which, according to the mean data for first formant frequencies of American English vowels compiled by Kent & Read (2001), falls in

the mid or high-mid vowel frequency range. Such formant values indicate that Russian learners of English, unlike native English speakers, significantly raised their tongue while producing low vowels, and that led to significant low vowel quality reductions. Moreover, Russian speakers reduced both low vowels—/ɑ/ and /æ/—to a single centralized vowel to an equal extent. English speakers, on the other hand, distinguished between the two vowels by having a higher F1 and thus a more open jaw position for the low back vowel /ɑ/ than for the low front vowel /æ/.

Across both speaker groups, the mean F1 was higher for vowels in UU syllables than for vowels in SS syllables. Thus, speakers in general tended to produce UU syllables with a more open jaw position and tongue lower in the oral cavity than SS syllables. Another finding was that the low back vowel /ɑ/ had an overall higher F1 and thus a lower tongue position than the low front vowel /æ/. This result is consistent with the mean data for first formant frequencies of American English vowels, provided in Kent and Read (2001), where F1 values for the low back vowel are indeed higher than for the low front vowel.

Second Formant (F2). Next, in order to analyze vowel second formant (F2) values that represent tongue advancement in the oral cavity, two repeated-measures 2 x 2 ANOVAs had to be performed separately for the low back vowel /ɑ/ and for the low front vowel /æ/. These two-way ANOVAs were done in place of a three-way ANOVA, due to the fact that /æ/ and /ɑ/ are expected to differ trivially in F2 as a function of their different phonological status in advancement: /æ/ is a front vowel and /ɑ/ is a back vowel, so higher F2 of /æ/ compared with /ɑ/ is naturally expected. In contrast, Vowel Type was a factor in F1 ANOVA analyses due to the fact that both /æ/ and /ɑ/ are phonologically low vowels; thus, differences in F1 for these two vowels are not trivially expected.

Figure 24 shows F2 values for Russian and English speaker-produced low back vowel /ɑ/ in SS and UU syllables. The first two-way ANOVA was conducted with Syllable Type (SS, UU) and Group (Russian, English) as within- and between-subjects independent variables, respectively, and F2 of the low back vowel /ɑ/ as the dependent variable. No main effects or interactions were found. First, there was no main effect of Group ($F(1,18) = 2.72, p = 0.117, \eta_p^2 = 0.131$), illustrating that both groups showed a comparable performance on F2 of the low back vowel. There was also no main effect of Syllable Type ($F(1,18) = 3.07, p = 0.097, \eta_p^2 = 0.146$), which means that the same tongue advancement was maintained for both SS and UU syllables. The lack of interaction between Syllable Type and Group ($F(1,18) = 1.43, p = 0.247, \eta_p^2 = 0.074$) indicates that both groups showed comparable performance on both syllable types.

Figure 24. F2 values for Russian and English speaker-produced low back vowel /ɑ/ in SS and UU syllables.

Figure 25 shows F2 values for Russian and English speaker-produced low front vowel /æ/ in SS and UU syllables. The second two-way ANOVA had syllable type (SS, UU) and group (Russian, English) as within- and between-subjects independent variables, respectively, and F2

of the low front vowel /æ/ as the dependent variable. Analyses revealed that there was a significant main effect of Group ($F(1,18) = 35.39, p < 0.001, \eta_p^2 = 0.663$); the average F2 for the vowel /æ/ for English speakers ($M = 1734$ Hz) was greater than F2 for /æ/ for Russian learners of English ($M = 1467$ Hz). Based on mean F2 of American English vowels by Kent and Read (2001), Russian speakers produced the low front vowel in a more central position in terms of tongue advancement, while native English speakers produced it in a significantly more frontal position.

Figure 25. F2 values for the low front vowel /æ/, as produced by Russian and English speakers for both syllable types (SS, UU).

No main effect was found for Syllable Type ($F(1,18) = 0.103, p = 0.75, \eta_p^2 = 0.006$), and there was no interaction between Syllable Type and Group ($F(1,18) = 0.59, p = 0.45, \eta_p^2 = 0.031$). This means that syllable type was not a factor, and both groups treated SS syllables relative to UU syllables similarly in terms of tongue advancement.

Overall, Russian speakers produced the low front vowel with their tongue in a more centralized position in the oral cavity than for native English speakers, as evidenced by the lower F2 values than for English speakers. Combined with the findings from the F1 analysis, the

formant analysis suggested that Russian speakers reduced /æ/ vowel quality to a “schwa”, both in terms of tongue height and advancement. Such reductions affected vowels in SS or UU syllables equally.

General Discussion and Conclusions

The objective of acoustic analyses presented in this chapter was to examine the phonetic implementation of two English lexical prominence degrees—UU and SS syllables—by non-native speakers, specifically, Russian learners of English. Russian speakers in English have been observed to reduce all full vowels in syllables other than PS, directly affecting the quality of SS and UU syllables. Thus, the acoustic analyses explored the phenomenon of vowel *over-reduction*—a characteristic which is common in speech of Russian L2 learner of English, but which has never been addressed in research before. To explore Russian speaker treatment of English SS and UU syllables, acoustic characteristics of Russian-accented and native English vowel productions in SS and UU syllables were measured in order to identify those segmental and suprasegmental aspects that deviated most from native English speaker productions. These differences then formed the basis for further perceptual explorations on how such deviations affect native listener speech processing described in the next chapter.

SS and UU syllable production in general has received limited attention in empirical research. No known studies have addressed non-native speaker implementation of SS syllables, and UU syllable production documented for Dutch learners of English has been addressed only indirectly (Braun et al., 2011). Moreover, whenever acquisition of segmental and suprasegmental aspects of stress degrees has been addressed, typically *insufficient* reduction of vowels in unstressed positions has been the focus (Braun et al., 2011; Flege & Bohn, 1989; Lee, Guion, &

Harada, 2006). Surprisingly, however, there are no studies which have addressed the problem of vowel over-reduction, which has been observed for Russian learners of English on SS and UU vowels specifically. Russian speakers of English constitute a relatively large group in the US, many of them in the academic environment both as students and international teaching assistants. Therefore, it was a goal of this acoustic study to obtain reliable acoustic data on Russian speaker production of vowels in English SS and UU syllables, which would provide evidence of lexical prominence acquisition patterns that might be dramatically different from those non-native patterns that had been documented in research literature before, such as insufficient vowel reduction.

The phonetic dimensions measured for Russian- and English-spoken SS, UU and PS syllables were vowel duration, intensity, and fundamental frequency, as well as vowel quality, measured in terms of vowel F1 and F2 values. Acoustic measurements were done separately for the low front vowel /ɑ/ and the low back vowel /æ/. Only low vowels were used in these analyses, since low vowels have the most potential to be reduced, according to Russian phonology (Avanesov, 1956), a fact which was expected to yield reduction effects which were most observable of all vowels.

Results showed that Russian speaker productions of both SS and UU syllables significantly differed from those of native English speakers in several aspects of the acoustic analyses. Not all phonetic dimensions were equally affected, however. The most striking differences between Russian learner of English and native English speaker SS and UU syllable productions were found in the temporal and vowel quality dimensions.

In terms of duration, Russian-spoken SS and UU vowels turned out to be only about half the duration of English-spoken SS and UU vowels, which is a dramatic contrast. Acoustic

differences in vowel quality between the two groups were of comparable magnitude. F1 values, which represent vowel height, showed that when producing low vowels in SS and UU syllables, native English speakers had an open jaw position and tongue low in the oral cavity, with F1 high. Russian speakers, on the other hand, had their jaw in a less open and tongue in a more centralized position, which lowered their F1 significantly. In terms of F1, Russian productions were closer to mid or high-mid vowels than low vowels; thus, vowel quality suffered significantly. F2, which represents tongue advancement in the oral cavity, has to be viewed separately for the low front vowel /æ/, which has higher F2 values due to more frontal tongue position, and for the low back vowel /ɑ/, which has lower F2 values and is characterized by a retracted tongue position. Russian speakers did not differ from the English group statistically on low back vowel position, despite the variability in how backed the tongue was. The more frontal production of the low front vowel /æ/, however, was replaced by a more centralized tongue position in Russian speaker productions. Paired with a half-open jaw position and short vowel duration, /æ/ in Russian speaker production becomes a truly reduced vowel that has very little in common with native-spoken low front vowel. Overall, the finding that vowel reductions were most extreme in vowel quality and temporal plane is not unusual, considering how closely related these two phonetic dimensions are, with short vowel duration affecting vowel quality due to articulatory constraints.

Vowel intensity and fundamental frequency analyses, on the other hand, showed that Russian-spoken UU and SS syllables were generally not different from English-spoken syllables on these two acoustic dimensions. These findings suggest that Russian learners were largely able to use the appropriate intensity and F_0 on these syllables. This was especially true regarding UU vowels: Russian learners of English produced these with native-like F_0 and intensity patterns. SS

syllable production, on the other hand, proved to be a more difficult case. Russian-spoken SS syllable intensity was significantly lower than that of English-spoken SS syllables, suggesting that Russian speakers failed to produce SS syllables with native-like intensity. Nevertheless, Russian-spoken SS syllables were of significantly higher intensity than their own UU vowels. What is more, Russian-spoken SS syllables were in fact identical in intensity to Russian-spoken PS and even English-spoken PS vowels. Thus, it can be argued that Russian speakers did use intensity cue on SS syllables, even if they did not demonstrate native-like performance.

So far, it could be claimed that Russian speakers have mastered native-like use of F_0 and intensity to emphasize stressed vowels in speech stream, even if those are not of primary importance. However, the fact that Russian speakers demonstrated low accuracy using temporal and vowel quality cues on SS vowels relative to native English speakers raises suspicions that the excellent performance, as evidenced by isolated SS syllable measurements, might not in fact be an accurate representation of Russian learners' F_0 and intensity cue use. It is possible that speakers in fact apply intensity and F_0 to all syllables in a linear fashion, instead of an alternating one that would let SS syllables achieve suprasegmental prominence. As mentioned in the previous sections, some speakers were observed to use pitch patterns in a gradually increasing, non-alternating manner that seemed detached from F_0 function to mark stressed syllables. Thus, unless syllables that stand between SS and PS syllables are also measured to confirm the alternating or linear use of intensity and F_0 , any further conclusions are not warranted.

So far, second language acquisition studies have shown that second language learners typically struggle with not sufficiently reducing unstressed vowels, while the opposite –vowel overreduction—has not been addressed. For example, Dutch and Spanish learners of English have been shown to use full vowels instead of reduced ones (Braun et al., 2011; Flege & Bohn,

1989). In Spanish, reduced vowels do not exist, while in Dutch reduced vowels never reduce to a schwa but only get slightly centralized. Flege and Bohn's (1989) and Braun et al.'s (2011) study results suggest that one's native language is responsible for the observed success, or lack of it, with second language phonological features. This appears to be the case with Russian learners of English as well. Russian speaker segmental and suprasegmental cue use in English closely mirrors Russian language phonological patterns that have been described in detail by Avanesov (1956) and Bondarko (1998), and which have been explored empirically by Padgett and Tabain (2005) and others. Transfer effects could account for Russian learners' success with PS in English, because English PS is a near-perfect match for Russian PS in that the two languages use nearly the same stress correlates to express prosodic prominence (Avanesov, 1956). The other two syllable types—SS and UU syllables—are more problematic. Transfer of Russian rhythmic patterns over to English would result precisely in the pattern observed in the speech of Russian learners of English—at least in terms of duration and vowel quality. Chapter 1 provided a detailed description of Russian rhythm pattern, where only one stressed syllable per multisyllabic word is allowed, with the rest of the syllables undergoing vowel reduction of some degree (Avanesov, 1956). The most extreme reductions take place in the second or third syllable before PS—corresponding to the position of English SS syllables, while the syllable immediately after PS is precisely the location of most UU syllables in English. As acoustic analyses on duration and vowel quality revealed, Russian speakers of English indeed reduced duration and quality of vowels in these syllables to a great extent, supporting native language transfer hypothesis.

Chapter 2 referred to various second language acquisition theories in order to predict and account for Russian learners' of English success with lexical degree acquisition. A review of such theories led to a prediction that Russian speakers would experience difficulties with UU and

SS syllable production. Some possible reasons for the expected low accuracy with these syllable types could be their absence from Russian phonological system, as well as their relatively weak phonological status in English. First of all, neither SS nor UU syllables are phonologically meaningful in Russian. Second, they do not appear to be phonologically contrastive in English, either, with a few exceptions. According to Jones (in Lehiste, 1970) SS is an intermediate category that is not used for phonological contrasts. That is, word meaning cannot be distinguished based on SS category alone in English. A few exceptions might exist, such as verb-adjective word pairs like “graduate” or “alternate”, where a reduced syllable contrasts with a SS syllable in the word-final position (e.g., /'grædʒuət/ vs. /'grædʒu,ert/). Word meanings are typically not distinguished based on UU syllables alone, either. A few exceptions can be found for UU syllables, such as minimal pairs “radar” and “raider”, or “audition” and “addition”. However, the very rare instances of lexical contrasts might not be sufficient to assume their phonological importance, thus preserving full vowels in SS and UU syllables might not be viewed as phonologically meaningful by Russian learners of English. Considering that fact that phonologically contrastive L2 features not used to signal phonological contrast in L1 are hypothesized to be difficult to perceive for English language learners (McAllister, Flege, & Piske, 2002), one can only imagine how much harder it is for the learner to perceive L2 features that are contrastive and phonologically meaningful neither in their L1 nor L2—which is the case of UU and SS syllables for Russian learners of English. Alternatively, it is also conceivable that L1—L2 differences in SS and UU syllable productions are noticed by Russian learners of English, but a conscious decision is made not to attend to them since they are not critical for communicative purposes.

Another possible explanation for Russian learners' of English non-native performance on SS and UU syllables could be the fact that, in general, English and Russian language rhythms might appear deceptively similar or even identical. The shared phonological features between Russian and English rhythms, such as stressed syllable realization and vowel reduction, might lead Russian learners of English into thinking that rhythms are similar enough and no further adjustments are needed. As most second language speech acquisition theories hold, similar L2 features are harder to acquire than dissimilar ones, because of smaller differences being less noticeable. This phenomenon has been addressed by Flege (1995) as "equivalence classification", or, specifically for acquisition of rhythm patterns, "rhythmic inertia in the face of subtle distinctions" by White and Mattys (2007, p. 518). White and Mattys provided evidence for that from another stress-based language that is very similar rhythmically to English—Dutch. Dutch learners of English in their study did not accommodate between L1 and L2, preferring to realize English phonological features just like Dutch. It is thus not unusual that similar results were obtained with Russian learners of English.

CHAPTER 3. PERCEPTUAL STUDY

English speakers are frequently exposed to English marked by various non-native accents, which can make communication difficult. Identifying those phonetic aspects in non-native speech that affect the ease of speech perception is of great importance for second language acquisition and teaching, as well as for psycholinguistic research. Since the production experiment revealed significant phonetic differences in the realization of SS and UU syllables between English native speaker and Russian learner of English groups, the goal of the following perceptual experiment was to identify the degree to which non-standard production of SS and UU syllables could affect native English listeners' speech perception performance.

The experiment intended to measure the amount of listener's lexical activation as a factor of the type of speech they were exposed to. This was done with the help of a lexical decision task, where listeners were presented with a series of letter strings and they had to make a decision whether each such string represented a real word or not. Research has shown that performance on this task is facilitated if listeners have already been exposed to a phonologically related word (called a "prime") shortly before (Norris et al., 2006; Tucker & Warner, 2007), e.g., people would respond faster that "pencil" is a word after hearing the same word "pencil" than after hearing an unrelated word "spy". The improvement in performance, such as accuracy and speed, due to prior experience with a phonologically related word is called priming (McNamara, 2005). A priming paradigm was selected due to its ability to maximally reflect the real-time underlying speech processing and listener's ease of accessing the mental lexicon when exposed to certain types of speech (Nicol, Swinney, Love, & Hald, 2006).

Specifically, the phonological priming paradigm is useful to assess whether the phonological form of a word that differs from its canonical form is able to prime listeners to the same extent as the canonical form or not (Norris et al., 2006). The phonological priming paradigm suits the purposes of the current study well, where the objective was to compare the extent of lexical activation of words produced with low accuracy UU and SS syllables to the lexical activation of the canonical form of a word produced by native English speakers.

Finally, priming can be measured in different modalities (auditory, visual). The current study used a cross-modal lexical priming paradigm (Nicol et al., 2006), where the activation of a visually presented word (target) was measured after an auditory presentation of a phonologically related word (prime). Listeners' response times and response accuracy to words produced with accurate and inaccurate SS and UU syllables was an indication of the role of SS and UU syllables in native listener speech processing and word recognition.

Methods

Design. The perceptual experiment used a phonological priming paradigm and a cross-modal lexical decision task. Participants first heard an auditory stimulus, and immediately afterwards saw a word appear on a computer screen. The lexical decision task required participants to make a decision whether the visually presented letter strings were real English words or not, and their responses to this question may be affected by the types of auditory stimuli they hear beforehand. A critical manipulation in the present experiment was that the auditory stimulus could have been spoken by a native speaker or by a non-native speaker. Moreover, the visually presented word might be a word which matched the auditory stimulus (an identity matching situation) or else it might be an unrelated word; based on previous research (Braun,

Lemhöfer, & Mani, 2011; Norris et al., 2006; Tucker & Warner, 2007), this difference was predicted to affect listeners' response times.

The priming paradigm in this study tests English native speakers' reaction time to a given visually presented target word with a UU or a SS syllable after having heard an identity prime spoken by (i) a native speaker with the SS or UU syllable accurately produced, (ii) a non-native speaker with SS and UU syllables inaccurately produced, (iii) a non-native speaker saying a word with artificially "corrected" SS and UU syllables, as well as after having heard an unrelated prime—a native speaker saying a word that has no particular phonological or semantic relation to the SS or UU-syllable containing words. Response times to these stimuli should show whether there are differences in activation of lexical representations depending on the speech type heard, and whether improper realization of SS and UU syllables affect native listeners' speech processing speed.

The perceptual experiment utilized a within-subjects design, with type of speech (native English, Russian Unmodified, Russian Modified (i.e., Russian with instrumentally "improved" SS and UU vowels), and Unrelated as control) and stress degree (SS, UU) as within-subjects factors. Lexical decision response time, measured in milliseconds and further expressed as the degree of priming, was the main dependent variable. Listeners were expected to recognize the visual stimulus as a word more quickly if they had just heard the same word presented auditorily; therefore, response times were overall expected to be shorter for the identity-matching conditions (native speaker productions, Russian Unmodified, and Russian Modified productions) than for non-matching conditions (Unrelated condition). Within the "matching" condition, the greatest facilitation in visual word recognition, and thus shortest lexical decision response times, should be shown by native English-spoken words. Thus, listeners were expected to be fast due to having

already pre-activated their phonological word form during the auditory presentation stage. On the other hand, when listeners first heard unrelated words, their lexical decision responses were expected to be longer due to the word not being pre-activated at the auditory stage. In this case, lexical activation started much later, specifically, at the visual presentation stage. Response times to Russian-spoken words depended on how big of a role SS and UU syllables played for lexical access. If the phonological form of Russian-spoken words was perceived as deviant enough to interfere with lexical access, response time might be just as great as for unrelated primes.

Russian-spoken words with instrumentally “improved” SS and UU vowels, on the other hand, were hypothesized to show more facilitation and result in shorter response times than unmodified Russian-spoken words, yet still less priming than native speaker productions due to the possible segmental deviations. The difference in facilitation between Unmodified and Modified Russian productions as primes will allow to directly measure the effect of inaccurately produced UU and SS syllables on speech processing, which is of importance for pronunciation instruction.

Participants. The participants in these experiments were 28 native speakers of American English, which was a different group than the participants recruited for the acoustic part of the study. All of them were undergraduate students at Bowling Green State University. Prior to participation in the experiments, participants were asked to complete a questionnaire on their linguistic background to ensure that they had no knowledge of Russian and limited previous exposure to Russian speakers that could aid them in recognizing Russian-accented speech, as well as no speech- or language-related problems (see APPNEDIX D). A standard hearing screening was administered prior to the experiment by a specialist at the BGSU Speech and Hearing Clinic to test for hearing impairments; all participants passed the test successfully.

Stimulus materials. As described in the Methods section, each critical trial in the priming study involved one of four types of auditory stimulus: (i) native English-spoken words, (ii) unmodified Russian-spoken words with low accuracy SS or UU syllables, (iii) modified Russian-spoken words with acoustically “improved” SS and UU syllables, and (iv) native English-spoken words that were unrelated to the visual target.

The third type of experimental stimulus was the only one that was acoustically modified. Modified productions of Russian were necessary in order to isolate the perceptual effects of accented SS and UU syllables from those of accented speech in general. In other words, if great response latencies were observed when perceiving Russian speech, they could be due to improperly produced SS or UU vowels, but could just as well be due to other accented sounds in the same word. Thus the overall degree of accentedness of Russian-produced speech in English (besides UU and SS syllables) was a potentially confounding factor that needed to be controlled for. To address these concerns, a special condition was created where only SS and UU syllables were acoustically improved in Russian original productions; increase or decrease of response time in response to these stimuli, when compared to un-modified Russian productions, would be a direct measure of the effect that solely SS and UU syllables exert on listener perception. If SS and UU syllables are critical for the ease of speech processing, RTs should be shorter for the acoustically modified Russian-spoken stimuli with properly realized target syllables than for the original Russian-spoken stimuli with improperly realized target syllables.

The acoustic modifications were executed with Praat acoustic analysis software (Boersma & Weenink, 2011). The deviant SS and UU vowels in Russian-accented speech were substituted with the same individual’s own productions of PS vowels; since Russian-spoken PS syllables contained full vowels, it was predicted that with certain suprasegmental adjustments these

vowels could be effectively substituted for SS or UU vowels. Thus for each reduced and thus deviant SS and UU vowel, a matching PS vowel of identical vowel type and phonetic environment was selected from the recordings of the same Russian speaker and spliced out. The recordings used for these purposes were mainly productions of PS syllables in 3-syllable words, but they also could have been any other full vowel in the recorded material that would be an acceptable substitute for the reduced SS and UU vowel in case the first PS syllable was unsatisfactory. The excised vowel was further spliced into the multisyllabic word production by the same Russian speaker, in place of the improperly produced SS or UU vowel. Cutting and splicing were done at zero crossings and with careful observation of the pitch period pattern that allowed to avoid clicks and other unnatural speech effects. Ensuring a smooth transition when splicing in vowels next to approximants that do not have clear vowel/consonant boundaries was particularly challenging, as slight discontinuities were immediately noticeable. A number of excised vowels were tested to find the best match. Additionally, a method was applied whereby a bigger portion of a vowel was spliced in, and pitch periods were gradually removed till the transition from a vowel to consonant, or vice versa, sounded natural.

Vowel F1 and F2 were not altered through resynthesis in Praat; rather, authentic full vowels were simply spliced in to replace the reduced ones. Perceptually maximally accurate non-native productions were selected. Vowel duration, intensity and fundamental frequency were the three parameters that were acoustically systematically modified for the excised vowels to match average values obtained from native speakers in the production study. At the same time, the perceived naturalness of the speech was the main consideration, and if there was a slight conflict between this criterion and the established standard values for speech modifications, it was resolved in favor of the overall naturalness of speech.

Intensity, measured in dB, for each “replacement” vowel needed to be adjusted to match the relative intensity ratio with the PS vowel of that word as close as possible to native speaker UU:PS or SS:PS ratios. Russian-spoken replacement vowels that were used instead of UU vowels were all of relatively high intensity, since they came from PS syllables. They had to be lowered by approximately 10 dB to match native speaker UU:PS vowel intensity ratio (e.g., if a PS vowel was 70 dB, then UU vowel was approximately 60 dB). Average intensity for vowels was modified by using the intensity scaling option in Praat and specifying the desired dB value.

Vowel duration and fundamental frequency for the corresponding vowels were changed by using the Pitch Synchronous OverLap-Add algorithm, or PSOLA (Moulines & Charpentier, 1990), in Praat software, which allows manipulation of fundamental frequency and duration. The manipulation editor displays a Pitch Tier and a Duration Tier within which modifications can be made by either spacing pitch period windows wider apart or closer together to change pitch, or discarding or repeating some windows to reduce or increase duration. Since English-spoken SS ($M = 116$ ms) vowels were slightly shorter in duration than English-spoken PS vowels ($M = 138$ ms) both according to raw and rate-normalized measures, the same SS to PS ratio was maintained also in the modified Russian productions: replacement vowels coming from PS syllables, intended as substitutes for the reduced SS syllables, were time-altered as needed to be roughly 20% shorter than Russian PS vowels in each word. On the other hand, native English-spoken UU vowels ($M = 175$ ms) were found by the acoustic study to be approximately 20% longer than their PS vowels ($M = 138$ ms), which meant that Russian PS replacement vowels that were used in place of the reduced UU vowels were time-altered to be approximately 20% longer relative to the Russian PS vowels in each word. The main criterion, however, remained the perceived naturalness of speech.

Fundamental frequency alterations were most needed to obtain native-like fundamental frequency values for UU vowels. UU vowels were significantly lower in fundamental frequency than PS vowels for both speaker groups. In order for Russian-spoken full PS vowels ($M = 180$ Hz) to match native speaker F_0 values for UU vowels ($M = 125$ Hz), according to both raw and pitch-normalized results of the acoustic study, they had to be lowered in fundamental frequency as needed to reach approximately 70% of the PS F_0 . In cases when the F_0 of the replacement vowel used was high, a large drop in F_0 was needed for the replacement vowel to serve as an UU vowel. Overly dramatic changes in fundamental frequency could result in an unnatural speech effect in PSOLA; therefore, in such cases it was decided to go with the maximum lowest F_0 before pitch got unnatural, despite the UU syllable being slightly higher in frequency than according to native-speaker PS:UU syllable proportions. Thus again, the naturalness of speech was the main criterion for selection of modification parameters.

Finally, all experimental stimuli, as well as filler items, were normalized for peak intensity (scaled to 0.99). Each word was copied into its own file, with silence portions before and after each word removed.

Four different counterbalanced lists were created using a Latin square design, where speech conditions and target words were systematically arranged across rows and columns. Participants were randomly assigned to each list. The experimental prime-target pairs on each list were 30 words from Table 1, half of which were SS-syllable-containing words, and the other half were UU-syllable-containing words (Appendix A). These 30 words were presented to each listener as auditory primes in three speech conditions: 10 as originally produced by American native speakers, 10 as originally produced by Russian speakers, and 10 modified productions of Russian speakers. Across all four lists, 30 words were Russian-spoken, 30 were English-spoken,

and 30 were modified. In each list, the same experimental word was presented auditorily as a prime in a different order and condition (native-produced, Russian-produced, modified Russian-produced), and the same stimulus word was never heard more than once. For the experimental stimuli, the auditorily presented prime and visually presented target always consisted of the same word (i.e., “prime-target match”), e.g. participants heard the word “*confirmation*” and shortly afterwards saw “*confirmation*” on the computer screen. The speech type in which this word was first heard—English-spoken, Russian-spoken, or modified Russian-spoken, was hypothesized to affect listeners’ response times when performing a lexical decision task. Such a case represents identity priming.

In addition, a total of 30 real-word auditory primes that were unrelated to the visual primes were included as control items; for example, listeners heard a prime “*policy*” and saw “*confirmation*”. The prime-target mismatch condition served as a baseline level, where long response times were expected due to no lexical activation of the target. Furthermore, in addition to the experimental and control stimuli, 120 prime-target filler pairs were created. Since the question that listeners had to answer about the visually presented word was “Is this a real word?”, half of the words had to be real words, and half had to be pseudo-words, which could be phonotactically possible words in English. Therefore, 75 filler items had pseudo-words as targets, for example, participants heard “*finance*” and saw a pseudo-word “*plaquast*” on their screens. A very small part of pseudo-word targets (12 items) had a partial phonological overlap with the auditory prime, for example, listeners heard “*process*” and saw “*pulcess*” on their screens. This was done in order to insure that the form overlap itself was not a reliable cue to whether the target was a word or not (Norris et al., 2006).

Overall, there were 75 *real-word* visual targets (30 experimental and 45 fillers) and 75 *pseudo-word* visual targets (all filler items). Out of the real-word visual targets, only experimental items were prime-target *matching* word pairs, while the rest (control and filler items) were prime-target *non-matching* pairs. The proportion of non-matching fillers was large so that listeners would not anticipate the repetition inherent in the priming paradigm, because “priming is a process that for the most part occurs with little awareness on the part of individual language users” (McDonough & Trofimovich, 2009). Thus the role of the large number of non-matching prime-target pairs was to ensure that participants did not actively monitor auditory primes and anticipate possible visual presentation of the same item, but rather to focus on the visually presented words and the lexical decision making task, without much reliance on the auditory information.

Listeners were expected to recognize the visual stimulus as a word more quickly if they had just heard the same word presented auditorily; therefore, response times were overall expected to be shorter for the identity-matching conditions (native speaker productions, Russian original, and Russian modified productions) than for non-matching conditions (control condition). Within the “matching” condition, the greatest facilitation in visual word recognition, and thus shortest response times should be shown by a native-speaker-produced prime with native suprasegmental patterns. Russian original word productions as primes were expected to yield the least facilitation and thus the longest response times from all experimental conditions due to the accented suprasegmental and segmental patterns. Russian modified productions as primes, on the other hand, were expected to show more facilitation and shorter response times than the Russian original productions due to the instrumentally “improved” suprasegmental patterns, yet still less priming than native speaker productions due to the possible segmental

deviations. The difference in facilitation between the original and modified Russian productions as primes will allow to assess the effect of inaccurately produced UU and SS syllables on speech processing, which is of importance for pronunciation instruction.

Procedure. A visual lexical decision task was used in the experiment; under this task listeners respond whether a letter string presented on the screen is a real word of English or not. Stimulus presentation and response collection was controlled by a computer using Labview computer software program by National Instruments (<http://www.ni.com/labview/>), which allows millisecond precision in response time (RT) measurements.

Listeners were tested individually in a quiet room in front of a computer monitor. Auditory stimuli (primes) were presented over headphones at a comfortable volume, and visually presented letter strings (targets) appeared at the offset of the auditory stimuli. The letter strings were centered in the middle of the screen, written in black capital letters against white background. Listeners were instructed to respond to the visual probe by hitting a “Yes” button if they thought the visually presented targets were real words, and a “No” button if they thought the targets were not real English words. In addition, they were asked to respond as quickly and accurately as possible.

The experiment began with a short practice block of 10 trials, which familiarized participants with the auditory and visual stimuli types that they would be exposed to, such as words and pseudo-words in identity-match and no-match situations in three speech conditions. The practice portion was followed by 4 blocks of 50 experimental trials in each block. Trials in the practice block were presented in the same order for all participants. In the experimental blocks, the order of stimulus presentation was randomized differently for each participant. Each

trial began with a warning tone, followed by an auditory prime after 200ms. During the playback of the prime, the screen was blank. A visual target appeared on the computer screen at the offset of the auditory stimulus, and remained on the screen until the participant pressed either the ‘Yes’ or ‘No’ button on the keyboard. The letter strings were presented one after another with a 3 second pause between them. The 50-trial blocks were separated by pauses that participants could end by pressing the space bar. If participants needed to stop in the middle of the trials for some reason, they were again instructed to press the space bar.

Response time was measured from the onset of the visual target to the moment when the participant clicked the “Yes” or “No” button. Response times only for correct lexical decisions were considered when calculating response times (McDonough & Trofimovich, 2009).

Responses that took longer than 2 seconds were counted as not accurate and were further excluded from analyses. Finally, response times were expressed as a “priming” effect, or the difference in RT between the Unrelated condition and the three test conditions. Specifically, priming was calculated by subtracting each participant’s average lexical decision response time for each item in English, Russian Unmodified, or Russian Modified condition from that same participant’s average for items in the Unrelated condition, obtaining three different values. If there is any identity priming for a given condition, then the difference between the identity priming condition and the control (Unrelated) context should be statistically greater than zero, which represented the baseline level with no lexical activation.

Results and Discussion

A one-way repeated measures analysis of variance (ANOVA) was conducted with the Type of Speech of the auditory prime (English, Russian Unmodified, Russian Modified) as the

within-subjects independent variable, and Priming as the dependent variable. ANOVA was conducted separately for words with SS syllables, and separately for words with UU syllables. Since the Type of Speech is a factor with three levels, post-hoc tests (with Bonferroni correction to maintain an alpha level of 0.05) were performed to identify which levels of the independent variable were significantly different from each other, and from the non-priming condition. Only correct responses were included in the analyses. The correct response rate was high (98.7%).

Secondary-stressed Syllables. Results of the statistical analyses for SS syllables (see Figure 26) showed that there was a main effect of the Type of Speech heard as a prime when data were organized by participant ($F_1(2, 54) = 4.052, p = 0.023, \eta_p^2 = 0.13$), but not when they were organized by item ($F_2(2, 28) = 1.655, p = 0.209, \eta_p^2 = 0.106$). Thus listeners' response times varied significantly as a factor of the type of speech they heard only in one type of analysis, which indicates that the observed effect might not be very reliable.

Figure 26. Priming effect (in ms) after hearing words that contained Secondary-stressed (SS) syllables in 3 conditions: (i) Unmodified Russian-spoken, (ii) Modified Russian-spoken with SS syllables acoustically altered, and (iii) native English-spoken (results of data organized by participant).

Post-hoc tests for by-participant analyses, adjusted using Bonferroni correction, showed that the degree of priming after hearing different speech types as primes was gradually increasing in the following order: Russian Unmodified prime < Russian Modified prime < English prime, but not all differences were statistically significant. The two types of speech that differed significantly in their ability to generate a priming effect were Russian Unmodified and English primes. Specifically, the observed priming was significantly greater when SS-syllable-containing words were spoken by English speakers ($M = 133$ ms) than when the same words were spoken by Russian speakers in the Unmodified condition ($M = 58$ ms) ($t(27) = 2.984, p = 0.006$). This means that unmodified Russian-accented speech in general was harder to process than speech produced by native speakers. However, this effect cannot be attributed solely to improperly produced SS syllables, as there were other potentially accented segments in the multisyllabic word that might have affected listeners' response time.

The critical question in this experiment was whether correction of inaccurately produced SS syllables facilitates speech processing. Analyses revealed that the priming effect in the Russian Modified condition ($M = 102$ ms) was the same as in the English condition ($M = 133$ ms) ($t(27) = 1.021, p = 0.316$), suggesting that acoustic correction of SS syllables led to native-like priming levels. At the same time, the Russian Modified condition ($M = 102$ ms) differed in priming from the Russian Unmodified condition ($M = 58$ ms) only marginally ($t(27) = 1.849, p = 0.075$). The difference between Russian Unmodified and Russian Modified speech priming—marginally significant in this case—is most critical for answering the main research questions, as it represents the total effect that SS syllable improvement can aid in rendering non-native speech more intelligible for native listeners, with confounds like overall accentedness of speech

controlled for. It can be concluded that SS syllable correction improved priming to a certain degree, but results regarding the role of SS syllable accuracy in speech processing are mixed.

Next, separate one-sample t-tests assessed the degree of priming provided by each of the three speech types by comparing reaction times to target items preceded by a phonologically *related* word against reaction times to the same items when preceded by phonologically *unrelated* words. Phonologically unrelated words generate no priming, and response times to such words served as the no-priming baseline level against which the degree of priming in the phonologically related conditions was assessed. Russian Unmodified words with low-accuracy SS syllables were found to generate weak priming in listeners: when compared to the no-priming baseline, the difference was marginally significant ($t(27) = 1.92, p = 0.065$). This was a surprising finding, considering the fact that, first, all speakers, although accented, appeared fairly intelligible, and second, that SS-syllable containing words generally do not have much lexical competition and are quite unique, making their recognition in the absence of correctly produced vowels still feasible.

Modified Russian words exhibited a significant effect of priming, compared to the baseline level ($t(27) = 3.72, p = 0.001$). English productions, as expected, were significantly different from the baseline level, demonstrating a substantial degree of priming on the part of the listeners ($t(27) = 5.752, p < 0.001$). All of the comparisons to the baseline level were also performed on the same data organized by item, and findings were the same.

Overall, there was mixed evidence for the importance of SS syllable accuracy in non-native speech patterns. First, the actual effect of SS syllable accuracy on native speaker lexical activation was demonstrated by the difference between priming levels of Unmodified Russian-spoken words and Modified Russian-spoken words. Since the difference was marginally

significant, it can be concluded that “correcting” SS syllables did improve speech processing to some degree, but their accuracy might not be critical—or at least not more critical than the accuracy of other segments in the word. Second, unlike words in the Russian Unmodified condition, Russian Modified words generated strong priming effects in listeners; thus, it can be concluded that acoustically “correcting” SS syllables in Russian speech did make the stimuli easier to process. Third, there was a great range variability in listeners’ priming performance for the three speech conditions that led to a significant overlap in priming levels. Specifically, variability in priming was the greatest in the Unmodified and Modified Russian conditions, which indicate that not all Russian-spoken words with improper SS syllables were equally incomprehensible, or that all improved SS syllables led to native-like priming. For example, the extent of priming to Modified Russian-spoken words equaled the degree of priming for the most intelligible exemplars of Russian-accented speech, and simultaneously also the degree of priming for the least intelligible exemplars of native English speech. Finally, only one analysis (by participants) showed any significant differences in priming between the types of speech, but not the other (by item). This suggests that the observed differences might be best described as trends or tendencies toward increased priming as a factor of SS syllable accuracy.

Unstressed-unreduced Syllables. A separate one-way repeated measures ANOVA was conducted for words with UU syllables as auditory primes. The independent within-subjects variable was the Type of Speech heard as a prime (English, Russian Unmodified, Russian Modified); the degree of priming was the dependent variable. Again, since the Type of Speech is a factor with three levels, additional post-hoc tests were conducted with Bonferroni correction to maintain an alpha level of 0.05.

The ANOVA revealed that the level of priming differed significantly as a function of the type of speech heard as a prime (Figure 27). Specifically, there was a main effect of the Type of Speech, both when organized by participant ($F_1(2, 54) = 21.817, p < 0.001, \eta_p^2 = 0.447$) as well as by item ($F_2(2, 28) = 14.975, p < 0.001, \eta_p^2 = 0.517$).

Figure 27. Priming effect (in ms) after hearing words that contained Unstressed-unreduced (UU) syllables in three speech conditions: (i) Unmodified Russian-spoken, (ii) Modified Russian-spoken with UU syllables acoustically altered, and (iii) native English-spoken (results of data organized by participant).

Post-hoc tests on the data organized by participant revealed that the level of priming in the Unmodified Russian condition ($M = 34$ ms) was significantly smaller than in the English condition ($M = 133$ ms) were heard ($t(27) = 5.142, p < 0.001$), or Modified Russian condition ($M = 138$ ms) ($t(27) = 5.249, p < 0.001$). The English and Modified Russian condition did not differ: priming triggered by native English-spoken primes and Modified Russian-spoken primes was found to be of similar magnitude ($t(27) = 0.400, p = 0.692$).

Next, in order to assess the absolute priming effect that each speech type is able to generate in listeners, separate one-sample t-tests compared priming in each condition to a

condition with no priming, such as hearing a phonologically *unrelated* word. Both English and Modified Russian conditions demonstrated a clear priming effect in that they were both significantly different from the baseline “zero priming” level—($t(27) = 6.034, p < 0.001$) and ($t(27) = 5.755, p < 0.001$), respectively. This is in sharp contrast to the Unmodified Russian condition, which was not statistically different from the zero-priming baseline level ($t(27) = 1.276, p = 0.213$).

The difference between priming levels of Unmodified Russian speech and Modified Russian speech should be taken as an indication of the individual effect of UU syllables on speech comprehensibility. This difference shows that by using a full vowel instead of a reduced one in UU syllables can lead to dramatic improvements in the speed with which speech is processed by native listeners. Moreover, modifying just UU syllables (Modified Russian condition) primed listeners just as much as unaccented speech (English condition).

The importance of UU syllable accuracy is evidenced not only by the priming levels of the three speech types relative to each other, but also their absolute priming relative to the zero-priming condition. Unmodified Russian words did not prime listeners at all, comparable to the no priming baseline condition with a phonologically unrelated word as a prime; after acoustically modifying the UU syllable (Modified Russian condition), the word showed a strong priming effect, which was significantly different from the baseline condition.

The fact that Russian-spoken words with UU syllables failed to prime listeners, but the same words with acoustically improved UU syllables created a dramatic priming effect indicates that reduction of UU syllables in this study was the main reason why Russian-spoken words were poorly processed by native listeners. Thus, this study has provided a convincing piece of

evidence that UU syllables are critical for speech intelligibility, and increasing their accuracy can facilitate lexical access to a great extent.

General Discussion and Conclusions

The current study investigated factors that might contribute to difficulties processing and understanding non-native speech, specifically, the role of inaccurate productions of SS and UU syllables. The goal of the experiment was to provide direct answers regarding the degree of difficulty that improper implementation of SS and UU syllables can cause for native English listeners, and the importance of these syllables in speech perception in general. To this end, listeners' response times during a lexical decision task using auditory primes consisting of native English-spoken words, Russian-spoken words, Russian-spoken words with corrected SS or UU syllables and Unrelated words were measured and compared to see how the speed of lexical activation might differ when hearing these three speech types. The non-native speech examined was English spoken by Russian learners of English; this group of speakers has been observed to reduce UU and SS syllables both in terms of duration and vowel quality. According to the findings of the acoustic experiment described in Chapter 2, productions of Russian learners of English exhibited considerable phonetic differences in the realizations of SS and UU syllables compared to English native speaker productions. Such differences might affect native English listeners' speech perception performance to a certain extent. Since neither SS nor UU syllables appear to be phonologically contrastive in English, their improper implementation by non-native speakers of English might not completely prevent word recognition, yet it might considerably affect the speed with which words are processed.

To capture the subtle effects of on-line speech processing, the perceptual experiment used a cross-modal identity priming task, where listeners had to make a lexical decision on visual words as fast as they could after hearing the same English word spoken in the three previously mentioned conditions: native English speech, Unmodified Russian speech, and Modified Russian speech with SS or UU syllables acoustically “improved”. Response times were measured and further expressed as a priming effect, or the difference between listeners’ response time to each identity prime condition and their response time to an unrelated prime. The existence and degree of priming was supposed to show whether improper realization of SS and UU syllables interfered with lexical activation by native speakers to any degree, which is a question of importance for second language pronunciation instruction.

Results showed that, first of all, there was significantly more priming when UU and SS syllable-containing words were spoken by native English speakers than when the same words were spoken by Russian learners of English. These findings indicate that it takes significantly more time to process Russian-accented speech in general than speech produced by native English talkers. This finding is consistent with Munro and Derwing’s (1995) study that used a sentence verification task in order to assess the effect of non-native English speech on speech processing time. Results of their study showed that Mandarin-spoken utterances in general required more time to evaluate than the utterances spoken by native English speakers.

As expected, English-spoken words in the current study showed substantial priming effects also compared to the no-priming baseline level, which was the condition where reaction time to a target was measured after hearing a phonologically unrelated word. Surprisingly, however, Unmodified Russian words with UU syllables did not significantly differ from the baseline level, suggesting that native speakers either failed to activate words upon hearing them,

or activation required a considerable amount of time. While less priming due to slowed-down process of lexical access was expected for Unmodified Russian-spoken words with inaccurate UU syllables as primes compared to native English-spoken primes, the complete lack of priming indicating a great degree of interference with native listener perception was an unexpected finding. The lack of priming observed could be accounted for by the fact that UU-syllable-containing words used in this study were relatively short, with the inaccurately produced UU syllable forming half of the disyllabic word (compared to words like “auditioning” or “idolatry”, where UU is the first syllable in a four-syllable long word). As a result, there was less lexical material for the listeners to rely on in the process of lexical perception. According to Pitt and Samuel (2006), long words produce stronger lexical activation than short words, since longer words provide more bottom-up (acoustic-phonetic) evidence and generally have less lexical competitors that need to be inhibited than short words.

Pitt and Samuel’s (2006) claim that long words would generate strong lexical activation was partially supported by the results of the current study. Words with low-accuracy SS syllables showed some priming in native listeners, despite the fact that they appeared to have few lexical competitors and their uniqueness point was relatively early in the word (e.g., “fabrication”, “population”, “magnification”). SS syllable-containing words were long multisyllabic words with four or five syllables in them, where the rest of the syllables could provide enough unique acoustic-phonetic material to compensate for the inaccuracy of one syllable. The latter argument, of course, is valid only if the rest of the segments are of acceptable quality, which generally seemed to be the case for Unmodified Russian productions in this study. Overall, the degree of difficulty that native English listeners experienced when processing Unmodified Russian words with low-accuracy SS syllables was relatively unexpected.

Most critical for the research questions of the current study were the comparisons in priming effects between Russian Unmodified and Russian Modified words, and Russian Unmodified and Russian Modified words against the zero-priming baseline level. Differences between these conditions reflect the actual degree of improvement in speech processing that correcting inaccurate productions of SS and UU syllables can lead to. The analyses revealed that after replacing the inaccurate target syllables with “correctly” produced ones the degree of priming increased dramatically for UU-syllable-containing words, and to a much lesser extent for SS-syllable-containing words. Such results indicated that the accuracy of SS syllables and UU syllables in the experimental words was not equally critical for native English listeners. UU syllables in this study appeared to play a more important role in lexical access than SS syllables. This is an important finding, since so far the role of UU syllable type has been undeservedly underexplored both in speech perception and second language research literature. Quite possibly, UU syllables have been underattended due to their perceived minor effect on speech processing and word recognition; this study, however, shows that UU syllables are more critical for lexical access than previously thought. Overall, the results of the current study substantially add to the knowledge about the importance of these syllables for processing of spoken words by native speakers of English.

Fear et al.’s (1995) perceptual study showed that listeners judged reduction of vowels in UU and SS syllables as relatively unacceptable, especially in SS syllables. In contrast, the results of the current priming study suggested that reducing UU vowels had a more detrimental effect on speech perception than reducing SS vowels. One explanation for the discrepancy is that acceptability ratings might not be a reliable indicator of listeners’ needs in the process of lexical access. An alternative explanation is the fact that Fear et al. used longer UU-syllable-containing

words (e.g., “authentic”, “automata”, “idolatry”) than the ones used in the current study, providing listeners with more acoustic-phonetic material to rely on when UU syllables were inaccurate. The study by Fear et al. also used a different location for UU syllables than in the current study, specifically, they examined UU syllables in pre-PS positions rather than post-PS positions. The location of UU syllable relative to the PS thus also could have affected the results somewhat.

It is important to note that the findings of SS and UU syllable importance are limited to the specific UU and SS syllable contexts used in this study, such as SS syllables occurring in pre-PS positions of multisyllabic words, and UU syllables occurring in post-PS positions of disyllabic words. These specific positions were selected to optimally reflect vowel reduction by Russian speakers of English; however, the specific position of the target syllable as well as word length are limiting in generalizing the current findings to other contexts.

Overall, based on the analyses, recommendations for selecting instructional priorities for Russian learner of English populations can be made. First of all, it is critical that pronunciation instructors are aware of the differences between Russian and English speech rhythms. While generally relatively close in their phonetic realization, the rhythms of these two languages feature certain suprasegmental and segmental differences. UU and SS syllables are an important feature of English stress and rhythm system, but they are not part of the Russian phonological system. These two syllable types might be overlooked by Russian learners of English as phonologically non-contrastive and not critical for meaning.

This study provided evidence that native listener speech perception process suffered significantly when UU syllables, and to a certain extent also SS syllables, were not produced correctly. Lexical activation of UU syllable-containing words was seriously impeded when UU

syllables were produced by Russian speakers with a temporally short vowel, equal in quality to a reduced, centralized vowel “schwa”. On the other hand, identical production of SS syllables by Russian speakers interfered with native English speech processing less and showed some lexical activation. When UU syllables were “improved” through instrumental modification, the processing speed of such words was just as fast as when processing native English-spoken words. Thus it is recommended that special attention be paid to words that contain UU syllables, since Russian students, especially at the earlier stages of English language acquisition, might be tempted to apply Russian stress and rhythm rules to English, affecting UU syllable quality. Working with SS syllables can also improve the speed with which words are recognized by native English listeners, but to much lesser extent. Overall, this study obtained mixed evidence for the critical role of SS syllables and more research is needed to elucidate the issue. Based on the results provided by the perceptual study, it is advised that pronunciation teaching instructors pay close attention to learners’ UU syllable production while administering the initial diagnostic test to Russian speakers, and, upon noticing inaccuracies, incorporate UU syllables in the pronunciation instruction syllabus.

CHAPTER 4. SUMMARY AND CONCLUSIONS

The current studies set out to explore the role of SS and UU syllables in speech perception as part of a larger L2 pronunciation research effort to identify those phonetic aspects that are critical for native listeners and whose accuracy is essential in non-native speech. The accuracy of SS and UU syllables is often overlooked by a specific English language learner group—native Russian speakers. Therefore, this study was conducted primarily with this learner population in mind. The results of these studies would be informative regarding whether and to what degree SS and UU syllable accuracy interferes with native English listener speech processing, and therefore directly applicable to pronunciation instruction. Moreover, the results are also of relevance for psycholinguistic research regarding the phonological role of SS and UU syllables, which so far have been largely underexplored degrees of lexical prominence.

In order to answer these questions, an acoustic study was first conducted to assess the differences in SS and UU syllable production by Russian learners of English and native English speakers. Russian speakers of English have been often observed to systematically reduce full vowels except those in PS syllables. Such reductions are consistent with the phonological stress and rhythm rules of Russian, where all syllables before and after PS undergo vowel reduction of some degree (Avanesov, 1956). Reduction effects are most dramatic for low vowels several syllables before the PS syllable, and immediately after PS, which are accordingly the specific contexts of interest for this study.

To assess Russian learner of English use of UU and SS syllables in English empirically, vowels in English and Russian-spoken syllables were measured in terms of F_0 , intensity, duration and F_1 and F_2 . Results showed that the most striking differences between the two speaker groups were in the duration and quality of vowels. Russian-spoken UU and SS syllables were only

approximately half the duration of English-spoken ones, and vowel quality suffered greatly in Russian productions. Instead of producing a low back or front vowel, Russian speakers significantly centralized the vowel, rendering it “schwa”-like. Vowel formant measures indicated that Russian speakers produced both vowel types with the tongue significantly higher in the oral cavity than for native English speakers, and tongue in a significantly more central position for the low front vowel. Differences in intensity and fundamental frequency, on the other hand, were minimal between the two groups.

The patterns observed in Russian speaker productions of English UU and SS syllables and rhythm in general mirror those patterns that are characteristic of the Russian rhythmic system. According to Bondarko (1998), vowel duration and vowel quality are the most important stress correlates in Russian rhythm and stress system. Both of these stress correlates are used to realize the stress of each word, which can only be one, according to Russian phonology (Avanesov, 1956). Unstressed syllables in Russian, on the other hand, are characterized by reduced vowel duration and quality. The current study showed that Russian speakers consistently reduced vowel duration and quality on syllables around the main stress also in English, suggesting that the observed effects are the result of a phonological transfer from learners’ L1 to L2. It has been suggested that the perception of non-native phonological structures in a second language are constrained by the phonological properties of one’s native language (Best, 1995; Flege, 1995). Numerous research studies have reported the profound effect that adult learners’ L1 phonological system can exert on the perception and production of L2 suprasegmentals (Braun et al., 2011; Cutler et al., 2007; Lehiste & Fox, 1992; McAllister, Flege, & Piske, 2002; Yu & Andruski, 2010).

The theoretical models and empirical research reviewed in Chapter 1 were used to make specific predictions regarding Russian speaker success with UU and SS syllable structures in English. The first two models that could predict the observed effects were the Speech Learning Model (Flege, 1995) and the Perceptual Assimilation Model (Best, 1995). Based on the tenets of these models, phonetic dissimilarities are noticed but similarities ignored by second language learners. Extrapolating the tenets of these models to suprasegmental learning, it was predicted that the many existing similarities between English and Russian language rhythms, such as the presence of lexical stress, stressed syllable realization and vowel reduction, could make the learners view L2 rhythm as very similar to L1 rhythm. This would prevent learners from noticing subtle differences due to equivalence classification and, as a result, learning would not happen. The current results are consistent with the models and confirm the findings of an earlier empirical research study by White and Mattys (2007) with Dutch learners of English, who failed to notice subtle rhythmic differences between L1 and L2 rhythms due to Dutch and English being rhythmically similar.

Russian learners' difficulty with English SS and UU syllable features was indirectly predicted also by the Feature Hypothesis (MacAllister, Flege, & Piske, 2002), which states that "L2 features not used to signal phonological features in L1 will be difficult to perceive for the L2 learner and this difficulty will be reflected in the learner's production" (p. 230). Given that PS and UU syllables are phonologically contrastive *neither* in learners' L1 *nor* L2, it seems reasonable to assume that learners will experience difficulties beyond those predicted by the Feature Hypothesis. Jones (as cited in Lehist, 1970) claimed that SS is not used for phonological contrasts; i.e., word meaning cannot be distinguished based on the SS category alone. The same seems to refer to UU syllables as well. While certain exceptions exist, overall,

neither SS nor UU syllables appear to be phonologically contrastive in English. Thus, it is possible that the observed reduction of SS and UU syllables is the result of a somewhat conscious decision on part of the learners to disregard structures that are not phonologically meaningful.

Finally, the Markedness Differential Hypothesis (Eckman, 2004) predicted that Russian learners will experience difficulty with SS syllables, but was less clear regarding UU syllable acquisition. The hypothesis holds that L2 phonological categories which differ from native language categories and that are less common in world's languages (more marked) will be difficult to learn, which might be the case with SS syllables. The markedness of UU syllables, however, is relatively unclear, not allowing specific predictions to be made. Overall, the current study only focused Russian speaker production of English syllable prominence degrees. Studies investigating Russian learner perception would be beneficial to explore the issue further and identify those mechanisms that second language learners use when exposed to new phonological features that are acoustically prominent but phonologically not contrastive in L2.

In the current study, a perceptual experiment further assessed the effect of temporal and vowel quality reductions on native speaker speech processing. A phonological priming experiment with a lexical decision task was conducted to implicitly assess the degree of lexical activation native speakers exhibited as a function of speech they heard. Listeners first heard either phonologically related native English-spoken words (English condition), Russian-spoken words with low-accuracy SS and UU syllables (Unmodified Russian condition), Russian-spoken words with accurate SS and UU syllables (Modified Russian condition), or phonologically unrelated English-spoken words. Upon visual presentation of the same word shortly afterwards, listeners were expected to be fast in responding that it was a real word after hearing native

English-spoken words. Faster responses were due to having pre-activated the word at the auditory stage of its presentation. Listeners were expected to be slow after hearing unrelated words due to no pre-activation. The Unmodified Russian condition was expected to yield longer response times than the English condition, but after acoustic modifications (Modified Russian condition) the response times were expected to approach the levels of native-spoken words. Response times were further expressed as “priming”, or the degree of lexical activation in response to the each speech type heard.

The priming results were largely consistent with these predictions; however, some aspects were unexpected. Unmodified Russian words with inaccurate UU syllables were found to trigger practically no priming in native listeners, indicating that words such as “abstract” or “impact”, produced with a short and centralized final vowel, were either not recognized or the speech recognition process was considerably delayed. Unmodified Russian words with inaccurate SS syllables, on the other hand, showed some priming, possibly due to being long multisyllabic words with relatively few lexical competitors that made it easier to recognize them even with inaccurate SS syllables. The degree of priming listeners demonstrated in the Unmodified Russian condition can only account for the effect of Russian-accented speech in general, but not the individual effect of SS and UU syllables.

The Modified Russian condition, however, was able to separate the effect of SS and UU syllable accuracy from the accuracy of the rest of phonetic segments within the same word by increasing *only* the SS and UU syllable accuracy and assessing the improvement in priming levels. Priming for Modified Russian words with acoustically “corrected” SS syllables did not yield strong improvement in priming levels, possibly due to a “ceiling” effect, since Unmodified Russian-spoken words with inaccurate SS syllables were somewhat recognizable in the first

place, perhaps due to their greater length (Pitt & Samuel, 2006). Thus, a conclusion can be made that SS syllable accuracy might not be as essential for perceptual recovery of word identity as previously thought. UU syllables, on the other hand, proved to be relatively more critical for speech processing and lexical access. In contrast to SS syllables, priming for Modified Russian words with acoustically “corrected” UU syllables were recognized by native listeners maximally quickly and reached native-English speech priming levels.

The detrimental effect that vowel quality and temporal reduction of UU syllables had on native English listener speech processing is an indication that Russian speakers of English should pay special attention to accurate implementation of this syllable type. The finding that UU syllable accuracy can undermine native listener lexical access is enlightening, since the role of this syllable type for word recognition might have been underestimated and overlooked by Russian speakers—despite the ubiquity of SS and UU syllable reductions in Russian-accented speech. The findings of this study are of relevance for pronunciation instructors working with Russian-speaking individuals. UU syllable inclusion in the pronunciation syllabus is recommended due to their demonstrated interference with speech processing under the conditions implemented in the perceptual study. The mixed results obtained for SS syllable importance suggest that this syllable type might be relatively less critical for speech processing than UU syllables; however, SS syllable accuracy could still facilitate speech perception. Given the time constraints for pronunciation teaching that most ESL instructors experience, it is recommended to include UU syllable accuracy among the pronunciation goals for Russian learners of English.

In the current pronunciation research and pedagogy, speaker intelligibility—rather than foreign accent elimination and sounding native-like—is viewed as the main goal of

pronunciation teaching (Derwing, 2010; Levis, 2005). The intelligibility principle, described by Levis (2005), implies that certain phonological features and pronunciation errors impair intelligibility of a speaker more than others. Traditionally, teachers have relied on their intuition to decide which errors are most detrimental for speech processing; however, for teachers to rely on their intuition is argued to be inefficient and unreliable. Practical research, on the other hand, can help instructors decide where to put the focus, therefore more empirical research is called for in order to identify the importance of various phonological structures for intelligibility (Derwing, 2010; Derwing & Munro, 2005; Hahn, 2004; Munro & Derwing, 2006). To determine which suprasegmental aspects of speech are most critical for intelligibility, isolating particular suprasegmental features for analysis would be helpful (Hahn, 2004). So far, for example, research on suprasegmentals has found that erroneous sentence stress affects native listener comprehension (Hahn, 2004); inaccurately placed or realized primary stress has been shown to be the source of unintelligibility and increased processing time (Bond & Small, 1983; Cutler & Clifton, 1984; Field, 2005; Small, Simon, & Goldberg, 1988). Not sufficient reduction of reduced vowels in unstressed syllables was found to increase processing time in studies using isolated word fragments as stimuli (Braun, Lemhöfer, & Mani, 2011). Finally, non-native speaker intelligibility can be affected by erroneous temporal patterns (Quené & van Delft, 2010; Tajima, Port, & Dalby, 1997) and overall speech rate (Anderson-Hsieh & Koehler, 1988; Kang, 2010; Munro & Derwing, 1998). The importance of UU and SS syllables for intelligibility and comprehensibility, on the other hand, has never been addressed, possibly due to the common assumption that inaccuracies in production of these syllable types would minimally affect word recognition. The current research, however, revealed that UU syllables can be critical for native English listener speech processing and lexical access, despite typically not being phonologically

contrastive. Overall, this study has contributed to the relatively limited body of empirical research on which suprasegmental aspects of speech are of greatest importance for accurate perception of non-native productions, and the results are therefore informative for pronunciation pedagogy.

The results of this study are relevant also for psycholinguistic research. The priming study findings suggested that listeners might not rely on SS syllables while processing speech as much as previously thought, in spite of the hypothesized important prosodic status of stressed syllables relative to unstressed. Improvement of SS syllable accuracy was expected to make the word substantially more recognizable, since SS and PS syllables might be those that listeners use as a code for retrieving words from memory (Grosjean & Gee, 1987). Yet, SS syllable modification improved speech processing only slightly, demonstrating that functionally they did not carry more weight than other syllables. Moreover, Unmodified Russian productions with inaccurate SS syllables were relatively recognizable by listeners in the “by participant” analyses and clearly recognizable when data were organized “by item”. For the latter, there was no significant difference in priming for English, Russian Modified and Russian Unmodified speech, suggesting that inaccurate productions of SS syllables did not prevent successful word recognition. Thus, this experiment did not provide evidence that SS syllables would serve as a “navigational guide” for accessing words in the mental lexicon, which is what word-level stressed syllables collectively are argued to do (Grosjean & Gee, 1987; Murphy, 2004; Murphy & Kandil, 2004), since words were somewhat recognizable even in the absence of correctly produced SS syllables.

It is possible that word length per se might partially account for the observed effects of UU and SS syllables on processing. Pitt and Samuel (2006) showed that long words provided

more acoustic-phonetic evidence and generated greater lexical activation that was available both earlier and for a longer time than shorter words. Word length effects in the current study can be seen in the great differences in lexical activation for UU and SS syllable-containing words, and the levels of activation after acoustically improving them. Each UU syllable was part of a disyllabic word; inaccuracy of an UU syllable would result in a single accurate syllable left in the word for listeners to use in speech processing. Words with SS syllables, on the other hand, were multisyllabic, typically consisting of 4-5 syllables. Inaccurate production of the SS syllable in such a word would still leave at least three accurately produced syllables; collectively, these other syllables might contain enough acoustic-phonetic information for the speaker to successfully recognize the word even when SS is inaccurate. Thus, word length is one possible reason why UU syllable inaccuracy might have shown stronger interference with speech processing than SS syllable inaccuracy.

It should be noted, however, that conclusions about the importance of SS and UU syllable accuracy for perception should be tempered based on those lexical contexts that were specifically addressed in this study. The varying word length of SS- and UU-containing words is one factor which constraints broader generalizations. While SS syllables always appear in the English language in long multisyllabic words, UU syllables are not restricted to disyllabic words as used in this study but may occur in English words of different lengths (e.g., “auditioning” or “idolatry”, where UU is the first syllable in a four-syllable long word). Only disyllabic words were selected for this study to match those phonological environments that would feature strongest vowel reductions in Russian, and the demonstrated effects on speech processing are therefore limited to such contexts.

Some other limitations of the study are the restrictions on the specific locations of target syllables within experimental words, as well as the restricted set of vowels in SS and UU syllables in experimental words selected for the experiments. Plag et al. (2011) made a clear distinction between SS syllables in pre-PS position (called *right-prominence* words, such as “violation”) and SS syllables in post-PS position (*left-prominence* words, such as “violate”). The study showed that in right-prominent words (as in this study), SS syllables were acoustically more prominent than in left-prominent words. Since the impact of inaccurate SS syllable reductions on speech processing might depend on the location of the syllable, the current findings should be limited to right-prominent words only. Moreover, the current study used only low vowels /ɑ/ and /æ/, since those were described as having most potential to be reduced in terms of vowel quality and duration in Russian. This study, however, did not test items with high or mid-high vowels in critical syllables on production and perception. The specific syllable locations, word lengths and syllable types were carefully selected to maximize the potential vowel reduction effect that Russian learners of English would exhibit in their speech. Having demonstrated intriguing results with unexpectedly large effect sizes in these particular contexts, this study calls for follow-up studies that will address those contexts that were outside the scope of this study to provide complete and generalizable results.

This research also does not make any claims about how widespread the observed SS and UU vowel reduction phenomena are in Russian learner of English populations. Similarly, it cannot answer the question of at what proficiency levels UU and SS vowel reduction effects are most pronounced, or how much exposure to native English is needed for SS and UU syllable productions to reach native-like accuracy. The sample size used in the acoustic study was relatively small, and variability in the English language learning background of the recruited

Russian participants was not controlled for. Nevertheless, the study identified significant vowel reductions as a common feature for all six Russian participants, regardless of the length of time spent learning English as a foreign language. Thus, there is solid evidence that vowel reduction does exist in the speech of Russian learners of English, even for more advanced speakers.

Further avenues for research can be identified based on this study. First, given the results of this study, follow-up studies with UU and SS syllable locations within words, word length and vowel types other than those used for this study would be beneficial to provide more evidence of the importance SS and UU syllables for word recognition. Second, evidence shown here suggests there is some benefit to producing SS syllables with accurate suprasegmentals, but the results are mixed. A repeated experiment with more participants and different sets of experimental items could make the results more clear-cut regarding the importance of SS syllables. Third, exploring Russian non-native speaker perception of SS and UU degrees in English would be a natural next step to explain the effects that were observed in the acoustic study. Such explorations would elucidate the issue of whether Russian speakers are perceptually attuned to their native language phonology, which features only one prominence in a lexical word, to such an extent that they do not perceive other prominence degrees in English words. Alternatively, learners might be able to detect differences in prominence, but at some lexical processing stage discard them as phonologically not meaningful and irrelevant for communicative purposes. An AX or AXB discrimination task with properly and improperly produced UU and SS syllables might provide answers to these questions.

Overall, the study has provided acoustic evidence for Russian learners of English reducing the quality and duration of those vowels that are normally never reduced, specifically, vowels in UU and SS syllables. This phenomenon can be referred to as vowel *over-reduction*,

and appears to be the result of transferring Russian stress and rhythm patterns over to English. Vowel over-reduction so far has not been addressed in second language research, and contrasts with a more common and more widely researched phenomenon—insufficient vowel reduction. Moreover, this study provided evidence that inaccurate production, or reduction of vowels in UU syllables, which is characteristic of speech of less advanced Russian learners of English, can have detrimental effects on native listener speech processing. SS syllable accuracy for speech processing was found to be less critical. Therefore, it is recommended that UU syllables are included in the pronunciation syllabus for Russian learners of English.

REFERENCES

- Abercrombie, D. (1967). *Elements of general phonetics*. Chicago, IL: Aldine Publishing Company.
- Anderson-Hsieh, J., & Koehler, K. (1988). The effects of speaking rate and foreign accent on native speaker comprehension. *Language Learning*, 38(4), 561–613.
- Anderson-Hsieh, J., Johnson, R., & Koehler, K. (1992). The relationship between native speaker judgments of nonnative pronunciation and deviance in segmentals, prosody, and syllable structure. *Language Learning*, 42(4), 529–555.
- Archibald, J. (1993). *Language learnability and L2 phonology: The acquisition of metrical parameters*. Dordrecht: Kluwer Academic Publishers.
- Avanesov, R. I. (1956). *Fonetika Sovremennogo Russkogo Literaturnogo Jazyka (Phonetics of Modern Russian Language)*. Moscow: Izdatelstvo Moskovskogo Universiteta.
- Baker, R. E., Baese-Berk, M. M, Bonnasse-Gahot, L., Kim M., Van Engen, K., & Bradlow A. (2011). Word durations in non-native English. *Journal of Phonetics*, 39(1), 1–17.
- Barnes, J. (2004, January). *Vowel Reduction in Russian: The Categorical and the Gradient*. Paper presented at the Annual Meeting of the Linguistic Society of America, Boston, MA.
- Barry, W. J., Andreeva, B., Russo, M., Dimitrova, S., & Kostadinova, T. (2003). Do rhythm measures tell us anything about language type? In M. Solé, D. Recasens, & J. Romero (Eds.), *Proceedings of the 15th International Congress of Phonetics Sciences* (pp. 2693–2696). Barcelona: Casual Productions.

- Beckman, M. E. & Edwards, J. (1994). Articulatory evidence for differentiating stress categories. In P. A. Keating (Ed.), *Papers in Laboratory Phonology III: Phonological Structure and Phonetic Form* (7-33). Cambridge: Cambridge University Press.
- Beckman, M. E. (1986). *Stress and non-stress accent*. Dordrecht: Foris.
- Best, C. T. (1995). A direct realist perspective on cross-language speech perception. In W. Strange (Ed.), *Speech perception and linguistic experience: Theoretical and methodological issues in cross-language speech research* (pp.167–200). New York, NY: Timonium.
- Best, C., & Tyler M. (2007). Nonnative and second-language speech perception: Commonalities and complementarities. In O.-S. Bohn & M. Munro (Eds.), *Language experience in second language speech learning: in honor of James Emil Flege* (pp. 13-34). Amsterdam: John Benjamins Publishing.
- Best, C., McRoberts, G., & Goodell, E. (2000). Discrimination of non-native consonant contrasts varying in perceptual assimilation to the listener's native phonological system. *Journal of the Acoustical Society of America*, 109(2), 775–794.
- Boersma, P., & Weenink, D. (2009). Praat: Doing phonetics by computer (Version 4.3.01) [Computer program]. Available from <http://www.praat.org/>.
- Bolinger, D. L. (1981). *Two kinds of vowels, two kinds of rhythm*. Bloomington, IN: Indiana University Linguistics Club.
- Bond, Z. S., & Small, L. H. (1983). Voicing, vowel and stress mispronunciations in continuous speech. *Perception and Psychophysics*, 34(5), 470–474.
- Bondarko, L. V. (1977). *Zvukovoi stroi sovremennogo russkogo jazyka (Phonetic structure of modern Russian language)*. Moscow: Proveshenie.

- Bondarko, L. V. (1998). *Fonetika sovremennogo russkogo jazyka (Phonetics of modern Russian language)*. St. Petersburg: University of St. Petersburg.
- Braun, B., Lemhöfer, K., & Cutler, A. (2008). English word stress as produced by English and Dutch speakers: The role of segmental and suprasegmental differences. *Proceedings of Interspeech 2008, Australia*, 1953–1953.
- Braun, B., Lemhöfer, K., & Mani, N. (2011). Perceiving unstressed vowels in foreign accented English. *Journal of the Acoustical Society of America*, 129(1), 376–387.
- Browman, C. P., & Goldstein, L. (1992). Articulatory Phonology: An Overview. *Phonetica*, 49, 155–180.
- Chen, Y., Robb, M. P., Gilbert, H. R., & Lerman, J. W. (2001). A study of sentence stress production in Mandarin speakers of American English. *Journal of the Acoustical Society of America*, 109(4), 1681–1690.
- Cooper, N., Cutler, A., & Wales, R. (2002). Constraints of lexical stress on lexical access in English: Evidence from native and nonnative listeners. *Language and Speech*, 45(3), 207–228.
- Crabtree, D. A., & Antrim, L. R. (1988). Guidelines for measuring reaction time. *Perceptual and Motor Skills*, 66(2), 363–370.
- Crosswhite, K. M. (2000). Vowel Reduction in Russian: A Unified Account of Standard, Dialectal, and “Dissimilative” Patterns. In K. M. Crosswhite & J. McDonough (Eds.), *University of Rochester Working Papers in the Language Sciences* (107–172). Retrieved from <http://www.bcs.rochester.edu/cls/working.html>.

- Cutler, A., & Clifton, C. E. (1984). The use of prosodic information in word recognition. In H. Bouma & D. G. Bouwhuis (Eds.), *Attention and performance X: Control of language processes* (pp. 183–196). Hillsdale, NJ: Erlbaum.
- Cutler, A., & Butterfield, S. (1992). Rhythmic cues to speech segmentation: Evidence from juncture misperception. *Journal of Memory & Language*, 31, 218–236.
- Cutler, A., & Norris, D. (1988). The role of strong syllables in segmentation for lexical access. *Journal of Experimental Psychology*, 14(1), 113–121.
- Cutler, A., Wales, R., Cooper, N., & Janssen, J. (2007). Dutch listeners' use of suprasegmental cues to English stress. In J. Trouvain, & W. J. Barry (Eds.), *Proceedings of the 16th International Congress of Phonetics Sciences (ICPhS 2007)* (pp. 1913–1916). Dudweiler: Pirrot.
- Dauer, R.M. (1983). Stress-timing and syllable-timing reanalyzed. *Journal of Phonetics*, 11(1), 51–62.
- Davenport, M., & Hannahs, S.J. (2010). *Introducing Phonetics and Phonology*. London: Hodder Education Publishers.
- Dellwo, V., & Wagner, P. (2003). Relations between language rhythm and speech rate. In M. Solé, D. Recasens, & J. Romero (Eds.), *Proceedings of the 15th International Congress of Phonetics Sciences* (pp. 471–474). Barcelona: Casual Productions.
- Derwing, T. M. (2010). Utopian goals for pronunciation teaching. In J. Levis & K. LeVelle (Eds.), *Proceedings of the 1st Pronunciation in Second Language Learning and Teaching Conference*, (pp. 24–37). Ames, IA: Iowa State University.
- Derwing, T. M., & Munro, M. J. (1997). Accent, intelligibility, and comprehensibility. *Studies in Second Language Acquisition*, 20, 1–16.

- Derwing, T. M., & Munro, M. J. (2005). Second language accent and pronunciation teaching: A research-based approach. *TESOL Quarterly*, 39(3), 379–397.
- Derwing, T. M., & Munro, M. J. (2009). Putting accent in its place: Rethinking obstacles to communication. *Language Teaching*, 42(4), 476–490.
- Derwing, T. M., & Rossiter, M.J. (2003). The effects of pronunciation instruction on the accuracy, fluency and complexity of L2 accented speech. *Applied Language Learning*, 13, 1-18.
- Derwing, T. M., Munro, M. J., & Wiebe, G. (1998). Evidence in Favor of a Broad Framework for Pronunciation Instruction. *Language Learning*, 48(3), 393–410.
- Dupoux, E., & Mehler, J. (1990). Monitoring the lexicon with normal and compressed speech: Frequency effects and the prelexical code. *Journal of Memory and Language*, 29, 316–335.
- Eckman, F. (2004). From phonemic differences to constraint rankings. *Studies in Second Language Acquisition*, 26, 513–549.
- Eckman, F. (2008). Typological markedness and second language phonology. In J. G. Hansen Edwards and M. L. Zampini (Eds.), *Phonology and second language acquisition* (pp. 95–115). Philadelphia: John Benjamins.
- Eek, A. (1987). The perception of word stress: A comparison of Estonian and Russian. In R. Channon & L. Shockey (Eds.), *In Honor of Ilse Lehist* (pp. 19–32). Dordrecht: Foris Publications.
- Fear, B. D., Cutler, A., & Butterfield, S. (1995). The strong/weak syllable distinction in English. *Journal of the Acoustical Society of America*, 97(3), 1893–1904.

- Field, J. (2005). Intelligibility and the listener: The role of lexical stress. *TESOL Quarterly*, 39(3), 399–423.
- Flege, J. (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, J. (1995). Second language speech learning: Theory, findings and problems. In W. Strange (Ed.), *Speech perception and linguistic experience: Theoretical and methodological issues* (pp. 233–277). Timonium, MD: York Press.
- Flege, J. E., & Bohn, O.S. (1989). An instrumental study of vowel reduction and stress placement in Spanish-accented English. *Studies in Second Language Acquisition*, 11, 35–62.
- Flege, J. E., MacKay, I. R. A., & Meador, D. (1999) Native Italian speakers' perception and production of English vowels. *Journal of the Acoustical Society of America*, 10, 2973–2987.
- Flege, J. E., Munro, M. J., & MacKay, I. (1995). Factors affecting strength of perceived foreign accent in a second language. *Journal of the Acoustical Society of America*, 97, 3125–3134.
- Fletcher J. (2010). The prosody of speech: Timing and rhythm. In W. Hardcastle, J. Laver, & F. Gibbon (Eds.), *The handbook of phonetic sciences* (pp. 523-602). United Kingdom: Wiley-Blackwell Publishing,
- Fry, D. B. (1955). Duration and intensity as physical correlates of linguistic stress. *Journal of the Acoustical Society of America*, 27, 765–768.
- Fry, D. B. (1958). Experiments in the perception of stress. *Language and Speech*, 1(2), 126–152.

- Grabe, E., & Low, E. L. (2002). Durational variability in speech and the rhythm class hypothesis. In C. Gussenhoven & N. Warner (Eds.), *Papers in Laboratory Phonology 7* (pp. 515–546). Berlin, New York: Mouton de Gruyter.
- Grosjean, F., & Gee, J. (1987). Prosodic structure and spoken word recognition. *Cognition*, 25, 135–155.
- Hahn, L. (2004). Primary stress and intelligibility: Research to motivate the teaching of suprasegmentals. *TESOL Quarterly*, 38 (2), 201–223.
- Hayes, B. (1984). The phonology of rhythm in English. *Linguistic Inquiry*, 15, 33–74.
- Jones, D., & Ward, D. (1969). *The phonetics of Russian*. Cambridge: Cambridge University Press.
- Kager, R. (2007). Feet and metrical stress. In P. de Lacy (Ed.), *The Cambridge handbook of phonology* (195–227). Cambridge, UK: Cambridge University Press.
- Kang, O. (2010). Relative salience of suprasegmental features on judgments of L2 comprehensibility and accentedness. *System*, 38, 301–315.
- Klatt, D. H. (1976). Linguistic uses of segmental duration in English: Acoustic and perceptual evidence. *Journal of the Acoustical Society of America*, 59(5), 1208–1221.
- Kondaurova, M. V., & Francis, A. L. (2008). The relationship between native allophonic experience with vowel duration and perception of the English tense/lax vowel contrast by Spanish and Russian listeners. *Journal of Acoustical Society of America*, 124(6), 3959–3971.
- Ladefoged, P. (2005). *A course in phonetics* (5th ed.). Boston: Heinle.

- Lee, B., Guion, S. G., & Harada, T. (2006). Acoustic analysis of the production of unstressed English vowels by early and late Korean and Japanese bilinguals. *Studies in Second Language Acquisition*, 28, 487–513.
- Lehiste, I. (1970). *Suprasegmentals*. Cambridge, MA: MIT Press.
- Lehiste, I. (1977). Isochrony reconsidered. *Journal of Phonetics*, 5(3), 253–263.
- Lehiste, I., & Fox, R. A. (1992). Perception of prominence by Estonian and English listeners. *Language and Speech*, 35(4), 419–434.
- Levis, J. M. (2005). Changing contexts and shifting paradigms in pronunciation teaching. *TESOL Quarterly*, 39(3), 369–377.
- Lieberman, P. (1960). Some acoustic correlates of word stress in American English. *Journal of Acoustical Society of America*, 32, 451–454.
- Major, R. (2008). Transfer in second language phonology: A review. In J. G. Hansen Edwards & M. L. Zampini (Eds.), *Phonology and second language acquisition* (pp. 63–94). Philadelphia: John Benjamins.
- Mareüil, P., & Vieru-Dimulescu, B. (2006). The contribution of prosody to the perception of foreign accent. *Phonetica*, 63, 247–267.
- Mattys, S. (2000). The Perception of Primary and Secondary Stress in English. *Perception and Psychophysics*, 62(2), 253–265.
- Mattys, S.L., Brooks, J., & Cooke, M. (2009). Recognizing speech under a processing load: Dissociating energetic from informational factors. *Cognitive Psychology*, 59, 203–243.
- McAllister, R., Flege, J. E., & Piske, T. (2002). The influence of L1 on the acquisition of Swedish quantity by native speakers of Spanish, English and Estonian. *Journal of Phonetics*, 30, 229–258.

- McDonough, K., & Trofimovich, P. (2009). *Using priming methods in second language research*. New York: Routledge.
- McNamara, T. P. (2005). *Semantic priming: Perspectives from memory and word recognition*. New York: Psychology Press.
- McQueen, J. M., & Cutler, A. (2010). Cognitive processes in speech perception. In W. J. Hardcastle, J. Laver, & F. E. Gibbon (Eds.), *The handbook of phonetic sciences* (pp. 489–520). Oxford: Blackwell.
- Meyer, D. E., & Schvaneveldt, R. W. (1971). Facilitation in recognizing pairs of words: Evidence of a dependence between retrieval operations. *Journal of Experimental Psychology*, 90, 227–234.
- Morley, J. (1991). The pronunciation component of teaching English to speakers of other languages. *TESOL Quarterly*, 25(3), 481–520.
- Moulines, E., & Charpentier, F. (1990). Pitch-synchronous waveform processing techniques for text-to-speech synthesis using diphones. *Speech Communication*, 9(5-6), 453–467.
- Munro, M. J., & Derwing, T. M. (1995a). Foreign accent, comprehensibility and intelligibility in the speech of second language learners. *Language Learning*, 45, 73–97.
- Munro, M. J., & Derwing, T. M. (1995b). Processing time, accent and comprehensibility in the perception of native and foreign accented speech. *Language and Speech*, 38, 289–306.
- Munro, M. J., & Derwing, T. M. (2006). The functional load principle in ESL pronunciation instruction: An exploratory study. *System*, 34, 520–531.
- Munro, M. J., & Derwing, T.M. (1998). The effects of speaking rate on listener evaluations of native and foreign-accented speech. *Language Learning*, 48(2), 159–182.

- Murphy, J. M. (2004). Attending to word-stress while learning new vocabulary. *English for Specific Purposes*, 23, 67–83.
- Murphy, J. M., & Kandil, M. (2004). Word-level stress patterns in the academic word list. *System*, 32, 61–74.
- Nazzi, T., & Ramus, F. (2003). Perception and acquisition of linguistic rhythm by infants. *Speech Communication*, 41, 233–243.
- Nicol, J., Swinney, D., Love, T., & Hald, L. (2006). The on-line study of sentence comprehension: An examination of dual task paradigms. *Journal of Psycholinguistic Research*, 35(3), 215–231.
- Norris, D., Cutler, A., McQueen, J. M., & Butterfield, S. (2006). Phonological and conceptual activation in speech comprehension. *Cognitive Psychology*, 53(2), 146–193.
- Norris, D., McQueen, J. M., & Cutler, A. (1995). Competition and segmentation in spoken word recognition. *Journal of Experimental Psychology: Learning, Memory and Cognition*, 21, 1209–1228.
- Odekar, A., Hallowell, B., Kruse, H., Moates, D. R., & Lee, C.-Y. (2009). Validity of eye movement methods and indices for capturing semantic (associative) priming effects. *Journal of Speech, Language, and Hearing Research*, 52, 31–48.
- Ohala, M., & Ohala, J. J. (1992). Phonetic universals and Hindi segment duration. In J. J. Ohala, T. Nearey, B. Derwing, M. Hodge, & G. Wiebe (Eds.), *Proceedings of International Conference on Spoken Language Processing* (pp. 831–834). Edmonton: University of Alberta.
- Ordin, M., & Setter, J. (2008). Objective indicators of rhythmic Russian-English transfer. In *Proceedings of the XX Session of the Russian Acoustical Society, Moscow*, 649–652.

- Padgett, J., & Tabain, M. (2005). Adaptive Dispersion Theory and phonological vowel reduction in Russian. *Phonetica*, 62, 14–54.
- Patel, A. D. (2008). *Music, language, and the brain*. NY: Oxford University Press.
- Peperkamp, S., & Dupoux, E. (2002). A typological study of stress ‘deafness’. In C. Gussenhoven & N. Warner (Eds.), *Laboratory phonology VII* (pp. 203–240). Berlin: Mouton de Gruyter.
- Peterson, G.E., & Lehiste, I. (1960). Duration of syllable nuclei in English. *Journal of the Acoustical Society of America*, 32, 693–703.
- Pike, K. (1945). *The Intonation of American English*. Ann Arbor: University of Michigan Press.
- Pitt, M. A., & Samuel, A. G. (1990). The use of rhythm in attending to speech. *Journal of Experimental Psychology: Human Perception and Performance*, 16, 564–573.
- Pitt, M.A., & Samuel, A.G. (2006). Word length and lexical activation: Longer is better. *Journal of Experimental Psychology: Human Perception and Performance*, 32, 1120–1135.
- Plag, I., Kunter G., & Schramm, M. (2011). Acoustic correlates of primary and secondary stress in North American English. *Journal of Phonetics*, 39(3), 362–374.
- Quené, H., & van Delft, L.E. (2010). Non-native durational patterns decrease speech intelligibility. *Speech Communication*, 52(11-12), 911–918.
- Ramus, F., Nespor, M., & Mehler, J. (1999) Correlates of linguistic rhythm in the speech signal. *Cognition*, 73(3), 265–292.
- Roach, P. (1982). On the distinction between “stress-timed” and “syllable-timed” languages. In D. Crystal (Ed.), *Linguistic controversies* (pp. 73–79). London: Edward Arnold.
- Schmidt, R. (2001). Attention. In P. Robinson (Ed.), *Cognition and second language instruction* (pp. 3–32). Cambridge: Cambridge University Press.

- Slowiaczek, L.M. (1990). Effects of lexical stress in auditory word recognition. *Language and Speech*, 33(1), 47–68.
- Small, L. H., Simon, S. D., & Goldberg, J. S. (1988). Lexical stress and lexical access – homographs versus nonhomographs. *Perception and Psychophysics*, 44(3), 272–280.
- Smith, M. R., Cutler, A., Butterfield, S., & Nimmo-Smith, I. (1989). The perception of rhythm and word boundaries in noise-masked speech. *Journal of Speech and Hearing*, 32, 912–920.
- Soto-Faraco, S., Sebastián-Gallés, N., & Cutler, A. (2001). Segmental and suprasegmental mismatch in lexical access. *Journal of Memory and Language*, 45, 412–432.
- Stevens, K. (1999). *Acoustic phonetics*. Cambridge, MA: The MIT Press.
- Stockmal, V., Markus, D., & Bond, D. (2005). Measures of native and non-native rhythm in a quantity language. *Language and Speech*, 48, 55–63.
- Tajima, K., Port, R., & Dalby, J. (1997). Effects of temporal correction on intelligibility of foreign accented English. *Journal of Phonetics*, 25, 1–24.
- Thomas, E. R. (2011). *Sociophonetics: An introduction*. New York: Palgrave Macmillan.
- Trofimovich, P., & Baker, W. (2006). Learning second-language suprasegmentals: Effect of L2 experience on prosody and fluency characteristics of L2 speech. *Studies in Second Language Acquisition*, 28, 1–30.
- Trubetzkoy, N. S. (1939/1969) *Principles of phonology* (C. A. M. Baltaxe, Trans.). Berkeley, CA: University of California Press.
- Tucker, B. V., & Warner, N. (2007). Inhibition of processing due to reduction of the American English flap. In J. Trouvain, & W. J. Barry (Eds.), *Proceedings of the 16th International Congress of Phonetics Sciences (ICPhS 2007)* (pp. 1949–1952). Dudweiler: Pirrot.

- Turk, A., & Shattuck-Hufnagel, S. (2007). Multiple targets of phrase-final lengthening in American English words. *Journal of Phonetics*, 35(4), 445–472.
- White, L., & Mattys S. L. (2007). Calibrating rhythm: First language and second language studies. *Journal of Phonetics*, 35, 501–522.
- White, L., & Mattys, S. L. (2007). Rhythmic typology and variation in first and second languages. In P. Prieto, J. Mascaró, & M.-J. Sole (Eds.), *Segmental and prosodic issues in romance phonology (Current issues in linguistic theory series)*. Amsterdam, Philadelphia: John Benjamins.
- White, L., Mattys, S. L., & Wiget, L. (in press). Language categorization by adults is based on sensitivity to durational cues, not rhythm class. *Journal of Memory and Language*.
- Wilson, E. O., & Spaulding, T. J. (2010). Effects of noise and speech intelligibility on listener comprehension and processing time of Korean-accented English. *Journal of Speech, Language, and Hearing Research*, 53, 1543–1554.
- Yu, V., & Andruski, J. (2010). A cross-language study of perception of lexical stress in English. *Journal of Psycholinguistic Research*, 39(4), 323–344.
- Zhang, Y., & Francis, A. (2010). The weighting of vowel quality in native and non-native listeners' perception of English lexical stress. *Journal of Phonetics*, 38, 260–271.
- Zhang, Y., Nissen, S. L., & Francis, A. L. (2008). Acoustic characteristics of English lexical stress produced by native Mandarin speakers. *Journal of the Acoustical Society of America* 123(6), 4498–4513.
- Zielinski, B. W. (2008). The listener: No longer the silent partner in reduced intelligibility. *System*, 36(1), 69–84.

APPENDIX A

		AUDITORY PRIME	VISUAL TARGET		
1	SS	application	application	match	real
2	SS	classification	classification	match	real
3	SS	collaboration	collaboration	match	real
4	SS	compatibility	compatibility	match	real
5	SS	condensation	condensation	match	real
6	SS	confiscation	confiscation	match	real
7	SS	domination	domination	match	real
8	SS	elaboration	elaboration	match	real
9	SS	fabrication	fabrication	match	real
10	SS	gratification	gratification	match	real
11	SS	magnification	magnification	match	real
12	SS	operational	operational	match	real
13	SS	optimistic	optimistic	match	real
14	SS	politician	politician	match	real
15	SS	population	population	match	real
1	UU	abstract	abstract	match	real
2	UU	compact	compact	match	real
3	UU	contact	contact	match	real
4	UU	contract	contract	match	real
5	UU	coupon	coupon	match	real
6	UU	format	format	match	real
7	UU	icon	icon	match	real
8	UU	impact	impact	match	real
9	UU	mascot	mascot	match	real
10	UU	neuron	neuron	match	real
11	UU	nylon	nylon	match	real
12	UU	python	python	match	real
13	UU	radar	radar	match	real
14	UU	robot	robot	match	real
15	UU	syntax	syntax	match	real
1	Control	repeatedly	application	no match	real
2	Control	multimillion	classification	no match	real
3	Control	portability	collaboration	no match	real
4	Control	comfortable	compatibility	no match	real
5	Control	conditioning	condensation	no match	real
6	Control	Californian	confiscation	no match	real
7	Control	correlation	domination	no match	real
8	Control	regulation	elaboration	no match	real

9	Control	complicated	fabrication	no match	real
10	Control	primitive	gratification	no match	real
11	Control	individual	magnification	no match	real
12	Control	attraction	operational	no match	real
13	Control	escalator	optimistic	no match	real
14	Control	revolution	politician	no match	real
15	Control	speculate	population	no match	real
16	Control	punish	abstract	no match	real
17	Control	sponsor	compact	no match	real
18	Control	speeding	contact	no match	real
19	Control	balloon	contract	no match	real
20	Control	pressure	coupon	no match	real
21	Control	flatten	format	no match	real
22	Control	luggage	icon	no match	real
23	Control	focus	impact	no match	real
24	Control	govern	mascot	no match	real
25	Control	spender	neuron	no match	real
26	Control	dial	nylon	no match	real
27	Control	scissors	python	no match	real
28	Control	failure	radar	no match	real
29	Control	purely	robot	no match	real
30	Control	mention	syntax	no match	real

A complete list of experimental items (SS—Secondary stressed syllables, UU—Unstressed-unreduced syllables) for the cross-modal priming experiment.

APPENDIX B

Demographic Information

Age: _____

Sex/Gender: *Male Female*

Nationality (select one or more):

____ *Russian*____ *Belorussian*____ *Ukrainian*____ *Other (please explain)* _____

What country and city/ town/ area are you from?

Is Russian your first language? *Yes No I am Bilingual**If No, what is your first language?* _____*If Bilingual, what is the other language?* _____

What nationality are your parents?

Mother _____

Father _____

What language do you speak in your family?

Did you attend a Russian school? *Yes No**If No, did you learn Russian at school? Yes (how many years?) _____ No*Do you use Russian daily? *Yes No*

How old were you when you came to the USA?

How long have you been in the USA?

At what age did you start studying English?

Do you have normal hearing? *Yes No**If No, please explain* _____Do you have any speech or language problems? *Yes No**If Yes, please explain* _____

Thank you!

For Experimenter Only:

Subject # _____ Date _____

Experiment _____ Time _____

APPENDIX C

Demographic Information

Age: _____

Sex/Gender: *Male Female*

Race/Ethnicity (select one or more):

☐ *Hispanic or Latino*☐ *American Indian or Alaska Native*☐ *Native Hawaiian or other Pacific Islander*☐ *Black or African American*☐ *Other (please explain)***For Experimenter Only:**

Subject # _____

Date _____

Experiment _____

Time _____

Do you have normal hearing? *Yes No**If No, please explain:* _____Do / Did you have any speech or language problems? *Yes No**If Yes, please explain* _____Is English your first language? *Yes No**If No, how old were you when you began learning English?* _____

What state and city/ town are you from? _____

Have you always lived in the Midwest? *Yes No**If No, please explain* _____

Have you ever noticed that you produce certain words/ sounds differently than your classmates?

*Yes No**If Yes, please give examples:* _____

Thank you!

APPENDIX D

Demographic Information

Age: _____

Sex/Gender: *Male Female*Handedness: *Right Left Ambidextrous*

Race/Ethnicity (select one or more):

____ *Hispanic or Latino*____ *American Indian or Alaska Native*____ *Native Hawaiian or other Pacific Islander*____ *Black or African American*____ *Other (please explain)* _____**For Experimenter Only:**

Subject # _____

Date _____

Experiment _____

Time _____

Do you have normal hearing? *Yes No**If No, please explain:* _____Do you have any speech or language problems? *Yes No**If Yes, please explain* _____Have you ever communicated in English with Russian native speakers? *Yes No**If Yes, when and how long?* _____

Do you ever hear Russian-accented English around you - e.g., in your school/ work/ home environment?

*Yes No**If Yes, can you describe the situation?* _____Have you ever visited or studied in Russia or any other Russian speaking country? *Yes No**If Yes, please describe when, for how long, and the purpose of your trip:*Do you ever use mass media where you can hear Russian-accented English? *Yes No**If Yes, what and how regularly* _____

When was the last time you heard Russian-accented English?

Do you think you are familiar with the Russian accent?

Thank you!

APPENDIX E

Bowling Green State University

Office of Research Compliance
309A University Hall
Bowling Green, OH 43403-0183
Phone: (419) 372-7716
E-mail: hsrb@bgsu.edu

HSRB MEMBERSHIP
2010-2011

Amy Morgan, HSRB Chair
Kinesiology
amorgan@bgsu.edu

Mary Hare, HSRB Vice Chair
Psychology
mlhare@bgsu.edu

D. Wayne Bell, M.D.
Wood Health Corp.
353-6225
speakingdoc@dacor.net

Cheryl Conley
Alzheimer's Assn., NW Ohio
conleyc@bgsu.edu

L. Fleming Fallon, Jr., M.D.
Public & Allied Health
ffallon@bgsu.edu

Rodney Gabel
Comm. Sciences & Disorders
rgabel@bgsu.edu

Hillary Harms
Office of Research Compliance
hsrb@bgsu.edu

Lesla Lockford
Theatre & Film
lockflo@bgsu.edu

Montana Miller
Popular Culture
montanm@bgsu.edu

Jeanne Novak
Intervention Services
jnovak@bgsu.edu

Erin Smith
Psychology
esmith@bgsu.edu

Ashutosh Sohoni
Family and Consumer Sciences
assohon@bgsu.edu

Marie Tisak
Psychology
mtisak@bgsu.edu

May 26, 2011

TO: Elina Banzina
CDIS

FROM: Hillary Harms, Ph.D.
HSRB Administrator

RE: HSRB Project No.: H11D260GX2

TITLE: *The Role of Secondary-stressed and Unstressed-unreduced Syllables in Speaker Intelligibility: An Acoustic and Perceptual Study with Russian Learners of English*

You have met the conditions for approval for your project involving human subjects. **As of May 26, 2011, your project has been granted final approval by the Human Subjects Review Board (HSRB). This approval expires on May 12, 2012.** You may proceed with subject recruitment and data collection.

The final approved version of the consent document(s) is attached. Consistent with federal OHRP guidance to IRBs, **the consent document(s) bearing the HSRB approval/expiration date stamp is the only valid version and you must use copies of the date-stamped document(s) in obtaining consent from research subjects.**

You are responsible to conduct the study as approved by the HSRB and to use only approved forms. If you seek to make any changes in your project activities or procedures, send a request for modifications to the HSRB via this office. Those changes must be approved by the HSRB prior to their implementation.

You have been approved to enroll 80 participants. If you want to enroll additional participants you must seek approval from the HSRB.

Good luck with your work. Let me know if this office or the HSRB can be of assistance as your project proceeds.

Comments/ Modifications:

Stamped original consent forms are coming to you via campus mail.

c: Dr. Lynne E. Hewitt

Research Category: EXEMPT #2

BOWLING GREEN STATE UNIVERSITY
Department of Communication Sciences and Disorders

Perceptual Study
Researcher: Elina Banzina

CONSENT TO ACT AS A RESEARCH PARTICIPANT

Dear BGSU Community Member,

We are conducting a research study to learn more about speech perception and production. To participate, you must be a native speaker of English, have normal hearing and be at least 18 years of age.

If you agree to participate, then the following will happen. You will be asked to participate in a session in which you listen to words in English, look at words on a computer screen, and make simple responses to them by pushing a button. The session will last for not more than an hour. All of the sounds will be presented at a comfortable listening level, and there are no known risks associated with any of the procedures used in this study other than those encountered in daily life.

There is no direct benefit to you from these procedures. However, the findings of this study will help better understand the process of speech perception and provide valuable information to language instruction.

In return for your participation, you will receive monetary compensation in the amount of \$15. Participation in this research is entirely voluntary, and you may refuse to participate or stop at any time without penalty. Your decision to participate or not will in no way impact your grades, class standing, or relationship to BGSU. We will end the experiment without your consent in the event of an equipment failure or other unforeseen circumstance. If we have to end the experiment, you will receive compensation based on the amount of time that you participated.

At the end of the study, you will also be asked to complete a short survey about your age, gender, linguistic background, and whether you have any speech-language related problems. All information you provide is confidential and all paper records of your research participation will not include your name and will be stored in a secure filing cabinet. All responses you make on the computer will be maintained on a password protected computer or secure backup storage device. All research records will be used solely for research purposes.

I have explained this study to you and answered your questions. If you have other questions, or wish to report a research-related problem, you may contact me by phone at (419) 372-4320 or by email at ebanzi@bgsu.edu, or you may contact Dr. Lynne Hewitt, project advisor, by phone at (419) 372-6031 or by email at lhewitt@bgsu.edu. You may also contact the Chair, Human Subjects Review Board, Bowling Green State University, by phone (419-372-7716) or by email (hsrb@bgsu.edu) if you have questions about your rights as a research participant.

By signing below, you are indicating that you have read the above consent form, agree to participate in this study, and are at least 18 years of age.

Participant's signature

Date

Participant's name (Please PRINT neatly)

200 Health Center Building
Bowling Green, OH 43403-0149

Phone: 419-372-2515
Fax: 419-372-8089

BGSU HSRB - APPROVED FOR USE
ID # H1D2606x2
EFFECTIVE 5-26-11
EXPIRES 5-26-12

BOWLING GREEN STATE UNIVERSITY
Department of Communication Sciences and Disorders

Acoustic Study
Researcher: Elina Banzina

CONSENT TO ACT AS A RESEARCH PARTICIPANT

Dear BGSU Community Member,

We are conducting a research study to learn more about speech perception and production. To participate, you must have normal hearing and be at least 18 years of age.

If you agree to participate, then the following will happen. You will be asked to read out loud words in English from a list provided to you, and you will be recorded while reading them. All recordings will be made in a sound-proof recording booth to ensure high recording quality. Your speech will be recorded for the purposes of acoustic analysis of the produced sounds and for creating materials for speech perception experiments that will follow afterwards. The recording session will last for not more than an hour.

There are no known risks associated with any of the procedures used in this study other than those encountered in daily life. Also, there is no direct benefit to you from these procedures. However, recordings of your speech will contribute to a deeper understanding of speech patterns and will inform language instruction.

In return for your participation in this project, you will receive monetary compensation in the amount of \$15. Participation in this research is entirely voluntary, and you may refuse to participate or stop at any time without penalty. Your decision to participate or not will in no way impact your grades, class standing, or relationship to BGSU. We will end the experiment without your consent in the event of an equipment failure or other unforeseen circumstance. If we have to end the experiment, you will receive compensation based on the amount of time that you participated.

At the end of the study, you will also be asked to complete a short survey about your age, gender, linguistic background, and whether you have any speech-language related problems. All information you provide is confidential and all paper records of your research participation will not include your name and will be stored in a secure filing cabinet. Recordings of your speech will be maintained on a password protected computer or secure backup storage device. All research records will be used solely for research purposes.

I have explained this study to you and answered your questions. If you have other questions, or wish to report a research-related problem, you may contact me by phone at (419) 372-4320 or by email at ebanzi@bgsu.edu, or you may contact Dr. Lynne Hewitt, project advisor, by phone at (419) 372-6031 or by email at lhewitt@bgsu.edu.

You may also contact the Chair, Human Subjects Review Board, Bowling Green State University, by phone (419-372-7716) or by email (hsrb@bgsu.edu) if you have questions about your rights as a research participant.

By signing below, you are indicating that you have read the above consent form, agree to participate in this study, and are at least 18 years of age.

Participant's signature

Date

Participant's name (Please PRINT neatly)

200 Health Center Building
Bowling Green, OH 43403-0149

Phone: 419-372-2515
Fax: 419-372-8089

BGSU HSRB - APPROVED FOR USE
ID # H11D2460672
EFFECTIVE 5.26.11
EXPIRES 5.12.12