J Psycholinguist Res (2016) 45:813–831
DOI 10.1007/s10936-015-9377-z

The Role of Secondary-Stressed and Unstressed–Unreduced Syllables in Word Recognition: Acoustic and Perceptual Studies with Russian Learners of English

Elina Banzina[1] · **Laura C. Dilley**[2] · **Lynne E. Hewitt**[3]

Published online: 17 May 2015

Abstract The importance of secondary-stressed (SS) and unstressed–unreduced (UU) syllable accuracy for spoken word recognition in English is as yet unclear. An acoustic study first investigated Russian learners' of English production of SS and UU syllables. Significant vowel quality and duration reductions in Russian-spoken SS and UU vowels were found, likely due to a transfer of native phonological features. Next, a cross-modal phonological priming technique combined with a lexical decision task assessed the effect of inaccurate SS and UU syllable productions on native American English listeners' speech processing. Inaccurate UU vowels led to significant inhibition of lexical access, while reduced SS vowels revealed less interference. The results have implications for understanding the role of SS and UU syllables for word recognition and English pronunciation instruction.

Keywords Secondary stress · Unstressed unreduced syllables · Vowel reduction · Speech intelligibility · Speech processing · English pronunciation · Russian accent · Word recognition

✉ Elina Banzina
elina.banzina@sseriga.edu

Laura C. Dilley
ldilley@msu.edu

Lynne E. Hewitt
lhewitt@bgsu.edu

[1] Department of Languages, Stockholm School of Economics in Riga, Strelnieku iela 4a, Riga 1010, Latvia

[2] Department of Communicative Sciences and Disorders, Michigan State University, 116 Oyer Center, East Lansing, MI 48824, USA

[3] Department of Communication Sciences and Disorders, Bowling Green State University, 238 Health and Human Services Building, Bowling Green, OH 43403, USA

Introduction

Identifying those phonetic elements that are most critical for successful spoken word recognition is a central issue in first and second language research. A variety of studies have investigated the relevance of specific segmental and suprasegmental characteristics for speech perception (Cooper et al. 2002; Derwing and Munro 1997; Hahn 2004; Quené and Delft 2010; Tajima et al. 1997), part of this attention devoted to assessing the importance of lexical prominence in speech processing (Cutler and Clifton 1984; Field 2005; Lehiste and Fox 1992; Small et al. 1988). In English, lexical prominence is achieved by a combination of vowel quality and stress, the latter expressed acoustically as greater intensity, higher fundamental frequency, and longer duration of a syllable relative to the surrounding syllables (Flege and Bohn 1989; Fry 1955; Lehiste 1970; Zhang et al. 2008). In English, four levels of lexical prominence have been identified: primary-stressed (PS) syllables (e.g., the first syllable in the word "garden"), secondary-stressed (SS) syllables (e.g., the second syllable in the word "imagination"), unstressed–unreduced (UU) syllables (e.g., the second syllable in the word "robot"), and reduced (R) syllables (e.g., the second syllable in the word "husband") (Beckman 1986; Beckman and Edwards 1994; Fear et al. 1995; Ladefoged 2005). Acoustic and perceptual studies have provided empirical evidence for the distinctiveness of the four prominence types in English, out of which UU and SS syllables as separate categories have been the most debatable. Fear et al. (1995) found that in terms of stress and vowel quality, UU syllable category stood separate from the rest. PS and SS were also found to belong to separate categories, but evidence regarding the specific acoustic correlates that distinguish them so far has been mixed: both Fear et al. (1995) and Mattys (2000) found durational differences between PS and SS syllables but no differences in intensity, while a study by Plag et al. (2011) reported no differences in PS and SS syllable durations but slight dissimilarities in F0 and intensity. Perceptual experiments show that listeners are capable of distinguishing between the lexical prominence degrees: in a study by Mattys (2000), native listeners were able to discriminate between PS and SS syllables based on two-syllable word fragments alone, while Fear et al. (1995) showed that native listeners viewed the reduction of UU vowels as generally more acceptable than the reduction of SS vowels, suggesting that SS and UU syllables belong to separate perceptual categories.

While a number of psycholinguistic studies have addressed the importance of lexical prominence degrees for speech processing and word recognition (Bond and Small 1983; Slowiaczek 1990; Small et al. 1988), almost all of the studies have focused on PS syllables. The effect of SS and UU syllables on speech perception, on the other hand, has received very limited attention in first and second language literature, despite a call for more second language (L2) pronunciation research that would elucidate those phonetic features that are most critical for intelligibility and would ultimately help identify priorities in L2 teaching and learning (Derwing and Munro 2005; Levis 2005).

A significant body of research on a variety of languages has shown that adult L2 learners often perceive and produce second language lexical prominence features differently than native speakers due to the native language (L1) influence (e.g., Chen et al. 2001; McAllister et al. 2002; Yu and Andruski 2010; Zhang and Francis 2010). For example, it has been documented that L2 speakers tend to produce insufficient vowel reduction if their native languages do not have phonological vowel reduction (Braun et al. 2011; Flege and Bohn 1989; Lee et al. 2006; Zhang et al. 2008). The opposite—vowel over-reduction, however, has not been reported by any studies, yet there is anecdotal evidence of this phenomenon in Russian learners' of English production, potentially a direct transfer from L1 Russian. Russian, a language of interest for the current study, exhibits a more complex and abundant

pattern of phonological vowel reduction than English. Russian allows only one stressed syllable per multisyllabic word, all surrounding syllables undergoing vowel reduction of some degree. Thus, unlike in English, SS in Russian occurs only in compounds and a few loanwords (Avanesov 1956). 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 and Tabain 2005). The three levels of stress in Russian are realized primarily in terms of vowel quality and vowel duration. In the syllable immediately preceding the stressed syllable (i.e., the "pre-stress" syllable), moderate vowel reduction occurs and results in a more limited vowel inventory. Extreme vowel reductions occur in all the other syllables before the main stress, or after it (in the so-called "post-stress" syllables), where the lower vowels /ɑ/ and /o/ undergo a dramatic reduction in vowel quality and turn into a "schwa" (Avanesov 1956; Bondarko 1998; Crosswhite 2000). Besides changes in vowel quality, equally critical vowel reductions take place also in the temporal domain. According to Bondarko (1977, 1998), the low vowels undergo three levels of temporal reduction, expressed as a ratio of 1:2:3, where 3 is the duration of the stressed syllable. An acoustic study by Padgett and Tabain (2005) confirmed the significant durational distinctions between low vowels in the first and second pre-stress syllables, except after palatalized consonants.

These extreme reduction positions in Russian correspond precisely to the location of SS and UU vowels in English, where 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/)); this arrangement presents an interesting opportunity for investigating vowel reduction in Russian-spoken English, and simultaneously assessing the role of SS and UU syllable accuracy in English listener perception. To date, there has been no research on SS and UU syllable production in English by Russian speakers, or speakers of any other language, that would verify vowel over-reduction; thus the transfer of Russian stress system to L2 English by Russian speakers and the subsequent temporal and vowel quality reductions remain unquantified and unestablished in a careful acoustic study. Similarly, as of yet no published studies have tested the importance of SS and UU syllable types in speech processing, and it is unclear whether or not such low-accuracy syllables would lead to inhibitory effects in the word recognition process.

In order to fill these gaps in research, first, an acoustic study was conducted to establish vowel over-reduction by Russian learners of English empirically, and a subsequent perceptual study was conducted to assess instrumentally the consequences of mispronunciation of SS and UU syllables on word recognition by English listeners. The results will shed light on the basic issues in psycholinguistic processing by identifying the importance of two previously under-researched syllable types for spoken word recognition, and would provide a clearer understanding of those phonetic features that are most critical for successful word recognition. Furthermore, the findings would be informative for setting priorities in pronunciation instruction for native Russian (and other) learners of English, as there are virtually no studies on the role of SS and UU syllable categories in second-language research and pedagogy.

Experiment 1: Acoustic Study

Participants

Six native speakers of Russian who speak English as their second language and six monolingual native speakers of American English were recruited and recorded for the acoustic study.

In each group, half of the participants were male, and half were female. The Russian participants were graduate students at a university in Ohio (average age = 26), and had learned English in their home country for an average of 13 years and resided in the United States for approximately 3 years at the time of the recording. The specific length of residence was selected to ensure that speakers had received sufficient exposure to American English yet their speech retained the specific non-native language patterns of interest for this study. The native speakers of American English were graduate and undergraduate students at the same university (average age = 23), all from Ohio. All participants were paid for their participation.

Materials

All experimental items were multisyllabic English words that contained either an UU, SS, or PS syllable. The words were systematically selected so that the target syllables—UU and SS—occupied positions within the multisyllabic words that would be most prone to the predicted vowel reductions by Russian speakers of English. Since the greatest vowel reduction in Russian occurs several syllables before the PS, and immediately after the PS, three different syllable types were explored: SS syllables in pre-PS position (not less than two syllables before PS, e.g., "com" in "compre'hension"), UU vowels in immediately post-PS position (one syllable after the PS, e.g., "bot" in "robot"), and PS syllables (e.g, the second syllable, "bo", in "ro'botic"). To control for word length effects, PS syllables were recorded in separate three-syllable words, the average between two- and four- to five-syllable-long UU and SS words, respectively.

Furthermore, to optimally test for the hypothesized vowel reduction effects, the study focused on the low back vowel /ɑ/, and the low front vowel /æ/. These vowel types undergo more extreme reductions and should demonstrate greater rhythm transfer effects than other vowels. Specifically, the low back /ɑ/ reduces to a "schwa" in Russian due to rising of the vowel space floor in unstressed syllables (Avanesov 1956; Padgett and Tabain 2005), whereas the low front English vowel /æ/, which is not part of the Russian sound inventory, could be perceived as a variation of /ɑ/ and similarly reduced to a schwa. For a complete list of experimental items, see Table 1 in the Appendix.

Items on each experimental list contained 20 words with SS, 20 words with UU, and 20 words with PS syllables, as well as 120 filler items. Items were carefully selected so that they could be used in the following priming task in Experiment 2. Words were presented in their citation form and in a random order to mask the repetitiveness of their orthographic and phonological structure. For words like "extract", where word stress location can alternate, accent marks were provided on the PS syllable to indicate whether the word was a noun or a verb. Finally, to control for variability in intonation, each word was presented orthographically with a capitalized initial letter and a period at the end to indicate a complete phrase.

Procedure

Recordings were made in a sound-proof booth with a Shure SM58 high-quality microphone at a sampling rate of 44.1 KHz. Prior to recordings, participants were first asked to read through the list and report any unclear words, and their understanding of the accent mark was checked. Any miss-stressed words prior to and during the recording were corrected withholding direct modeling on the part of the experimenter and only providing indirect corrective feedback. After a short practice trial, participants proceeded to the experimental trial, where they were instructed to first read two words from printed list silently and then

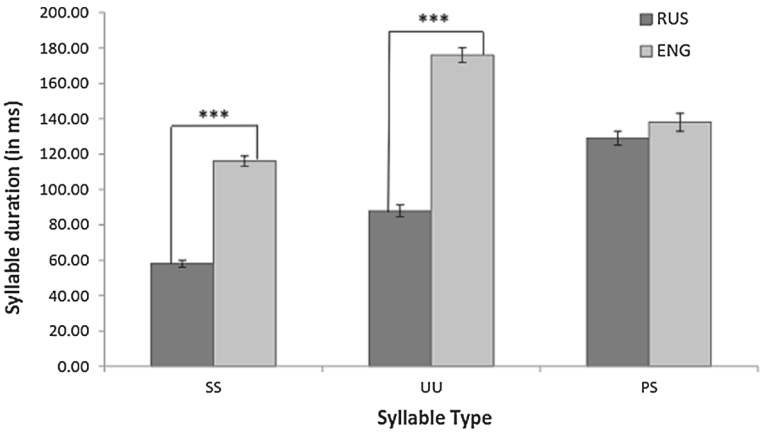

Fig. 1 Average durations of SS, UU and PS syllables in seconds, as spoken by Russian learners of English and native English speakers

say them out loud from memory twice; only the second repetition of each word was used as a more natural, "non-read" production.

Analysis

Acoustic measurements were obtained and analyzed using Praat 5.0.32 acoustic analysis software (Boersma and Weenink 2011). The acoustic analyses involved (i) formant, (ii) duration, (iii) fundamental frequency, and (iv) intensity measures for SS, UU and PS vowels. Vowel F1 and F2 measurements were made at vowels' midpoints in the spectrogram, using formant track readings provided by the LPC analysis in Praat; alternatively, an FFT analysis was used for more reliable results. Duration measurements were obtained from the linked spectrogram and waveform displays. Fundamental frequency was estimated with the help of the autocorrelation method, and 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. For more robust results and greater reliability, data were analyzed both on raw and normalized measures; both measures yielded identical findings.

Results

Separate $2 \times 2 \times 3$ mixed measures ANOVAs were conducted on duration, intensity, and F0 with Group (Russian, English) as a between-subjects independent variable, and Vowel Type (/ɑ/, /æ/) and Syllable Type (SS, UU, PS) as within-subjects independent variables.

Duration

As shown in Fig. 1, the ANOVA revealed a significant main effect of Group ($F(1, 18) = 679.95$, $p < 0.001$), with vowel durations produced by Russian speakers ($M = 90$ ms) being on average only 64 % of the duration of those produced by native English speakers ($M = 140$ ms). No main effect of Vowel Type was found ($F(1, 18) = 2.89$, $p = 0.106$), but there was a main effect of Syllable Type ($F(2, 36) = 94.37$, $p < 0.001$) and a significant interaction between Group and Syllable Type ($F(2, 36) = 53.54$, $p < 0.001$). Post hoc t tests with Bonferroni correction revealed that there was a statistically significant difference

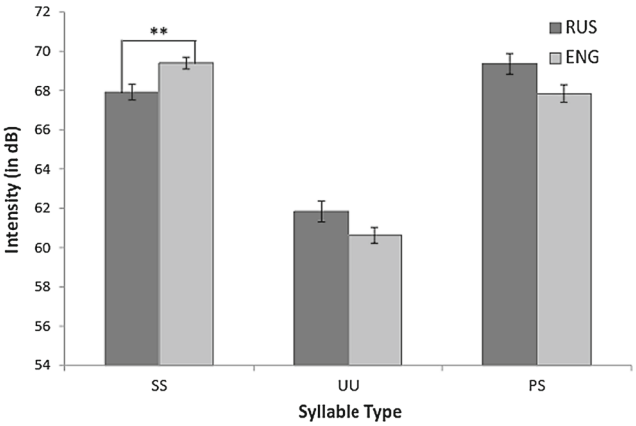

Fig. 2 Intensity values (in dB) for secondary-stressed (SS), unstressed–unreduced (UU) and primary-stressed (PS) syllables for Russian learners of English and native English speakers

between Russian-spoken ($M = 58$ ms) and English-spoken SS syllables ($M = 116$ ms), with English speaker productions approximately double the duration of Russian speakers' productions ($t(38) = 15.68$, $p < 0.001$). There was similarly a statistically significant difference between Russian-spoken ($M = 88$ ms) and English-spoken UU syllables ($M = 175$ ms), with the duration of English-spoken syllables about twice that of Russian-spoken syllables ($t(38) = 15.15$, $p < 0.001$). No significant difference was found between Russian-spoken and English-spoken PS syllables ($t(38) = 1.55$, $p = 0.128$).

Intensity

Figure 2 presents intensity values for native English and Russian learner of English groups. Statistical analyses showed no significant main effect of Group ($F(1, 18) = 1.95$, $p = 0.18$), but there was a significant interaction between Syllable Type and Group ($F(2, 36) = 14.80$, $p < 0.001$). Post-hoc t tests, adjusted with Bonferroni correction, revealed that Russian-spoken and English-spoken UU syllables were of comparable intensity ($t(38) = 1.82$, $p = 0.77$), however, SS-syllable intensity was significantly lower for Russian ($M = 67.91$ dB) than English speaker ($M = 69.40$ dB) groups ($t(38) = 2.97$, $p = 0.005$), and PS syllable intensity was marginally significantly higher for Russian ($M = 69.35$ dB) than English speaker (M = 67.84 dB) groups ($t(38) = 2.19$, $p = 0.035$). Post-hoc tests conducted within the Russian speaker group showed that SS syllable intensity took an intermediate position between PS syllables, albeit marginal (t(19) = 2.255, $p = 0.036$), and UU syllables (t(19) = 10.154, $p < 0.001$). Within the English speaker group, this order was somewhat reversed, as SS syllables were found to be significantly more intense than PS syllables (t(19) = 2.797, $p = 0.011$), followed by UU syllables as less intense than PS (t(19) = 11.916, $p < 0.001$).

Fundamental Frequency

Figure 3 shows F0 data. There was a main effect of Syllable Type ($F(2, 36) = 615.382$, $p < 0.001$), where PS syllables had the highest and UU syllables had the lowest F0, and a main effect of Vowel Type ($F(1, 18) = 20.837$, $p < 0.001$), where the low back vowel /ɑ/ had an overall higher frequency than the low front vowel /æ/. However, there was no main effect of

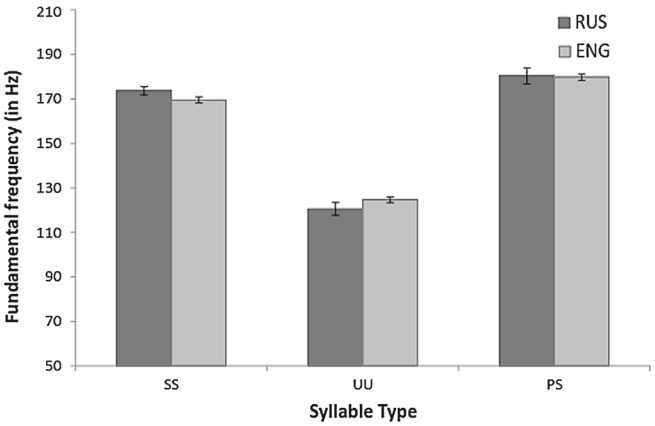

Fig. 3 F0 values (in Hz) for secondary-stressed (SS), unstressed–unreduced (UU) and primary-stressed (PS) syllables for Russian learners of English, and native English speakers

Group ($F(1, 18) = 0.021$, $p = 0.886$); the overall F0 across SS, UU and PS syllables and vowel types was maintained the same by both English and Russian speaker groups. Finally, no significant interactions were found.

F1 and F2 Frequencies

F1 and F2 frequency values were submitted to two separate mixed-measures ANOVAs with Group (Russian, English) as a between-subjects independent variable, and Vowel Type (/ɑ/, /æ/) and Syllable Type (SS, UU) as within-subjects independent variables. Figures 4 and 5 show differences between Russian- and English-spoken vowels, graphed separately for /æ/ and /ɑ/ vowel types. Analysis on F1 yielded significant main effects for all three factors: Group ($F(1, 18) = 629.08$, $p < 0.001$), where native ($M = 839$ Hz) and Russian ($M = 580$ Hz) speakers of English differed from each other in terms of F1 frequency across syllable and vowel types, Syllable Type ($F(1, 18) = 15.67$, $p = 0.001$), and Vowel Type ($F(1, 18) = 7.38$, $p = 0.014$).

A significant interaction was found between Vowel Type and Group ($F(1, 18) = 5.95$, $p = 0.025$). Post-hoc paired samples t tests showed that English speakers' F1 was higher for the low back vowel /ɑ/ than for the low front vowel /æ/ ($t(19) = 3.335$, $p = 0.003$), consistent with the mean data for F1 frequencies of American English vowels (Kent and Read 2002), while Russian speakers maintained the same F1 for both vowel types ($t(19) = 0.125$, $p = 0.902$).

Two repeated-measures 2 × 2 ANOVAs were performed to analyze F2 values separately for the /ɑ/ and /æ/ vowels. The first two-way ANOVA with the low back vowel /ɑ/ as the dependent variable found no main effects or interactions: there was no main effect of Group ($F(1, 18) = 2.72$, $p = 0.117$) or Syllable Type ($F(1, 18) = 3.07$, $p = 0.097$), and no interaction between Syllable Type and Group ($F(1, 18) = 1.43$, $p = 0.247$). The second two-way ANOVA had F2 of the low front vowel /æ/ as the dependent variable. Analyses showed a significant main effect of Group ($F(1, 18) = 35.39$, $p < 0.001$); 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). No other main effects or interactions were found.

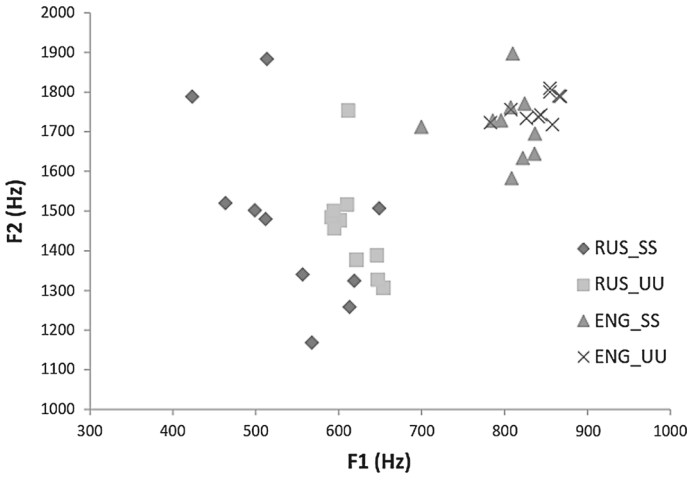

Fig. 4 F1 and F2 values for the low front vowel /æ/ in SS and UU syllables, as produced by native Russian and native English speaker groups

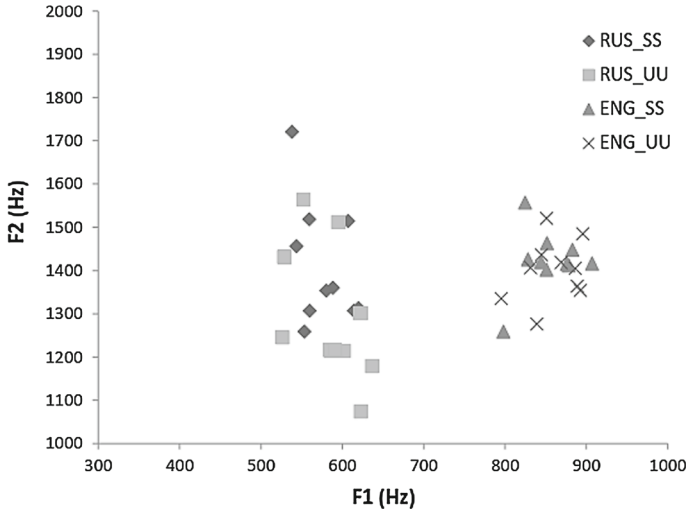

Fig. 5 F1 and F2 values for the low back vowel /ɑ/ in SS and UU syllables, as produced by native Russian and native English speaker groups

Discussion

The acoustic analyses revealed significant phonetic differences in the realization of SS and UU syllables between native English speakers and Russian learners of English. One of the most striking differences concerned UU and SS syllable duration. While Russian speakers were able to realize PS vowels in a native-like fashion, Russian-produced SS and UU vowels were only about half the duration of those produced by English speakers. Thus, Russian learners of English consistently failed to apply temporal patterns in a native-like manner when producing SS and UU syllables.

In terms of intensity, Russian learners of English largely managed to attain native-like PS, SS and UU levels. UU syllable production was similar for both groups in that this syllable type was of lowest intensity. Findings regarding SS and PS syllable intensity, however, are less clear-cut. Contrary to expectations, English-spoken SS syllables were found to be significantly greater in intensity than PS syllables, and could be potentially attributed to item presentation in their citation form in a list format. Thus, while Russian-spoken SS syllables fell short of native speaker levels, in reality their intensity was relatively high: Russian-spoken SS syllables only marginally differed from their own PS syllables, which, in turn, were greater in intensity than English-spoken PS syllables. Along with the fact that Russian-spoken SS syllables significantly surpassed their UU syllable intensity, these results collectively show that Russian speakers did use the intensity cue to mark SS syllable prominence to a certain extent.

The overall fundamental frequency across SS, UU and PS syllables and vowel types was maintained the same by both English and Russian speaker groups, suggesting that Russian speakers were able to use F0 in English in a native-like manner. Nevertheless, these results should be treated with caution, in that the present analyses focused solely on speakers' performance on certain isolated syllables rather than the whole pitch pattern. Since the F0 of the intervening unstressed syllables was not measured, there is a chance that Russian speakers applied higher pitch continuously to all syllables from SS to PS, unstressed ones included: during recordings, idiosyncratic F0 use by Russian learners of English was anecdotally observed, possibly a reflection of its relatively low importance as a stress cue (Bondarko 1977, 1998).

Finally, vowel quality analyses, determined by F1 and F2 measures, revealed significant differences between the two speaker groups. The average F1 measured for Russian learners of English was 580 Hz, which, according to the mean F1 frequencies of American English vowels compiled by Kent and Read (2002), falls in the mid or high-mid vowel frequency range. 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/.

F2 analyses showed that both groups demonstrated comparable performance in the production of the low back vowel /ɑ/. In contrast, F2 for the low front vowel /æ/ for English speakers was higher than for Russian learners of English, indicating that Russian speakers' tongue position was significantly less frontal than native English speakers'. Combined with F1 analysis findings, these results suggest that Russian speakers reduced /æ/ quality to a pure "schwa", both in terms of tongue height and advancement.

The evident temporal and vowel quality differences between English and Russian productions thus provide the very first empirical evidence for the anecdotally observed yet previously undocumented phenomenon of vowel over-reduction by Russian learners of English.

Experiment 2: Perceptual Study

The goal of Experiment 2 was to determine the degree to which the non-standard production of SS and UU syllables established in Experiment 1 affects native English listeners' word recognition performance. This study did not attempt to make any direct comparisons of UU versus SS syllable importance but rather compare UU syllables to themselves, and SS syllables to themselves, under different conditions. The perceptual experiment used a

cross-modal lexical decision task in a phonological priming paradigm to measure the degree of listeners' lexical activation as a function of the type of speech—native, non-native, or "corrected" non-native—they were exposed to. The phonological priming paradigm was selected due to its ability to maximally reflect real-time underlying speech processing and listeners' ease of accessing the mental lexicon when exposed to certain types of speech (Nicol et al. 2006).The lexical decision task required participants to respond whether visually presented letter strings were real English words or not, and the priming paradigm assessed changes in response speed and accuracy based on the auditory stimuli, or primes, participants heard beforehand: native-spoken words with accurate SS or UU syllables, Russian-spoken words with inaccurate SS and UU syllables, Russian-spoken words with artificially "corrected" SS and UU syllables (i.e., syllables which had been manipulated to have acoustic characteristics closer to those of native English speakers), or native-spoken unrelated words. Upon presentation of the same word visually shortly afterwards, listeners were expected to be fast in responding that it was a real word after hearing native English-spoken words due to having pre-activated the word at the auditory stage of its presentation; slow after hearing unrelated words; and slow after hearing Russian-spoken words with inaccurate SS and UU vowels due to no pre-activation. Finally, response times to Russian-spoken words with "corrected" or "modified" vowels depended on how large a role the accuracy of SS and UU syllables played for lexical access. Since stressed syllables are phonetically more prominent than unstressed, listeners might regard SS syllables as carrying more prosodic weight and therefore attend to SS syllable accuracy more than the accuracy of UU syllables that lack the stress feature.

Design

The perceptual experiment utilized a within-subjects design, with type of speech (English-spoken, Russian-spoken, modified Russian-spoken, and unrelated) and type of stress (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. A priming paradigm was combined with a cross-modal lexical task, where the activation of a visually presented word (target) was measured after an auditory presentation of a phonologically related word (prime).

Participants

The participants were 28 native speakers of American English, all of them undergraduate students at a university in Ohio. They were paid a small fee for their participation. Participants were asked to complete a questionnaire on their linguistic background to ensure that they had no knowledge of Russian or previous exposure to Russian speakers. A standard hearing screening was administered prior to the experiment; all participants passed the test successfully.

Materials

Each critical trial involved one of four types of auditory stimuli, produced by various speakers from Experiment 1: (i) native English-spoken words, (ii) Russian-spoken words with unmodified SS or UU syllables, (iii) Russian-spoken words with acoustically "modified" SS and UU syllables, and (iv) native English-spoken words that were unrelated to the visual target. The third type of experimental stimulus, "modified" productions of Russian, were necessary to isolate the perceptual effects of inaccurate SS and UU syllables from those

of Russian-accented speech in general. By acoustically correcting SS and UU syllables in Russian original productions, any changes in response times (RTs) 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.

Acoustic modifications were executed with Praat acoustic analysis software (Boersma and Weenink 2011). Each deviant SS and UU vowel in Russian-accented speech was substituted with a matching PS vowel of identical vowel type and phonetic environment from the recordings of the same Russian speaker, excised and spliced in place of the improperly produced vowel. Cutting and splicing were done at zero crossings and with careful observation of the pitch period pattern in order to avoid clicks and other unnatural speech effects. Vowel F1 and F2 were not altered. Vowel duration, intensity, and F0, on the other hand, were acoustically systematically modified for the excised vowels to match the average values obtained from native speakers in the acoustic study, yet the naturalness of speech was the most important criterion. Intensity was modified by using the intensity scaling option in Praat and specifying the desired dB value. Vowel duration and F0 for the corresponding vowels were manipulated by using the Pitch Synchronous OverLap-Add algorithm, or PSOLA (Moulines and Charpentier 1990), in Praat software. All experimental stimuli, as well as filler items, were normalized for peak intensity (scaled to 0.99).

Four different counterbalanced lists were created. The experimental prime-target pairs on each list were 15 SS-syllable-containing and 15 UU-syllable-containing words from the acoustic study, 10 of which were heard by each listener as English-spoken, 10 as Russian-spoken, and 10 as modified Russian-spoken. Each experimental word was heard only once by each participant. For the experimental stimuli, the auditory prime and visual target matched, e.g., participants heard the word "confirmation" and shortly afterwards saw "confirmation" on the computer screen. Next, a control condition was created with 30 unrelated native-spoken auditory primes, e.g., listeners heard "policy" and saw "confirmation". Additionally, 120 prime-target filler pairs with 45 real and 75 phonotactically legitimate pseudo-words were created, e.g., participants heard "finance" and saw "plaquast" on their screens. All control and filler items were prime-target non-matching pairs, which ensured that listeners would not anticipate the repetition inherent in the priming paradigm. Finally, half of the visual targets were real words and half were pseudo-words.

Procedure

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.

The experiment began with a short practice block, followed by 4 blocks of 50 experimental trials in each block. The order of stimulus presentation was randomized. Each trial began with a warning tone, followed by an auditory prime after 200 ms, presented over headphones. A visual target appeared in the middle of the computer screen at the offset of the auditory stimulus, and listeners had to respond maximally quickly and accurately whether the letter string was a real word by pressing either the 'Yes' or 'No' button. Visual targets were separated by 3 s pauses. RT was measured from the onset of the visual target to the moment the response button was clicked, and only correct lexical decisions, not longer than 2 s, were considered. Response times were expressed as a "priming" effect, or the difference in RT between the experimental conditions and the non-priming control condition.

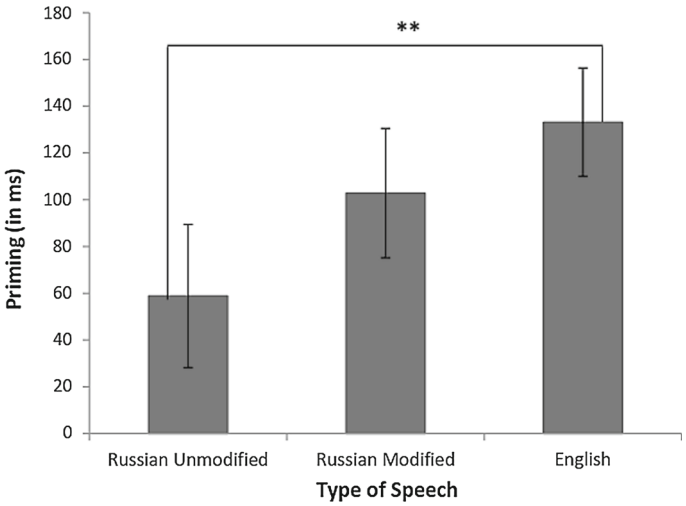

Fig. 6 Priming effect (in ms) after hearing words that contained secondary-stressed (SS) syllables in three conditions: (i) unmodified Russian-spoken, (ii) modified Russian-spoken, and (iii) native English-spoken

Results

A one-way repeated measures ANOVA was conducted separately for words with SS and UU syllables. Post-hoc tests with Bonferroni correction were performed to determine which experimental conditions differed significantly from each other, and from the control condition. Analyses included only correct responses (98.7 %).

Secondary-Stressed Syllables

Figure 6 shows priming results for words with SS syllables. Statistical analyses revealed a main effect of the Type of Speech (native, non-native, modified non-native, unrelated) heard as a prime when data were organized by participant ($F_1(2, 54) = 4.052$, $p = 0.023$), but not when they were organized by item ($F_2(2, 28) = 1.655$, $p = 0.209$), which indicates that the observed effect might not be very reliable. Post-hoc tests for by-participant analyses showed that the degree of priming after hearing different speech types 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 presented a clear case in their ability to generate priming effects were Russian Unmodified and English primes. As expected, the observed priming was significantly greater when SS-syllable-containing words were spoken by English speakers ($M = 133$ ms) than when they were spoken by Russian speakers in the Unmodified condition ($M = 58$ ms)($t(27) = 2.984, p = 0.006$). Regarding the Russian Modified condition ($M = 102$ ms), priming was not significantly different from the English condition ($M = 133$ ms)($t(27) = 1.021, p = 0.316$), however, at the same time only marginally different from the Russian Unmodified ($M = 58$ ms) condition ($t(27) = 1.849$, $p = 0.075$).

Separate one-sample t tests assessed the degree of priming generated by each of the three speech types relative to the non-priming control condition. The difference between priming in Russian Unmodified and Unrelated word conditions was marginally significant

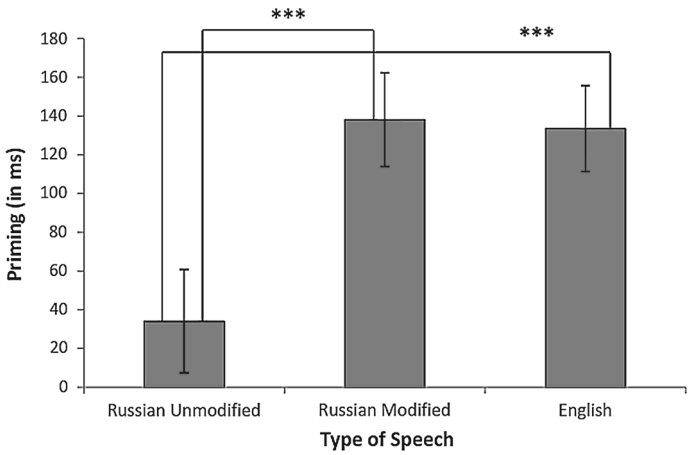

Fig. 7 Priming effect (in ms) after hearing words that contained unstressed–unreduced (UU) syllables in three speech conditions: (i) unmodified Russian-spoken, (ii) modified Russian-spoken, and (iii) native English-spoken

($t(27) = 1.92$, $p = 0.065$). Modified Russian words exhibited a significant effect of priming, compared to the control condition ($t(27) = 3.72$, $p = 0.001$), and English productions demonstrated a substantial degree of priming ($t(27) = 5.752$, $p < 0.001$). All of the comparisons to the control condition were also performed on the same data organized by item, and the findings were identical.

Unstressed–Unreduced Syllables

Figure 7 shows the degree of priming for words with UU syllables. A one-way repeated measures ANOVA revealed that the level of priming differed significantly as a function of the type of speech heard as a prime, both when data were organized by participant ($F_1(2, 54) = 21.817$, $p < 0.001$) and by item ($F_2(2, 28) = 14.975$, $p < 0.001$).

Post-hoc tests on the data organized by participant showed that the level of priming in the Unmodified Russian condition ($M = 34$ ms) was significantly smaller than in the English ($M = 133$ ms)($t(27)$=5.142, $p < 0.001$), or Modified Russian conditions ($M = 138$ ms)($t(27) = 5.249$, $p < 0.001$). 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, separate one-sample t tests assessed the priming of each speech type relative to the non-priming control condition. Both English and Modified Russian conditions demonstrated a clear priming effect in that they were both significantly different from the control condition—($t(27) = 6.034$, $p < 0.001$) 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 control condition ($t(27) = 1.276$, $p = 0.213$).

Discussion

The priming experiments investigated the role of SS and UU syllable accuracy on native listener speech processing. In line with expectations, analyses showed that priming in general was significantly greater for English-spoken than Russian-spoken words, presumably due to accented speech effects. The difference in priming between Russian Unmodified and Russian Modified speech, on the other hand, was an indication of the effect that solely SS and UU syllable accuracy had on native perception. Overall, the data presented mixed evidence for the importance of SS syllables. The difference between the priming levels of Unmodified and Modified Russian-spoken words was only marginally significant, yet, compared to the control condition, Russian Modified words generated strong priming effects in listeners, while words in the Russian Unmodified condition with low-accuracy SS syllables showed only weak priming. Another contributor to the mixed evidence was the great range of variability in priming for the Unmodified and Modified Russian conditions, which indicates that not all Russian-spoken words with improper SS syllables were equally incomprehensible, or that all corrected SS syllables led to native-like priming. Finally, only analyses by participant showed significant differences in priming between the types of speech, but not by item. This suggests that, overall, the observed differences might be best described as trends toward increased priming with SS syllable accuracy.

Results regarding UU syllables contrast starkly with those of SS syllables. Analyses revealed significantly more priming for English-spoken words, and Russian-spoken words with "modified" UU syllables than unmodified Russian-spoken words. In fact, after UU syllable modification Russian-spoken words were recognized just as quickly as English-spoken. Moreover, both English and Modified Russian conditions showed strong priming relative to the control condition, while the Unmodified Russian condition showed no priming at all. This finding has provided evidence that, contrary to predictions, inaccurate UU syllables may affect speech intelligibility and hinder lexical access.

General Discussion

The goals of this study were, first, to document the possible over-reduction of vowels in the speech of Russian learners of English, which is a previously undocumented phenomenon in second language research; and second, to investigate the role of SS and UU syllable accuracy in spoken word recognition—so far an underexplored topic. To this end, an acoustic study first assessed differences in SS and UU syllable production by Russian learners of English and native English speakers by measuring vowels in terms of F0, intensity, duration, F1 and F2. A subsequent perceptual experiment further explored the effect of temporal and vowel quality reductions in Russian speech on native speaker speech processing; it used a phonological priming method in a lexical decision task to implicitly assess the degree of lexical activation native speakers exhibited in response to such speech.

Results of the acoustic study showed that the most striking differences between native-spoken and Russian-accented English were in the temporal and vowel quality dimensions; differences in intensity and fundamental frequency were less pronounced and inconclusive. Specifically, Russian-spoken UU and SS syllables were approximately only half the duration of English-spoken ones. Similarly, vowel quality in Russian-spoken UU and SS syllables was significantly reduced: vowel formant measures indicated that Russian speakers produced both vowel types with the jaw significantly less open and tongue higher in the oral cavity than native English speakers. Specifically for the low front vowel /æ/, the tongue was also in

a significantly more central position, thus in Russian speaker production it became a truly reduced vowel that has very little in common with the native-spoken low front vowels.

The following priming study provided only partial support for the predictions that there would be strong lexical activation for English-spoken words, slightly less activation for Russian-spoken words with "corrected" SS and UU syllables, and very little activation for unmodified Russian-spoken words. Russian-spoken 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 short and centralized final vowels, were either not recognized or produced considerable delays in the speech recognition process. After acoustically correcting UU syllables, Russian-spoken words were recognized by native listeners maximally quickly and reached native-English speech priming levels. This is an important finding that clearly identifies inaccurate UU syllable production as one of the main contributors to unintelligibility. In contrast, Russian-spoken words with SS syllables triggered some lexical activation even when SS syllables were inaccurate, possibly due to being long multisyllabic words with relatively few lexical competitors. Unsurprisingly, correcting 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 largely recognizable in the first place. Thus, based on the current findings, UU syllables appear to be critical for speech processing and lexical access, while SS syllable accuracy might not be as essential for word identity as previously thought.

The acoustic study contributes uniquely to second language research. So far, studies have shown that L2 learners typically struggle with under-reduction of unstressed vowels due to experience with the L1 phonological system (Braun et al. 2011; Flege and Bohn 1989). The opposite—vowel over-reduction—to date has not been addressed in second language research literature. Russian participants' systematic reductions of vowel duration and quality before and after PS in this study closely mirror patterns of the Russian rhythmic system, where only one lexically prominent syllable per word is phonologically permitted, and the rest undergo primarily temporal and vowel quality reductions (Avanesov 1956; Bondarko 1998). Additionally, the relatively weak phonological status of SS and UU syllables in English could also play a role: according to Jones (in Lehiste 1970) SS is not used for phonological contrasts, i.e., word meaning cannot be distinguished based on SS category alone, with a few exceptions such as verb-adjective word pairs like "graduate" or "alternate". Word meanings are typically not distinguished based on UU syllables alone, either, with a few exceptions, such as minimal pairs "radar" and "raider", or "audition" and "addition". Consequently, Russian speakers might not view preserving full vowels in SS and UU syllables as phonologically meaningful, which contributes to learners' difficulty perceiving and producing such features accurately in a second language. White and Mattys (2007) showed in a study with Dutch speakers of English that subtle rhythmic differences are not noticed by second language learners if language rhythms appear deceptively similar or even identical. The shared phonetic features between Russian and English rhythms, such as stressed syllable realization and vowel reduction, might lead Russian learners of English to a conclusion that due to the similarities no further adjustments are necessary.

Findings from the current perceptual experiment are relevant for speech perception research, and informative for selecting goals in second language pronunciation instruction. The perceptual experiment did not provide evidence that stressed syllables, SS syllables in this case, have an important prosodic status such that they serve as "navigational guides" for accessing words in the mental lexicon (Grosjean and Gee 1987; Murphy 2004; Murphy and Kandil 2004). Russian-spoken words, regardless of whether SS syllables were inaccurate or "corrected", were relatively recognizable by listeners. This finding may be attributed to

the word length effect (Pitt and Samuel 2006), whereby long multisyllabic words with SS syllables provide enough unique acoustic-phonetic material to compensate for the inaccuracy of a single syllable. In contrast, an inaccuracy of a syllable in disyllabic words such as UU-syllable-containing words resulted in a single accurate syllable left in the word for listeners to rely on, and subsequently led to stronger interference with speech processing than SS syllable inaccuracy. Thus the importance of these syllable types in practical terms appears to be intrinsically and uniquely tied to the overall length of the word that comprises them, and less concerned with their prosodic status.

The perceptual study has answered the call for more empirical research that would identify the importance of various phonetic features for intelligibility and subsequently L2 instruction (Derwing 2010; Derwing and Munro 2005; Levis 2005; Munro and Derwing 2006). So far, for example, research on suprasegmentals has found that comprehension or speech processing time has been affected by erroneous sentence stress (Hahn 2004), inaccurately placed or realized primary stress (Bond and Small 1983; Cutler and Clifton 1984; Field 2005; Small et al. 1988), insufficient reduction of reduced vowels in unstressed syllables (Braun et al. 2011), or overall speech rate (Anderson-Hsieh and Koehler 1988; Kang 2010; Munro and Derwing 1998). The importance of UU and SS syllable accuracy for intelligibility, on the other hand, has not been previously addressed in published work. The detrimental effects of improper UU syllable realization on native listener comprehension carries important implications for second language pedagogy. It is strongly recommended that UU syllables be included in the pronunciation syllabus for Russian learners of English, and that Russian speakers of English pay special attention to accurate implementation of this syllable type in their speech. The mixed results obtained for SS syllables would suggest that their accuracy might be relatively less critical for intelligibility.

Finally, it should be noted that conclusions about the importance of SS and UU syllable accuracy for perception should be limited to those lexical contexts that were specifically addressed in this study. The word length of UU-containing words is a factor that constrains broader generalizations. Disyllabic words with UU syllables were selected for this study to match those phonetic environments that feature the strongest vowel reductions in Russian, although UU syllables can also occur in English words of greater length (e.g., the first syllable in "auditioning" or "idolatry"). SS syllables, on the other hand, always appear in the English language in long multisyllabic words. Some other limitations of the study are the specific locations of target syllables. Plag et al. (2011) found acoustically important distinctions between SS syllables in pre- and post-PS positions; since this study addressed SS syllables in pre-PS positions only, study findings should be limited to these contexts. Finally, the current study used only low vowels /ɑ/ and /æ/ as having most potential to be reduced in terms of vowel quality and duration in Russian. The specific syllable locations, word lengths and syllable types were optimally selected to maximize the potential vowel reduction effect that Russian learners of English would exhibit in their speech, but might prevent from generalizing the current findings to other contexts. Having demonstrated interesting results in these particular contexts, this study calls for follow-up studies that would address those contexts that were outside the scope of this study to provide complete and generalizable results.

Appendix

Experimental Stimuli for the Acoustic Study

See Table 1.

Table 1 Experimental stimuli for the acoustic study

	Secondary-stressed (SS)	Unstressed–unreduced (UU)	Primary-stressed (PS)
/ɑ/	opti' mistic	'robot	'optimist
	ope'rational	'radar	'politics
	domi'nation	'neuron	'popular
	poli'tician	'python	'hospital
	popu'lation	'nylon	'dominate
	popu'larity	'moron	ro'botic
	conden'sation	'icon	i'conic
	confis'cation	'mascot	'rockery
	accomo'dation	'coupon	'cottony
	hospi'tality	'shamrock	res'pondent
/æ/	appli'cation	'syntax	syn'tactic
	grati fi'cation	'abstract	abs'traction
	acci'dental	'contract	ex'tracted
	compati'bility	'borax	con'traction
	elabo'ration	'extract	com'paction
	magnifi'cation	'combat	'gratitude
	collabo'ration	'format	'magnetize
	mascu'linity	'impact	'accident
	classifi'cation	'contact	'labyrinth
	fabri'cation	'compact	'classify

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

References

Anderson-Hsieh, J., & Koehler, K. (1988). The effects of speaking rate and foreign accent on native speaker comprehension. *Language Learning*, *38*(4), 561–613.

Avanesov, R. I. (1956). *Fonetika Sovremennogo Russkogo Literaturnogo Jazyka (Phonetics of modern Russian language)*. Moscow: Izdatelstvo Moskovskogo Universiteta.

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* (pp. 7–33). Cambridge: Cambridge University Press.

Beckman, M. E. (1986). *Stress and non-stress accent*. Dordrecht: Foris.

Boersma, P., & Weenink, D. (2011). Praat: Doing phonetics by computer (Version 4.3.01) [Computer program]. http://www.praat.org/

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., & Mani, N. (2011). Perceiving unstressed vowels in foreign accented English. *Journal of the Acoustical Society of America*, *129*(1), 376–387.

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.

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* (pp. 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.

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.

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. 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.

Fry, D. B. (1955). Duration and intensity as physical correlates of linguistic stress. *Journal of the Acoustical Society of America*, *27*, 765–768.

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.

Kang, O. (2010). Relative salience of suprasegmental features on judgments of L2 comprehensibility and accentedness. *System*, *38*, 301–315.

Kent, R. D., & Read, C. (2002). *Acoustic analysis of speech* (2nd ed.). San Diego, CA: Singular.

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., & 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.

Mattys, S. (2000). The perception of primary and secondary stress in English. *Perception and Psychophysics*, *62*(2), 253–265.

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.

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. (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.

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.

Padgett, J., & Tabain, M. (2005). Adaptive dispersion theory and phonological vowel reduction in Russian. *Phonetica*, *62*, 14–54.

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.

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.

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

White, L., & Mattys, S. L. (2007). Calibrating rhythm: First language and second language studies. *Journal of Phonetics*, *35*, 501–522.

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.