Journal of Cognitive Psychology, 2015
Vol. 27, No. 8, 944–966, http://dx.doi.org/10.1080/20445911.2015.1058267

Predicting stress patterns in an unpredictable stress language: The use of non-lexical sources of evidence for stress assignment in Russian

Olessia Jouravlev and Stephen J. Lupker

Department of Psychology, University of Western Ontario, London, ON, Canada N6A 5C2

(*Received 22 September 2014; accepted 29 May 2015*)

The main goal of this research was to examine how readers of Russian assign stress to disyllabic words. In particular, we tested the claim that the process of stress assignment in Russian can only be accomplished lexically. Eleven potential non-lexical sources of evidence for stress in Russian were examined in regression and factorial studies. In Study 1, onset complexity, coda complexity, the orthography of the first syllable (CVC1), of the second syllable (CVC2), and of the ending of the second syllable (VC2) were found to be probabilistically associated with stress in Russian disyllables. In Studies 2 and 3, it was shown that Russian speakers do use 3 of these cues (CVC1, CVC2, and VC2) when making stress-assignment decisions. These results provide evidence against the idea that the nature of stress in the Russian language is so unpredictable that stress assignment can only be accomplished lexically. These results also suggest that any successful model of stress assignment in Russian needs to contain mechanisms allowing these 3 orthographic cues to play a role in the stress-assignment process.

Keywords: Binary logistic regression; Coda complexity; Lexical stress; Linear mixed-effects model; Onset complexity; Orthography; Russian; Stress cues; Syllable; Word beginning; Word ending; Word reading.

INTRODUCTION

Lexical stress is used to distinguish between meanings of words in many languages and, hence, its proper placement in polysyllabic words is crucial for successful communication. In the field of reading research, several computational models explaining the principles of lexical stress assignment within the dual-route, connectionist, or probabilistic inferential frameworks have been proposed (Arciuli, Monaghan, & Seva, 2010; Jouravlev & Lupker, in press; Perry, Ziegler, & Zorzi, 2010; Rastle & Coltheart, 2000; Seva, Monaghan, & Arciuli, 2009). Despite the fact that all these models differ in the way they conceptualise the process of stress assignment in reading, they do have one point in common: Readers can assign lexical stress by computing it based on stress rules or non-lexical sources of evidence for stress (stress cues) present in the language.

The ability of readers to predict what syllable in a written word should be stressed based on different

Correspondence concerning this article should be currently addressed to Olessia Jouravlev, McGovern Institute for Brain Research, Massachusetts Institute of Technology, Cambridge, MA, USA. E-mail: olessiaj@mit.edu
The authors would like to thank Deb Jared, Marc Joanisse, Ken McRae, Victor Kuperman, and an anonymous reviewers for their extremely helpful comments on earlier drafts of this manuscript.
No potential conflict of interest was reported by the authors.
This research was supported by Natural Sciences and Engineering Research Council of Canada Grant A6333 to Stephen J. Lupker.

non-lexical aspects of the word has been successfully demonstrated in a number of languages (e.g., Arciuli et al., 2010; Burani & Arduino, 2004). More specifically, previously identified sources of evidence for stress include diacritics (Gutierrez-Palma & Palma-Reyes, 2008; Protopapas, 2006), the orthographic complexity of word onsets and codas (Kelly, Morris, & Verrekia, 1998; Kelly, 2004), the orthography of word endings and beginnings (Colombo, 1992; Jouravlev & Lupker, 2014), morphology (Rastle & Coltheart, 2000), and grammatical category (Arciuli & Cupples, 2006; Kelly & Bock, 1988).

Prior research has, therefore, provided us with a plethora of knowledge about possible cues for computing lexical stress patterns in reading. However, due to the nature of factorial experimental designs, the predominantly used technique in the earlier cited studies, it has been difficult to investigate multiple sources of evidence for stress within a single data set, making it impossible to assess the strength of reliance readers place on those cues relative to each other. Further, it is likely that there are other cues to stress assignment that have not yet been discovered. Factorial techniques do not provide any obvious way of discovering those cues. Thus, in the present research, in addition to using a factorial manipulation (in our final study), the approach taken was to examine the effect of a wide range of potential stress cues simultaneously. Doing so involved both examining the stress patterns in a large corpus of disyllabic words as well as analysing a large set of disyllabic word naming data. The present work is limited to disyllables only (with that decision having been made in order to simplify the design of the studies conducted as a part of this research project). Therefore, the conclusions that we draw based on the empirical investigations reported here do not necessarily generalise to all polysyllabic words. Nevertheless, we believe that examining stress assignment in disyllables is the appropriate first step towards a full understanding of the mechanisms of stress assignment in polysyllabic word reading.

The main purpose of the present studies was to extend the investigation of potential mechanisms of stress assignment to the Russian language. Considering the importance of lexical stress in distinguishing meanings of numerous stress minimal pairs present in Russian (e.g., *за́мок* ([zámok],[1] "castle")—*замо́к* ([zamók], "lock"); *му́ка* ([múka], "burden")—*мука́* ([muká], "flour)), and the very high rate of stress-assignment errors demonstrated by native speakers of this language (Jouravlev & Lupker, 2014), understanding the mechanisms of stress assignment in Russian is an important endeavour in and of itself. Russian is transparent in its mapping of orthography to phonology; however, it has a complicated system of lexical stress that neither is explicitly marked in the orthography (i.e., diacritics are not used in texts for adult native readers) nor does it conform to any clear implicit rules. Another feature of the Russian stress system is that, unlike many previously investigated languages that have more frequently occurring stress patterns (also known as regular stress patterns; Colombo, 1992; Kelly et al., 1998), in Russian disyllables there is no overall regular stress pattern (Jouravlev & Lupker, 2014). Further, Russian lexical stress is flexible, that is, any syllable in a word may be stressed with inflected forms of the same lexeme often having different stress patterns (e.g., *рука́* [ruká], singular for "hand"—*ру́ки* [rúki], plural for "hand").

There are a number of linguistic theories that attempt to describe the types of alternations of stress patterns across inflected forms of the same lexeme present in Russian (Halle, 1997; Melvold, 1989; Zaliznjak, 1985). The common idea shared by these theories is that Russian morphemes are stored in the lexicon along with the corresponding information about their "accented" (i.e., stressed) or "unaccented" (i.e., unstressed) status. A word's stress is computed using a set of complicated rules that govern how "accented/unaccented" information retrieved for each morpheme from lexical memory is combined for a particular word. Thus, the above-mentioned linguistic theories of stress essentially posit that stress assignment is completed primarily via lexical mechanisms in Russian. In fact, the view that there are no useful non-lexical cues to stress assignment in Russian, meaning that speakers of Russian can only assign lexical stress to words by retrieving stress information from lexical memory, is very common in linguists and psycholinguists (Molczanow, Domahs, Knaus, & Wiese, 2013; Zsiga, 2013).

The suggestion that stress assignment in Russian reading is a completely lexical phenomenon stands in contrast, however, to recently published behavioural findings by Jouravlev and Lupker (2014), who investigated the effect of stress regularity (overall and at the level of grammatical category) and of the stress consistency of the word's ending (defined as an association between the orthography

[1] Here and in all subsequent examples, Roman transliterations of any Russian words are given in square brackets.

of the word's ending and one of the potential stress patterns) on naming performance for readers of Russian. As those results showed, while there were no processing differences for trochaically stressed (first-syllable stress) vs. iambically stressed (second-syllable stress) words overall or in separate analyses of nouns and verbs (presumably due to the absence of a regular stress pattern for words in those grammatical categories), trochaically stressed adjectives were named faster and more accurately than iambically stressed adjectives (presumably due to the fact that trochaic stress is a regular stress pattern for Russian adjectives). The effect of the stress consistency of the word's ending was also successfully demonstrated, with words that contained endings representative of those words' stress patterns enjoying a processing advantage over words that had endings associated with a stress pattern different than the pattern of the to-be-named word.

Jouravlev and Lupker's (2014) results do indicate that the process of stress assignment in Russian is not completely a lexically based one and that readers of Russian do use non-lexical information (in particular, the knowledge of a regular stress pattern for certain types of words and knowledge of the consistency with which the orthography of the word's ending maps onto one of the stress patterns) in assigning stress to polysyllabic words. In the present research, we continued investigating the mechanisms of stress assignment in Russian by considering a greater range of sources of evidence for stress. We begin the discussion with a consideration of the sources of evidence for stress that have been suggested by prior research.

SOURCES OF EVIDENCE FOR STRESS

An orthographic element that explicitly indicates what syllable in a word should be stressed is known as a diacritic. In many languages (including Russian), the use of diacritics is optional and unlikely to occur in texts created for adult skilled readers. However, even in languages in which these stress marks are obligatory, the processing advantage for words that contain diacritics is unclear. For example, there is some evidence that diacritics are useful stress cues in Spanish (Gutierrez-Palma & Palma-Reyes, 2008), but are often ignored in Greek (Protopapas, 2006; Protopapas, Gerakaki, & Alexandri, 2006).

Stress patterns can also be marked in the orthography implicitly via associations that exist between orthographic components and stress patterns. One of the associations of this type is that of stress and the orthographic complexity of words' onsets (Kelly, 2004) and codas (Kelly et al., 1998). For example, disyllabic English words with orthographically complex onsets are more likely to be trochaically stressed, whereas disyllabic words with simple onsets are more likely to be iambically stressed (Kelly, 2004). Further, words with complex codas tend to have iambic stress, whereas words with simple codas tend to have trochaic stress (Kelly et al., 1998). Most importantly, the behaviour of speakers of English in naming and recognising words has been shown to be impacted by these stress cues (Kelly, 2004; Kelly et al., 1998).

Another source of evidence for stress patterns is the orthography of word beginnings and endings (Arciuli & Cupples, 2006, 2007). In this research, words with the same orthographic component in their structure are defined to be "neighbours" (e.g., *mark-et*, *brack-et*, *cad-et*). Neighbours that have the same stress pattern are called "stress friends" (*market*: *bracket*), whereas neighbours with different stress patterns are called "stress enemies" (*market*: *cadet*). A word that has mainly "stress friends" is called consistent, whereas a word with mainly "stress enemies" is called inconsistent. A strong effect of spelling-to-stress consistency of word endings has been demonstrated in Italian (Burani & Arduino, 2004), in English (Arciuli & Cupples, 2006; Arciuli et al., 2010), and in Russian (Jouravlev & Lupker, 2014). Note, however, that in some studies, the impact of spelling-to-stress consistency of word endings has been demonstrated only when readers name words with a less frequent stress pattern (Colombo, 1992; Jouravlev & Lupker, 2014, when examining naming of Russian adjectives). In English, it has also been shown that readers can also use word beginnings as useful cues (Arciuli & Cupples, 2007; Arciuli et al., 2010) to stress assignment.

Another potential source of evidence for stress patterns present in some languages is morphology (Lagerberg, 1999; Protopapas, Gerakaki, & Alexandri, 2007), a cue that some researchers consider to be an important source of information about stress (Rastle & Coltheart, 2000). Nevertheless, it seems unlikely that morphology would be a primary stress cue as this cue is only available for polymorphemic words and, therefore, cannot explain how readers manage to assign stress to monomorphemic words.

Finally, the grammatical category of a word may be a source of evidence for stress patterns. For instance, in English, trochaic stress is more typical in disyllabic nouns, whereas disyllabic verbs often exhibit iambic stress. Evidence for the impact of grammatical category on stress assignment has been provided in naming and lexical decision tasks (Arciuli & Cupples, 2004, 2006; Kelly & Bock, 1988). Arciuli and Cupples (2006), however, proposed that the relationship between grammatical category and lexical stress might be artifactual. The word's orthography might be cuing its grammatical category and its lexical stress at the same time, essentially independently of one another. Therefore, the correlation between grammatical category and stress pattern might be an artifact of the relationship between each of these factors and some set of orthographic cues.

As clear from this review, researchers have proposed a number of sources of evidence that readers may use in assigning stress to polysyllabic words. Some of these cues are available for polymorphemic words only (e.g., stress relevant affixes) or are present in a limited number of languages (e.g., diacritics). Stress cues that have been shown to influence the processing of polysyllables in a large number of languages and that are relevant for all words regardless of their morphological status appear to derive from the orthography of the word (in particular, the orthographic complexity of word onsets and codas and the orthography of word beginnings and word endings). Finally, there is some suggestion that the grammatical category of a word is a potential stress cue although, at this point, it is unclear whether grammatical category cues stress directly or indirectly via the relationships between orthographic cues and both grammatical category and stress patterns.

THE PRESENT RESEARCH

The present research extended the previous investigations of the use of stress cues in a number of ways. Almost all previous studies used factorial experimental designs to test the effect of each potential stress cue independent of all other cues. While factorial studies can allow researchers to conclude that a manipulated variable produced an effect, this approach also has its limitations. For example, the stimuli in the various conditions need to be matched on other dimensions, including the existence or nonexistence of other potential stress cues that might impact processing. Beyond the difficulty involved in creating stimulus sets equated on all relevant dimensions, there is the real potential that doing so will produce sets of very atypical items. Further, in factorial studies, item selection is somewhat biased as researchers mainly use items characterised by extreme values of an examined independent continuous variable which is artificially transformed into a binary variable for experimental purposes. For instance, in examining the effect of spelling-to-stress consistency of word endings, researchers compare the performance of readers on highly consistent versus highly inconsistent words only rather than using the whole range of the spelling-to-stress consistency measure. Finally, as noted earlier, any factorial investigation can only focus on the specific cue(s) selected for examination. The above-mentioned limitations can be overcome to some degree with a regression design that allows for a more exploratory type of investigation. However, one should acknowledge that the regression approach is not without problems as well due to a number of statistical issues such as the collinearity of predictors (Tabachnick & Fidell, 2007). Because both factorial and regression designs have some advantages and disadvantages, some researchers have suggested combining them (Chateau & Jared, 2003; Treiman, Mullenix, Bijeljac-Babic, & Richmond-Welty, 1995). Therefore, in the present research, we used both regression and factorial approaches to examine potential sources of evidence for stress in Russian disyllabic words.

The potential sources of evidence for stress that we examined were Grammatical Category, Log Frequency, Length, Word Onset Complexity, Word Coda Complexity, and a set of six orthographic components. The orthographic components were the First Syllable (referred to as CVC1), the Beginning of the First Syllable (CV1), the Ending of the First Syllable (VC1), the Second Syllable (CVC2), the Beginning of the Second Syllable (CV2), and the Ending of the Second Syllable (VC2). An orthographic component is an abstract structural element across all words of the same syllabic length in a language. For example, the orthographic component CVC1 refers to a structural component of the first syllable of all disyllabic words. Further, this component has different orthographic identities (further referred to as spellings) in different words. For instance, the spelling of CVC1 is *мар-* [mar-] in the word *маркер* [marker], but *кар-* [kar-] in the word *карман* [karman].

Do note that the symbols C and V refer not just to single consonants or vowels, but rather to all letters of that type before an instance of the next type is encountered (e.g., in the word *стрóйка* [strójka], *стро-* [stro-] is the CV1 component). Another thing to note is that not all Russian words conform to the CVC1.CVC2 pattern. While a nucleus vowel is a required component in any syllable, consonants in onset and coda positions may be omitted. In these cases, "dummy" onsets and codas (indicated by an asterisk) were introduced for purposes of our analysis. For example, the following orthographic components were identified for the word *áрка* [árka]: **ар* (CVC1), **а* (CV1), *ар* (VC1), *ка** (CVC2), *ка* (CV2), and *а**(VC2). An example of the division of a word into the orthographic components examined here is presented for the word *мáркер* [márker] in Figure 1.

The main empirical goal of the present studies was to determine whether probabilistic associative relations exist between the examined variables and stress patterns in the Russian language (Study 1) and whether readers of Russian use knowledge of these associations in their stress-assignment performance (Studies 2 and 3). The finding of an associative link between Grammatical Category and Stress would mean that words of different grammatical categories tend to have particular stress patterns (cf. trochaic stress in English nouns and iambic stress in English verbs). Log Frequency could be of theoretical interest if certain stress patterns are more likely to occur in words that readers encounter frequently or, alternatively, in words that occur rarely in the language. The identification of Word Length as a reliable stress cue would mean that certain stress patterns tend to occur in longer than in shorter words of the language or vice versa. Word Onset (Coda) Complexity could also be an important variable capable of predicting stress if words with complex onsets (codas) are more likely to have certain stress patterns than words with simple onsets (codas).

Finally, with respect to the question of the usefulness of various orthographic components, we attempted to determine whether certain structural components of a word (e.g., CVC1 or VC2) reliably provide spelling information that might help readers in their stress-assignment decisions. Essentially, we assessed whether spelling-to-stress consistency patterns are found for a number of such components in Russian. If spellings of a particular component can help readers in assigning stress in a large number of words in a language, then, this component can be considered as a reliable source of evidence for stress. If so, readers may be likely to rely on spelling information provided by that component in their language as, in most cases, the spelling of the component would provide correct information about stress. Therefore, stress-assignment performance should be improved when the spelling of that component cues a stress pattern that is consistent with the actual stress that the word has and it might be disrupted if the particular identity of that component cues a stress pattern that is inconsistent with the actual stress that a word has. On the other hand, if spellings of a particular component can help readers in assigning stress in only a small number of words (or not at all), then, readers would, presumably, not rely on the spelling information provided by that component. Therefore, stress-assignment performance would not be impacted by the consistency/inconsistency of spellings of that component. Essentially, in order to assess the reliability of an orthographic component as a stress cue, we examined how often particular spellings of this component might predict stress patterns correctly across all disyllabic words of the Russian language. If particular spellings of a component predict stress patterns across many words in a given language, then, this component would be considered a reliable stress cue which may be used by readers/speakers of the language. On the other hand, if particular spellings of a component are not capable of predicting stress patterns across words in a language, then, this component would not be a reliable stress cue.

The choice of Grammatical Category, Word Onset Complexity, Word Coda Complexity, CV1, and VC2 components as potential predictors of stress was, to some extent, empirically driven (Arciuli & Cupples, 2004; Jouravlev & Lupker, 2014; Kelly et al., 1998; Kelly, 2004). The other variables examined have not previously been investigated as sources of evidence for stress. Their inclusion was driven by our intuition and a desire to explore as many potential stress cues as possible. With respect to the orthographic components of a word that we examined, two involved syllables themselves because syllables have been

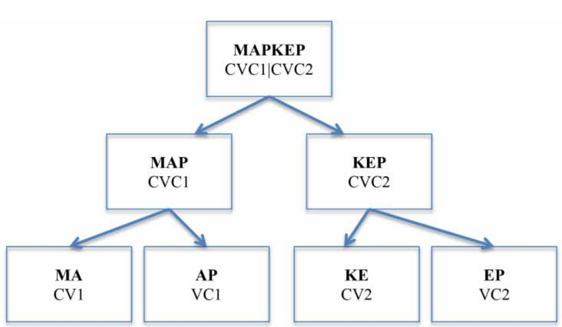

Figure 1. The division of the word МАРКЕР [MARKER] into six orthographic components for calculating spelling-to-stress consistency.

suggested to play an important role in visual word recognition previously (Carreiras & Perea, 2002; Conrad & Jacobs, 2004), and four involved components of syllables. Two components of syllables (i.e., CV1 and VC2) have been investigated previously (Arciuli & Cupples, 2006, 2007; Jouravlev & Lupker, 2014). To obtain a broader picture of the impact of various orthographic components on the process of stress assignment, CV2 and VC1 components were also included as potential predictors of stress patterns.

In the first regression study (Study 1), the predictability of stress patterns in a corpus of disyllabic words based on a number of potential cues was assessed. In addition to examining relationships existing in this normative corpus, we felt that it would also be useful to investigate real language behaviour for a number of reasons. First, there is the possibility that, in most cases, not all sources of evidence for stress present in the language are needed/used by readers in assigning stress. A number of sources of evidence present in a word might provide the same cue to the word's stress pattern, and, thus, given this redundancy, one or more of them may be ignored (i.e., stress-assignment decisions may be based on a subset of the sources of reliable evidence for stress present in the language). This subset is likely to be those sources of evidence for stress that are the most reliable in the language. A second reason for investigating the correlation between sources of evidence for stress and the stress patterns actually produced by readers is that real language usage might deviate significantly from the canonical, prescribed usage reflected in the dictionaries. Although language users are generally aware of the correct stress patterns that most words have in their native language, their actual performance is not perfect. They assign incorrect stress patterns to some words on some occasions and/or they can have difficulty in deciding quickly what stress pattern to apply to a word, thus, requiring additional time for naming such words. In order to investigate what sources of evidence for stress might cause differences in the ease of processing of disyllabic words, in Study 2, we collected naming latencies and stress-assignment accuracy data for a large number of disyllabic words and ran another set of regression analyses assessing the effects of many potential sources of evidence for stress on behavioural performance. Finally, in order to validate the sources of evidence for stress identified in the regression studies, we further examined those sources in a factorial design in Study 3 with a new set of words that were not used in the regression analysis of Study 2. The examination of a large set of potential cues to stress in regression and factorial studies is, in some sense, a new methodological approach with its goal here being identifying non-lexical sources of evidence for stress in Russian.

Another novel aspect of this research is the language of the stimuli/participants. Prior research on stress assignment has mainly been conducted in English and Italian, languages that have stress patterns that are highly frequent in the language. For example, in English, about 80% of disyllabic words are trochaically stressed (CELEX database: Baayen, Pipenbrock, & Gulikers, 1993). Readers of English might, therefore, rely to a reasonable extent on this distributional information in making stress-assignment decisions. In this research, we investigated stress cues in Russian disyllables, for which there does not appear to be a more frequent stress pattern (Jouravlev & Lupker, 2014). As has been already noted, the task of assigning stress to words is so complicated in Russian that it might be the case that readers of this language always complete this task by accessing lexical representations of words and retrieving their stress patterns from memory. Alternatively, despite the seeming complexity of the stress system in Russian, the identification of stress patterns in reading might proceed in a way similar to that in languages with more predictable stress; specifically, the stress-assignment process might also involve computations based on information that is present in the orthography of words. A demonstration of the use of this type of information in the process of stress assignment in Russian would serve as strong evidence for the universal nature of non-lexical mechanisms of stress-assignment in reading.

STUDY 1: BINARY LOGISTIC REGRESSION OF A SET OF PREDICTORS ON STRESS PATTERNS IN A CORPUS OF RUSSIAN DISYLLABIC WORDS

Method

Materials. A corpus of 13,945 disyllabic Russian words was compiled for the present study. Only words with a word frequency of at least one per million words (as per the Russian National Corpus; http://ruscorpora.ru) were included.[2] The compiled corpus contained polymorphemic as well as monomorphemic words because polymorphemic words are quite common in highly inflectional languages such as Russian. The words were used as items in a binary logistic regression. The binary dependent variable was the stress pattern of the word coded as "0" (trochaic stress pattern) or "1" (iambic stress pattern) based on the Dictionary of Russian Lexical Stress (Zarva, 2001). The length measure was the number of letters in the word. The onset complexity of each word was calculated by counting the number of consonants in the word onset positions. The coda complexity of each word was calculated in a similar way, by counting the number of consonants in the word ending positions.

A set of six orthographic components that correspond to syllables and sub-syllabic units as potential sources of evidence for stress was also examined. As there is a debate about what constitutes a syllable in Russian linguistics (Avanesov, 1956; Scherba, 1957), we decided to investigate empirically the method of syllabification preferred by Russian speakers and, then, to use this information to aid us in deciding on the appropriate syllabification of our words. Doing so involved an online survey in which 20 native speakers of Russian indicated how they would divide 200 disyllabic words into syllables by typing in the first and second syllables for each word. One half of the words had one intervocalic consonant in their structure (e.g., *каток* [ka**t**ok], *ротик* [ro**t**ik], *домик* [do**m**ik]). A reader might establish the syllable division for these words in one of the two ways: (1) by maximising the coda of the first syllable (*кат-ок* [kat-ok]) or (2) by maximising the onset of the second syllable (*ка-ток* [ka-tok]). The other half of the words used had complex intervocalic consonant clusters in their structure (e.g., *маска* [ma**sk**a], *близкий* [bli**zk**ij], *выслать* [vi**sl**at']). For these words, a reader might establish the syllable division in one of the three ways: (1) by maximising the coda of the first syllable (*маск-а* [mask-a]), (2) by maximising the onset of the second syllable (*ма-ска* [ma-ska]), or (3) by splitting a consonant cluster between the coda of the first syllable and the onset of the second syllable (*мас-ка* [mas-ka]).

The results of this survey showed that the Russian readers consistently syllabified disyllabic words containing one intervocalic consonant by maximising the onset of the second syllable (97%), which is consistent with the Maximal Onset Principle, a widely recognised principle of syllabification in contemporary linguistics (Giegerich, 1992). Therefore, in our analyses, words with one intervocalic consonant in their structure were divided into syllables in such a way that an intervocalic consonant was always assigned to the beginning of the second syllable (e.g., *ка-ток* [ka-tok], *ро-тик* [ro-tik], *до-мик* [do-mik]). For words that have complex intervocalic consonant clusters in their structure, readers preferred to split the consonant cluster between the coda of the first syllable and the onset of the second syllable (78% of responses), followed by maximising the onset of the second syllable (19% of responses), and, finally, by maximising the coda of the first syllable (3%). Based on these findings, in our analyses, words with complex intervocalic consonant clusters were divided into syllabic units by splitting a consonant cluster between the coda of the first syllable and the onset of the second syllable, except in a few instances in which the created syllables had illegal onsets and/or codas or violated morphemic divisions (e.g., the word *выслать* [vislat'], meaning "to send away") was divided into the prefix *вы-* [vi-], "away") + the root *-слать* [-slat'], "to send")).

Next, the four sub-syllabic orthographic components were identified in each word. The beginning of the first syllable (CV1) corresponded to all initial consonants of the first syllable preceding the vowel plus the vowel of that syllable. The beginning of the second syllable (CV2) corresponded to all initial consonants of the second syllable preceding the vowel plus the vowel of that syllable. The ending of the first syllable (VC1) corresponded to the vowel of the first syllable plus all consonants of that syllable following the vowel. The ending of the second syllable (VC2) corresponded to the vowel of the second syllable plus all consonants of that syllable following the vowel.[3]

An orthographic component of a word may be considered to be a potentially significant source of evidence for stress if the presence of spelling-to-stress consistency for this component is confirmed. In other words, for the majority of words in the corpus, the stress pattern cued by the spellings of this component should, in fact, correspond to the correct stress pattern that these words have.

[2] Because there is no available electronic database that contains relevant information about various lexical characteristics of Russian words, we had to check those characteristics manually for each word. Therefore, for practical purposes, only words with a frequency of more than 1 per million words were included in the database created for this study.

[3] The VC1 and VC2 components, as defined in the present work (i.e., components that comprise the syllable's vowel and the following consonants), are referred to in the previous literature as rimes (Treiman & Kessler, 1995).

Therefore, for every word in the corpus, we calculated spelling-to-stress consistency measures, in particular spelling-to-trochaic stress measures,[4] for the six orthographic components under consideration that were, then, used as predictors of trochaic stress in the corpus of all words included in the binary logistic regression. This method for calculating spelling-to-stress measures was analogous to that used by Treiman et al. (1995) for calculating spelling-to-sound consistency.

Words having the orthographic component under consideration were defined as the words comprising the target's neighbourhood (e.g., all words with VC2 "-et" form a neighbourhood).[5] In calculating the type consistency measure, the proportion of words with trochaic stress in the neighbourhood was calculated by dividing the number of words in the neighbourhood having trochaic stress by the total number of words in the neighbourhood. High values mean that in the specified orthographic neighbourhood, the number of words with trochaic stress is higher than the number of words with iambic stress, whereas low values mean that there are more words with iambic than trochaic stress in the neighbourhood. Values near .50 indicate that the neighbourhood has approximately equal percentages of the two stress types (i.e., it is an inconsistent neighbourhood). In addition to type consistency measures, token consistency measures, corresponding to the summed frequency of words with trochaic stress in an orthographic neighbourhood divided by the summed frequency of all words in that orthographic neighbourhood, were calculated.[6]

Descriptive statistics for the examined predictors and correlations between each predictor and stress patterns in the corpus of Russian disyllabic words are given in Table 1. The results of the bivariate correlations between all of the examined predictor variables in the corpus of Russian disyllabic words are provided in Table 2.

Results and discussion

A set of binary logistic regressions was run to predict stress patterns for words in the corpus using combinations of 11 predictors: *Grammatical Category* (coded using dummy coding), *Log Frequency*, *Length*, *Onset Complexity*, *Coda Complexity*, *CVC1*, *CV1*, *VC1*, *CVC2*, *CV2*, and *VC2*. The goal was to find a model with a minimum number of factors that would still have high predictive power. In other words, the goal was to find a balance between the simplicity of the model and its goodness of fit. The goodness of fit of a model was assessed using the Akaike information criterion (AIC), the Bayesian information criterion (BIC), and the deviance information criterion (DIC). A model that minimises AIC, BIC, and DIC is a preferred choice. Further, to select amongst competing models a likelihood ratio test was computed. The analysis was conducted using the R package *lme4* (Bates & Maechler, 2010). Due to the fact that for the six examined orthographic components spelling-to-stress consistency could be measured based on type or token count (i.e., weighted by word frequency), two separate sets of regressions were conducted. The measures of goodness of fit of the full and the final models are given in Table 3.

In the analysis where the six orthographic components reflecting type-based consistency measures were used in the equation (subsequently referred to as a *type-based analysis*), the final simplified model fit the data better than a null model with intercept only, χ^2 (6) = 9,862.23, $p < .001$, and as well as the full model, χ^2 (7) = 12.56, $p = .08$. The final model had six (non-lexical) predictors of stress: *Onset Complexity* ($z = -2.38$, $p = .02$), *Coda Complexity*

[4] Therefore, a word in a completely trochaic neighbourhood would get a 1.0 trochaic stress consistency value, while a word in a completely iambic neighbourhood would get a 0.0 trochaic stress consistency value.

[5] In calculating spelling-to-stress consistency measures, most researchers consider only words of the same syllabic length (Arciuli & Cupples, 2006; Arciuli, Monaghan, & Seva, 2010), although in some work in Italian all words of the language regardless of their number of syllables were taken into consideration during the computation of consistency measures (Burani & Arduino, 2004). No empirical investigation has been conducted to determine which approach provides a better reflection of the processes that take place during lexical stress assignment in word reading. We decided to use the former way of calculating spelling-to-stress consistencies as this approach appears to be more consistent with the architecture of the CDP++ (Perry et al., 2010), a model that proposes that "the processing system has information about the number of syllables of an orthographic input before stress pattern information is computed (non-lexically)."

[6] It should be noted that both the type and token comparisons have limitations due to the fact that only words with a frequency of more than 1 per million words were included in the database used in the regression analyses. Potentially, this action may have affected the type analysis more than the token analysis because each word in the neighbourhood is assumed to be equally important in any type analysis, whereas words with very low frequencies will have very little impact in a token analysis. Nonetheless, as will be seen, the type analysis appeared to be the more reliable analysis in our studies.

TABLE 1
Descriptive characteristics of the corpus of words (Study 1) and their correlations with stress patterns

Word characteristics	*Descriptive characteristics for* Trochaic words	Iambic words	Total	r_{pb} *between a variable and stress patterns (coded as 0 and 1)*
Grammatical category				
Nouns	4,680	3,885	8,565	.005
Verbs	1,100	1,844	2,944	.142**
Adjectives	1,707	401	2,108	−.221**
Other	162	166	328	.001
Log frequency	1.87 (1.82)	1.74 (2.02)	1.82 (1.84)	−.035**
Length	5.67 (1.11)	5.67 (1.05)	5.67 (1.08)	.003
Onset complexity	1.38 (.67)	1.24 (.66)	1.32 (.67)	−.102**
Coda complexity	.7 (.72)	.96 (.72)	.82 (.73)	.152**
Stress consistency				
CVC1-Type/CVC1-Token	.73 (.27)/.70 (.25)	.65 (.26)/.62 (.22)	.69 (.29)/.66 (.32)	−.591**/−.549**
CV1-Type/CV1-Token	.64 (.19)/.62 (.20)	.51 (.19)/.54 (.17)	.63 (.20)/.58 (.21)	−.393**/−.361**
VC1-Type/VC1-Token	.62 (.19)/.59 (.17)	.48 (.15)/.50 (.12)	.56 (.16)/.55 (.18)	−.315**/−.289**
CVC2-Type/CVC2-Token	.72 (.20)/.72 (.22)	.65 (.28)/.66 (.30)	.69 (.31)/.69 (.32)	−.615**/−.585**
CV2-Type/CV2-Token	.62 (.15)/.61 (.15)	.52 (.20)/.50 (.19)	.59 (.18)/.56 (.18)	−.365**/−.341**
VC2-Type/VC2-Token	.66 (.18)/.65 (.18)	.59 (.24)/.59 (.25)	.63 (.25)/.62 (.24)	−.499**/−.485**

Notes: Trochaic words refer to disyllabic words with stress on their first syllable. Iambic words refer to disyllabic words with stress on their second syllable. Total refers to all disyllabic words of the corpus ($N = 13{,}945$). The distribution of words across grammatical categories is given in total counts of words belonging to each grammatical category in the corpus. For log frequency, length, onset complexity, coda complexity, and spelling-to-stress consistency measures, means and standard deviations (in brackets) are provided. All spelling-to-stress consistency measures reflect mapping of spelling of the corresponding component onto the correct stress patterns that words in the corpus have (i.e., trochaic stress for the trochaically stressed words and iambic stress for the iambically stressed words). **Point-biserial correlation (r_{pb}) is significant at $p < .01$.

($z = 4.94$, $p < .001$), *CVC1* ($z = -46.54$, $p < .001$), *CV2* ($z = 5.14$, $p < .001$), *CVC2* ($z = -32.99$, $p < .001$), and *VC2* ($z = -2.11$, $p = .04$).[7]

In the analysis where the six orthographic components reflecting token-based consistency measures were used in the equation (subsequently referred to as a *token-based analysis*), the final simplified model fit the data better than the null model, $\chi^2(9) = 7{,}803.21$, $p < .001$, and as well as the full model, $\chi^2(4) = 8.83$, $p = .07$. That model contained nine (non-lexical) predictors: *Onset Complexity* ($z = -2.49$, $p = .02$), *Coda Complexity* ($z = 7.66$, $p < .001$), *Length* ($z = 3.59$, $p = .003$), *CVC1* ($z = -36.58$, $p < .001$), *CV1* ($z = -2.57$, $p = .01$), *VC1* ($z = -2.69$, $p = .01$), *CVC2* ($z = -38.67$, $p < .001$), *CV2* ($z = 3.70$, $p = .01$), and *VC2* ($z = -19.99$, $p < .001$).

Despite the fact that the final model obtained in the type-based analysis had fewer predictors (six vs. nine), this model provided a significantly better fit to the data than the final model obtained in the token-based analysis, $\chi^2(3) = 994.04$, $p < .001$. Therefore, the six non-lexical variables that were significant predictors of stress patterns in the type-based model were taken to be relevant stress cues. (Note that the former six predictors are a proper subset of the latter nine predictors.) The following significant associations between stress patterns and predictor variables in the corpus of disyllabic words were identified: (1) words with complex onsets had trochaic stress more often than words with simple onsets; (2) words with complex codas had iambic stress more often than words with simple codas; and (3) the spelling information provided by four components (CVC1, CVC2, CV2, and VC2) was strongly associated with stress patterns.

This latter pattern means that, in a large number of words in the Russian disyllabic corpus, there are associations between spellings of the CVC1, CVC2, CV2, and VC2 components and stress patterns. Thus, readers can rely on spelling information provided by these components in assigning stress to disyllabic words. On the other hand, the number of words in the corpus with spellings of the CV1 and VC1 components cuing stress patterns was small,

[7] Recall that trochaic stress was coded as "0" and iambic stress was coded as "1". Therefore, a negative z value in this analysis denote a significant association between a predictor variable and trochaic stress, while a positive z value represent a significant relationship between a predictor variable and iambic stress.

TABLE 2
Correlations of predictor variables (Pearson's r) in the corpus of disyllabic words used in Study 1

	Noun	*Verb*	*Adjective*	*Other*	*LogFreq*	*Length*	*Onset*	*Coda*	*CVC1*	*CV1*	*VC1*	*CVC2*	*CV2*	*VC2*
Noun	1	−.65	−.53	−.09	−.04	−.27	−.11	−.30	−.06 (−.05)	−.02 (−.02)	−.07 (−.05)	.01 (.01)	−.06 (−.03)	.02 (.02)
Verb	−.65	1	−.22	−.11	−.07	.21	.10	.37	−.15 (−.15)	−.07 (−.07)	−.12 (−.13)	−.29 (−.28)	−.23 (−.24)	−.34 (−.33)
Adjective	−.53	−.22	1	−.07	.01	.17	.05	.01	.27 (.25)	.13 (.12)	.25 (.23)	.31 (.29)	.35 (.32)	.35 (.35)
Other	−.09	−.11	−.07	1	.03	.15	.03	.05	−.04 (−.04)	−.03 (−.02)	−.08 (−.06)	.09 (.07)	−.08 (−.07)	.07 (.06)
LogFreq	−.04	−.07	.01	.03	1	−.17	−.03	−.17	−.01 (−.01)	−.03 (−.04)	−.05 (−.05)	.06 (.06)	−.01 (.01)	.09 (.09)
Length	−.27	.21	.17	.15	−.17	1	.50	.58	.13 (.13)	.09 (.08)	.17 (.17)	−.04 (−.05)	.06 (.04)	−.08 (−.07)
Onset	−.11	.10	.05	.03	−.03	.50	1	−.04	.18 (.15)	.26 (.22)	.03 (.02)	.09 (.09)	.08 (.06)	.09 (.08)
Coda	−.30	.37	.01	.05	−.17	.58	−.04	1	−.12 (−.10)	−.10 (−.08)	−.07 (−.07)	−.29 (−.29)	−.15 (−.15)	−.36 (−.33)
CVC1	−.06 (−.05)	−.15 (−.15)	.27 (.25)	−.04 (−.04)	−.01 (−.01)	.13 (.13)	.18 (.15)	−.12 (−.10)	1 (.93)	.62 (.59)	.51 (.46)	.37 (.34)	.31 (.30)	.30 (.29)
CV1	−.02 (−.02)	−.07 (−.07)	.13 (.12)	−.03 (−.02)	−.03 (−.04)	.09 (.08)	.26 (.22)	−.10 (−.08)	.62 (.59)	1 (.92)	.30 (.28)	.23 (.21)	.17 (.17)	.20 (.17)
VC1	−.07 (−.05)	−.12 (−.13)	.25 (.23)	−.08 (−.06)	−.05 (−.05)	.17 (.17)	.03 (.02)	−.07 (−.07)	.51 (.46)	.30 (.28)	1 (.92)	.22 (.22)	.25 (.23)	.20 (.20)
CVC2	.01 (.01)	−.29 (−.28)	.31 (.29)	.09 (.07)	.06 (.06)	−.04 (−.05)	.09 (.09)	−.29 (−.29)	.37 (.34)	.23 (.21)	.22 (.22)	1 (.95)	.56 (.53)	.78 (.78)
CV2	−.06 (−.03)	−.23 (−.24)	.35 (.32)	−.08 (−.07)	−.01 (.01)	.06 (.04)	.08 (.06)	−.15 (−.15)	.31 (.30)	.17 (.17)	.25 (.23)	.56 (.53)	1 (.93)	.43 (.44)
VC2	.02 (.02)	−.34 (−.33)	.35 (.35)	.07 (.06)	.09 (.09)	−.08 (−.07)	.09 (.08)	−.36 (−.33)	.30 (.29)	.20 (.17)	.20 (.20)	.78 (.78)	.43 (.44)	1 (.97)

Notes: Because type counts were better predictors when considering the six orthographic components, correlations with type counts are shown on top in each cell with the correlations involving token counts shown in parenthesis. Note that the correlations between type and token counts are between .92 and .97.

TABLE 3
Measures of goodness of fit for the full and the final models from the binary logistic regression analysis of stress patterns in the corpus of Russian disyllabic words (Study 1)

Model type	AIC	BIC	DIC
A. Goodness of fit measures			
Full model (11 factors)	8856	8953	8830
Final model (6 factors)	8850	8913	8827
B. Goodness of fit measures			
Full model (11 factors)	9585	9954	9831
Final model (9 factors)	9558	9926	9829

Notes: *A*, measures for the models obtained in the type-based analysis; *B*, measures for the models obtained in the token-based analysis. A model that minimises AIK, BIK, and DIC is the preferred choice.

suggesting that readers are unlikely to rely on the information provided by the spellings of these two components in making their decisions about stress patterns.

Expectedly, spellings of CVC1, CVC2, and VC2 cuing trochaic stress were associated with a trochaic stress pattern and spellings cuing iambic stress were associated with an iambic stress pattern. The CV2 component, which was also identified as a significant predictor of stress, showed, on the other hand, a strange pattern of association between spellings and stress patterns. More specifically, as noted, for each component, spelling-to-trochaic stress consistency measures (with higher values meaning that the spelling of the component cues trochaic stress and lower values meaning that the spelling of the component cues iambic stress) were entered into the equation as predictors of trochaic (coded as "0") or iambic stress (coded as "1"). Thus, we expected that any significant correlations for the orthographic components should be in the negative direction. This was, in fact, what we observed for CVC1, CVC2, and VC2, but not for CV2. For CV2, in contrast, the correlation was positive. Thus, according to this analysis, spellings of CV2 cuing trochaic stress (i.e., that had high spelling-to-trochaic stress consistency values) were associated with iambic stress and spellings of CV2 cuing iambic stress (i.e., that had lower spelling-to-trochaic stress consistency values) were associated with trochaic stress. A pattern such as this, of course, does not make much sense. We believe that the reversed relationship with stress patterns that CV2 demonstrated is likely the result of a suppression effect due to the presence in the equation of other variable(s) that are strongly correlated with

TABLE 4
Descriptive characteristics of the set of words used in Study 2 and their correlations with stress patterns

	Descriptive characteristics for			
Word characteristics	*Trochaic words*	*Iambic words*	*Total*	r_{pb} *between a variable and stress patterns (coded as 0 and 1)*
Grammatical category				
Nouns	158	123	281	.004
Verbs	44	62	106	.170**
Adjectives	66	22	88	−.172**
Other	2	3	5	.003
Log frequency	1.30 (1.13)	1.10 (1.01)	1.21 (1.08)	−.107**
Length	5.58 (1.02)	5.57 (0.99)	5.58 (1.01)	.009
Onset complexity	1.42 (.67)	1.20 (.63)	1.32 (.65)	−.148**
Coda complexity	.60 (.63)	.95 (.61)	.75 (.62)	.247**
Stress consistency				
CVC1-Type/CVC1-Token	.72 (.23)/.70 (.23)	.58 (.20)/.57 (.23)	.66 (.22)/.64 (.23)	−.490**/−.449**
CV1-Type/CV1-Token	.60 (.19)/.61 (.19)	.52 (.17)/.47 (.18)	.56 (.18)/.55 (.19)	−.330**/−.301**
VC1-Type/VC1-Token	.58 (.15)/.58 (.16)	.53 (.10)/.53 (.11)	.56 (.13)/.56 (.15)	−.224**/−.218**
CVC2-Type/CVC2-Token	.74 (.20)/.72 (.22)	.71 (.26)/.70 (.28)	.73 (.22)/.71 (.26)	−.653**/−.630**
CV2-Type/CV2-Token	.58 (.12)/.60 (.14)	.57 (.18)/.55 (.21)	.58 (.16)/.57 (.17)	−.351**/−.343**
VC2-Type/VC2-Token	.65 (.18)/.63 (.19)	.63 (.23)/.63 (.21)	.64 (.21)/.63 (.20)	−.507**/−.503**

Notes: Trochaic words refer to disyllabic words with stress on their first syllable. Iambic words refer to disyllabic words with stress on their second syllable. Total refers to all disyllabic words of the set (N = 480). The distribution of words across grammatical categories is given in total counts of words belonging to each grammatical category in the corpus. For log frequency, length, onset complexity, coda complexity, and spelling-to-stress consistency measures, means and standard deviations (in brackets) are provided. All spelling-to-stress consistency measures reflect mapping of spelling of the corresponding component onto the correct stress patterns that words in the corpus have (i.e., trochaic stress for the trochaically stressed words and iambic stress for the iambically stressed words).
**Point-biserial correlation (r_{pb}) is significant at $p < .01$.

stress patterns and that act as suppressor(s) of CV2 (Wurm & Fisicaro, 2014). The idea that we have observed a suppression effect for the CV2 variable is supported by the fact that the direction of the bivariate correlation between CV2 and stress patterns was negative, as would be expected (see Table 1). Such behaviour of the CV2 variable (i.e., a positive regression coefficient in the multiple regression equation despite the fact that the binary correlation between the predictor and the outcome variable is negative) appears to be a classical illustration of the suppression effect in action (Friedman & Wall, 2005; Maassen & Bakker, 2001).

STUDY 2: GENERALISED LINEAR MIXED-EFFECTS REGRESSION OF A SET OF PREDICTORS ON STRESS-ASSIGNMENT PERFORMANCE BY NATIVE SPEAKERS OF RUSSIAN.

In order to conclude that any particular cue is an important source of evidence for stress, it is necessary to demonstrate that this cue is not only of high validity (i.e., strongly associated with a stress pattern), the goal of Study 1, but also of high utility (i.e., that it impacts readers' performance in word naming). To assess the utility of the stress cues, we applied a generalised linear mixed-effects model with the set of predictors used in Study 1 to response latencies and stress-assignment performance of readers naming disyllabic words. Thus, we investigated whether some sources of evidence for stress are in associative probabilistic relations with stress patterns not only in a normative language corpus, but also in observable word naming performance of readers, performance that might, in fact, deviate from the standards and norms reflected in the corpus.

Method

Participants. Thirty-four undergraduate students from Altay State University (Barnaul, Russia) took part in this experiment for a small monetary remuneration (age 17–23; $M = 19$). All were native speakers of Russian. None of the participants reported high proficiency in any second language.

Materials. A set of 500 disyllabic words (see Appendix A) was randomly selected from the database created for Study 1. This set included both polymorphemic and monomorphemic words.[8] Out of 500 selected words, 20 had ambiguous stress (also known as minimal stress pairs) as they corresponded to two lexical items that differed in stress pattern only (e.g., *со́ски* [sóski]—plural for "pacifier" vs. *соски́* [soskí]—plural for "nipple"). These words might be the source of stress-assignment confusion for readers; therefore, we decided to exclude them from the study.

To make sure that the selected set of words is representative of the full corpus of Russian disyllabic words from the point of view of associations existing between the identified cues and stress patterns, we ran a binary logistic regression on these words that was similar to that carried out in Study 1 (see Table 4 for the descriptive statistics for the examined predictors and correlations between each predictor and stress patterns in the selected set of Russian disyllabic words). The results of this analysis showed that five variables were individually associated with stress patterns in the corpus of 480 words: *Onset Complexity* ($z = -1.93$, $p = .05$), *Coda Complexity* ($z = 1.98$, $p = .05$), *CVC1* ($z = -5.07$, $p < .001$), *CVC2* ($z = -5.46$, $p < .001$), and *VC2* ($z = -1.80$, $p = .06$). These stress cues were also reported to have high validity in predicting stress patterns in Study 1, where a corpus of 13,943 words was analysed. The only variable that was a significant predictor of stress in Study 1 that did not make it into the final model in this analysis of the much smaller corpus was CV2. However, remember that in Study 1, the CV2 variable demonstrated an unexpected reversed relationship with stress patterns, implying that the appearance of this component in many words with trochaic stress was associated with a higher likelihood that a word with that pattern had iambic stress. As discussed, this fact implies that this factor's inclusion in the set of predictors appears to have been the result of a suppression effect. Most importantly, overall, the results of the binary logistic regression on the corpus of 480 selected words suggest that this corpus is representative of the much larger corpus of Russian disyllabic words in that the same stress cues have high validity.

[8] In this research, morphology as a potential stress cue was not examined directly. In Russian, there are a few derivational morphemes, mainly of foreign origin (e.g., -изм (-ism) as in фашизм (fascism), атеизм (atheism)) that may be predictive of assigned stress patterns; however, the majority of derivational and inflectional morphemes are not associated with a single stress pattern and, thus, are likely not highly reliable sources of evidence for stress.

Procedure

Participants were instructed to read aloud words presented on a screen as quickly and as accurately as possible. Instructions and stimuli were presented using the DMDX display system (Forster & Forster, 2003). Participants' response times defined as the elapsed time between the word's presentation and the triggering of the voice-key as well as complete vocal responses were recorded. The list of 480 words used as stimuli in Study 2 was mixed with another 520 disyllabic words, some of which were to be analysed in Study 3. The total of 1,000 words was presented in 4 blocks of trials. Every participant named all blocks of trials. The order of blocks and of words within blocks was randomised for each participant. Each trial started with the presentation of a fixation point for 500 ms. The target word in upper case appeared in white on a black background (Courier New, 12 font) for 2,000 ms or until the participant responded. The intertrial interval was 1,000 ms.

Results and discussion

Using CheckVocal (Protopapas, 2007), the first author and two other native speakers of Russian who were unaware of the purpose of the experiment marked stress patterns (i.e., trochaic vs. iambic) that participants assigned to words. There were no cases that were treated by these markers as ambiguous from the point of view of stress pattern assignment. A pronunciation of a word with a trochaic stress was coded as "0"; whereas a pronunciation of a word with an iambic stress was coded as "1."

First of all, we examined whether stress patterns assigned by readers to words were any different from stress patterns that these words have according to dictionaries. Although there was a strong positive correlation between behavioural and normative data (which was expected as readers do know these words and their correct stress patterns), the correlation was not perfect, r (16,318) = .85, $p < .001$. Participants incorrectly assigned iambic stress to trochaically stressed words in approximately 6% of cases and trochaic stress to iambically stressed words in 11% of cases. Thus, stress-assignment performance of readers of Russian did, in fact, deviate to a certain degree from the norms described in dictionaries. Study 2 allowed us to investigate what factors might have produced such deviations from the norms in the behavioural data as well as to assess how the examined factors impacted the speed of naming responses.

In the analysis of stress patterns assigned by readers, that categorical variable (i.e., which pattern was assigned by the majority of the readers) was the dependent measure in a generalised linear mixed-effects model (GLMM), with *Subjects* and *Items* considered as random crossed factors and *Grammatical Category*, *Log Frequency*, *Length*, *Onset Complexity*, *Coda Complexity*, *CVC1*, *CV1*, *VC1*, *CVC2*, *CV2*, and *VC2* considered as fixed factors. In the analysis of naming latencies, the characteristics of the initial phoneme (vowel; bilabial-occlusive; labiodental-fricative; labiodental-sonorant; alveolar-occlusive; alveolar-affricate; alveolar-fricative; alveolar-sonorant; velar-fricative; glottal-occlusive; and glottal-fricative) which might have an impact on the triggering of the voice-key was also included as an additional fixed factor. The characteristics of the initial phoneme are not of theoretical interest in this study, but that factor was included in order to remove any variance associated with it. Note also that in the analysis of naming latencies, spelling-to-stress consistency measures of orthographic components were computed differently from that in the analysis of stress patterns assigned by readers (where the computation was of the spelling-to-trochaic stress consistency for all words). In particular, for words with a trochaic stress pattern, spelling-to-trochaic stress consistency was computed, and for words with an iambic stress pattern, spelling-to-iambic stress consistency was computed.

Similar to Study 1, two separate sets of GLMMs were conducted. In the first set (the type-based analyses), six orthographic components reflecting spelling-to-stress consistency measures based on type counts were entered into regression equations along with the rest of the variables. In the second set of regressions (the token-based analyses), six orthographic components reflecting spelling-to-stress consistency measures based on token counts were entered into the regression equation along with the rest of the variables. The analyses were exploratory in nature as the goal was to find a model with a minimum possible number of factors that would fit the data well. The procedures were identical to those used in Study 1. The goodness of fit of a model was assessed using the same measures and likelihood ratio tests as in Study 1 (see Table 5).

Further, due to the fact that in prior research, some studies reported significant interactions of spelling-to-stress consistency of word endings and lexical frequency (i.e., a consistency effect is present for low, but not for high frequency words;

TABLE 5
Measures of goodness of fit for the full and the final models from the generalised linear mixed-effects model analysis of stress pattern assignment performance of native speakers of Russian naming 480 disyllabic words (Study 2)

A. Goodness of fit measures			
Model type	AIC	BIC	DIC
Full model (11 factors)	7478	7585	7451
Final model (4 factors)	7474	7532	7451
B. Goodness of fit measures			
Model type	AIC	BIC	DIC
Full model (11 factors)	7586	7698	7558
Final model (6 factors)	7490	7535	7480

Notes: *A*, measures for the models obtained in the type-based analysis; *B*, measures for the models obtained in the token-based analysis. A model that minimises AIK, BIK, and DIC is the preferred choice.

Colombo, 1992), we decided to explore whether such relations would be observed in our data set. Therefore, additional sets of GLMMs with interactions of *Log Frequency* and *CVC1, CV1, VC1, CVC2, CV2, and VC2* entered as fixed factors among the other predictor variables of interest (that were described earlier) were run. These models were fitted both to stress pattern assignment and response latency data following the same procedure as the models without the interaction factors.

Analysis of stress patterns assigned by readers. The simplified model for the type-based analysis had four predictors, although only three of them were individually significant predictors of stress patterns assigned to words: *VC1* ($z = -1.50$, $p = .13$), *VC2* ($z = -2.43$, $p = .02$), *CVC2* ($z = -9.47$, $p < .001$), and *CVC1* ($z = -5.46$, $p < .001$). This model's fit to the data was as good as that of the full model, χ^2 (8) = 7.85, $p = .45$, and significantly better than that of the null model, χ^2 (4) = 444.57, $p < .001$.

In the simplified model for the token-based analysis, three factors were significant predictors of stress pattern assignment: *VC2*, $z = -5.14$, $p < .001$, *CVC2*, $z = -9.09$, $p < .001$, and *CVC1*, $z = -9.55$, $p < .001$. This model could explain the data as well as the full model, χ^2 (9) = 12.80, $p = .17$, and significantly better than the null model, χ^2 (3) = 152.61, $p < .001$. Finally, the simplified model for the type-based analysis fit the data significantly better than the model for the token-based analysis, χ^2 (1) = 19.95, $p < .001$.

The same three factors remained significant predictors of stress pattern assignment performance in the models with interaction effects entered into the equations (significant predictors of stress in the type-based analysis: *VC2*, $z = -4.03$, $p < .001$, *CVC2*, $z = -5.14$, $p < .001$, and *CVC1*, $z = -9.37$, $p < .001$; significant predictors of stress in the token-based analysis: *VC2*, $z = -5.34$, $p < .001$, *CVC2*, $z = -4.48$, $p < .001$, and *CVC1*, $z = -9.50$, $p < .001$). In addition, there were significant interactions of *CVC2* and *Log Frequency* (type-based analysis: $z = -1.95$, $p = .05$; token-based analysis: $z = -2.81$, $p = .01$) and of *VC2* and *Log Frequency* (type-based analysis: $z = -2.11$, $p = .04$; token-based analysis: $z = -2.22$, $p = .03$), suggesting that the effect of these orthographic components on stress pattern assignment is most evident for words of low frequency. To conclude, it appears that three factors (CVC1, CVC2, and VC2) are clearly in associative relationships with stress pattern assignments of readers naming disyllabic words and that two of those factors may interact with frequency.

Analysis of naming latencies. The simplified model obtained in the type-based analysis had six fixed factors, with all of them being significant predictors of naming latencies: *Log Frequency* (t (16,810) = −8.82, $p < .001$), *Length* (t (16,810) = 3.22, $p = .001$), *Onset Complexity* (t (16,810) = 5.24, $p < .001$), *CVC1* (t (16,810) = −4.89, $p < .001$), *CVC2* (t (16,810) = −1.97, $p = .05$), and *VC2* (t (16,810) = −2.90, $p = .01$). This model's fit to the data was as good as that of the full model, χ^2 (16) = 9.80, $p = .88$, and significantly better than that of the null model, χ^2 (6) = 149.43, $p < .001$.

The simplified model obtained in the token-based analysis had the same six predictors: *Log Frequency* (t (16,810) = −7.08, $p < .001$), *Length* (t (16,810) = 3.01, $p = .002$), *Onset Complexity* (t (16,810) = 5.09, $p < .001$), *CVC1* (t (16,810) = −4.74, $p < .001$), *CVC2* (t (16,810) = −2.03, $p = .05$), and *VC2* (t (16,810) = −1.89, $p = .06$). This model's fit to the data was as good as that of the full model, χ^2 (16) = 2.32, $p = .99$, and significantly better than that of the null model, χ^2 (6) = 106.08, $p < .001$. There was no significant difference in the goodness of fit of the final models obtained in the type- versus token-based analyses, χ^2 (1) = 0.09, $p = .76$.

The same set of variables was identified as predictors of response latencies in the analysis that included interaction effects (significant predictors of response latencies in the type-based analysis: *Log Frequency*, t (16,810) = −1.93, $p = .05$; *Length*, t (16,810) = 3.21, $p = .01$; *Onset Complexity*, t (16,810) = 4.90, $p < .001$; *CVC1*, t (16,810) = −3.18, $p = .01$; *CVC2*: t (16,810) = −1.90, $p = .06$; *VC2*: t (16,810) = −2.37, $p = .02$; significant predictors of response latencies in the token-based analysis: *Log Frequency*, $t\,(16{,}810) = -2.01$, $p = .04$; *Length*, $t\,(16{,}810) = 3.26$, $p = .01$; *Onset Complexity*, $t\,(16{,}810) = 5.13$, $p < .001$; *CVC1*, $t\ (16{,}810) = -1.96$, $p = .05$; *CVC2*: $t\ (16{,}810) = -2.02$, $p = .04$; *VC2*: $t\ (16{,}810) = -2.03$, $p = .05$). In addition, in the model for the type-based analysis, there was a marginally significant interaction of *CV1* and *Log Frequency*, $t\ (16{,}810) = -1.91$, $p = .06$, whereas in the model for the token-based analysis, there was a significant interaction of *CVC1* and *Log Frequency*, $t\ (16{,}810) = -2.55$, $p = .01$. These interactions suggest that the impact of the CV1 and CVC1 orthographic components on speed of naming of disyllabic words varies with the frequency of those words with the effect of the factor being larger for low frequency words.

In conclusion, along with variables such as Log Frequency, Length, and Onset Complexity that are known to have an impact on naming latencies (Bijeljac-Babic, Millogo, Farioli, & Grainger, 2004; Brysbaert et al., 2011; Rastle & Davis, 2002), spelling-to-stress consistencies of three orthographic components that we identified as strong predictors of stress pattern assignment performance (CVC1, CVC2, and VC2) were also predictive of the speed of processing of disyllabic words. Thus, words with CVC1, CVC2, and/or VC2 components that signalled stress patterns inconsistent with these words' actual stress patterns were associated with both more stress-assignment errors and longer response times than words with CVC1, CVC2, and/or VC2 components that signalled stress patterns consistent with the words' actual stress patterns.[9]

STUDY 3: FACTORIAL INVESTIGATION OF THE IMPACT OF CVC1, CVC2, AND VC2 ON WORD NAMING

To assess whether the hypothesised relationships between CVC1, CVC2, and VC2 and stress patterns generalise to words beyond the data set used in Study 2, a further validation of each variable as a source of evidence for stress was undertaken using a factorial experimental design.[10] To do so, we selected words with CVC1, CVC2, or VC2 spellings that were highly consistent or highly inconsistent with the actual stress patterns that these words have. If readers, indeed, pay attention to the spellings of each of these components as sources of evidence for stress when making stress-assignment decisions, then, we should observe faster and more accurate performance for readers on words with highly consistent, as opposed to inconsistent, spellings of CVC1, CVC2, or VC2. The observation of such a pattern of results would mean that CVC1, CVC2, and VC2 are, indeed, valid and reliable sources of evidence for stress in Russian.

Method

Participants. The participants were the same 34 individuals who participated in Study 2.

Materials. Three testing sets, with 40 disyllabic words in each set, were created (see Appendix B). None of the words selected for Study 3 were used in Study 2. In each set, spelling-to-stress consistencies (both type-based and token-based) of one orthographic component were manipulated (consistent vs. inconsistent), whereas spelling-to-stress consistencies of two others (both type-based and token-based) were controlled. Further, in each set, words with consistent versus inconsistent spelling-to-stress mappings were matched on stress patterns, grammatical category, length, logarithmic word frequency, onset and coda complexity, and in a word-by-word manner, on initial phoneme characteristics. Due to the fact that, in Study 2, some evidence for the impact of frequency on the spelling-to-stress consistency effect for the CV1 component on the speed of naming was found (i.e., an interaction of CV1 and Log Frequency in the token-based analysis), all items were matched on the CV1 variable. The mean characteristics of the word sets are shown in Table 6.

[9] As noted, Russian is transparent in its mapping of orthography to segmental phonology, but quite opaque in its lexical stress system. This characteristic of Russian was apparent in the distribution of types of errors in Study 2. Overall, the Study 2 error rate in word naming was 9.72%, with 94.13% of these errors being stress-assignment errors. The initial analysis reported for Study 2 (the analysis of stress patterns assigned by readers) was essentially an analysis of stress-assignment errors. No analysis of the other types of errors was undertaken due to their extremely low number.

[10] Although VC1 contributed significantly to the final model in the type consistency analysis of Study 2, because it was not significant in the token consistency analysis in Study 2 nor did it emerge as a significant predictor of stress in the type consistency analysis in Study 1, the argument that it might be a viable cue to stress assignment in Russian is weak. Therefore, VC1 consistency was not one of the variables investigated in Study 3.

Procedure

Because the naming data for these words were collected at the same time as the naming data in Study 2, the procedure was identical to that of Study 2.

Results and discussion

Responses were marked using the same procedure as in Study 2. Latencies were analysed using a linear mixed-effects model, whereas error rates were analysed using a generalised linear mixed-effects model.[11] In these analyses, *Subjects* and *Items* were considered crossed random factors, and *Stress Consistency* (consistent vs. inconsistent) of *CVC1, CVC2*, or *VC2* (depending on the testing set) was considered a fixed factor. Two separate models (with type-based or token-based Stress Consistency) were fit to the data of each set. Note that in Study 3 spelling-to-correct stress consistency measures were computed and used as fixed factors in the models. In particular, for words with a trochaic stress pattern, it was a spelling-to-trochaic stress consistency measure and for words with an iambic stress pattern, it was a spelling-to-iambic stress consistency measure. Therefore, low consistency values in Study 3 indicate that the word is inconsistent with its neighbourhood. The analysis was conducted using the R package *lme4* (Bates & Maechler, 2010). The mean latencies and percentage of errors are shown in Table 7.

CVC1: In the model with the type-based Stress Consistency of CVC1 entered as a fixed factor, there was a significant main effect of Stress Consistency of CVC1 both in the accuracy ($z = -4.10$, $p < .001$) and in the latency data (t (1,286) = −2.42, $p = .02$). Further, the pattern of results remained the same in the model using the token-based Stress Consistency of CVC1: there was a significant main effect of Stress Consistency of CVC1 in the accuracy ($z = -4.61$, $p < .001$) and in the latency data (t (1,286) = −2.16, $p = .03$). The participants were faster and more accurate in naming words with consistent ($M = 636$ ms ($SD = 49$), 2% incorrect responses) versus inconsistent ($M = 680$ ms ($SD = 56$), 9% incorrect responses) mappings of the orthography of CVC1 onto stress patterns.

[11] Similar to Study 2, errors in Study 3 were mainly stress assignment errors. Overall, the Study 3 error rate in word naming was 5.33% with 97.55% of these errors being stress-assignment errors.

CVC2: There was a significant effect of Stress Consistency of CVC2 on the accuracy of performance both in the type-based ($z = -2.73$, $p = .01$) and in the token-based analyses ($z = -2.65$, $p = .01$). The effect of Stress Consistency of CVC2 on response latency was significant in the token-based analysis (t (1,286) = −1.93, $p = .05$) and marginal in the type-based analysis (t (1,286) = −1.79, $p = .07$). The participants were faster and more accurate in naming words with consistent ($M = 633$ ms ($SD = 58$), 4% incorrect responses) versus inconsistent ($M = 667$ ms ($SD = 43$), 12% incorrect responses) mappings of the orthography of CVC2 onto stress patterns.

VC2: When the type-based Stress Consistency of VC2 was entered as a fixed factor in the model, the main effect of Stress Consistency of VC2 was significant in the analysis of the accuracy data ($z = -2.69$, $p = .01$) and marginal in the analysis of the latency data (t (1,286) = −1.81, $p = .07$). In the model with the token-based Stress Consistency of VC2, the results were identical (errors: $z = -3.09$, $p < .001$; latencies: t (1,286) = −1.72, $p = .09$). These results showed that the participants were more accurate and slightly faster in naming words with consistent ($M = 635$ ms ($SD = 40$), 4% incorrect responses) versus inconsistent ($M = 667$ ms ($SD = 63$), 11% incorrect responses) mappings of the orthography of VC2 onto stress patterns.

GENERAL DISCUSSION

The main purpose of the present research was to identify a set of cues that are in predictive relationships with (a) stress patterns that Russian disyllables have in the corpus and (b) stress patterns that native speakers of Russian assign to disyllabic words. The 11 potential stress cues that were examined in the regression (Studies 1 and 2) and factorial (Study 3) studies were Grammatical Category, Log Frequency, Word Length, Word Onset Complexity, Word Coda Complexity, and the following orthographic components: CVC1, CV1, VC1, CVC2, CV2, and VC2. In Study 1, a set of cues that are probabilistically associated with stress patterns in Russian were discovered. Then, in Studies 2 and 3, we assessed the strength of the relationship between these cues and the actual performance of speakers of Russian naming disyllabic words.

The results of Study 1 showed that in the corpus of Russian disyllabic words there were six variables that were in strong associative relationships with stress patterns: Word Onset Complexity, Word

TABLE 6
Characteristics of the words used as stimuli in Study 3

	Characteristics					
Condition	*Length*	*Log frequency*	*CVC1 consistency*	*CVC2 consistency*	*VC2 consistency*	*CV1 consistency*
CVC1 consistent	6.00	1.10	.94(.92)	.60(.59)	.59(.57)	.57(.62)
CVC1 inconsistent	5.65	1.23	.21(.17)	.67(.71)	.54(.54)	.63(.60)
CVC2 consistent	5.45	1.32	.69(.72)	.85(.86)	.56(.56)	.60(.58)
CVC2 inconsistent	5.15	1.21	.68(.69)	.22(.19)	.50(.51)	.54(.58)
VC2 consistent	5.60	1.08	.62(.62)	.60(.58)	.78(.76)	.65(.59)
VC2 inconsistent	5.50	1.29	.66(.67)	.54(.53)	.24(.25)	.60(.63)

Note: Consistency measures provided are based on both type and token (in brackets) counts.

Coda Complexity and the orthographic components CVC1, CVC2, CV2, and VC2. The following relationships between stress cues and stress patterns in Russian disyllables were identified.

First, disyllabic words with complex onsets are more likely to have a trochaic than an iambic stress pattern, whereas the presence of complex codas in words appears to be associated with an iambic rather than a trochaic stress pattern. Evidence that the complexity of words' onsets and codas is related to stress patterns in the same way as observed in the present data has been previously provided in English (Kelly, 2004; Kelly et al., 1998). Further, there is corroborating support for there being an association between word onsets and stress patterns in Russian (Ryan, 2014). Thus, the results from both Russian and English studies demonstrate that a syllable with a more complex structure is likely to be stressed in a disyllabic word. In linguistics, this effect of the complexity of onsets and codas on stress pattern assignment can be ascribed to the concept of syllabic weight (Gordon, 2006). The idea is simply that so called "heavy" syllables (i.e., syllables having more segments) tend to be more likely to be stressed than "light" syllables (i.e., syllables having fewer segments). Traditionally, the concept of syllabic weight referred only to the structure of the rhyme (i.e., nucleus vowel and coda), ignoring the onset (Hyman, 1985). More recently, there have been demonstrations that the onset's structure also contributes to the syllabic weight (Gordon, 2005; Ryan, 2014).

The other four variables that were predictive of stress patterns in the corpus were the orthographic components CVC1, CVC2, CV2, and VC2. For three of these components (CVC1, CVC2, and VC2), the nature of the neighbourhoods created by different versions of that component was significantly predictive of the stress of a word containing that specific component (e.g., in English, the existence of the unstressed—der in murder, fodder, order, border, etc. would predict that words ending in—der should be assigned trochaic stress). On the other hand, for the orthographic component CV2, there was an unexpected reversed relationship between the consistency measure and the stress pattern. More specifically, the analysis indicated the specific versions of that component that had neighbourhoods having mainly trochaic stress were predictive of an iambic stress assignment. This result is, of course, completely counterintuitive and virtually impossible to interpret from the perspective of any cognitive mechanisms, suggesting that it is likely an artefact of the multiple regression analysis. Therefore, although CV2 as a variable was a significant predictor of stress in the final equation, it seems unlikely that CV2 is a valid stress cue in Russian.

Out of the three orthographic components that were found to be significant predictors of stress in Russian, only the association between VC2 and stress patterns has been previously investigated. That is, VC2 is essentially the orthographic component previously referred to as word ending (defined as the letters corresponding to the vowel

TABLE 7
Mean RTs (in ms) and error rates of native speakers of Russian naming words with consistent vs. inconsistent spelling-to-stress mappings of CVC1, CVC2, and VC2 (Study 3)

Word type	*RTs (in ms)*	*Error rate (%)*
CVC1 consistent	636	2
CVC1 inconsistent	680	9
CVC2 consistent	633	4
CVC2 inconsistent	667	12
VC2 consistent	635	4
VC2 inconsistent	666	11

Note: Words were consistent/inconsistent with the actual stress patterns that these words have based both on type and token counts.

of the second syllable and all following consonants). There is consistent evidence that there are probabilistic relations between word endings and stress patterns in English (Arciuli & Cupples, 2006, 2007; Arciuli et al., 2010). The present research further documents the existence of a relationship between VC2 and stress patterns in Russian, a language with a stress system quite different from that in English. Interestingly, Russian is also different from English in that 31% of Russian words have VC2 components that contain only nucleus vowels but no codas, which seems to be a much larger proportion than in English. Thus, the orthography of the VC2 component might be thought of as being more generic in Russian compared to English. However, as the results of the present research and the research by Jouravlev and Lupker (2014) showed, VC2 is, in fact, specific enough in Russian to be a valid cue to stress in disyllabic words. Further, VC2 remains valid even when the impact of CVC2, a component that partially overlaps with VC2, is accounted for. Spelling-to-stress consistency measures of the first syllable (CVC1) and of the second syllable (CVC2) have not been previously studied as stress cues in any other languages. The fact that such associations were found in Russian might act as an incentive for further examinations of the relationship between CVC1/CVC2 and stress patterns in other languages.

Another orthographic component that has been previously shown to be probabilistically related to stress patterns in English is word beginnings (Arciuli et al., 2010). Word beginnings (defined as all letters up to and including the vowel of the first syllable) correspond to the CV1 component in our analysis. In none of the analyses reported in this paper, was CV1 a significant predictor of stress patterns in Russian (although note that there was a significant interaction of CV1 and Log Frequency in the analysis of response latencies in Study 2). The obvious question is why. One possibility is that, in contrast to English, Russian word beginnings actually are not effective stress cues. A second possibility is that they are not effective cues in either Russian or English and that their apparent impact in prior research was actually due to the impact of a confounding factor such as the orthography of the first syllable (i.e., CVC1). A third and more likely possibility is that associative relationships between word beginnings and stress patterns do exist not only in English but also in Russian, however, the regression analysis (Studies 1 and 2) did not identify them as such due to the fact that we included so many factors, including CVC1 consistency, among our predictors. For predictors that are correlated and provide partially overlapping information, a regression model will select the most strongly correlated cue first which can lead to it disregarding other less informative and partially overlapping cues. Indeed, as seen from the correlational analysis of spelling-to-stress consistencies of CV1 (word beginning) and CVC1 in the present corpus, these two variables are strongly related, r (13,943) = .62, $p < .001$, and, as a result, most of the variance in stress patterns that CV1 explains can also be explained by CVC1. Essentially, the analysis may have robbed CV1 of much of its variance.

In thinking about which of these three possibilities might be correct, it is potentially worth noting that if, as in the study done in English (Arciuli et al., 2010), our logistic regression model had contained CV1 but not CVC1 as a potential stress cue, CV1 would have been regarded as a significant predictor of stress patterns in Russian (type-based analysis: $z = -2.4$, $p = .02$; token-based analysis: $z = -2.6$, $p = .01$) although the model with CV1 but not CVC1 in it does fit the data significantly worse than the model that has both CV1 and CVC1 (type-based analysis: χ^2 (1) = 2,318.78, $p < .001$; token-based analysis: χ^2 (1) = 2,298.11, $p < .001$). Further, it is also worth noting, as just mentioned, that there was a significant interaction of CV1 and Log Frequency in one of the analysis of Study 2 which supports the idea that the orthography of word beginnings does matter in Russian at least when low frequency words are processed.

The same type of situation also exists between VC2 (word ending) and CVC2 (r (13,943) = .78, $p < .001$). Despite being correlated, however, each of these variables apparently accounts for a substantial amount of individual variance in stress patterns and, therefore, both VC2 and CVC2 remained in the regression equation as significant predictors of stress. The implication is that VC2 (word endings) are likely stronger cues to stress than CV1 (word beginnings). A similar argument concerning the special status of word endings (VC2) compared to word beginnings (CV1) has been previously made by Arciuli et al. (2010).

Another variable that has been previously suggested to act as a cue to lexical stress is grammatical category (Arciuli & Cupples, 2004; 2006; Jouravlev & Lupker, 2014; Kelly & Bock, 1988). In Russian, most disyllabic adjectives have trochaic stress (81% of the adjectives in the corpus of Study 1). Thus, the information about the distribution of stress patterns across words of this grammatical category, indeed, might assist readers in their stress-assignment decisions, and that is exactly what Jouravlev and Lupker (2014) demonstrated in their prior research. In particular, they found that trochaically stressed adjectives were named faster and more accurately than iambically stressed adjectives. Note, however, that this evidence was obtained in one of the planned comparisons where adjectives were analysed separately as a group. In the main analysis of Jouravlev and Lupker (2014), there was no evidence for a significant interaction of stress type and grammatical category. Similarly, in the present research, we did find that adjectives were strongly correlated with stress patterns (r (13,943) = −.22, $p < .001$), however, in the binary logistic regression of Study 1 that involved all the predictor variables, grammatical category was not identified as a significant predictor of stress in Russian disyllables. The most likely explanation of this result is that, for all other grammatical categories in Russian, there is no stress pattern that occurs more frequently than the other pattern (see Table 1 for the distribution of trochaic vs. iambic stress patterns for all grammatical categories) and because adjectives comprise only about 15% of the corpus of disyllables, information about grammatical category would be helpful in assigning stress for only a small subset of words in the corpus. Therefore, while grammatical category may be a cue to stress assignment in Russian, evidence supporting that claim would be expected to be rather weak.

To summarise, in Study 1, we singled out a set of five variables that are probabilistically associated with stress patterns in the corpus of Russian disyllabic words: Onset Complexity, Coda Complexity, and the orthographic components CVC1, CVC2, and VC2. The goal of Study 2 was to assess whether all of these stress cues or, alternatively, only a particular subset of these cues are used by readers in assigning stress. This goal was accomplished by examining stress-assignment performance and naming latencies of speakers of Russian. Those individuals were found to make use of the existence of probabilistic relationships between stress patterns and orthographies for only three out of the five components: CVC1, CVC2, and VC2. If the stress pattern cued by the CVC1, CVC2, and/or VC2 in a word was the one that a word had, participants were faster to name that word and less likely to assign it the incorrect stress pattern. On the other hand, if the CVC1, CVC2, and/or VC2 of a word cued an incorrect stress pattern, readers were often misled, causing them to slow down and to make stress-assignment errors.

These findings of associative relationships between the orthography of the three components and the ease of stress pattern assignment were further corroborated, using a factorial design, in Study 3 where participants named Russian disyllabic words with different degrees of spelling-to-stress consistency of CVC1, CVC2, or VC2. The results of Study 3 showed that participants were more accurate in assigning stress patterns to words with CVC1s, CVC2s, or VC2s that had consistent versus inconsistent spelling-to-stress mappings. The reliance of readers on the orthography of VC2 (word endings) in making decisions about the stress pattern to be applied to a polysyllabic word has been previously demonstrated in English (Arciuli & Cupples, 2006, 2007; Arciuli et al., 2010) and in Russian (Jouravlev & Lupker, 2014). Therefore, it is likely that the word ending is indeed a strong cue to stress that is of relevance in a number of languages. The present study is the first one to demonstrate that the orthography of syllables (CVC1 and CVC2) also appears to be used by readers of Russian in the process of stress assignment.

The results of Study 2 also indicated that readers of Russian do not make use of all potential stress cues that are present in the Russian language. Although word onset complexity and word coda complexity were significant predictors of stress patterns in the corpus of words used as stimuli in Study 2, the participants' responses seemed to be driven mainly by the other three potentially useful cues to stress (i.e., CVC1, CVC2, and VC2). The absence of effects of word onset complexity and word coda complexity on stress assignment in Russian stands in contrast to the results reported in English (Kelly, 2004; Kelly & Bock, 1988). There are, again, three potential reasons for this difference. These factors could matter in English but not in Russian. They may not matter in either language and evidence of their impact in prior research was due to a confound with other factors. Finally, they may matter in both languages, but their impact was obscured in the present research by the large number of predictors of stress assignment being analysed. Consistent with the first of these possibilities, it does appear that the impact of CVC1, CVC2, and VC2 consistencies is quite strong and, potentially, sufficient for proper stress assignment in Russian. Hence, using the complexity of word onsets or word codas as predictors of stress patterns may not be necessary for readers/speakers of that language.

Overall, Studies 2 and 3 demonstrated that stress-assignment errors occur quite often in Russian (8%

of words in Study 2 and 7% of words in Study 3 were mis-stressed). For some of these words, stress-assignment errors were committed by more than half of our participants, suggesting that there may be some non-lexical information in these commonly mis-stressed words that cues incorrect stress patterns and leads the readers astray. To assess this possibility, we examined the descriptive characteristics of a subset of words from Studies 2 and 3, for which at least 50% of responses were incorrect. Twenty-seven such words were identified (out of 600 words used in Studies 2 and 3). While this particular subset of words was comparable to that of the whole set from which it has been drawn on such characteristics as length (M (subset) = 5.55 vs. M (whole set) = 5.56), onset complexity (M (subset) = 1.24 vs. M (whole set) = 1.40), coda complexity (M (subset) = .49 vs. M (whole set) = .60), and spelling-to-stress consistency of CV1 (M (subset) = .47 vs. M (whole set) = .56), VC1 (M (subset) = .49 vs. M (whole set) = .53), and CV2 (M (subset) = .52 vs. M (whole set) = .56); the subset of commonly mis-stressed words differed from the whole set in log frequency (M (subset) = .48 vs. M (whole set) = 1.21) and spelling-to-stress consistencies of CVC1 (M (subset) = .46 vs. M (whole set) = .67), CVC2 (M (subset) = .49 vs. M (whole set) = .73), and VC2 (M (subset) = .50 vs. M (whole set) = .65). Thus, most commonly mis-stressed words are words of low frequency that have spellings of CVC1, CVC2, and VC2 cuing incorrect stress patterns. This observation further corroborates our conclusion that CVC1, CVC2, and VC2 components play an active role as cues to stress in Russian disyllables.

While the role of lexical stress during word recognition in Russian has not been examined extensively yet, there is an ample empirical evidence from other Indo-European languages (Cooper, Cutler, & Wales, 2002; Soto-Faraco, Sebastián-Gallés, & Cutler, 2001; Sulpizio & McQueen, 2012), suggesting that lexical stress does play an active part in the process of lexical disambiguation. The ability of lexical stress to constrain lexical activation is likely to be a universal phenomenon and, thus, it likely also occurs for Russian speakers. An interesting question for future research is, then, how speakers of this language manage to communicate effectively considering their relatively high rate of stress-assignment errors. The point to realise here is that human speech is inevitably noisy, meaning that various errors, including errors of stress assignment, are often present. The process of language acquisition entails the mastering of certain noise handling abilities that allow speakers and listeners to overcome various types of speech errors efficiently. What the mechanisms for handling stress-assignment errors are would appear to be a fruitful area for future research.

To conclude, with respect to Russian we appear to have identified three non-lexical sources of evidence for stress assignment that have high validity (i.e., strong probabilistic associations between the cues and stress patterns exist in the language) and high utility (i.e., readers appear to possess sufficient knowledge of these probabilistic associations between cues and stress patterns in order for this knowledge to impact their behaviour). Therefore, the orthography of the first syllable of a word (CVC1), of the second syllable of a word (CVC2), and of the ending of the second syllable of a word (VC2) appear to be important sources of evidence for determining stress assignment when naming Russian disyllabic words. Although at the moment we can make conclusions about the effect of these variables on stress assignment in Russian disyllables only and do not assume that the same sources of evidence for stress necessarily impact performance in any other languages, in the future, such a possibility needs to be investigated empirically. Overall, the findings of the present research should be considered by modellers of the process of lexical stress assignment in general and of lexical stress assignment in Russian in particular.

Our final point to note about these results is that they provide strong evidence against the idea that stress assignment in Russian is accomplished only by retrieving stress information from a word's lexical representation. If the lexical retrieval hypothesis were in fact true, none of the non-lexical variables investigated here would have been a significant predictor of stress patterns or, more importantly, of stress pattern assignment in the regression studies and there should not have been a stress consistency effect in the factorial study. In contrast, the present studies demonstrated that there are probabilistic, associative connections between non-lexical cues and stress patterns in Russian and that native speakers of Russian do utilise at least some of this non-lexical information about stress in naming disyllabic Russian words. This conclusion does not, of course, indicate that Russian readers/speakers do not also rely on the specific retrieval of lexically stored stress information in the process of stress assignment (e.g., Jouravlev & Lupker, in press). In fact, such reliance may even be greater in Russian than in some other languages in which word stress is more predictable. The point is merely that assigning stress in Russian is a process that involves considerably more than retrieving stress information from lexical memory.

REFERENCES

Arciuli, J., & Cupples, L. (2004). The effects of stress typicality during spoken word recognition by native and non-native speakers: Evidence from onset-gating. *Memory & Cognition*, *32*, 21–30.

Arciuli, J., & Cupples, L. (2006). The processing of lexical stress during visual word recognition: Typicality effects and orthographic correlates. *The Quarterly Journal of Experimental Psychology*, *59*, 920–948.

Arciuli, J., & Cupples, L. (2007). Would you rather "embert a cudsert" or "cudsert an embert"? How spelling patterns at the beginning of English disyllables can cue grammatical category. In A. C. Schalley & D. Khlentzos (Eds.), *Language and cognitive structure* (pp. 213–237). Amsterdam: John Benjamins.

Arciuli, J., Monaghan, P., & Seva, N. (2010). Learning to assign lexical stress during reading aloud: Corpus, behavioural and computational investigations. *Journal of Memory and Language*, *63*, 180–196.

Avanesov, R. I. (1956). *The phonology of the standard Russian language*. Moscow: Moscow State University.

Baayen, R. H., Pipenbrock, R., & Gulikers, L. (1993). *The CELEX lexical database (CD-ROM). Linguistic data consortium*. Philadelphia, PA: University of Pennsylvania.

Bates, D., & Maechler, M. (2010). *Package 'lme4'*. Reference manual for the package. Retrieved from: http://cran.r-project.org/web/packages/lme4/lme4.pdf

Bijeljac-Babic, R., Millogo, V., Farioli, F., & Grainger, J. (2004). A developmental investigation of word length effects in reading using a new on-line word identification paradigm. *Reading and Writing*, *17*, 411–431.

Brysbaert, M., Buchmeier, M., Conrad, M., Jacobs, A. M., Bölte, J., & Böhl, A. (2011). The word frequency effect: A review of recent developments and implications for the choice of frequency estimates in German. *Experimental Psychology*, *58*, 412–424.

Burani, C., & Arduino, L. S. (2004). Stress regularity or consistency? Reading aloud Italian polysyllables with different stress patterns. *Brain and Language*, *90*, 318–325.

Carreiras, M., & Perea, M. (2002). Masked priming effects with syllabic neighbors in the lexical decision task. *Journal of Experimental Psychology: Human Perception and Performance*, *28*, 1228–1242.

Chateau, D., & Jared, D. (2003). Spelling–sound consistency effects in disyllabic word naming. *Journal of Memory and Language*, *48*, 255–280.

Colombo, L. (1992). Lexical stress and its interaction with frequency in word pronunciation. *Journal of Experimental Psychology: Human Perception and Performance*, *18*, 987–1003.

Conrad, M., & Jacobs, A. M. (2004). Replicating syllable frequency effects in Spanish in German: One more challenge to computational models of visual word recognition. *Language and Cognitive Processes*, *19*, 369–390.

Cooper, N., Cutler, A., & Wales, R. (2002). Constraints of lexical stress on lexical access in English: Evidence from native and non-native listeners. *Language and Speech*, *45*, 207–228.

Forster, K. I., & Forster, J. C. (2003). DMDX: A windows display program with millisecond accuracy. *Behavior Research Methods, Instruments, & Computers*, *35*, 116–124.

Friedman, L., & Wall, M. (2005). Graphical views of suppression and multicollinearity in multiple linear regression. *The American Statistician*, *59*, 127–136.

Giegerich, H. J. (1992). *English phonology: An introduction*. Cambridge: Cambridge University Press.

Gordon, M. (2005). A perceptually-driven account of onset-sensitive stress. *Natural Language and Linguistic Theory*, *23*, 595–653.

Gordon, M. (2006). *Syllable weight: Phonetics, phonology, typology*. New York: Routledge.

Gutierrez-Palma, N., & Palma-Reyes, A. (2008). On the use of lexical stress in reading Spanish. *Reading and Writing*, *21*, 645–660.

Halle, M. (1997). On stress and accent in Indo-European. *Language*, *73*(2), 275–313.

Hyman, L. (1985). *A theory of phonological weight*. Dordrecht: Foris.

Jouravlev, O., & Lupker, S. J. (2014). Stress consistency and stress regularity effects in Russian. *Language, Cognition and Neuroscience*, *29*, 605–619.

Jouravlev, O., & Lupker, S. J. (in press). Lexical stress assignment as a problem of probabilistic inference making. *Psychonomic Bulletin and Review*. doi:10.3758/s13423-015-0802-y

Kelly, M. (2004). Word onset patterns and lexical stress in English. *Journal of Memory and Language*, *50*, 231–244.

Kelly, M., & Bock, J. (1988). Stress in time. *Journal of Experimental Psychology: Human Perception and Performance*, *14*, 389–403.

Kelly, M., Morris, J., & Verrekia, L. (1998). Orthographic cues to lexical stress: Effects on naming and lexical decision. *Memory & Cognition*, *26*, 822–832.

Lagerberg, R. (1999). *Stress and suffixation in modern Russian: The development of uniform syllable stress*. Nottingham: Astra Press.

Maassen, G. H., & Bakker, A. B. (2001). Suppressor variables in path models: Definitions and interpretations. *Sociological Methods and Research*, *30*, 241–270.

Melvold, J. L. (1989). *Structure and stress in the phonology of Russian*. ProQuest, UMI Dissertations Publishing.

Molczanow, J., Domahs, U., Knaus, J., & Wiese, R. (2013). The lexical representation of word stress in Russian: Evidence from event-related potentials. *The Mental Lexicon*, *8*, 164–194.

Perry, C., Ziegler, J. C., & Zorzi, M. (2010). Beyond single syllables: Large-scale modeling of reading aloud with the connectionist dual process (CDP++) model. *Cognitive Psychology*, *61*, 106–151.

Protopapas, A. (2006). On the usefulness of stress diacritics in reading Greek. *Reading and Writing*, *19*, 171–198.

Protopapas, A. (2007). CheckVocal: A program to facilitate checking the accuracy and response time of vocal responses from DMDX. *Behavior Research Methods*, *39*, 859–862.

Protopapas, A., Gerakaki, S., & Alexandri, S. (2006). Lexical and default stress assignment in reading Greek. *Journal of Research in Reading*, *29*, 418–432.

Protopapas, A., Gerakaki, S., & Alexandri, S. (2007). Sources of information for stress assignment in reading Greek. *Applied Psycholinguistics*, *28*, 695–720.

Rastle, K., & Coltheart, M. (2000). Lexical and nonlexical print-to-sound translation of disyllabic words and nonwords. *Journal of Memory and Language*, *42*, 342–364.

Rastle, K., & Davis, M. H. (2002). On the complexities of measuring naming. *Journal of Experimental Psychology: Human Perception and Performance*, *28* (2), 307–314.

Ryan, K. M. (2014). Onsets contribute to syllable weight: Statistical evidence from stress and meter. *Language*, *90*, 309–341.

Scherba, L. V. (1957). *Select papers on the Russian language.* Moscow: Ministry of Education Publishing House.

Seva, N., Monaghan, P., & Arciuli, J. (2009). Stressing what is important: Orthographic cues and lexical stress assignment. *Journal of Neurolinguistics*, *22*, 237–249.

Soto-Faraco, S., Sebastián-Gallés, N., & Cutler, A. (2001). Segmental and suprasegmental mismatch in lexical access. *Journal of Memory and Language*, *45*, 412–432.

Sulpizio, S., & McQueen, J. M. (2012). Italians use abstract knowledge about lexical stress during spoken-word recognition. *Journal of Memory and Language*, *66*, 177–193.

Tabachnick, B. G., & Fidell, L. S. (2007). *Using multivariate statistics.* Boston, MA: Pearson/Allyn & Bacon.

Treiman, R., & Kessler, B. (1995). In defense of an onset-rime syllable structure for English. *Language and Speech*, *38*, 127–142.

Treiman, R., Mullenix, J., Bijeljac-Babic, R., & Richmond-Welty, E. D. (1995). The special role of rimes in the description, use, and acquisition of English orthography. *Journal of Experimental Psychology: General*, *124*, 107–136.

Wurm, L. H., & Fisicaro, S. A. (2014). What residualizing predictors in regression analyses does (and what it does not do). *Journal of Memory and Language*, *72*, 37–48.

Zaliznjak, A. (1985). *From pro-slavonic to Russian accentuation* [Ot proslavyanskoj akzentuazii k russkoj]. Moscow: Prosvesheniye.

Zarva, M. (2001). *The dictionary of Russian lexical stress.* Moscow: ENAS.

Zsiga, E. (2013). *The sounds of language: An introduction to phonetics and phonology.* New York, NY: Wiley Blackwell.

APPENDIX A: RUSSIAN DISYLLABIC WORDS USED AS STIMULI IN STUDY 2

краёв [krajov], сдают [sdajut], забав [zabav], кабин [kabin], чабан [chaban], агат [agat], зажат [zazhat], сажал [sazhal], найдёт [najd'ot], талон [talon], канав [kanav], канат [kanat], царит [tzarit], засад [zasad], красив [krasiv], распад

[raspad], фасон [fason], застал [zastal], кастет [kast'et], фашизм [fashism], плевал [pl'eval], беглец [b'egl'etz], ведёт [v'ed'ot], шедевр [sh'ed'evr], телец [t'el'etz], белья [b'el'ja], беречь [b'er'ech'], берёшь [b'er'osh'], вершат [v'ershat], сверлить [sv'erlit'], песков [p'eskov], цветки [tzv'etki], жрецов [zhr'etzov], печаль [p'echal'], трещал [tr'eshal], клиент [klijent], сиять [sijat'], прибрать [pribrat'], кидал [kidal], лизнуть [liznut'], длину [dlinu], иным [inim], пинать [pinat'], чинил [chinil], пират [pirat], стирал [stiral], цитат [tzitat], лицей [litz'ej], причуд [prichud], пробник [probnik], содрать [sodrat'], вождю [vozhd'u], дожде [dozhd'e], колёс [kol'os], пролив [proliv], стволы [stvoli], толчков [tolchkov], болтлив [boltliv], блондин [blondin], ворчун [vorchun], горят [gor'at], дворце [dvortz'e], морщин [morshin], горбат [gorbat], корсет [kors'et], бросал [brosal], проспект [prosp'ekt], гостил [gostil], костров [kostrov], мотал [motal], отёк [ot'ok], отчёт [otch'ot], брошюр [brosh'ur], вуаль [vual'], дубрав [dubrav], убрал [ubral], угар [ugar], дружин [druzhin], уныл [unil], купать [kupat'], журчат [zhurchat], труслив [trusliv], грустить [grustit'], состав [sostav], крылом [krilom], рычаг [richag], рычат [richat], смычок [smichok], этап [etap], юнцов [juntzov], плясал [pl'asal], стряхнуть [str'ahnut'], мячом [m'achom], терплю [t'erpl'u], зверьё [zv'er'jo], кривых [krivih], родить [rodit'], контракт [kontrakt], попеть [pop'et'], трудов [trudov], дружить [druzhit'], разгул [razgul], скала [skala], запор [zapor], застыл [zastil], лежит [l'ezhit], терять [t'er'at'], вестей [v'est'ej], смешна [sm'eshna], листу [listu], возврат [vozvrat], колец [kol'etz], молил [molil], пролезть [prol'ezt'], жонглёр [zhongl'or], отцы [ottzi], сумел [sum'el], мычат [michat], ключа [klyucha], газет [gaz'et], рванул [rvanul], спастись [spastis'], дебош [d'ebosh], слезай [sl'ezaj], неправ [n'eprav], дефект [d'ef'ect], спешил [sp'eshil], визжал [vizzhal], огнём [ogn'om], смолчать [smolchat], котлы [kotli], жуют [zhujut], губам [gubam], кури [kuri], глухой [gluhoj], брала [brala], антенн [ant'enn], редис [r'edis], седле [s'edl'e], сбежал [sb'ezhal], месил [m'esil], взирать [vzirat'], лице [litz'e], войной [vojnoj], гонец [gon'etz], вопил [vopil], копну [kopnu], борца [bortza], глубин [glubin], шумок [shumok], суров [surov], юлит [julit], равнин [ravnin], вредить [vr'edit'], брезент [br'ez'ent], ремнём [r'emn'om], слепых [sl'epih], весам [v'esam], стола [stola], катков [katkov], еде [jed'e], подвод [podvod], холоп [holop], багор [bagor], кора [kora], красу [krasu], един [jedin], плести [pl'esti], жилом [zhilom], зачат [zachat], страшны [strashni], равны [ravni], полос [polos], ору [oru], крутом [krutom], зерне [z'ern'e], стопам [stopam], волхвы [volhvi], плоту [plotu], звонить [zvonit'], ясна [jasna], станка [stanka], сонет [son'et], искра [iskra], вирши [virshi], мордвы [mordvi], мысок [misok], мани [mani], порты [porti], велось [v'elos'], лифтов [liftov], банщик [banshik], отрок [otrok], строёв [strojov], слога [sloga], балах [balah], сеном [s'enom], бород [borod], угли [ugli], складу [skladu], пахал [pahal], копны [kopni], ирод [irod], гонят [gon'at], схватит [shvatit], пристав [pristav], росчерк [roscherk], ведал [v'edal], гибок [gibok], стилю [stil'u], лопал [lopal], рощу [roshu], грызло [grizhlo], клеток [kl'etok], платном [platnom], статуй [statuj], должным [dolzhnim], пуху [puhu], провод [provod], саржа [sarzha], скверны [skv'erni], пошлой [poshloj], пляски [pl'aski], валят [val'at], самых [samih], дарит [darit],

воин [voin], лобном [lobnom], дому [domu], тощей [toshej], сума [suma], пятом [p'atom], мазал [mazal], фраза [fraza], парты [parti], хате [hat'e], влезли [vl'ezli], веком [v'ekom], сельской [s'el'skoj], сценой [stz'enoj], пшённый [psh'onnij], мехом [m'ehom], чехов [ch'ehov], близок [blizok], взводе [vzvod'e], рода [roda], тройка [trojka], скромной [skromnoj], носишь [nosish'], доску [dosku], отрасль [otrasl'], груду [grudu], скуле [skul'e], крутят [krut'at], рухлядь [ruhl'ad'], мыса [misa], слышал [slishal], бабьей [bab'ej], впала [vpala], массе [mass'e], страстных [strastnih], беден [b'ed'en], средне [sr'edn'e], сделал [sd'elal], вечным [v'echnim], мило [milo], вводит [vvodit], рока [roka], томы [tomi], тонут [tonut], щуки [shuki], трупу [trupu], шкурки [shkurki], взятки [vz'atki], жаден [zhad'en], драли [drali], пара [para], бегло [b'eglo], верных [v'ernih], игры [igri], хилый [hilij], листья [list'ja], стоек [stojek], мощным [moshnim], мужем [muzh'em], узких [uzkih], стычка [stichka], тяжба [t'azhba], связях [sv'az'ah], лаять [lajat'], грабли [grabli], главном [glavnom], давший [davshij], равен [rav'en], травля [travl'a], ладно [ladno], стадий [stadij], свадьбе [svad'b'e], влажной [vlazhnoj], жаждет [zhazhd'et], зайцев [zajtz'ev], гайки [gajki], плакал [plakal], жалки [zhalki], залпы [zalpi], мальчик [mal'chik], пальма [pal'ma], грамот [gramot], мамин [mamin], рамка [ramka], шрамом [shramom], бланки [blanki], странным [strannim], бантик [bantik], шапках [shapkah], жарят [zhar'at], баржи [barzhi], пасхе [pash'e], каски [kaski], часто [chasto], краток [kratok], кратком [kratkom], трата [trata], хваткой [hvatkoj], дачей [dach'ej], плача [placha], башня [bashn'a], слаще [slash'e], неба [n'eba], левом [l'evom], брёвна [br'ovna], дёготь [d'ogot'], бедность [b'ednost'], мёда [m'oda], режу [r'ezhu], свежых [sv'ezhih], лезло [l'ezlo], грёзы [gr'ozi], ездил [jezdil], ездят [jezd'at], рейсы [r'ejsi], смелых [sm'elih], целость [tz'elost'], жёлтых [zh'oltih], бремя [br'em'a], бренность [br'ennost'], женских [zh'enskih], ленты [l'enti], цепью [tz'ep'ju], дерзок [d'erzok], мера [m'era], сера [s'era], сербы [s'erbi], держим [d'erzhim], чёртов [ch'ortov], кресле [kr'esl'e], песен [p'es'en], всплески [vspl'eski], честен [ch'est'en], светский [sv'etskij], шефом [sh'efom], видном [vidnom], книжках [knizhkah], снизит [snizit], вскрикнул [vskriknul], фильмом [fil'mom], зримый [zrimij], льдину [l'dinu], принцы [printzi], миски [miski], спискам [spiskam], мистер [mist'er], битых [bitih], нитке [nitk'e], плитки [plitki], ритмом [ritmom], биться [bit's'a], птички [ptichki], вишни [vishni], пробке [probk'e], коврик [kovrik], повод [povod], бродим [brodim], водном [vodnom], входом [vhodom], лодку [lodku], сода [soda], сходят [shod'at], дрожжи [drozhzhi], кожей [kozh'ej], позы [pozi], стройность [strojnost'], сроки [sroki], молод [molod], громом [gromom], ломом [lomom], тонна [tonna], тонок [tonok], топот [topot], створки [stvorki], спорта [sporta], сносно [snosno], воском [voskom], просьбе [pros'b'e], кротость [krotost'], сотни [sotni], смотришь [smotrish'], вздохов [vzdohov], почка [pochka], будни [budni], кружат [kruzhat], грузный [gruzhnij], булка [bulka], ткнула [tknula], дума [duma], пунктам [punktam], хмурым [hmurim], куртка [kurtka], вкусно [vkusno], мутной [mutnoj], скучном [skuchnom], штучки [shtuchki], пушке [pushk'e], взрыву [vzrivu], ссылка [ssilka], выплат [viplat], сытный [sitnij], вышка [vishka], крышам [krisham], пышно [pishno], клюнуть [kl'unut'], нюхал [n'uhal], вязкий [v'azkij], сняли [sn'ali], пряный [pr'annij], мяса [m'asa], зятя [z'at'a], кляча [kl'acha], прячешь [pr'ach'esh'], печать [p'echat'], сводный [svodnij], людный [l'udnij], стержнем [st'erzhn'em], ляжешь [l'azh'esh'], пачкой [pachkoj], бункер [bunk'er], риска [riska], встречных [vstr'echnih], чашек [chash'ek], пачек [pach'ek], внучек [vnuch'ek], квасом [kvasom], ставишь [stavish'], съезда [s'jezda], пешку [p'eshku], школой [shkoloj], скачет [skach'et], ножны [nozhni], валим [valim], банда [banda], прессы [pr'essi], ждало [zhdalo], вольной [vol'noj], эра [era], лыжи [lizhi], речке [r'echk'e], чая [chaja], ахнул [ahnul], мылом [milom], кукиш [kukish], кремом [kr'emom], скажут [skazhut], храмом [hramom], сбили [sbili], дверцы [dv'ertzi], пуска [puska], мозга [mozga], кашлял [kashl'al], гадов [gadov], перьях [p'er'jah]

APPENDIX B: RUSSIAN DISYLLABIC WORDS USED AS STIMULI IN STUDY 3

CVC1: ценность [tzennost'], фольга [fol'ga], дамских [damskih], жёлтым [zh'oltim], точка [tochka], пивко [pivko], струну [strunu], ведьму [v'ed'mu], толчки [tolchki], альбом [al'bom], монтаж [montazh], бруски [bruski], звёздах [zv'ozdah], устным [ustnim], робким [robkim], помни [pomni], торчат [torchat], ползти [polzti], ярок [jarok], станков [stankov], звери [zv'eri], скрипок [skripok], гнили [gnili], подлой [podloj], складной [skladnoj], пыльца [pil'tza], принял [prin'al], минут [minut], тоннель [tonn'el'], жестов [zh'estov], явить [javit'], акций [aktzij], пошлин [poshlin], девчат [d'evchat], тиски [tiski], высок [visok], резком [r'ezkom], ядром [jadrom], цапля [tzapl'a], умны [umni]

CVC2: фирмы [firmi], лужок [luzhok], нищей [nish'ej], койках [kojkah], путей [put'ej], глотать [glotat'], дрянью [dr'an'ju], дырки [dirki], ручка [ruchka], шипел [shipel], спички [spichki], чулан [chulan], валун [valun], жилец [zhil'etz], чужом [chuzhom], доллар [dollar], пищей [pish'ej], истцом [isttzom], выпив [vipiv], грибов [gribov], спискам [spiskam], дружке [druzhk'e], жука [zhuka], выжил [vizhil], грубой [gruboj], былом [bilom], кишка [kishka], детях [d'et'ah], вылет [vil'et], выпей [vip'ej], грозе [groz'e], щекам [sh'ekam], тётей [t'ot'ej], рукой [rukoj], имам [imam], рыбин [ribin], лотки [lotki], чулки [chulki], чистил [chistil], каплю [kapl'u]

VC2: стула [stula], грызла [grizla], лирик [lirik], праздник [prazdnik], кисло [kislo], номер [nom'er], верен [v'er'en], обществ [obsh'estv], устных [ustnih], дую [duju], сетям [s'et'am], месить [m'esit'], сбежать [sb'ezhat'], навар [navar], сустав [sustav], борец [bor'etz], копал [kopal], полям [pol'am], гробниц [grobnitz], пружин [pruzhin], лоцман [lotzman], опыт [opit], звонишь [zvonish'], таять [tajat'], нажим [nazhim], ночник [nochnik], письмо [pis'mo], седло [s'edlo], гербом [g'erbom], денщик [d'enshik], служат [sluzhat], пьяниц [p'janitz], ландыш [landish], красе [kras'e], слепым [sl'epim], копна [kopna], весел [v'es'el], ахал [ahal], кресел [kr'es'el], топал [topal]