Russian Sentence Corpus

Benchmark measures of eye movements in reading in Russian

Anna K. Laurinavichyute*[1,2], Irina A. Sekerina[3], Svetlana Alexeeva[4,1], Kristine Bagdasaryan[1] and Reinhold Kliegl[2]

[1]National Research University Higher School of Economics, Russian Federation

[2] University of Potsdam

[3] College of Staten Island and The Graduate Center of The City University of New York

[4] St.-Petersburg State University

* 105066 Russian Federation, Moscow, Staraya Basmannaya, house 24/1 c.1 National Research University Higher School of Economics; +79197253152; alaurinavichute@hse.ru

Abstract

The article introduces the new corpus of eye-movements in silent reading – the Russian Sentence Corpus (RSC). Russian uses Cyrillic script that has not yet been investigated in cross-linguistic eye-movement research. As in every language studied so far, we have confirmed the expected effects of low-level parameters, such as word length, frequency, and predictability, on the eye-movements of skilled Russian readers. These findings allow us to add Slavic languages using Cyrillic script (exemplified by Russian) to the growing number of languages with different orthographies ranging from the Roman-based European languages to logographic Asian ones whose basic eye-movement benchmarks conform to the universal comparative science of reading (Share 2008). We additionally report basic descriptive corpus statistics, and three exploratory investigations of the effects of Russian morphology on the basic eye-movement measures that illustrate the kinds of questions researchers can answer using the RSC. The annotated corpus is freely available from the project page at Open Science Framework: https://osf.io/x5q2r/.

Keywords: reading, eye movements, Russian, ambiguity, part of speech, corpus

1. Introduction

Eye movements in reading have been a research topic since Huey (1908) while psycholinguistic research started in the 1970s. Since then measures of eye movements have been the most widely used behavioral data in empirical linguistic research, ranging from testing cognitive models of eye movement control in reading (Rayner 2009) to core questions in the psycholinguistic theory such as timing of processing difficulties in complex sentences, and interaction between attention and the eye movements in language production and comprehension (Rayner 1998; Rayner et al. 2006). Eye movements have been recorded during reading of single words, sentences, paragraphs and whole texts in languages with different orthographies. Their analysis allows us to establish the fundamental characteristics of eye movements within and across languages that are referred to as *eye-movement benchmarks*. The reading materials together with the eye-movement benchmarks collected from individuals reading these materials constitute corpora of eye movements that started to appear in the past 20 years.

Eye-movement corpora are an indispensable tool for basic research in cognitive psychology and psycholinguistics. First, they serve as a source of data for establishing the basic benchmarks of eye movements while reading in languages with typologically diverse orthographies and grammars and constitute an important testing ground for models of eye movements in reading, e.g., the *E-Z Reader* model (Reichle et al. 1998) and the *SWIFT* model (Engbert et al. 2002). Second, eye movements reflect typical linguistic behavior, i.e., silent reading process, and serve as a natural material to evaluate theories of language processing in psycholinguistics. For example, Gibson's (2000) Dependency Locality theory was tested on eye-movement data in English (Demberg & Keller 2008) and Hindi (Husain et al. 2015), the Entropy Rate principle (Genzel & Charniak 2002) was tested on the English corpus (Keller 2004), and the Surprisal account (Hale 2001) was confirmed for the Potsdam Sentence Corpus (Boston et al. 2008). Finally, eye-movement-while-reading corpora are the necessary control data to study acquisition of literacy in unskilled (Ashby et al. 2005) and bilingual adults (Cop et al. 2016), developmental and acquired reading difficulties in children with and without learning disabilities (Tiffin-Richards & Schroeder 2015) and adults with cognitive

impairments, such as aphasia (Ablinger et al. 2014) and Alzheimer's disease (Crawford et al. 2015).

The basic benchmarks of eye-movement control in reading include measures related to the fixation probabilities and fixation durations. They were first established for reading in English, a language with the Roman-based alphabetic script and deep orthography, in the 1908 (Huey 1908, see also Tinker 1958). Follow-up studies revealed that these benchmarks vary depending on the lexical characteristics of words, i.e., their frequency, length, and predictability from context. They also determine the probability of a word being fixated or skipped, the expected number of fixations on it, and the probability of regression to it later. In recent years, other factors, e.g., word familiarity, age-of-acquisition, polysemy, plausibility, have been added to the inventory of characteristics that influence eye movements in reading.

In the 1990s, as psycholinguistics in general started to rapidly expand from English into other languages, it became clear that the focus on the English language in reading research was slowing down the development and empirical testing of “a universal science of reading” (Share 2008: 584). Eye movements in reading in other Roman script-based languages, namely, French, German, Dutch, and Finnish, that have more transparent orthography, but more complex morphology, often differ from English. Thus, it was found that eye-movement benchmarks were affected by parafoveal word familiarity in French (Kennedy & Pynthe 2005), word position in the sentence in Dutch and Spanish (Kuperman, Bertram, & Baayen 2010a; Fernández et al. 2014), and complex derivational and inflectional morphology in Finnish (Hyönä et al. 1995).

Recently, there has been a virtual explosion of comparative cross-linguistic research on reading in the typologically diverse languages with non-Roman orthographies, such as Chinese (Bai et al. 2008; 2009, Tsai et al. 2012; Yan et al. 2006), Japanese (Sainio et al. 2007), Korean (Kim et al. 2012), Hebrew (Pollatsek et al. 1981), Thai (Winskel et al. 2009), Hindi (Husain et al. 2015), Arabic (Paterson et al. 2015), Urdu (Paterson et al. 2014), and Uighur (Yan et al. 2014). Their visual, orthographic, lexical, and sentence-level characteristics required modification of existing models of reading and psycholinguistic theories. For example, it was found that in non-spaced logographic scripts, such as traditional Chinese, the average saccade length is much shorter (2-3

character spaces) than in the spaced scripts (8). However, in unspaced scripts, readers are able to direct their eyes towards the preferred viewing location (close to the middle of the word), just as in spaced scripts where the between-word spaces can be used to estimate word length (Yan et al. 2010; for similar results in Thai, see Winskel et al. 2009).

Nonetheless, even if we take all the studied European Roman script-based, Arabic, and Asian logographic languages together, their number remains very small compared to the world's 80 writing systems. What is inconspicuously absent in the abovementioned research is languages that use the Cyrillic orthography, namely, five major Slavic languages (Russian, Ukrainian, Belarusian, Serbian, and Bulgarian) and more than 100 languages from other language families whose newly established writing systems were based on Cyrillic alphabet, i.e., indigenous languages of the former Soviet Union (Lewis 1972). The languages that use Cyrillic script are typologically very diverse: they belong to such language families as Slavic, Turkic (Tatar, Kyrgyz), Caucasian (Abkhaz, Adyghe), Mongolic (Mongolian), etc. Their omission in cross-linguistic research on eye movements in reading is a sizable lacuna in the comparative science of reading that should be universal (Share 2008).

In this paper, we will focus on Russian as a representative Slavic language that uses the Cyrillic alphabet, with more than 160 million speakers in the Russian Federation alone. The transparency of its writing system puts it in the middle of the continuum, between shallow (Finnish) and deep (English) orthographies. Several characteristics of Russian, especially its phonology (e.g., non-systematic stress patterns, conditional pronunciation in the form of vowel reduction and consonant assimilation, complex syllable structure, and long polymorphemic words) as well as morphology (rich inflectional and derivational morphology), are of considerable interest for comparative reading research. We introduce *the Russian Sentence Corpus (RSC)* that is the first systematic corpus of basic benchmarks of eye movements in reading in Russian by skilled young adults that extends the existing eye-movement corpora of European Roman-based and Asian logographic languages to include Cyrillic.

2. Towards a common protocol for cross-linguistic eye-movement corpora

Despite the fact that there are several cross-linguistic corpora of eye movements in reading, they are difficult to compare because of discrepancies in stimuli materials, data-collecting methods, and statistical analysis techniques. This is one of the reasons of why cross-linguistic progress has been so slow. The solution is to develop a common protocol that provides guidelines for creating a set of reading materials that are tightly controlled along several design manipulations that influence eye movements in other languages.

2.1 Eye-movement corpus for English (Schilling et al. 1998)

Schilling et al. (1998) constructed 48 English sentences containing either one of 24 low- or one of 24 high-frequency target words closely matched in length and preceded by a neutral sentence context. The goal was to compare frequency effects in lexical-decision reaction times, isolated word naming, and various measures of fixation durations during sentence reading. Reichle et al. (1998) added cloze predictabilities (measure of how successfully a word can be guessed based on the previous context, see Section 2.2) for all Schilling et al.'s words and then used the fixation durations and probabilities to fit the parameters of the E-Z Reader model. The Schilling data were also used to test the first version of the SWIFT computational model of eye-movement control (Engbert et al. 2002; Engbert & Kliegl 2001).

The successful fit of several computational models to the same data motivated an extension of this approach to other languages to systematically test both universal and language-specific characteristics that may affect eye movements in reading and language comprehension. The idea was to design similar materials across languages regardless of the type of orthography and create a protocol that could be flexible enough to choose language-specific grammatical features and structures.

2.2 Eye-movement corpus for German: Potsdam Sentence Corpus (Kliegl et al. 2004)

Kliegl et al. (2004) expanded the Schilling et al.'s (1998) protocol that resulted in creation of the *Potsdam Sentence Corpus* (*PSC*). The initial step in the protocol was selection of *target words* by orthogonally manipulating their three lexical characteristics in a 2 x 3 x 2 design: part of speech, length in characters, and frequency. There were only two parts of speech, nouns and verbs. Length in characters had three levels: short (3-4 characters), medium (5-7), and long (8–12). Frequency was either high, >50 items per million (ipm) or low, 1–4 ipm. Twelve target words were selected for each cell of this between-items design resulting in a total number of 144. Next, a novel sentence was created around each target word in such a way as to provide natural context for it, with the restriction that the target word was never in the sentence-initial or sentence-final position (e.g., *Die meisten* ***Hamster*** *bleiben bei Tag in ihrem Häuschen* 'Most **hamsters** stay in their houses during the day'; the target word is in bold). The 144 sentences ranged in length from five to eleven words, with the total number of 1138 words in the PSC. Grammatical structures of the sentences were simple and represented a variety of syntactic constructions characteristic of German, but they were not parametrically manipulated. The protocol allows for testing hypotheses about eye-movement control during reading (a) for all words in the sentences, and simultaneously (b) for target words with tightly controlled characteristics (namely, length and frequency) that are embedded in the sentences.

The second step for the PSC was collection of predictability norms for all words in 144 sentences using the cloze task. The predictability norming study preceded data collection for the PSC and was conducted with a separate group of 264 participants, resulting in 83 predictions for each word. Participants started with a blank screen and were asked to type any word. The script then would replace the word typed by the participant by the first actual word from one of the 144 sentences (e.g., *Die* …), and the participant had to guess the second word. At the beginning of the sentence, the participants' chance of guessing the actual word was close to zero, but it improved as they approached the end of the sentence.

The third step was to collect eye-movement data and extract the benchmarks from them using monolingual skilled German readers after they read the 144 sentences. Statistical analysis of eye movements was conducted first for the 144 target words and

then for remaining 994 words in the corpus (the first and last words of each sentence were excluded from the analysis). The dependent measures became the basic benchmarks of eye movements in German and were of three types—fixation durations, probabilities of skipping or fixating words, and probabilities of regression saccades (see Section 3.1.4). The basic benchmarks of eye movements in reading in German are presented in comparison to those in Russian in Section 3.2 (Table 2). In recent years, two additional extensions of the PSC have been added: PSC2 includes data of 85,000 predictions for 1230 words for the original 144 sentences (Laubrock & Kliegl 2015) and PSC3 crossed frequency with predictability within otherwise identical sentential frames (Dambacher et al. 2009; 2012). The benchmarks of eye movements in reading in German from the PSC have been successfully used to fit and test predictions of later versions of the SWIFT model (Engbert et al. 2005; Risse et al. 2014; Schad & Engbert 2012).

2.3 Other eye-movement corpora based on the PSC protocol

The main parameters of words that influence eye movements — frequency, length, and predictability — are universal in that they affect eye movements in the same direction in all languages regardless of orthography, but the differences between scripts (e.g., orthographic transparency) should yield predictable differences in the size of effects. This prediction has been tested in several studies that followed the PSC protocol in a variety of languages. These include French (Kennedy & Pynthe 2005), Dutch (*GECO Corpus* of Cop et al. 2016; Kuperman et al. 2010a), Argentinian Spanish (*Bahia Blanca*, Fernández et al. 2014), Chinese (Bai et al. 2008; Li et al. 2014; Yan et al. 2006; Yan et al. 2010), Japanese (Sainio et al. 2007), Thai (Winskel et al. 2009), Hindi (Husain et al. 2015), and Uighur (Yan et al. 2014). There is also one study with the same sentences read by Chinese, English, and Finnish participants in their respective language (Liversedge et al. 2016).

Regardless of the language, the basic benchmarks that are reported in the literature seem to hold in every language studied: average fixation duration ranges from 220 to 250 ms and reading times increase with words' length and decrease with their frequency. Saccade length and saccade landing position depend more strongly on the

writing system. The average saccade length is the longest in alphabetic languages that use Latin script (8 characters), shorter in Hebrew (5) and the shortest in Chinese (2–3). The single fixation position is more likely to be at the beginning or middle of the word for Chinese and Japanese, and at the middle for alphabetic languages. In Uighur, an agglutinative language that relies on heavy use of suffixes, landing position is also influenced by the number of suffixes (Yan et al. 2014), suggesting that morphological structure of parafoveal words influences saccade programs (also found in Finnish, Hyönä et al. 2017).

For Russian, several studies using eye tracking while reading have already been conducted, but they explored specific theoretical issues in low-level eye movements or in sentence processing (Alexeeva & Slioussar 2017; Anisimov, Fedorova, & Latanov 2014; Bezrukikh & Ivanov 2012; Chernova 2015; Jouravlev & Jared 2016). They aimed to answer questions unrelated to particular properties of the Cyrillic alphabet or reading strategies in Russian. In Section 3, we describe our study whose goal was to identify basic benchmarks in reading in Russian and create the Russian Sentence Corpus following the PSC protocol. To do so, we investigate the effects of length, frequency, and predictability on eye movements and test a few hypotheses about factors that may be specific for reading in Russian.

Following the PSC and other eye-movement corpora based on it, the RSC materials represent isolated sentences. The majority of previously published corpora used isolated sentences, and only a few employed coherent texts (e.g., newspaper articles in Kennedy & Pynthe 2005, short narratives in Husain et al. 2015, novel reading in Cop et al. 2016). The obvious advantage of using full texts is higher ecological validity because such a setup closely resembles natural reading. Coherent texts may be especially interesting when studying local predictability and contextual effects. However, one particular genre may be not characteristic of the texts and genres found in the language. In contrast, isolated sentences selected from different texts and genres are more representative of the variability in the language. From a methodological point of view, isolated sentences are also easy to fit on one line on the screen and therefore avoid line wrap-up and line switch effects. Presenting material on one line mitigates the problem of runaway fixations that are registered in the vertical space between two lines of text. Thus

full-text corpora that closely resemble natural reading are most useful as a second step in the reading research after basic benchmarks have been identified in an isolated sentence-based corpus.

3. The present study: Russian Sentence Corpus (RSC)

The design of the RSC followed the PSC protocol (Section 2.2) with data from 96 monolingual skilled readers of Russian.

3.1 Method

3.1.1 *Participants*

There were three groups of participants in the current study, all monolingual Russian-speaking adults. Group 1 (n = 215) provided acceptability judgments for the corpus sentences, Group 2 (n = 750) participated in the predictability norming study, and Group 3 (n = 96) read the corpus sentences. Their eye movements were used to calculate the basic benchmarks for reading in Russian and together with the materials constitute the Russian Sentence Corpus (RSC).

Group 3 that provided data for the main part of the study, i.e., reading the sentences from the RSC, consisted of 96 participants (66 women and 30 men, M_{Age} = 24, range 18–80). They volunteered for the study and did not receive any compensation for taking part in it. The study was carried out in accordance with the ethical principles of psychologists and code of conduct of the American Psychological Association and was approved by the local Institutional Review Board. All participants gave written informed consent in Russian in accordance with the Declaration of Helsinki. The study took between 25 and 40 minutes.

3.1.2 *Design and materials*

The materials were designed following the PSC protocol (Kliegl et al. 2004; 2006) as described above in Section 2.2, with an important modification: in contrast to the PSC, for which the sentences were created by the experimenters around the target words,

Russian sentences were randomly selected from the Russian National Corpus (Ruscorpora.ru). Using existing sentences increases the ecological validity of the study and potentially allows for a more natural contextual embedding of the words into sentences, which might influence the strategies of the readers.

First, we randomly selected 144 target words from the *StimulStat* database (stimul.cognitivestudies.ru, Alexeeva et al. 2017) using the pre-defined criteria for a modified 3 x 3 x 2 design that is manipulating word's part of speech, length, and frequency. We increased the number of levels for the part-of-speech variable from 2 to 3 by adding adjectives (e.g., *узкой* 'narrow-FEM.INSTR.SG') in addition to nouns (e.g., *страницы* 'page-FEM.GEN.SG') and verbs (e.g., *заварил* 'brewed-MASC.PAST.SG'). Each length-frequency design cell contained 12 nouns, six verbs and six adjectives, except for the short words where we had to increase the number of nouns to 16 and decrease the number of verbs and adjectives because 3–4 letter verbs and adjectives (e.g., *всей* 'entire-FEM.GEN.SG', *жить* 'to live-INF') are rare in Russian. This affected four of the design cells. The length variable had three levels: short (3–4 characters), medium (5–7), and long (8–10). Frequency was either high (> 50 ipm) or low (<10 ipm). For selection of the target words, we used lemma length and lemma frequency information taken from Lyashevskaya and Sharov (2009).

Using the resulting list of 144 target words we extracted sentences from the Russian National Corpus that included target words in such a way that their position ranged from the third from the beginning to the third from the end of the sentence. We aimed at representing diverse types of syntactic structures typical for Russian including declarative, exclamatory, and interrogative sentences and sentences with non-canonical word orders, but did not manipulate the grammatical structure parametrically. We replaced complex lexical items with simpler ones and shortened the sentences when they exceeded the preset maximum length of 13 words (for details, see Table 1). Example (1) illustrates how one such long and lexically complex original sentence (1a) from the Russian National Corpus was adapted for the RSC (1b) (the target word ***лёд*** 'ice-MASC.NOM.SG' is in bold).

(1) a. В болотах млел ещё жёлтый кислый **лёд**, но на берегах уже появилась из-под снега прошлогодняя трава и груды торфа.

‘The yellow sour **ice** was still melting in the marshes, but the grass from last year and piles of peat already appeared on the river banks.’

b. На болотах оставался ещё **лёд**, но на берегах реки появилась трава.

‘**The ice** remained on the marshes, but the grass appeared on the river banks.’

A representative set of 13 sentences is provided in Appendix 1.

Second, the 144 selected sentences were subjected to acceptability norming. We used the web-based service *Virtualexs* (https://virtualexs.ru/) designed to conduct online surveys in Russia. Participants (n = 215) read each sentence online and were asked to judge its acceptability on a Likert scale ranging from 1 “*totally unacceptable*” to 5 “*perfectly acceptable*”. The four sentences with mean scores below 3 were modified by our research team.

Third, the 144 modified sentences were used in a predictability norming study (see Section 2.2), with one technical modification: We collected the norms online and did not pose any restrictions on the number of sentences each participant guessed. We included data from every participant that made more than 20 guesses out of 1362 words in the corpus.

The resulting set of 144 sentences was then morphologically annotated. First, an automated annotation was performed using the Mystem algorithm (https://tech.yandex.ru/mystem/): the lemma was identified, tagged for part-of-speech information and for morphological features (animacy, number, gender, and case for nouns; transitivity, tense, mood, number, gender, and aspect for verbs, etc.). Possible ambiguity between parts of speech and morphological features was noted. Two trained linguists independently reviewed the results of the automated annotation and, if necessary, disambiguated or corrected them.

The main and final step was to collect eye movements from 96 monolingual Russian-speaking participants as they read the entire RSC and to calculate the basic

benchmarks of eye movements in reading in Russian described in Sections 3.1.3 and 3.2, respectively.

3.1.3 *Procedure*

Sentences were presented in the middle of a 24-inch ASUS VG248QE monitor (resolution: 1920 x 1080 pix, response time: 1 ms, frame rate: 144 Hz, font face: 22pt Courier New) controlled by a ThinkStation computer. The presentation of the materials and recording of the eye movements were implemented by *Experiment Builder* (SR Research Ltd.). Participants were tested individually with the Eyelink 1000+ desktop mount eye-tracker using a chin rest. They were seated at a comfortable distance of 55 cm from the camera and 90 cm from the monitor. In this setup, one character subtended 0.29° visual angle. Only the right eye was tracked, at 1000 Hz rate. Calibration consisting of 9 points was performed before the beginning of the experiment and after every 15 sentences afterwards.

Each trial began with a fixation point at the position of the first letter of the first word in the sentence. If the participant fixated it for at least 500 ms, the sentence presentation automatically commenced; otherwise, after 2 s, 9-point calibration was repeated. Sentences were presented in one line in the middle of the screen against light gray background. After finishing reading the sentence, participants were instructed to look at the red dot in the lower right-hand corner of the screen. To ensure that participants read the sentences for comprehension, 33% of them were followed by an easy three-choice comprehension question; the response was recorded from a mouse click. Accuracy was always above 80%. The program advanced to the next trial after a 1-s delay.

3.1.4 *Data Analyses*

Data from all participants regardless of their accuracy in answering comprehension questions were included. Eye-movement data were split into fixations and saccades based on the algorithm from the *Data Viewer* package (SR Research Ltd). The first and last words in every sentence were excluded from the analyses. The analyses were modeled on the ones used for the PSC in German (Kliegl et al. 2004); however, we used (generalized)

linear mixed models [(G)LMMs] instead of repeated-measure multiple regressions using R (R Core Team 2016) and *ggplot2* (Wickham 2016). (G)LMMs were estimated with *lme4* package version 1.1-8 (Bates et al. 2015), partial effects were modeled with *remef* package (Hohenstein & Kliegl, NA), and the comparison table for (G)LMM outcomes (Table 6) was created with *sjPlot* package (Lüdecke 2017).

The (G)LMMs included varying intercepts for participants, sentences, and individual words. Fixed effects were estimated for the following variables: (a) centered and scaled word form length (linear and quadratic trends), (b) logarithm (base 10) of word form frequency (as taken from the StimulStat database), and (c) logit-transformed predictability. The effects of the variables were estimated for nine dependent variables: four measures of reading time (i-iv) and five probabilities relating to skipping, fixating or regression to or from words (viii-ix):

i. first fixation duration (FFD);

ii. single fixation duration (SFD);

iii. gaze duration (GD);

iv. total reading time (TT);

v. probability of skipping the word (P0);

vi. probability of fixating the word only once (P1);

vii. probability of fixating the word more than once (P2);

viii. probability of regression to the previous words from the current word (RO);

ix. probability of regressing back to the word from the following words (RG).

To ensure the normal distribution of model residuals, durations (FF, SF, GD, and TT) were log-transformed. Binary dependent variables (P0, P1, P2, RO, and RG) were fit with GLMMs with a logistic link function. There was no excessive collinearity of model predictors, as the variance inflation factor (VIF) for all of them was less than 5.

The sentences, eye-movement data, and the script used for the analyses reported below are available at the Open Science Framework project page: https://osf.io/x5q2r/.

3.2 Replication Results: Similarities between the RSC and PSC

3.2.1 *RSC: Descriptive characteristics of the materials*

Table 1 presents the comparison of the descriptive characteristics of the materials (for all sentences, corpus words, and target words) from the RSC with the PSC. The Russian sentences were longer than the German ones; therefore, the RSC contains 224 more words than the PSC. As Russian possesses a number of highly frequent short words (1- and 2-character long), there were many more short words in the RSC, but the proportion of short words (1-4 characters) was lower in the RSC (35%) compared to the PSC (41%). Word frequency distribution was also different across corpora: the RSC had 61% low (1–100 ipm), 16% average, and 23% high-frequency words (PSC: 45%, 24%, and 30%, respectively). Word predictability was measured as the number of correct guesses divided by the total number of guesses, and the distribution was quite comparable in both corpora: the RSC had 65% words with low predictability, 9% with average, and 23% with high (PSC: 66%, 11%, and 26%). The part-of-speech composition for the entire RSC was 468 nouns (34%), 282 verbs (21%), 126 adjectives (9%), 52 adverbs (4%), and 434 (32%) pronouns and function words (no data is available for the PSC).

Table 1. Descriptive statistics of the RSC and PSC.

	Russian Sentence Corpus (this study)		**Potsdam Sentence Corpus (Kliegl et al. 2004)**
# of sentences	144		144
Sentence length	Range: 5–13 words, *M* = 9		Range: 5–11 words, *M* = 7.9
# of words	1362 words 1218 (without first and last words)		1138 words 850 (without first and last words)
Word length	Range: 1–16 *M* = 5.7, *Mdn* = 6 (all words) *M* = 6.3, *Mdn* = 6 (target words)		Range: 2–20 *M* = 5.5, *Mdn* = 5 (all words) M = 7 †
Guesses per word	20–151		83
Word length (characters)	All words: 1 – 102	Target words: –	All words: 1 – 0

	Russian Sentence Corpus (this study)			**Potsdam Sentence Corpus (Kliegl et al. 2004)**	
	2 – 88	–		2 – 54	
	3 – 99	3 – 13		3 – 222	
	4 – 141	4 – 30		4 – 134	
	5 – 163	5 – 20		5 – 147	
	6 – 151	6 – 14		6 – 129	
	7 – 154	7 – 19		7 – 92	
	8 – 101	8 – 15		8 – 72	
	9 – 77	9 – 14		9 – 66	
	10 – 71	10 – 19		10 – 20	
	11 – 40	–		11 – 25	
	12 – 17	–		12 – 16	
	13 – 9	–		13 – 17	
	13+ – 5	–		13+ – 7	
Word frequencies					
all words:	class 1 (1–10 ipm)	404		class 1 (1–10 ipm[††])	242
	class 2 (11–100 ipm)	340		class 2 (11–100 ipm)	207
	class 3 (101–1000)	192		class 3 (101–1000)	242
	class 4 (1001–10 000)	151		class 4 (1001–10 000)	227
	class 5 (10 001–max)	131		class 5 (10 001–max)	76
target words:	class 1 (1–10 ipm)	89			
	class 2 (11–100 ipm)	49			
	class 3 (101–1000)	5			
	class 4 (1001–10 000)	1			
	class 5 (10 001–max)	0			
Predictability (%)					
all words:	*M* = 16%, *Mdn* = 1%			*M* = 18%, *Mdn* = 5%	
target words:	*M* = 10%, *Mdn* = 0%				
Predictability					
all words:	class 1 (-2.553 to -1.5)	663		class 1 (-2.553 to -1.5)	506
	class 2 (-1.5 to -1.0)	139		class 2 (-1.5 to -1.0)	111
	class 3 (-1.0 to - 0.5)	115		class 3 (-1.0 to - 0.5)	114
	class 4 (-0.5 to 0)	120		class 4 (-0.5 to 0)	88
	class 5 (0 to 2.553)	181		class 5 (0 to 2.553)	175

		Russian Sentence Corpus (this study)	Potsdam Sentence Corpus (Kliegl et al. 2004)
target words:	class 1 (-2.553 to -1.5)	102	
	class 2 (-1.5 to -1.0)	12	
	class 3 (-1.0 to - 0.5)	11	
	class 4 (-0.5 to 0)	8	
	class 5 (0 to 2.553)	11	

† Empty cells for the PSC mean that no data are reported in Kliegl et al. (2004; 2006)

†† items per million

3.2.2 *RSC: Benchmark statistics of eye movements in reading in Russian*

Descriptive characteristics

The entire RSC consisted of 1362 words, with the first and last words of every sentence excluded from the analysis resulting in 1218 words. Figure 1 presents four average fixation duration measures (measures i-iv) and their confidence intervals as a function of word's length, frequency, and predictability (Fig. 1A-1C). Means (*SD*) aggregated by participants are as follows: SFD (blue line) – 228 (26) ms, FFD (lilac line) – 217 (23) ms, GD (green line) – 259 (42) ms, TT (red line) – 318 (79) ms.

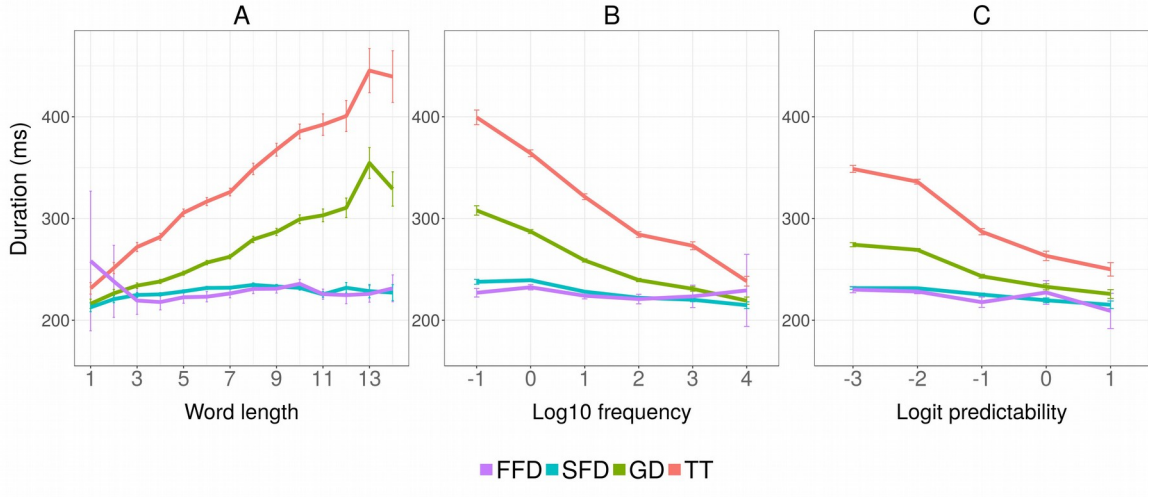

Figure 1. All analyzed corpus words in the RSC (*n* = 1218): mean RT and 95% CI for four fixation duration measures (FFD, SFD, GD, TT) as a function of word length (A), log-transformed frequency (B), and logit-transformed predictability (C).

Figure 2 illustrates mean proportions and confidence intervals of skipping (P0) or fixating the corpus word (P1 and P2, measures v-vii) as a function of word's length (A), frequency (B), and predictability (C). One third of all the corpus words in the RSC were skipped (34%), and this rate is consistent with 30–35% skipping rate reported for English (Rayner 1998). Half of the words were fixated once (56%), which is, again, highly consistent with the rate of single fixations reported for German, 57% (Heister, Würzner, & Kliegl 2012). The remaining 9% of words were fixated two or more times. The means are different from the model predictions in Table 2 because the intercept of the model represents predictions for words of average length, frequency of 1 ipm and 50% predictability, while the mean skipping rate provided here is computed over all corpus words.

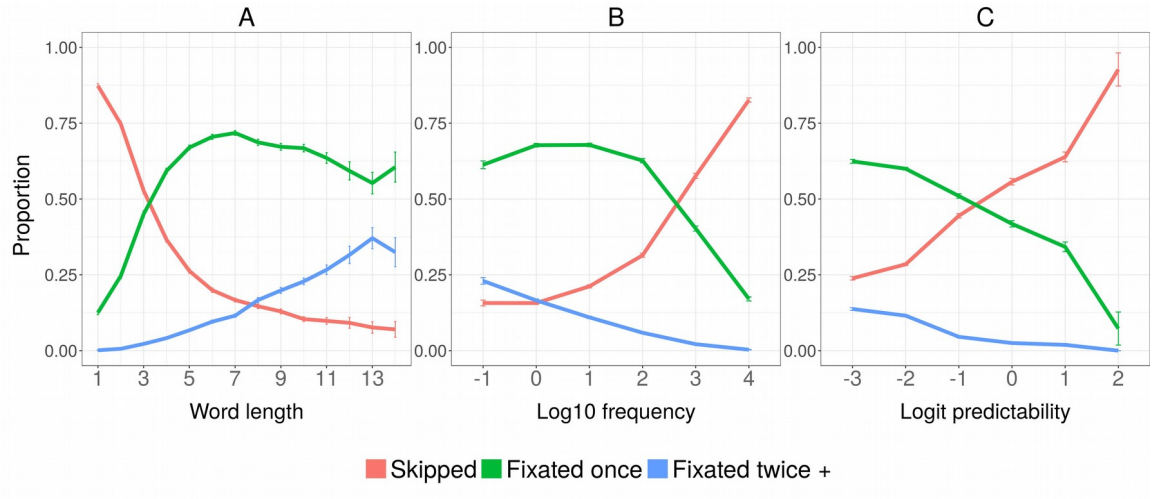

Figure 2. All analyzed corpus words in the RSC (*n* = 1218): P0, P1 and P2 as a function of (A) word length, (B) log-transformed word frequency, and (C) logit-transformed predictability.

Finally, for the saccade measures (RO and RG, viii-ix), 13% of the corpus words were regressed to from the following regions, and 17% of words served as an origin of a regressive saccade. Similar to other alphabetic languages, the average saccade length in the RSC spans 8 character spaces, with the saccades landing on the first half of the word closer to the word center (0.43 of the word's length, where zero represents the beginning, and 1 the end of the word).

Comparison with the PSC

Table 2 summarizes the comparison between the RSC and the PSC. The analysis of all corpus words (top part of Table 2) shows that most of the basic effects reported in the PSC for German were also replicated in the RSC for Russian, with a few differences (differences between the RSC and PSC that manifested in presence/absence of a certain effect or in its direction are shaded in grey). The first such difference is that in Russian, but not in German, P1 increases with the increase in word's length and predictability. The explanation may be trivial: in Russian, if a word is not fixated once, it is more likely to be skipped than to be fixated more than once (see Fig. 2), while in German the opposite is true. This means that in RSC we are comparing the words that were fixated once with those that were skipped, and longer words are more often fixated than skipped. The second difference is less clear: higher predictability increases P1 in Russian, but this effect was not significant in German. Theoretically, higher predictability should increase the probability of skipping, and this trend is present in the analysis of the target words in Russian. It is possible that the fact that as a word's predictability increases, its probability of being fixated also increases is due to the lower correlation between the word's length and frequency, or word position in the sentence (compared to German), yielding better statistical power for this positive predictability effect in Russian.

Table 2. All corpus (*n*=1218) and target (*n*=144) words (controlled for length and frequency) in the RSC: Basic benchmarks of eye movements in reading in Russian as compared to German. The cells in which the effects pattern differently between the RCS and PCS are shaded in grey. Significant effects are in bold. The first four fixation duration measures (FF, SF, GD, TT) are in ms, the rest are probabilities.

	Russian Sentence Corpus (this study)					**Potsdam Sentence Corpus (Kliegl et al. 2004)**				
	All Corpus Words									
	Constant	Length	Length^2	Freq	Pred	Constant	Length	Length^2	Freq	Pred
FF	222	2.8	-0.1	**-4.5**	0.5	207	1.5	0.3	**-5.4**	-3.2
SF	232	**2.7**	-0.5	**-9.5**	**-4.5**	210	**3.3**	0.5	**-6.3**	**-5.3**
GD	231	**7**	-0.5	**-14.8**	**-6**	241	**8.5**	0.5	**-11.8**	**-10.3**
TT	283	**11.3**	-1.33	**-18.7**	**-13.3**	245	**7.4**	-0.1	**-14.5**	**-17.5**

P0	17.8	**-5.4**	**0.56**	**1.8**	0	9.1	**-3.4**	**0.8**	**2.3**	1.8
P1	68.8	**4.7**	**-0.9**	-0.2	**0.6**	74.1	-0.6	**-1.0**	0.7	0.9
P2+	4.9	**2.0**	**-0.08**	**-0.99**	**-1.1**	17.0	**4.1**	0.2	**-2.5**	**-2.6**
RO	17.8	**-1.9**	0.9	**-1.3**	**2.9**	12.5	-0.6	-0.3	-0.8	0.0
RG	7	**-1.3**	**0.7**	-0.2	**-1.4**	0.4	**-1.0**	-0.1	-0.7	**-3.7**
Target Words										
FF	207	**2**	**-0.15**	-1.5	-0.9	214	0.4	**0.3**	**-4.5**	1.2
SF	225	**1**	**-0.17**	**-5**	-0.7	213	**1.6**	**0.6**	**-4.1**	-0.5
GD	251	**5**	0.2	**-8.6**	**-3**	247	**9.1**	**2.0**	**-8.1**	-3.5
TT	276	**9**	0	**-13**	**-7**	253	**9.1**	**1.9**	**-9.5**	**-7.3**
P0	14.9	**-9.8**	**5.7**	**2.5**	**1.68**	7.1	**-4.6**	**0.8**	**3.0**	**1.7**
P1	70.4	**9.6**	**-8.3**	0.1	-1.2	76.2	0.4	**-1.4**	**-1.8**	-0.1
P2+	6.7	**10.6**	**-1.3**	**-1.7**	**-0.9**	16.4	**4.2**	**0.5**	**-1.2**	**-1.7**
RO	14.7	**-2.8**	0	-1.1	0	7.3	0.0	0.0	**-0.6**	**-1.8**
RG	5.5	**-1.4**	**0.6**	-0.1	**-1.2**	0.2	**-1.0**	**0.1**	0.5	**-2.5**

Finally, for the regression measures, the probability that a word is the origin of a regression (RO) does not depend on any of the parameters in German, while in Russian, it increases with word predictability and decreases when word length and frequency increase. However, only the length effect remains constant for the target words, so the frequency and predictability influence might once again have to do with the length and frequency correlation in all corpus words. We leave the explanation of this pattern of results for future research.

For the target words (*n* = 144; Table 2, bottom part), controlled for length and frequency, the relationships between the basic word parameters and dependent measures are also very close to those of the PSC in German: As frequency and predictability of the target word increase, the reading times decrease (all measures), and as the target word length increases, the reading times also increase.

There were some minor differences in the timing of these effects. First, in Russian, the target word length affects all fixation duration measures (i.e., FFD, SFD, GD, and TT) whereas FFD was not affected in German. Second, predictability in Russian affects both GD and TT, but only TT in German. These effects might have a trivial explanation: in the analysis by Kliegl et al. (2004), data from 65 participants were

included, while the materials of the RSC were read by 96 participants. It is possible that higher statistical power allowed us to detect the effects of smaller size in the ‘earlier’ duration measures. Third, in Russian, FFD and P1 do not depend on word frequency; in German, frequency affects all eye-movement measures.

The most notable difference between the two corpora with respect to the target words is the influence of the square of word's length (*length*^{2} in Table 2, which exaggerates the difference between short and long words): in German, increase in the length^2 leads to an increase in FFD, SFD, GD, TT, and in P2+ while in Russian, the opposite is true. That is, in Russian, longer words do not attract longer fixation durations; moreover, there is a tendency for fixation durations to get shorter for longer words. At the moment, pending future exploration of the RSC, we hypothesize that it has to do with predictability of morphological marking in Russian. Longer words contain more affixes and because they can be anticipated in the sentential context, skilled readers take advantage of this anticipatory information by spending less time on longer words with affixes. An alternative explanation concerns reading proficiency: Kuperman and Van Dyke (2011) demonstrated that for more proficient readers, the correlation between the word’s length and reading time was weaker than for lower-skilled readers; that difference between readers was most apparent in reading times for longer words. As the majority of our sample were skilled readers (i.e., university students), the difference between corpora might be explained by individual differences between readers, and not languages.

The impact of the previous and upcoming words on single fixation durations

Finally, to see how the properties of the previous and the upcoming words influence the SFD on the current word in Russian and German, we compared the data from the RSC with multiple regressions analysis from Kliegl et al. (2006). The most notable differences between the corpora concern the effect of the previous, current, and upcoming words’ length on SFD are in Table 3.

Table 3. All analyzed corpus words (*n* = 1218): Predicted single fixation duration (SFD, measured in ms) as a function of the previous, current, and the next words’ frequency,

predictability, and length. The cells in which there are significant differences between the RCS and PCS are shaded in grey.[†]

	Russian Sentence Corpus (this study)	**Potsdam Sentence Corpus (Kliegl et al. 2006)**
	Current Word	
Constant	**225**	**208**
Frequency	**-5.5**	**-4.6**
Predictability	**-1.7**	**-5**
Length	0.05[††]	**55**
	Previous Word	
Frequency	**-1.23**	**-5.1**
Predictability	**1.93**	**-1.6**
Length	1.05	**15**
	Following Word	
Frequency	**-3.6**	**-2.5**
Predictability	**3.2**	**1.9**
Length	**-5.7**	-1.2
Incoming saccade amplitude	**2.11**	**4.7**
Saccade landing position	**-2.8**	**-7.7**

[†] Some of the models reported in Kliegl et al. (2006) included additional predictors and interactions between those predictors.

[††] In Kliegl et al. (2006), word length was reciprocally transformed. In our analysis, it was centered and scaled. It follows that the effect size estimates are on different scales and not directly comparable.

The current word. In contrast to well-established length effects in English and German, in Russian, the current word's length does not affect SFD. One possible explanation is that the word that was fixated once was already anticipated before the saccade was launched to it; in this case, the single fixation serves to check whether the prediction was correct and does not require the reader to fully process the word. Or again, the individuals that read the RSC were more proficient readers who could quickly recognize whole word forms, which lead to their reading times being less affected by word length. Finally, the relationship between frequency, predictability, and single fixation duration was as

predicted: as in other languages, increase in frequency and predictability decreases SFD on the word.

The previous word (n-1). The previous word's length does not affect SFD in Russian, in contrast to German. Another difference to German concerns predictability: increase in predictability of the previous word increases rather than decreases FFD on the current word in Russian. This might be explained by more predictable words being skipped more often as fixations following word skipping are known to be longer.

The upcoming word (n+1). We also found that in Russian, but not in German, the increase in length of the upcoming word decreases reading times on the current one. We tentatively attribute these faster reading times to the distributed word processing: Russian readers process the upcoming word parafoveally when it is short (thus spending more time fixating the current word), and in the fovea, when it is long (thus making a saccade to the upcoming word and spending less time on the current word). This strategy confirms the other replicated effects that speak in favor of the distributed word processing: both the negative $n + 1$ frequency and positive $n + 1$ predictability effects that were previously found (Fernandez et al. 2013; Kliegl et al. 2006; Laubrock & Kliegl 2015; Schad et al. 2012) were significant. Although the idea of distributivity of lexical processing across several words during reading is debated (Rayner et al. 2007), at least for Russian, the negative $n + 1$ frequency and positive $n + 1$ predictability effects, as well as the negative $n + 1$ length effects, all strongly support distributed lexical processing.

4. Novel results: Exploitation of the RSC

To demonstrate a broader range of potential applications of the RSC, we used it in three small exploratory investigations of how eye movements in reading in Russian are influenced by the most prominent characteristic of the Russian language, i.e., morphology. The analyses reported below used LMMs that were based on two sets of predictors: the ones used for comparison between RSC and PSC (i.e., length, frequency, and predictability of the previous, current, and the upcoming words, as well as the

amplitude of the incoming saccade and the saccade landing position (see Table 3 in Section 3.2.2) and three novel morphological predictors, namely, *the part-of-speech (PoS)* category, *morphosyntactic ambiguity*, and *morphological word form* (base versus non-base). We also controlled for the relative position of the word in the sentence, an important predictor of reading speed (Kuperman et al. 2010b). The comparison between the models is presented in Table 4 below. The full summary of the models is presented in Table 5 in Appendix 2.

Table 4. Comparison between the basic model (Section 3.2.2) and models including additional parameters of interest. Each cell contains the value of the Akaike information criterion (AIC) for the given model. Cells with predictors added to the basic model contain the results of comparison to the simpler model in the row above: the value of the χ^2 statistic and the corresponding *p*-value. Significant improvements over simpler model are in bold.

	FFD	**SFD**	**GD**	**TT**
Basic model	1348	-716	39732	75667
+ Word position in the sentence	**1333**	**-898**	**39606**	**75664**
	$\chi^2(1) = 16.9$	$\chi^2(1) = 184$	$\chi^2(1) = 127$	$\chi^2(1) = 104$
	All *p*s < 0.0001			
+ Part of speech	1339	**-902**	**39598**	**75547**
	$\chi^2(4) = 2.25$,	$\chi^2(4) = 12$,	$\chi^2(4) = 16$,	$\chi^2(4) = 25$,
	$p > 0.05$	$p = 0.017$	$p = 0.003$	$p < 0.0001$
+ Morphosyntactic ambiguity	1341	-902	39599	75549
	$\chi^2(1) = 0.12$	$\chi^2(1) = 1.75$	$\chi^2(1) = 1.15$	$\chi^2(1) = 0.16$
	All *p*s > 0.05			
+ Base/non-base word form	**1338**	-903	39598	**75546**
	$\chi^2(1) = 4$,	$\chi^2(1) = 3.5$,	$\chi^2(1) = 3.4$,	$\chi^2(1) = 4.7$,
	$p = 0.044$	$p > 0.05$	$p > 0.05$	$p = 0.03$

4.1 Part-of-Speech (PoS) category

Research on lexical processing has found that verbs are often more difficult to process than nouns: they are acquired later (Bassano 2000), take longer time to produce (Szekely et al. 2005), induce higher processing-based activation in neuroimaging studies (Crepaldi et al. 2011), and are more impaired in naming than nouns in aphasia (Jonkers & Bastiaanse 1996; Mätzig et al. 2009). We hypothesized that verbs should be read slower than nouns in Russian. Indeed, Figure 3 shows that the fixation durations are longer for the verbs than for the other parts of speech in the RSC.

Adding the part of speech predictor significantly improved the fit of the models for SFD, GD, and TT (see Table 4), and statistical analysis confirmed that verbs are read slower than nouns in the GD and TT measures (see Table 5, Appendix 2). Words belonging to the other parts of speech (i.e., adjectives, adverbs, and function words) did not differ significantly from the verbs in any of the eye-tracking measures: the numerical difference in mean reading times is most likely accounted for by low-level parameters, such as frequency, length, and predictability. The difference between nouns and verbs, however, cannot be fully explained by these parameters. Thus, our findings confirm that verb processing requires more effort than noun processing and do so in one of the most ecologically valid setups, i.e., when verbs and nouns are embedded into natural sentences.

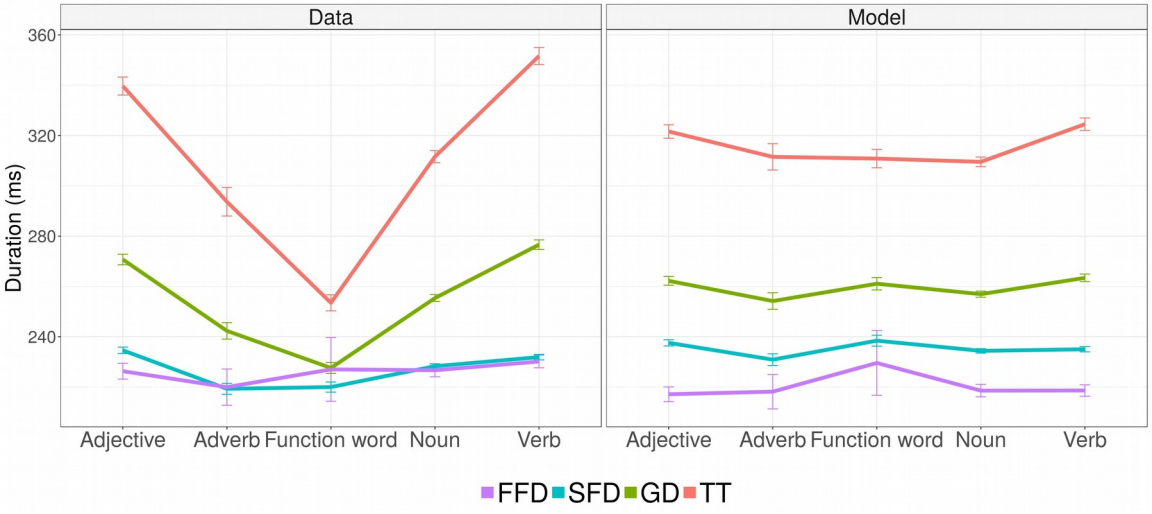

Figure 3. All corpus words in the RSC (*n* = 1218): mean RT and 95% CI for four duration measures as a function of the part-of speech category (adjectives, adverbs, function words, nouns,

and verbs). The left panel shows the empirical means; the right panel, partial effects of the mixed-effects model.

4.2 Morphosyntactic ambiguity

Research on lexical ambiguity in English has revealed that reading times increase at ambiguous words, if two meanings of the word are equally probable or if the context favors the less frequent meaning. But to the best of our knowledge, it is not known whether morphosyntactic ambiguity would influence reading times in the same way. Morphosyntactic ambiguity in the form of case syncretism on the noun (and its modifiers) is ubiquitous in Russian because it has a very elaborate nominal system with six grammatical cases and three declension classes. One morpheme (e.g., *-i*) can represent different cases as well as be used to convey syncretic information about the grammatical case, gender, and number (Baerman et al. 2005: Ch. 5). In the RSC, 35% of all words were ambiguous with respect to morphosyntactic form. For example, the word *аварии* ‘car accident(s)’ is morphosyntactically ambiguous between the ‘car accident-PREP/DAT/GEN.SG’ and ‘car accidents-NOM/ACC.PL’. Within the sentence, the majority of these ambiguous morphosyntatic forms is disambiguated by context, but we hypothesized that they might be processed slower, just as lexical ambiguity. Figure 4 demonstrates that the morphosyntactically ambiguous word forms in the RSC were read numerically slower than the unambiguous ones.

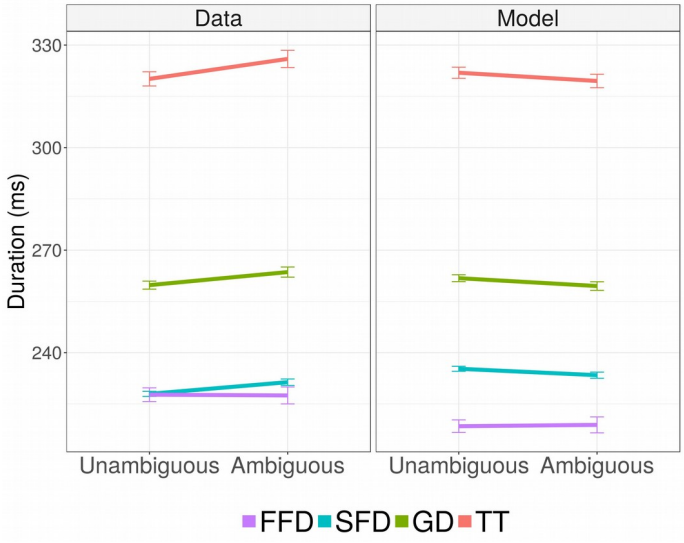

Figure 4. All analyzed corpus words (*n* = 1218): Mean RT and CI as a function of morphosyntactic ambiguity. The left panel shows the empirical means; the right panel, partial effects of the mixed-effects model.

However, adding the morphosyntactic ambiguity as a predictor did not improve fit of any of the time duration models (see Table 4). It follows that in the LMMs, there is no evidence for a difference in reading times between morphosyntactically ambiguous and unambiguous word forms in the RSC (see Table 5, Appendix 2). We attribute this apparent divergence between the means and the model estimates to the fact that the model accounts for the influence of the previous, current and upcoming words' length, frequency, and predictability. The observed difference in the mean reading times between the ambiguous and unambiguous word forms may be better explained by these parameters.

4.3 Base vs. non-base word form

The last question we explored concerned reading times for words in their base form (corresponding to the dictionary form, e.g., the NOM.SG case for nouns and the infinitive for verbs) as compared to their non-base forms (other cases for nouns and conjugated forms for verbs). Russian nouns have 12 inflectional forms (six cases x two numbers). Russian verbs belong to two conjugational classes and bear grammatical markings for person and number (as well as gender in the past tense). In addition, Russian grammatical

markers are always syncretic (see Section 4.2). This proliferation of inflected forms to a greater degree is also found in Finnish; however, Finnish is an agglutinative language in which one morphological marker corresponds to one grammatical feature and it does not display morphosyntactic ambiguity in the form of syncretism, the way Russian does.

Reading in Finnish has been studied extensively, and, in particular, a lot of attention has been paid to reading of inflected and compound word forms. Hyönä et al. (1995) found that in reading isolated Finnish words, inflected forms attracted longer first and second fixation durations than words in their base form. We were interested to see if the same effect is present in Russian for words in their base vs. non-base form (note that most base forms are inflected in Russian in contrast to Finnish). In the RSC, 34% of all words were in their base form, and the mean fixation durations were higher for the non-base form words (Fig. 5).

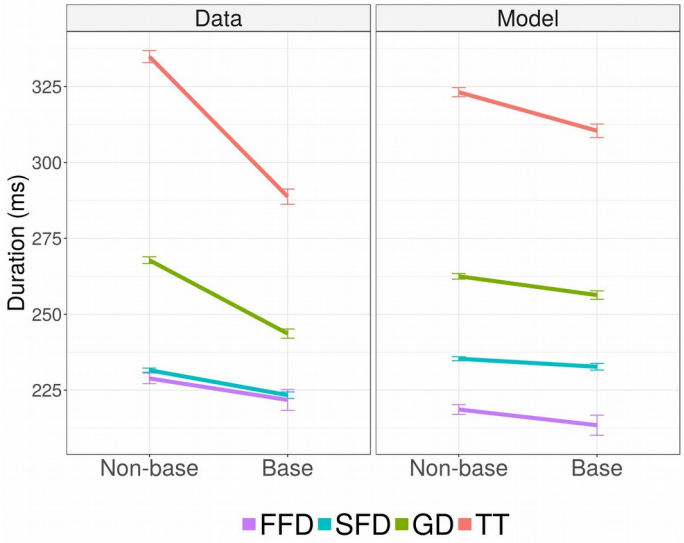

Figure 5. All corpus words (N=1218): mean RT 95% CI as a function of base/non-base word form. The left panel shows the empirical means; the right panel, partial effects of the mixed-effects model.

Adding the predictor differentiating base and non-base word forms significantly improved the fit of the models for SFD and TT (see Table 4). Non-base word forms indeed took longer to read, and the effect was significant in the SFD and TT measures (see Table 5, Appendix 2). However, given that no other lexical measures influenced SFD, the influence of word form on SFD is likely to be a Type I error. We leave this

intriguing question of whether base word forms are easier to process universally for our future investigation of the morphological factors in Russian.

5. Conclusion

The main goal of this article was to introduce the new Russian Sentence Corpus of eye-movements during sentence reading in a Slavic language with a Cyrillic script, i.e., Russian, which has not yet been investigated in cross-linguistic eye-movement research. As in every language studied so far, we have confirmed the expected effects of low-level parameters, such as word length, frequency, and predictability, on the eye-movements of skilled Russian readers. The findings from our study allow us to add Cyrillic-based Slavic languages to the growing number of languages with different orthographies ranging from the Roman-based European languages to logographic Asian ones whose eye-movement benchmarks confirm the universality of basic benchmarks in reading (Share 2008). We have also established descriptive corpus statistics for reading in Russian in the form of the average saccade length, landing site, fixation duration measures, probabilities of skipping and fixating words, as well as proportions of regressions, in reading of natural sentences. Finally, we have conducted three simple exploratory investigations of the effects of morphology on the basic eye-movement measures in Russian that illustrate the kinds of questions researchers can answer using the RSC.

We are confident that the RSC will be of particular use to the researchers interested in morphological processing because of rich inflectional and derivational morphology characteristics not only of Russian but of most Slavic languages. The novel feature of the RSC is its full morphological annotation, namely, full specification of the morphemes that comprise each word. Currently the Russian Sentence Corpus has the following levels of annotation: *i*) morpheme annotation (number and identity of word's affixes, annotated manually based on the Word formation dictionary by A.N. Tikhonov (2003), *ii*) disambiguated morphological annotation (part of speech, grammatical characteristics for each part of speech) performed with mystem2 (https://tech.yandex.ru/mystem/) and validated manually, *iii*) syntactic annotation in the terms of dependency grammar (according to the Universal Dependencies guidelines

http://universaldependencies.org), *iv*) phonetic stress annotation, and *v*) semantic annotation, i.e. the number of meanings according to Efremova (2000). The annotated corpus is freely available at https://osf.io/x5q2r/.

The effects of morphosyntactic information on eye movements in reading in fusional languages with pervasive syncretism like Russian differ from many Indo-European and agglutinative languages and await to be explored which may well result in modification of the existing theories of reading.

Funding statement

The study has been funded by the Center for Language and Brain NRU Higher School of Economics, RF Government grant, ag. № 14.641.31.0004. The first author was also supported by the DAAD funding program “Research Grants—Doctoral Programmes in Germany.”

References

Ablinger, I., Huber, W., & Radach, R. (2014). Eye movement analyses indicate the underlying reading strategy in the recovery of lexical readers. *Aphasiology, 28*(6), 640-657.

Ashby, J., Rayner, K., & Clifton, C. (2005). Eye movements of highly skilled and average readers: Differential effects of frequency and predictability. *The Quarterly Journal of Experimental Psychology Section A, 58*(6), 1065-1086.

Alexeeva, S., Slioussar, N., & Chernova, D. (2017). StimulStat: A lexical database for Russian. *Behavior Research Methods*, FirstView, 1-11.

Alexeeva S.V., & Slioussar, N. A. (2017). Effekt dliny pri parafoveal'noj obrabotke slov vo vrema chtenia. (in Russian). [‘The effect of length in parafoveal processing of word in reading.] *Tomsk State University Journal*. *Filologia, 45*, 5-29.

Anisimov, V. N., Fedorova, O. V., & Latanov, A. V. (2014). Eye movement parameters in reading sentences with syntactic ambiguities in Russian. *Human Physiology, 40*(5), 521-531.

Bai, X., Yan, G., Liversedge, S. P., Zang, C., & Rayner, K. (2008). Reading spaced and unspaced Chinese text: evidence from eye movements. *Journal of Experimental Psychology: Human Perception and Performance*, *34*(5), 1277.

Baerman, M., Brown, D., & Corbett, G. G. (2005). *The Syntax-Morphology Interface: A Study of Syncretism.* NY: Cambridge University Press.

Bassano, D. (2000). Early development of nouns and verbs in French: Exploring the interface between lexicon and grammar. *Journal of Child Language*, *27*(3), 521-559.

Bates, D., Maechler, M., Bolker, B., Walker, S. (2015). Fitting linear mixed-effects models using *lme4*. *Journal of Statistical Software*, *67*(1), 1-48.

Bezrukikh, M. M., & Ivanov, V. V. (2012). Eye movements during reading as an indicator of development of reading skill. *Fiziologiia Cheloveka*, *39*(1), 83-93.

Boston, M. F., Hale, J. T., Kliegl, R., Patil, U., & Vasishth, S. (2008). Parsing costs as predictors of reading difficulty: An evaluation using the Potsdam sentence corpus. *Journal of Eye Movement Research*, *2*(1), 1–12.

Chernova, D. A. (2015). Eye-tracking study of attachment ambiguity resolution in Russian. *Voprosy Psikholingvistiki*, *26*, 256-267.

Clifton, C., Jr., Staub, A., & Rayner, K. (2007). Eye movements in reading words and sentences. In R. P. G. van Gompel, M. H. Fischer, W. S. Murray, & R. L. Hill (Eds.), *Eye Movements: A Window on Mind and Brain.* (pp. 341-369). Amsterdam: Elsevier.

Cop, U., Dirix, N., Drieghe, D., & Duyck, W. (2017). Presenting GECO: An eyetracking corpus of monolingual and bilingual sentence reading. *Behavior Research Methods, 49*(2), 602-615..

Crawford, T. J., Devereaux, A., Higham, S., & Kelly, C. (2015). The disengagement of visual attention in Alzheimer's disease: A longitudinal eye-tracking study. *Frontiers in Aging Neuroscience, 7*. ArtID: 118.

Crepaldi, D., Berlingeri, M., Paulesu, E., & Luzzatti, C. (2011). A place for nouns and a place for verbs? A critical review of neurocognitive data on grammatical-class effects. *Brain and language*, *116*(1), 33-49.

Dambacher, M., Rolfs, M., Göllner, K., Kliegl, R., & Jacobs, A. (2009). Event-related potentials reveal rapid verification of predicted visual input. *PLoS One,4*, e5047, 1-8.

Dambacher, M., Dimigen, O., Braun, M., Wille, K., Jacobs, A.M., & Kliegl, R. (2012). Stimulus onset asynchrony and the time course of word recognition: Effects of frequency and predictability on event-related potentials. *Neuropsychologia, 50(8),* 1852-1870.

Demberg, V., & Keller, F. (2008). Data from eye-tracking corpora as evidence for theories of syntactic processing complexity. *Cognition, 109*(2), 193-210.

Efremova T.F. (2000). *Novyj tolkovo-slovoobrazovatel'nyj slovar' russkogo jazyka*. (in Russian) ['The new Russian language dictionary'].

Engbert, R. & Kliegl, R. (2001). Mathematical models of eye movements in reading: A possible role for autonomous saccades. *Biological Cybernetics, 85*, 77-87.

Engbert, R., Longtin, A., & Kliegl, R. (2002). A dynamical model of saccade generation in reading based on spatially distributed lexical processing. *Vision Research, 42*, 621-636.

Engbert, R., Nuthmann, A., Richter, E.M., & Kliegl, R. (2005). SWIFT: A dynamical model of saccade generation during reading. *Psychological Review, 112,* 777-813.

Fernández, G., Shalom, D. E., Kliegl, R., & Sigman, M. (2014). Eye movements during reading proverbs and regular sentences: The incoming word predictability effect. *Language, Cognition and Neuroscience*, *29*(3), 260-273.

Genzel, D., & Charniak, E. (2002). Entropy rate constancy in text. In *Proceedings of the 40th Annual Meeting on Association for computational linguistics* (pp. 199-206). Association for Computational Linguistics.

Gibson, E. (2000). The dependency locality theory: A distance-based theory of linguistic complexity. In A. Marantz, Y. Miyashita, & W. O'Neil (Eds.), *Image, language, brain:*

Papers from the first mind articulation project symposium (pp. 94-126). Cambridge, MA: The MIT Press.

Hale, J. (2001, June). A probabilistic early parser as a psycholinguistic model. In *Proceedings of the second meeting of the North American Chapter of the Association for Computational Linguistics on Language Technologies* (pp. 1-8). Association for Computational Linguistics.

Heister, J., Würzner, K.M., & Kliegl, R. (2012). Analysing large datasets of eye movements during reading. In J.S. Adelman (ed.), *Visual Word Recognition. Vol. 2: Meaning and Context, Individuals and Development* (pp. 102-130). Hove, UK: Psychology Press.

Hohenstein, S. & Kliegl, R. (NA). *remef: Remove Partial Effects*. R package version 1.0.6.9000. https://github.com/hohenstein/remef/

Husain, S., Vasishth, S., & Srinivasan, N. (2015). Integration and prediction difficulty in Hindi sentence comprehension: Evidence from an eye-tracking corpus. *Journal of Eye Movement Research, 8*(2), 1-12.

Huey, E. B. (1908). *Basic Studies on Reading*. New York, NY: Basic Books.

Hyönä, J., & Bertram, R. (2011). Optimal viewing position effects in reading Finnish. *Vision Research, 51*, 1279–1287.

Hyönä, J., Laine, M., & Niemi, J. (1995). Effects of a word's morphological complexity on readers' eye fixation patterns. In J. M. Findlay, R. Walker, & R. W. Kentridge (Eds.), *Eye Movement Research: Mechanisms, Processes and Applications.* (pp. 445-452). Amsterdam: Elsevier.

Hyönä, J., Yan, M., & Vainio, S. (2017). Morphological structure influences the initial landing position in words during reading Finnish. *The Quarterly Journal of Experimental Psychology*, FirstView.

Jonkers, R., & Bastiaanse, R. (1996). The influence of instrumentality and transitivity on action naming in Broca's and anomic aphasia. *Brain and Language*, *55*(1), 37-39.

Jouravlev, O., & Jared, D. (2016). Cross-script orthographic and phonological preview benefits. *The Quarterly Journal of Experimental Psychology*, 1-10. DOI:10.1080/17470218.2016.1226906

Keller, F. (2004). The Entropy Rate Principle as a predictor of processing effort: An evaluation against eye-tracking data. In *Proceedings of the Conference on Empirical Methods in Natural Language Processing. Barcelona 2004* (pp. 317-324).

Kennedy, A., & Pynthe, J. (2005). Parafoveal-on-foveal effects in normal reading. *Vision Research, 45*, 153-168.

Kim, Y. S., Radach, R., & Vorstius, C. (2012). Eye movements and parafoveal processing during reading in Korean. *Reading & Writing*, *25*, 1053–1078.

Kliegl, R. (2007). Toward a perceptual-span theory of distributed processing in reading: A reply to Rayner, Pollatsek, Drieghe, Slattery, and Reichle (2007). *Journal of Experimental Psychology: General*, *136*(3), 530.

Kliegl, R., Grabner, E., Rolfs, M., & Engbert, R. (2004). Length, frequency, and predictability effects of words on eye movements in reading. *European Journal of Cognitive Psychology*, *16*(1/2), 262-284.

Kliegl, R., Nuthmann, A., & Engbert, R. (2006). Tracking the mind during reading: The influence of past, present, and future words on fixation durations. *Journal of Experimental Psychology: General*, *135*(1), 12-35.

Kuperman, V., & Van Dyke, J. A. (2011). Effects of individual differences in verbal skills on eye-movement patterns during sentence reading. *Journal of Memory and Language*, 65(1), 42-73.

Kuperman, V., Bertram, R., & Baayen, R. H. (2010a). Processing trade-offs in the reading of Dutch derived words. *Journal of Memory and Language*, *62*(2), 83-97.

Kuperman, V., Dambacher, M., Nuthmann, A., & Kliegl, R. (2010b). The effect of word position on eye-movements in sentence and paragraph reading. *The Quarterly Journal of Experimental Psychology*, *63*(9), 1838-1857.

Laubrock, J., & Kliegl, R. (2015). The eye-voice span during reading aloud. *Frontiers in psychology*, *6*.

Lewis, G. E. (1972). *Multilingualism in the Soviet Union: Aspects of Language Policy and Its Implementation*. The Hague: Mouton De Gruyter.

Li, X., Bicknell, K., Liu, P., Wei, W., & Rayner, K. (2014). Reading is fundamentally similar across disparate writing systems: A systematic characterization of how words and characters influence eye movements in Chinese reading. *Journal of Experimental Psychology: General*, *143(2)*, 895-913.

Li, X., Liu, P., & Rayner, K. (2011). Eye movement guidance in Chinese reading: Is there a preferred viewing location? *Vision Research*, *51*, 1146–1156.

Liversedge, S.P., Drieghe, D., Li, X., Yan, G., Bai, X., Hyönä, J. (2016). Universality in eye movements and reading: A trilingual investigation. *Cognition*, *147*, 1-20.

Lüdecke, D. (2017). sjPlot: Data Visualization for Statistics in Social Science_. R package version 2.3.3, https://CRAN.R-project.org/package=sjPlot.

Lyashevskaya, O. N., & Sharov, S. A. (2009). *Chastotnyj Slovar' Sovremennogo Russkogo Jazyka (na Materialakh Natsional'nogo Korpusa Russkogo Jazyka*. (in Russian) ['*Frequency Dictionary of Modern Russian* (*based on the materials of the Russian National Corpus)'*]. Moscow: Azbukovnik.

Mätzig, S., Druks, J., Masterson, J., & Vigliocco, G. (2009). Noun and verb differences in picture naming: Past studies and new evidence. *Cortex*, *45*(6), 738-758.

Natsional'nyi korpus russkogo jazyka. [Russian National Corpus]. http://www.ruscorpora.ru/

Nuthmann, A., Engbert, R., & Kliegl, R. (2005). Mislocated fixations during reading and the inverted optimal viewing position effect. *Vision Research*, *45*(7), 2201-2217.

Paterson, K. B., Almabruk, A. A., McGowan, V. A., White, S. J., & Jordan, T. R. (2015). Effects of word length on eye movement control: the evidence from Arabic. *Psychonomic Bulletin & Review*, *22*(5), 1443-1450.

Paterson, K. B., McGowan, V. A., White, S. J., Malik, S., Abedipour, L., & Jordan, T. R. (2014). Reading direction and the central perceptual span in Urdu and English. *PloS One*, 9(2), e88358.

Pollatsek, A., Bolozky, S., Well, A. D., & Rayner, K. (1981). Asymmetries in the perceptual span for Israeli readers. *Brain and Language*, *14*(1), 174-180.

Radach, R., & Kennedy, A. (2004). Theoretical perspectives on eye movements in reading: Past controversies, current issues, and an agenda for future research. *European Journal of Cognitive Psychology*, *16*(1/2), 3-26.

R Core Team. (2016). *R: A language and environment for statistical computing [Computer software manual]*. Vienna, Austria. Retrieved from https://www.R-project.org/

Rayner, K. (1998). Eye movements in reading and information processing: 20 years of research. *Psychological Bulletin*, *124*(3), 372.

Rayner, K. (2009). Eye movements and attention in reading, scene perception, and visual search. *The Quarterly Journal of Experimental Psychology*, *62*(8), 1457-1506.

Rayner, K., Chace, K. H., Slattery, T. J., & Ashby, A. (2006). Eye movements as reflections of comprehension processes in reading. *Scientific Studies of Reading*, *10*(3), 241-255.

Rayner, K., Pollatsek, A., Drieghe, D., Slattery, T. J., & Reichle, E. D. (2007). Tracking the mind during reading via eye movements: Comments on Kliegl, Nuthmann, and Engbert (2006). *Journal of Experimental Psychology: General*, *136*(3), 520-529.

Reichle, E. D., Pollatsek, A., Fisher, D. L., & Rayner, K. (1998). Toward a model of eye movement control in reading. *Psychological Review*, *105*, 125–157.

Risse, S., Hohenstein, S., Kliegl, R., & Engbert, R. (2014). A theoretical analysis of the perceptual span based on SWIFT simulations of the n+2 boundary paradigm. *Visual Cognition, 22,* 283-308.

Sainio, M., Hyönä, J., Bingushi, K., & Bertram, R. (2007). The role of interword spacing in reading Japanese: An eye movement study. *Vision Research*, *47*(20), 2575-2584.

Schad, D. J., & Engbert, R. (2012). The zoom lens of attention: Simulating shuffled versus normal text reading using the SWIFT model. *Visual Cognition, 20(4-5),* 391-421.

Schad, D. J., Nuthmann, A., & Engbert, R. (2012). Your mind wanders weakly, your mind wanders deeply: objective measures reveal mindless reading at different levels. *Cognition, 125*(2), 179-194.

Schilling, H. E., Rayner, K., & Chumbley, J. I. (1998). Comparing naming, lexical decision, and eye fixation times: Word frequency effects and individual differences. *Memory & Cognition, 26*(6), 1270-1281.

Share, D. (2008). On the Anglocentricities of current reading research and practice: The perils of overreliance on an "Outlier" Orthography. *Psychological Bulletin, 134*(4), 584-615.

Szekely, A., D'Amico, S., Devescovi, A., Federmeier, K., Herron, D., Iyer, G., & Bates, E. (2005). Timed action and object naming. *Cortex*, *41*(1), 7-25.

Tiffin-Richards, S. P., & Schroeder, S. (2015). Word length and frequency effects on children's eye movements during silent reading. *Vision Research, 113*, 33-43.

Tikhonov, A.N. (2003). *Slovoobrasovatel'nyj slovar' russkogo yazyka v dvux tomax*. (in Russian) ['Word formation dictionary of Russian in two volumes']. Astrel'.

Tinker, M. A. (1958). Recent studies of eye movements in reading. *Psychological Bulletin*, *55*(4), 215.

Tsai, J. L., Kliegl, R., & Yan, M. (2012). Parafoveal semantic information extraction in traditional Chinese reading. *Acta Psychologica*, *141*(1), 17-23.

Vasishth, S., von der Malsburg, T., & Engelmann, (2013). What eye movements can tell us about sentence comprehension? *Wiley Interdisciplinary Review, 4*(2), 125-134.

Wickham, H. ggplot2: Elegant Graphics for Data Analysis. Springer-Verlag New York, 2016.

Winskel, H., Radach, R., & Luksaneeyanawin, S. (2009). Eye movements when reading spaced and unspaced Thai and English: A comparison of Thai–English bilinguals and English monolinguals. *Journal of Memory and Language*, *61*(3), 339-351.

Yan, G., Tian, H., Bai, X., & Rayner, K. (2006). The effect of word and character frequency on the eye movements of Chinese readers. *British Journal of Psychology, 97*, 259-268.

Yan, M., Richter, E. M., Shu, H., & Kliegl, R. (2009). Readers of Chinese extract semantic information from parafoveal words. *Psychonomic Bulletin & Review*, *16*(3), 561-566.

Yan, M., Kliegl, R., Richer, E. M., Nuthmann, A., & Shu, H. (2010). Flexible saccade-target selection in Chinese reading. *The Quarterly Journal of Experimental Psychology, 63*(4), 705-725.

Yan, M., Zhu, W., Shu, H., Yusupu, R., Miao, D., Krügel, A., & Kliegl, R. (2014). Eye movements guided by morphological structure: Evidence from the Uighur language. *Cognition, 132*, 181-215.

Appendix 1. Sample sentences from the RSC

Target words are in bold.

1. Не поручайте **мужу** ухаживать за рыбками в аквариуме, он обязательно забудет.
   Do not entrust the aquarium fish to your **husband**, he is certain to forget all about them.
2. Сделав мне знак помолчать, он приложил **ухо** к двери.
   Prompting me to keep silent, he pressed his **ear** to the door.
3. Дорога ведет в глухой **лес**, петляя по склонам.
   The road leads to the dense **forest**, winding along the slopes.
4. Мне было лень идти на стоянку и сметать **снег** с машины.
   I was too lazy to go to the parking lot and clean the **snow** off the car.
5. Если мы позволим этим людям **уйти**, наши проекты станут гораздо беднее.
   If we let these people **go**, our projects will be impoverished.
6. Тема эта в то время была **новой** для многих.
   This topic was **new** for many people at that time.
7. Зоопарк — это кусочек другого мира, находящийся в самом **центре** нашего района.
   Zoo is a piece of some other world right in the **center** of our district.
8. Чтобы придать объем **тонким** волосам, нанесите на них лечебную маску.
   To add volume to **thin** hair, apply the healing mask.
9. Судя по огромному **расходу** воды, они купали слонов.
   Judging by the enormous water **consumption**, they bathed elephants.
10. Володя каким-то образом **узнал** то, чего ему не надо было знать.
    Volodya somehow **learned** what he should not have learned.
11. Мне нравится сын коллеги, **который**[1] недавно заходил в наш отдел.
    I like the son of the colleague **who** recently stopped by in our department.
12. Зачем ему звонить, если откликается **спокойный** женский голос?
    Why call him, if a **calm** female voice answers the phone?
13. Но четыре года я не мог себя **заставить** сделать это.
    But for years ago I couldn't **make** myself do it.

Appendix 2

Table 5. Summary of the LMMs for the duration measures: FFD, SFD, GD, and TT. N+1 and N-1 represent the next and the previous words. The intercept represents the log-transformed mean

1 “Который” is a pronoun that takes adjective declension.

duration of the corresponding measure, the predictors – adjustments to the intercept per unit change.

	log FFD			log SFD			log GD			log TT		
	Estimate	*std. Error*	*p*	*Estimate*	*std. Error*	*p*	*Estimate*	*std. Error*	*p*	*Estimate*	*std. Error*	*p*
Fixed Effects												
(Intercept)	5.302	0.026	<.001	5.301	0.018	<.001	5.385	0.023	<.001	5.499	0.033	<.001
Log frequency	-0.007	0.004	.138	-0.023	0.003	<.001	-0.035	0.004	<.001	-0.048	0.005	<.001
Logit predictability	-0.012	0.005	.013	-0.015	0.002	<.001	-0.022	0.003	<.001	-0.036	0.004	<.001
Length	0.009	0.003	.001	0.003	0.001	.034	0.018	0.002	<.001	0.028	0.003	<.001
Length squared	-0.001	0.000	.034	-0.001	0.000	.002	0.001	0.000	.006	0.000	0.000	.653
n+1 length	-0.013	0.005	.011	-0.025	0.003	<.001	-0.033	0.004	<.001	-0.041	0.005	<.001
n+1 log frequency	-0.002	0.004	.594	-0.012	0.002	<.001	-0.016	0.003	<.001	-0.021	0.004	<.001
n+1 predictability	-0.006	0.004	.090	0.003	0.002	.137	0.002	0.003	.520	-0.008	0.004	.027
n-1 length	-0.001	0.005	.922	0.005	0.003	.056	-0.001	0.004	.710	-0.007	0.005	.151
n-1 log frequency	-0.002	0.004	.571	-0.003	0.002	.131	0.001	0.003	.842	0.001	0.004	.694
n-1 predictability	-0.002	0.004	.564	-0.002	0.002	.402	-0.003	0.003	.223	-0.011	0.004	.003
Word position in the sentence	0.071	0.017	<.001	0.130	0.009	<.001	0.137	0.012	<.001	0.171	0.017	<.001
Adjective	-0.005	0.012	.658	0.009	0.008	.220	0.000	0.010	.972	-0.000	0.015	.996
Adverb	0.002	0.020	.932	-0.019	0.012	.107	-0.026	0.016	.099	-0.030	0.023	.186
Function word	0.038	0.025	.126	0.016	0.012	.181	0.007	0.015	.649	-0.014	0.022	.532
Noun	0.000	0.009	.975	-0.005	0.006	.378	-0.017	0.008	.031	-0.039	0.011	<.001
Ambiguity	0.002	0.009	.852	-0.007	0.006	.242	-0.010	0.007	.191	-0.007	0.011	.489
Base form	-0.021	0.010	.048	-0.009	0.006	.141	-0.015	0.008	.069	-0.025	0.012	.031
Incoming saccade amplitude	-0.001	0.001	.367	0.009	0.000	<.001	0.006	0.000	<.001	0.006	0.000	<.001
Random Effects												
σ^2	0.065			0.056			0.101			0.171		
τ_{00}, word	0.002			0.002			0.005			0.010		
τ_{00}, sentence	0.000			0.001			0.001			0.003		
τ_{00}, participant	0.008			0.014			0.022			0.043		
Nwords	749			778			778			778		
Nsentences	144			144			144			144		
Nparticipants	96			96			96			96		
Observations	8746			55772			68725			68725		
R^2 / Ω_0^2	.176 / .170			.243 / .243			.259 / .259			.301 / .300		