INTERNATIONAL LITERACY ASSOCIATION

READING RESEARCH QUARTERLY

Monolingual and bilingual reading processes in Russian: An exploratory scanpath analysis

Journal:	*Reading Research Quarterly*
Manuscript ID	RRQ-2020-05-0111.R4
Manuscript Type:	Article
Keywords:	Fluency, Language learners, Methodological perspectives
Level of Learners:	2-Childhood, 4-Adolescence, 5-College / university, 6-Adult
Theoretical Perspectives:	Linguistic, Cognitive
Areas of Expertise:	Bilingualism, Cognitive Processes, Eye Movement, Fluency, Reading Strategies
Method/Analysis:	Cluster Analysis, Psycholinguistic, Mixed Methods

Abstract:	In the present study, we used a scanpath approach to investigate reading processes along with factors that can shape these processes in monolingual Russian-speaking adults, 8-year-old children, and bilingual Russian-speaking readers. We found that monolingual adults' eye-movement patterns exhibited a fluent scanpath reading process representing effortless processing of the written material: They read straight from left to right at a fast pace, skip words, and regress rarely. Both high-proficiency Heritage Speakers' and children's eye-movement patterns exhibited an intermediate scanpath reading process, characterized by a slower pace, longer fixations, an absence of word skipping, and short regressive saccades. L2 learners as well as low-proficiency Heritage Speakers exhibited a beginner reading process that involves the slowest pace, even longer fixations, no word skipping, and frequent re-reading of the whole sentence and of particular words. We suggest that, unlike intermediate readers who use the respective process to resolve local processing difficulties (e.g., word recognition failure), beginner readers, in addition, experience global-level challenges in semantic and morphosyntactic information integration. Proficiency in Russian for Heritage Speakers and comprehension scores for L2 learners were the only individual difference factors predictive of the scanpath reading process adopted by bilingual speakers. Overall, the scanpath analysis revealed qualitative differences in scanpath reading processes among various groups of readers and thus adds a qualitative dimension to the conventional quantitative evaluation of word-level eye-tracking measures.

Monolingual and bilingual reading processes in Russian: An exploratory scanpath analysis

Olga Parshina[1], Irina A. Sekerina[1,2], Anastasiya Lopukhina[1] and Titus von der Malsburg[3,4]

[1]HSE University, Moscow, Russian Federation
[2]College of Staten Island and The Graduate Center of The City University of New York, New York, NY, USA
[3]University of Stuttgart, Stuttgart, Germany
[4]Massachusetts Institute of Technology, Cambridge, MA, USA

Abstract

In the present study, we used a scanpath approach to investigate reading processes along with factors that can shape these processes in monolingual Russian-speaking adults, 8-year-old children, and bilingual Russian-speaking readers. We found that monolingual adults' eye-movement patterns exhibited a *fluent* scanpath reading process representing effortless processing of the written material: They read straight from left to right at a fast pace, skip words, and regress rarely. Both high-proficiency Heritage Speakers' and children's eye-movement patterns exhibited an *intermediate* scanpath reading process, characterized by a slower pace, longer fixations, an absence of word skipping, and short regressive saccades. L2 learners as well as low-proficiency Heritage Speakers exhibited a *beginner* reading process that involves the slowest pace, even longer fixations, no word skipping, and frequent re-reading of the whole sentence and of particular words. We suggest that, unlike intermediate readers who use the respective process to resolve local processing difficulties (e.g., word recognition failure), beginner readers, in addition, experience global-level challenges in semantic and morphosyntactic information integration. Proficiency in Russian for Heritage Speakers and comprehension scores for L2 learners were the only individual difference factors predictive of the scanpath reading process adopted by bilingual speakers. Overall, the scanpath analysis revealed qualitative differences in scanpath reading processes among various groups of readers and thus adds a qualitative dimension to the conventional quantitative evaluation of word-level eye-tracking measures.

Introduction

Over the last decade, eye-tracking has become a widely used methodology in bilingual language reading research (Cop, Dirix, Drieghe, & Duyck, 2017; Cop, Drieghe, & Duyck, 2015; Dirix, Vander Beken, De Bruyne, Brysbaert, & Duyck, 2020; Kang, 2014; Roberts & Siyanova-Chanturia, 2013; Schmidtke & Moro, in press; Whitford & Titone, 2012). Silent reading without metalinguistic tasks allows us to study written language comprehension in real time, and therefore, in laboratory settings, reflects the closest approach obtainable to the 'natural' reading behavior of bilingual readers. Conventional *local quantitative* characteristics of eye movements in reading, e.g., fixation durations and counts, skipping, and regression probabilities, reveal differences between bilingual and monolingual reading behavior, calculated on a word-by-word basis. The current study focuses on a previously unexplored approach to bilingual reading: *global qualitative* differences between bilingual and monolingual readers which we investigated using a *scanpath* approach to eye movements in reading.

Scanpaths are sequences of eye movements or *gaze trajectories,* that extend beyond the word level to the sentence and discourse levels. Specifically, scanpaths constitute sequences of eye gaze positions determined by the x- and y-coordinates and reflect the precise time frame of each fixation within a sentence or text. In contrast to local word-by-word eye-movement measures in reading (for review see Clifton, Staub, & Rayner, 2007; Rayner, Chace, Slattery, & Ashby, 2006; Vasishth, von der Malsburg, & Engelmann, 2013), scanpaths inform us about global eye movements at the sentential level. For instance, a scanpath analysis of regressive saccades allows us to identify not just whether a regression has occurred or not, but also to distinguish between different types of regression (e.g., single short leftward saccades versus multiple leftward saccades for 'reverse reading'). Furthermore, reading left-to-right or right-to-left can produce similar fixation durations even though they represent different cognitive processes. Scanpaths provide the ability to distinguish between these two reading behaviors because they take the whole global sequence of eye-movement events into account. Furthermore, scanpaths allow differentiation between various types of word skipping instances (e.g., skipping followed by re-reading, skipping only in the second or third pass reading, absence of skipping) and between fixation distributions (e.g., increased fixation durations during first-pass reading and decreased durations throughout subsequent re-readings). Overall, scanpath analyses can render a much more detailed and coherent picture of the qualitative differences between groups of readers than conventional, local eye-tracking metrics alone (see the Data Analysis section for a more concrete description of the scanpath measure).

This global qualitative approach to bilingual reading presents a unique opportunity to bridge the theories of visual lexical access and written language comprehension in bilingualism. The lexical access models of Dijkstra and colleagues (i.e. the *Bilingual Interactive Activation Plus*, Dijkstra & Van Heuven, 2002; and *Multilink,* Dijkstra et al., 2019) predict delayed lexical access in reading in bilinguals with low L2 proficiency levels. It is manifested in 'early-stage' processing measures, specifically, less word skipping and longer fixation durations (Roberts & Siyanova-Chanturia, 2013; Whitford, Pivneva, & Titone, 2016). These measures, however, do not address the question of whether delayed lexical access has repercussions for parsing of the entire sentence. The scanpath approach, on the other hand, is ideal for detecting any such implications in global eye-movement behavior as it can identify eye-movement patterns at the later stages of sentence comprehension and 'visualize' connections to the eye-movement patterns at the early stages of processing (e.g., a specific regression pattern to a problematic word). At the same time, a scanpath approach can inform theories of bilingual sentence processing such as the

Shallow Structure Hypothesis (Clahsen & Felser, 2006), the *good-enough parsing account* (Ferreira, Bailey, & Ferraro, 2002) or the accounts that emphasize the failure in memory retrieval operations (Cunnings, 2017; Van Dyke, Johns, & Kukona, 2014). While the goal of this exploratory study is not to tackle any specific theory, the scanpath approach will allow us visually and qualitatively to evaluate the word recognition and sentence processing stages simultaneously. The scanpath approach therefore provides us with a novel tool to investigate reading behavior from the perspective of both the bilingual lexical access and the language comprehension theories in parallel.

The scanpath approach can also play a critical role in distinguishing various types of global gaze trajectories that are frequently characteristic of a participant or a group of participants: We refer to these types as *scanpath reading processes*. We define this term as a recurrent pattern of gaze trajectories at sentence level, characterized by the specific time frames and locations of fixations, word skipping, and re-readings of the words, phrases, or sentences. This is in contrast to *reading comprehension strategies* (e.g., McNamara, 2007) that reflect readers' metacognitive efforts to understand the text and that is extensively investigated in educational research on bilingual reading (e.g., Comer, 2012; K.S. Goodman, 1979; Y.M. Goodman, 1996; Jiménez, García, & Pearson, 1996; Sheorey & Mokhtari, 2001; Stevenson, Schoonen, & Glopper, 2003). While these concepts are likely related, they focus on different aspects underlying comprehension strategies vs. overt reading behavior.

In this study, we focus specifically on the scanpath reading processes exhibited by a group of readers that can enable us to distinguish their eye-movement reading behavior from that of other groups. Thus, we establish and qualitatively describe three separate scanpath reading processes that occur in silent, uninterrupted reading without metacognitive tasks in two groups (adults and children) of monolingual, and two groups of bilingual readers—Heritage Speakers (HSs; bilinguals acquiring the minority language at home and then switching to the majority language in later childhood), and second language (L2) learners (bilinguals learning the second language through formal education)—when they read isolated Russian sentences. The two main questions we address are: 1) What scanpath reading processes do readers engage in when reading simple sentences in Russian? 2) How does group membership (i.e. monolingual adults, children, second language learners, or Heritage Speakers) determine which scanpath reading processes a reader engages in to decode written sentences efficiently?

We start with an overview of recent eye-tracking research that introduces the scanpath approach. Then we discuss what is already known about scanpaths in bilingual reading and about eye-movement characteristics in reading by monolingual adults, children, Heritage Speakers, and L2 learners of Russian. We conclude the introduction with our hypotheses and specific research questions as well as presenting some basic information concerning the Cyrillic writing system and the complexities it creates for bilingual speakers who learn to read in Russian.

Written language comprehension: A scanpath approach

Monolingual readers. Since the seminal work of Yarbus (1967) researchers have repeatedly used specific eye-movement sequences (i.e. scanpaths) to test written language comprehension. One instance is the interpretation of temporarily ambiguous sentences (Frazier & Rayner, 1982; Meseguer, Careirras, & Clifton, 2002; Mitchell, Shen, Green, & Hodgson, 2008). The focus of the analyses in these studies was on individual local regressive saccades (i.e. backward eye movements to revisit particular sentence regions) with the goal of investigating how participants recovered from the ambiguity once disambiguating information became available. This research, however, did not consider more complex and global-level spatio-temporal eye-movement patterns which limited the authors' ability to draw conclusions. To address the limitation of the early approaches von der Malsburg and colleagues (von der Malsburg & Vasishth, 2011, 2013; von der Malsburg, Kliegl, & Vasishth, 2015) devised and applied a new scanpath analysis technique that took into account the full spatial and temporal characteristics of the eye-movement sequences over the entire sentence in order to detect the scanpath reading patterns (called 'strategies' in these studies).

Von der Malsburg and Vasishth (2011) started by re-examining the data in Meseguer et al. (2002) and used a scanpath clustering procedure that revealed functionally different types of eye-movement sequences: *re-reading* of the whole sentence (this pattern had gone unnoticed in the more conventional analysis used in the original study by Meseguer et al.) and *checking* (rapid saccades from the end of the sentence to the disambiguating word). Von der Malsburg and Vasishth also reported that participants varied considerably in their choice of reanalysis strategies. They hypothesized that individual differences, such as working memory capacity, might play some role in determining the strategy preference (cf. 'Time Out hypothesis' in Mitchell et al., 2008). In a follow-up study with Spanish-speaking participants (von der Malsburg & Vasishth, 2013), the same type of analysis confirmed the previously found scanpath categories and revealed an additional pattern of rapid backward saccades to re-read individual words after the participants encountered the disambiguating region. This study also measured the working memory ability of the participants as a potential predictor of their reading strategy. Somewhat surprisingly, participants with high working memory capacity produced more regressive eye movements in response to disambiguation than did low-capacity readers, indicating that they had greater difficulty processing temporarily ambiguous sentences. The authors interpreted this in accordance with the good-enough parsing account (Ferreira et al., 2002)—unlike high memory-capacity readers, low-capacity participants did not immediately commit to one of the two available sentence interpretations, thus leaving the sentence interpretation initially underspecified. As a result, low-capacity readers did not have to reanalyze the sentence when they encountered the disambiguating word (see also Dirix et al., 2020; Hyönä, Lorch, & Kaakinen, 2002 for an investigation of the relationship between memory and reading patterns at text-level).

More recently, von der Malsburg et al. (2015) confirmed the sensitivity of the newly developed (2011) scanpath analysis to factors that are known to influence conventional eye-tracking measures, such as *word length* (Clifton et al., 2007), *syntactic processing difficulty* (Boston, Hale, Kliegl, Patil, & Vasishth, 2008; Boston, Hale, Vasishth, & Kliegl, 2011; Demberg & Keller, 2008), and the *reader's age* (Kliegl, Grabner, Rolfs, & Engbert, 2004; Whitford & Titone, 2017). Collectively, the results of these studies suggested that the scanpath approach might also be a useful method for investigating global reading processes in different populations of participants, including not only skilled monolingual adult readers but also developing readers, both monolingual and bilingual.

L2 learners. To the best of our knowledge, the only study that has examined bilingual reading using a scanpath approach is the recent study by Godfroid et al. (2015). Twenty L2 learners of English, and 20 native speakers, performed a grammaticality judgment task of 68 sentences in which half of the sentences contained violations in various grammatical structures. The task was administered under two conditions, with and without time limits. Using iterative visual inspection, Godfroid and colleagues identified three major reading patterns that their participants used to process the sentences, regardless of the group or grammaticality: (1) no regressions, (2) an incomplete reading of the sentence followed by a regression to the beginning of the sentence, and (3) re-reading of large portions of the sentence. Based on the finding that time pressure effects were observed only in the L2 group, Godfroid and colleagues suggested that untimed reading in bilingual participants might tap into their explicit knowledge (vs. implicit in timed tasks) and represent a controlled eye-movement behavior that participants employed to achieve the most efficient sentence comprehension.

Heritage Speakers. At present, we are not aware of any studies investigating scanpaths in heritage language reading. Heritage Speakers are quite different from typical L2 bilinguals (such as the L2 participants in Godfroid et al., 2015). The first difference concerns the timing and the environment of language acquisition—while L2 learners acquire the language in school and university settings, typically after puberty, Heritage Speakers learn the language at home from their caregivers in a similar way to how monolingual children do. In addition, the mode of acquisition also differs: Heritage Speakers mostly learn language in an auditory modality (listening, speaking), while L2 learners receive formal instruction both in auditory and visual modalities (reading and writing).

Despite Heritage Speakers' generally sound command of spoken language, they are rarely taught literacy at home and often do not read in their heritage language, especially in the case of different orthographies between the dominant (e.g., English) and heritage language (e.g., Russian). According to *the divergent attainment* trajectory of heritage language development (Benmamoun, Montrul, & Polinsky, 2013; Montrul, 2008; Scontras, Fuchs, & Polinsky, 2015; Polinsky & Scontras, 2020), the competence growth in heritage languages slows down and eventually stops after the switch to the dominant language. As the switch typically occurs when children start school and the input in the dominant language increases, the divergent attainment trajectory suggests that Heritage Speakers' language skills 'freeze' at the age of their entry to school. Thus, the main prediction is that the language abilities of adult Heritage Speakers should resemble those of school-age monolingual children (for review of reading in children, see Blythe & Joseph, 2011). With respect to literacy and reading skills specifically, the divergent attainment trajectory predicts more similarities in global reading patterns between adult Heritage Speakers and monolingual children than between Heritage Speakers and L2 learners.

Our recent study (Parshina, Laurinavichyute, & Sekerina, 2021) in which we investigated conventional eye-tracking measures of reading in Cyrillic by adult Heritage Speakers and L2 learners of Russian who live in the USA confirmed these predictions. We compared the eye movements of bilinguals with those observed in monolingual 8-year-old children (Korneev, Akhutina, & Matveeva, 2017) and adults (Laurinavichyute, Sekerina, Alexeeva, Bagdasaryan, & Kliegl, 2019). We found that Heritage Speakers exhibited quantitatively different eye-movement characteristics (i.e. longer mean fixation durations and lower probability of skipping words, but higher rates of regressions and multiple fixations on words). High-proficiency Heritage Speakers resembled monolingual children the most, while low-proficiency Heritage Speakers were at a disadvantage and read on a par with unbalanced L2 learners, suggesting that early exposure to heritage language, alone, did not necessarily facilitate literacy acquisition.

Present Study

The present study takes a step further from Parshina et al. (2021) by using the scanpath approach to investigate *qualitative* differences and similarities in global eye-movement patterns among the four groups of Russian speakers (as opposed to the quantitative evaluation of word-level eye-tracking measures in Parshina et al.). While these groups are the same as in Parshina et al., to be able to compare the qualitative characteristics of the global eye-movement patterns between groups it was critical that all participants read the same sentences (they had read a different set of sentences in the previous study). To that end, while the current analysis includes some reanalyzed data (see Method section for details), we also collected a large set of new data in order to include Heritage Speakers of various proficiency levels as well as monolingual adults in the study.

The study has three goals. The first one is to identify scanpath patterns reflecting the scanpath reading processes that are common among Russian readers regardless of the speaker's group membership (i.e. monolingual, child, L2 learner or Heritage Speaker groups). Accordingly, our second goal is to investigate whether group membership predicts if the reader will engage in a specific scanpath reading process. Based on our previous findings, monolingual adults are expected to engage in qualitatively different scanpath reading processes from the other three groups of readers. For Heritage Speakers, we expect that high-proficiency HS readers will exhibit scanpath reading processes that are similar to those of children, and that low-proficiency HS readers will engage in the same process as low-proficiency L2 learners. Our third goal is specific to bilingual readers: to uncover the effects of various demographic and reading performance factors (e.g., age of arrival, exposure to non-dominant language, comprehension abilities, reading fluency) on scanpath reading processes, as such factors had been identified as strong predictors of reading performance in previous research (for review, see Koda, 2005, 2007).

We capitalized on the fact that there are many young adults in New York (where we collected our data) who are either Heritage Speakers of Russian or are learning Russian as L2s, so we asked our participants to read simple, unambiguous Russian sentences of the type that are appropriate for monolingual children who are learning to read in Russian. Despite the simplicity of the materials in this study, the difference in scripts (Cyrillic vs. Latin) and the morphological principle of Russian orthography (explained below) present some additional challenges for literacy acquisition in Russian for bilingual readers, beyond the typical difficulties associated with learning a second language (i.e. vocabulary and knowledge of grammar).

Although the Russian orthography is characterized as *shallow* with almost one-to-one correspondence between graphemes and phonemes, there are some irregularities. For example, the vowel position in a word dictates its pronunciation, which is not directly reflected in the orthography (e.g., the phoneme /o/ has multiple allophones, e.g., [o] and [ə], depending on the stress on the word). Another example is consonant assimilation in which the pronunciation of a consonant depends on the position of the letter in the word (e.g., *lodka* is pronounced as [lótkə]), as well as the quality of the preceding or following sounds. Such discrepancies between pronunciation and spelling are due to *the morphological principle* of Russian orthography: The spelling of the morphemes stays invariant regardless of the phonological laws used in speech. This phenomenon requires readers to have both morphological awareness and knowledge of orthographic patterns. Typically, to avoid delays in literacy development, Russian children as early as those in the second grade receive instruction in the morphemic analyses of words (Kerek & Niemi, 2012) and master decoding skills by the fourth grade (Rakhlin, Kornilov, & Grigorenko, 2017). The situation, of course, looks quite different for adult Heritage Speakers who typically do not receive formal instruction in reading and writing in their heritage language and, as with L2 learners, struggle with the differences in scripts between their two languages.

Method

Participants

There were 120 participants distributed across four groups: 30 monolingual Russian-speaking adults (13 women, M_{Age} = 23.3, Range 19–28), 30 monolingual Russian-speaking children in the 2nd grade (11 girls, M_{Age} = 8.5, Range 8–9), 30 English-dominant Heritage Speakers of Russian (14 women, M_{Age} = 17.5, Range 13–24; $M_{Age\ of\ Arrival}$ = 4.3 years), and 30 English-dominant L2 learners of Russian (21 women, M_{Age} = 21.2, Range 16–43). The data for 17 Heritage Speakers and all the L2 learners (*n*=30) come from Parshina et al. (2021) and were reanalyzed for the present study, using the scanpath approach. The rest of the data were new, namely 30 monolingual speakers and 13 high-proficiency Heritage Speakers. We also collected new data for the children (*n*=30) in order to rule out general difficulties in reading and other cognitive processing abilities. To achieve this, we assessed the children's reading ability using the *Standardized Assessment of Reading Skills* (*SARS*; Kornev, 1997), as well their non-verbal fluid intelligence (*Raven's Colored Progressive Matrices* [Raven, 2004]) before the corpus-reading task.

Participants were recruited from three sites: an urban university in New York City (Heritage Speakers and L2 learners), and a public school and a university in Moscow, Russia (children and monolingual adults, respectively). None of the Heritage Speakers in this or the previous study (Parshina et al., 2021) had more than 4 weeks of formal instruction in Russian. All the monolingual adult participants were skilled readers of Russian (i.e. university undergraduates), reported Russian as their native language, and did not identify as speaking any other language fluently. Before the start of the study, all the participants (over 18 years old) and the parents of the children signed their informed consent or an assent form (minor participants under 18 years old) and filled out a language background questionnaire, administered in English or Russian (see Table 1 for the bilingual participant characteristics).

<Insert Table 1 about here>

Design and materials

Reading assessments. Heritage Speakers often vary considerably in their heritage language reading ability due to different extents of reading exposure. As noted above, it is necessary to consider the proficiency level of the bilingual speakers when conducting any kind of analysis as it is a highly influential factor. Operationally, we defined proficiency in bilingual reading in Russian as a set of scores in the *Russian Oral Reading Fluency* (*ORF-Rus*) test. In this task, participants read out loud the short text *Kak ja lovil rakov* "How I was Catching Crayfish" (202 words; Kornev, 1997). The text is intended for monolingual Russian primary school students and measures the speed and quality of reading, as well as comprehension in Russian (not included in the final score calculation). The sentences in the text include a wide range of grammatical constructions in Russian (e.g., relative clauses, passives, null object), tenses, different types of word order (SVO, VSO, OVS) and contain words of different frequencies.

Next, for the bilingual participants we administered a parallel standardized English task, *English Oral Reading Fluency* (*ORF-Eng,* Woodcock WRMT 3rd edition, 2011), to rule out general reading difficulties in the dominant language. The participants were asked to read out loud a short text (217 words) in English in which sentences gradually increase in complexity.

The performance on both the *Russian* and *English Oral Reading Fluency* tests was scored on the basis of the formula for calculating the final score in the *English Oral Reading Fluency* test provided in the Woodcock Reading Mastery Tests manual. This formula includes such factors as total reading time, number of words in the passage, and number of errors (omissions, mispronunciations, word substitutions, hesitations, repetitions, and transpositions). Table 1 provides the oral fluency scores both in Russian and English.

All monolingual adult participants were classified as fluent in reading in Russian, based on the following data: 1) their performance in the *Russian Oral Reading Fluency* test was at ceiling for all participants; 2) Russian was the native language for all the participants and none of them were bilingual according to the language background questionnaire.

Data for scanpath analysis. The materials were 30 sentences from the children's version of the *Russian Sentence Corpus* (Korneev et al., 2017), appropriate for 8-year-old monolingual Russian-speaking children, as illustrated in Examples (1)–(2).

(1) Stimulus item	В магазине	Андрей	купил	молоко,	сметану,	творог.
Transliteration	v magazine	Andrey	kupil	moloko	smetanu	tvorog
Gloss	in store	Andrey	bought	milk	sour cream	cottage cheese
Translation	'In the store Andrey bought milk, sour cream, cottage cheese.'					

(2) Stimulus item	Недалеко	был сложен	стог сена,	рядом	стояли	грабли.

Transliteration	nedaleko	byl slozhen	stog sena	ryadom	stoyali	grabli
Gloss	nearby	was stacked	haystack	next to	stood	rake
Translation	'A haystack was stacked nearby, a rake was next to it.'					

The sentences were presented to our participants in isolation for silent reading. They represented diverse types of grammatical structures typical of the Russian language (canonical and non-canonical word orders, passive and active voice constructions, sentences with null subjects, relative clauses, etc.). All words in the text corpus were annotated for length and frequency (Lyashevskaya & Sharov, 2009). Table 2 presents the descriptive characteristics of all the corpus words and sentences.

<Insert Table 2 about here>

Procedure

All sentences were presented in Ubuntu Mono Normal black font, size 22 pt, on a light gray background programmed in *Experiment Builder* (SR Research Ltd.). We used a BenQ XL2411Z 144Hz monitor (resolution: 1920 x 1080 pix) controlled by a ThinkStation computer. Eye movements were recorded using an EyeLink 1000+ desktop mount eye-tracker with a chin rest. The right eye was tracked, at 1000 Hz rate.

The experiment started with a 9-point calibration procedure repeated after every 15 sentences. Stimuli appeared on the screen in randomized order. Each sentence was followed by a multiple-choice question to ensure the participants were paying attention and reading for comprehension. For example, the sentence in (1) was followed by the question *“What did Andrey buy in the store?”,* with *“bread”* and *“milk”* as possible answers.

Sentence presentation was as follows. First, we performed drift correction where the participants fixated a black dot on the left edge of the screen. The dot disappeared after it had been fixated for 500 ms, followed by the presentation of the sentence in which the first letter of the first word appeared in the position of the black dot. After the participants finished reading the sentence, they looked at a red dot in the lower right-hand corner of the screen. After 500 ms, the multiple-choice question appeared. As soon as the participants clicked on one of the options, the presentation of the next sentence began with the drift correction. Overall, participants took approximately 30 minutes to complete the task, ranging from 10 minutes for monolingual adult and child participants, to 40 minutes for low-proficiency bilinguals.

Data analysis

The first stage of our analysis served to identify the scanpath reading processes occurring in our data set. It consisted of the following steps:

1. Plotting the scanpaths for each sentence and participant for visual inspection;
2. Calculating the pair-wise scanpath dissimilarity scores for each sentence;
3. Fitting a map of the scanpaths for each sentence (multi-dimensional scaling);
4. Conducting cluster analysis for each sentence using the map as the input;
5. Identifying a prototypical reading process of each detected cluster.

In the second stage, we investigated the factors that predict which of the scanpath reading processes a participant adopted for a given sentence. The materials and the script used for the analyses reported below, as well as all supplementary materials, are available at the Open Science Framework project page https://osf.io/9z7yv/.

The overall comprehension accuracy was high, i.e. 91% (*SD* =28). Average accuracy in answering the comprehension questions was almost at ceiling for three out of the four groups: 95% for the monolingual adults (*SD* =21), 99% for the monolingual children (*SD* =10), and 91% for the Heritage Speakers (*SD* =29). The L2 learners' accuracy was the lowest, i.e. 85% (*SD* =36). Only sentences with correct answers to the comprehension questions were included in the scanpath analysis.

1. Plotting scanpaths for each sentence and participant. First, to get a general idea of how the participants read the sentences, we created plots of the scanpaths for every participant and every sentence by using the x-coordinate of each fixation and its time within the trial (the fixations on y-coordinates remained largely constant since the sentences were presented without line breaks). Figure 1 shows the scanpaths recorded for Example (1), which was the sentence that elicited the least diverse scanpaths (participants generally read it in a similar way, according to the scanpath similarity measure by von der Malsburg & Vasishth, 2011). Each plot shows the scanpath for one participant (with the participant number at the top, coded by color: *ML1-30* – monolingual, *CH1-30* – child, *HS1-30* – Heritage Speaker, *L21-30* – L2 learner). The x-axis shows the words and the y-axis indicates the trial time in seconds (see file S1 for the scanpath plots for all the sentences in the Supplementary Materials).

<Insert Figure 1 about here>

Visual inspection of Figure 1 reveals that some participants read slower (i.e. the scanpath line is extended vertically e.g., CH26, HS4), made more regressions (i.e. the scanpath line returns to the previous words in the sentence, e.g., HS24, HS29), or skipped words more often (i.e. scanpath line is flat across three or more words, e.g., ML18, ML28) than others. Figure 2 shows the scanpaths for Example (2), which elicited the most variable gaze patterns. In Figure 2, we can see that the scanpaths vary in their characteristics from participant to participant: some participants skip a lot (ML26), or read slowly but without skipping or long regressions (CH18); some read very slowly and they produced many regressions (HS8, ML17); others re-read longer passages (CH30, L24).

<Insert Figure 2 about here>

2. Calculating the scanpath dissimilarity scores. Next*,* for each sentence, we calculated the pair-wise dissimilarities of all the scanpaths. This measure calculates the difference between two scanpaths as a function of the spatio-temporal differences between their matched (i.e. sequentially aligned) fixations, where the x- and y-coordinates and durations of the fixations are represented as continuous variables (see von der Malsburg & Vasishth, 2011 for a detailed description of the dissimilarity measure). For example, if two matched fixations have the same x- and y-coordinates, then their dissimilarity is equal to the difference in their durations. If, on the other hand, two matched fixations are far away from each other, the difference between them is the sum of their durations. The rationale for that is the following: If these two fixations are long, it means that the spatial disparity between them lasted longer (readers looked at different things longer) and, therefore, these two fixations add more dissimilarity than two fixations with short durations. When the distance between two matched fixations is in the medium range (not too far nor too close), the dissimilarity score is a weighted sum of the difference and the sum of the fixation durations. We calculated the scanpath dissimilarity scores using the software package *scanpath* for R (von der Malsburg et al., 2015). Having obtained the dissimilarity scores for each pair of scanpaths recorded for a sentence, we then calculated the average dissimilarity among the scanpaths for each sentence and determined which sentences had elicited the most and which the least diverse scanpaths in the corpus.

This difference in scanpath variability between Examples (1) and (2) is to be expected in light of their syntactic and semantic characteristics. Example (1) (Figure 1) has a canonical SVO word order and includes high-frequency words (M_{freq} = 4250 ipm). On the other hand, Example (2) (Figure 2) has a non-canonical OVS word order, passive voice, and words of lower frequency (M_{freq} = 407 ipm). Therefore, in contrast to Example (1), the participants fixated the words more (CH18, ML17) and re-read them (L24) as well as the entire sentence (CH30, HS8, L24) multiple times in Example (2).

3. Fitting maps of scanpaths. Next, to visualize the variability of the scanpaths across all the corpus sentences, multi-dimensional scaling (Kruskal, 1964) was used to calculate a map of the scanpaths for every sentence (Figure 3). On these maps, every scanpath (one from each participant) is represented as a point (i.e. triangles – monolinguals, circles – children, squares – Heritage Speakers, diamonds – L2 learners). Similar scanpaths are located closer to each other, i.e. the participants read in the same way (Figure 3A). The further away the scanpath is from the gravity center of the map, the more unusual or irregular was the reading pattern compared to those of the other participants (Figure 3B). More difficult sentences typically elicit more variable, and hence irregular, scanpaths (cf. Figure 3A for Example (1) – Figure 3B for Example (2), see also von der Malsburg et al., 2015).

<Insert Figure 3 about here>

The goodness of fit of a map depends on the number of its dimensions (for more detailed specification of the procedure for fitting maps, see von der Malsburg & Vasishth, 2011, 2013, von der Malsburg et al. 2015). A large number of dimensions results in a map with more degrees of freedom and a smaller percentage of variability that cannot be explained by the map. Therefore, high-dimensional maps more faithfully represent the similarities among the scanpaths, however, they also risk overfitting the data. The number of dimensions of our maps was therefore set to 6, such that the percentage of unexplained variance by the maps was, on average, 11% (*SD* =1.17) which indicates a reasonably good fit (Kruskal, 1964). Maps of scanpaths were fitted using the *isoMDS* function in the R package MASS (Venables & Ripley, 2002). [See Figure S2 for the scanpath maps for all the sentences in the Supplementary Materials.]

The map dimensions can correspond to interpretable scanpath features. For instance, examining Figure 2 suggests that dimension 1 in Figure 3B corresponds to the variability in reading speed, while dimension 2 represents the amount of re-reading. For other sentences, the dimensions can capture different aspects of the scanpath variance. Thus, hypotheses tests are needed to investigate which features dominate the structure of a given scanpath space.

4. Cluster analyses for each sentence. The goal of the cluster analysis was to identify the categories of scanpaths that represent qualitatively different scanpath reading processes (if any). The clusters for each sentence were identified by applying Gaussian mixture modeling (*Mclust* package, Fraley & Raftery, 2007), where all the parameters of the clusters (e.g., position, variance and rotation) were allowed to vary freely. The benefit of using Gaussian mixture modeling is the ability of the procedure to detect clusters even if they overlap (vs. *k*-means clustering) based on the distributional properties of the data.

Models were fitted for numbers of Gaussians fixed at 3 to avoid overfitting the data (i.e. capturing random variation in reading patterns) and to prevent clusters that capture the tails of slightly non-Gaussian distributions [See Figure S3 for maps of the scanpath clusters for all the sentences in the Supplementary Materials].

We can overlay the derived clusters directly onto the participants' scanpaths to obtain a general idea of which features are characteristic of a cluster (i.e. reading times, regressions, skipping, number of reading passes, etc.). Figure 4 presents the scanpaths for Example (2), color-coded by cluster. Cluster 1 (magenta) includes scanpaths that are close to the *'baseline'* as characterized by regular left-to-right reading, short fixations, frequent word skipping, and few regressions. Cluster 2 (brown) is also characterized by regular left-to-right reading, but has increased fixation durations, backward saccades to re-read individual words, some sentence re-reading, and instances of word skipping. Finally, Cluster 3 (blue) is characterized by considerably increased reading times, virtually no word skipping, and frequent re-reading of the sentence, occasionally multiple times.

<Insert Figure 4 about here>

5. Identifying prototypical scanpath reading processes. Finally, we identified the '*prototypical'* scanpath reading process for each cluster represented by the scanpath that is the closest to each cluster's center of gravity. In other words, in each sentence and each cluster within that sentence, we identified the scanpath with the shortest distance to the center of the respective cluster (using the map of clusters for distance calculations, see Figure S3 in the Supplementary Materials). As a result, we were able to identify three prototypical scanpath reading processes adopted by Russian readers: a) a *fluent* scanpath reading process (Cluster 1); b) an *intermediate* scanpath reading process (Cluster 2); and c) a *beginner* scanpath reading process (Cluster 3). Figure 5 presents these three scanpath reading processes for Example (2).

[See File S4 for the prototypical scanpath reading processes for all the sentences in the Supplementary Materials.] Note that, at this point, these processes are descriptive in nature. As can be seen in Figure 3, there is a continuum of scanpaths and the clusters identified in the cluster analysis serve to concisely characterize and summarize that continuum. Across all 30 sentences in the corpus, our participants were classified as follows: 27.2% (*SD = 4.8*) exhibited the fluent scanpath reading process, 49.3% (*SD = 6.1*) preferred the intermediate scanpath reading process, and 23.5% (*SD = 7.2*) followed predominantly the beginner scanpath reading process.

<Insert Figure 5 about here>

Results

Recall that our study had three goals: (1) to identify scanpath patterns reflecting scanpath reading processes that are common among Russian readers regardless of the speaker's group membership; (2) to investigate whether group membership predicts engagement in a specific scanpath reading process; and (3) specific to bilingual readers: to uncover the effects of various demographic and reading performance factors on scanpath reading processes.

Common scanpath readings processes

All data analyses were performed in R (version 3.5.1; R Core Team, 2018). For the (generalized) linear mixed effects models, we used *lme4* (1.1-13) and *sjPlot* package 2.8.3 (for data visualization and the computation of *p*-values; Lüdecke, 2017). Throughout the analysis, all (G)LMMs included random intercepts for sentences and readers.

Table 3 (top panel) presents the means and standard deviations for some canonical eye-movement measures for each of the three scanpath reading processes. All *p*-values for differences between scanpath process (as assessed using series of linear mixed-effect models that included process as a fixed predictor and eye-movement measure as an outcome) were less than .001 (see Table A1 for estimates and corresponding *p*-values in the Appendix). Table 3 (bottom panel) presents the percentage distribution of the 120 participants for each of the scanpath reading processes across all 30 sentences (i.e. the distribution of groups in each scanpath reading process, together, comprise 100%).

<Insert Table 3 about here>

Although all three scanpath reading processes were found in all four groups of Russian readers, Table 3 (bottom panel) reveals that for each process there was a group that used this process most frequently. Figure 6 shows the counts of instances of scanpath reading processes that each reader contributed (the first 2 rows are children, the next 2 rows are Heritage Speakers, followed by L2 learners and, in the last 2 rows, monolingual adults). For every scanpath reading process, there were readers that showed a strong preference for it. Some readers did not use the fluent or the beginner processes at all, but the intermediate scanpath reading process was found in most participants with the exception of 5 of the monolingual readers.

<Insert Figure 6 about here>

Group preferences for scanpath reading processes

After we had determined the three scanpath reading processes that are common in all four groups, we investigated the question of which group (i.e. monolingual adults, children, Heritage Speakers, and L2 learners) is characterized by which preferred process. We suggested that while monolinguals would mostly adopt the fluent scanpath reading process, the children's and

Heritage Speakers' scanpath reading processes would overlap. We also predicted that proficiency would have an effect, in that low-proficiency Heritage Speakers would be clustered together with the L2 learners.

To test these predictions, we ran a binomial mixed-effect model for each of the three scanpath reading processes ('1' = scanpath of a participant belongs to the process, '0' = scanpath does not belong), and each group where group membership was dummy-coded as a binary variable ('1' = participant is a group member; '0' = participant is not a group member). Thus, the model estimates the probability of a participant in a specific group to exhibit a particular scanpath reading process when compared to all other readers not belonging to the same group. The effect sizes (Cohen's *d*) were calculated using the *lme.dscore* function from the *EMAtools* package for R (Kleiman, 2017; R code example in Appendix B).

The results of the generalized linear modeling are presented in Table 4 ($N_{participants}$ = 120, $N_{sentences}$ = 30, observations: 3477). They indicate that, in comparison to other readers, monolingual speakers exhibited a high probability of adopting the fluent scanpath reading process. Accordingly, it was highly unlikely that monolinguals would rely on the intermediate or beginner processes. Monolingual children read sentences by following the intermediate process more often than other readers. Heritage Speakers did not show a strong preference for either the intermediate or beginner scanpath reading process, whereas, out of all the groups, L2 learners had a higher probability of engaging in the beginner scanpath reading process than did the other participants, although they also produced many scanpaths in the intermediate process category.

<Insert Table 4 about here>

Demographic and reading performance factors in bilingual readers

For each of the two groups of bilingual readers, we tested the impact of two demographic factors on their reading processes, i.e. age of arrival in the United States and daily exposure to Russian, in addition to two reading performance factors, i.e. self-estimated comprehension ability in Russian and proficiency in reading (as defined by the ORF-Rus test), along with their scores for the Oral Reading Fluency assessment in English. Two predictors in the model were statistically significant (see Tables A2 and A3 in the Appendix for full statistical analysis and Appendix C for the R code example).

First, what matters for Heritage Speakers ($N_{participants}$ = 30, $N_{sentences}$ = 30, observations: 896) in engaging in one of the three processes was their proficiency level in reading in Russian (ORF-Rus). Higher proficiency led to a reliance on fluent reading ($\beta = 2.6$, $SE = .57$, $d = 1.6$, $p < .001$), whereas lower proficiency resulted in a beginner scanpath reading process ($\beta = -1.5$, $SE = .41$, $d = -1.2$, $p < .001$). For L2 learners ($N_{participants}$ = 30, $N_{sentences}$ = 30, observations: 814), the probability of engaging in the fluent scanpath reading process increased with higher self-estimated comprehension scores ($\beta = 2.0$, $SE = .73$, $d = .71$, $p = .018$).

Discussion

In this study, we applied a scanpath approach to establish and define global eye-movement patterns that comprise the scanpath reading processes in bilingual and monolingual speakers of Russian. The cluster analysis that grouped similar gaze trajectories in the reading of 30 sentences allowed us to identify three such processes based on their distinct scanpath patterns: 1) *fluent*; 2) *intermediate*; and 3) *beginner* (see Table 5 for the conceptual comparison of the eye-movement characteristics of these processes). We also found that the group membership of our participants (monolingual adults, children, Heritage Speakers, and L2 learners) strongly correlated with the clustering of their gaze trajectories into one or another common scanpath reading process. Finally, we established that, out of the demographic and reading performance factors that represented individual differences in our bilingual readers, only the proficiency, for Heritage Speakers, and the comprehension scores, for L2 learners, affected which scanpath reading process they adopted.

<Insert Table 5 about here>

The *fluent* scanpath reading process is characterized by straight left-to-right reading that includes short fixation durations, a high word skipping probability, and an absence of long regressions and sentence re-readings. We suggest that this process is a characteristic of participants who generally do not experience any difficulties in lexical access or morphosyntactic processing in reading (e.g., monolingual adults) while comprehending simple Russian sentences.

The primary characteristics of the second, *intermediate* scanpath reading process are fixation durations that are twice as long, higher rates of producing short leftwards saccades, a lower probability of word skipping, and an absence of sentence re-readings. Previous studies suggest that short regressions to the beginning of the current or previous word (i.e. word re-reading) can be the result of the reader's need to perform a local 'targeted repair', namely, to come back to the area where the processing difficulty occurred (e.g., Frazier & Rayner, 1982; Meseguer et al., 2002). While our sentences did not include experimentally created ambiguities, we propose that short leftward regressions serve the same function, as they help resolve local and lower-level processing difficulties such as, for example, word recognition failure (Bicknell & Levy, 2011). The natural tendency of readers to avoid such failures also triggers a 'careful' reading pattern, which is characterized by the absence of skipping and slower total reading times.

This reading behavior might be optimal for readers who have insufficient exposure to the language (e.g., L2 learners: Gollan, Montoya, Cera, & Sandoval, 2008; Schmidtke & Moro, in press; Whitford & Titone, 2012, 2016) or are unfamiliar with reading materials (e.g., children or adult poor-readers: Barnes & Kim, 2016; Kuperman & Van Dyke, 2011).

Finally, the third *beginner* scanpath reading process is characterized by the longest fixation durations, almost no word skipping, very long total reading times, and re-reading of the whole sentence, sometimes multiple times. We suggest that this scanpath reading process characterizes readers who not only experience delays in visual word recognition, but also struggle with global challenges in semantic and morphosyntactic information integration. These readers often re-parse the sentence from scratch after the first-pass reading. Following von der Malsburg and Vasishth (2011, 2013) as well as the good-enough parsing account (Ferreira et al., 2002), it is possible that these readers have difficulties with incremental interpretation of the whole sentence during the first-pass reading. Instead, they assemble lexicosemantic information from individual words in a piecemeal fashion and only integrate it with the syntactic structure during the second or even third re-readings (evident through word skipping and faster reading times in the second and subsequent re-readings).

The beginner scanpath reading process might be the most viable way to allocate limited cognitive resources (attention, working memory, and the decoding of visual information) and reduce cognitive load during written language comprehension which is, undoubtedly, a challenging task for these readers (for review on cognitive automaticity in L2 language processing see Segalowitz & Hulstijn, 2009). We suggest that the beginner scanpath reading process is the preferred one for those readers who are either at the very first stages of literacy acquisition (e.g., pre-school children) or have just started to acquire a new (second) language, especially if that language is dissimilar from the dominant language (e.g., in phonology, grammar, or orthography).

In general, our results confirm the predictions of the divergent attainment trajectory of heritage language development (Benmamoun et al., 2013; Montrul, 2008, Scontras et al., 2015; Polinsky & Scontras, 2020) as well as findings from previous reading studies (Cop et al., 2015; Parshina et al., 2021). The scanpaths of monolingual speakers were consistent with the fluent scanpath reading process, whereas those of Heritage Speakers depended on their proficiency. The higher the proficiency, the higher the probability that their gaze trajectories were more 'advanced', i.e. similar to those of monolingual children, who showed a stronger tendency towards the intermediate scanpath reading process compared to other readers (or in some cases, to those of monolingual adults who read following the fluent scanpath reading process). Low-proficiency Heritage Speakers, on the other hand, read on a par with L2 learners of Russian, who engage in the beginner scanpath reading process more often than any other group.

The absence of the effects of other demographic factors in Heritage Speakers, besides proficiency (i.e. age of arrival and daily exposure to Russian), on their ability to move from the beginner to the intermediate scanpath reading process suggests that, despite their early exposure to their heritage language in its spoken modality, their reading skills (as opposed to auditory comprehension, production, phonology, grammar, and vocabulary knowledge) do not seem to benefit from acquiring their heritage language in the family, at least not to the extent that it could be detected in the present study. These readers exhibit reading behavior typical of unbalanced L2 learners who started to learn the second language later in adulthood and in a classroom setting. This conclusion supports the previously reported results of there being no advantage for Heritage Speakers in literacy acquisition (Ke, 1998; Xiao, 2006; Zhang & Koda, 2018).

Not surprisingly, our L2 participants were different from the other three groups of readers, in that they showed the strongest reliance of any group on the beginner scanpath reading process. Very few of them produced scanpaths (3%) that follow the fluent scanpath reading process, and the ability to engage in it was predicted by their self-estimated comprehension scores. We hypothesize that L2 learners' engagement in the beginner scanpath reading process is due to the difficulties they experience with the grapheme-to-phoneme decoding process while reading in Russian (Comer & Murphy-Lee, 2004; Comer, 2012). The time that L2 learners spend on this process (which is automatized in more proficient readers) delays visual word recognition and makes information integration from the entire sentence challenging (Gor, 2017). As a result, L2 learners have to use a global-level remedy for comprehension difficulties, i.e. they re-read the sentence multiple times. It is also likely that the cognitive resources allocated to reading in L2 are limited by working memory. Thus, we hypothesize that the most distinct characteristic of the beginner scanpath reading process, namely, the re-reading of simple sentences, is a way to re-allocate the resources and give the parser a "fresh start" after all initial difficulties have been resolved during the first-pass reading.

In contrast to Heritage Speakers, it was not the proficiency, but their self-estimated comprehension ability that was a significant predictor of reading fluency for L2 learners. Specifically, higher comprehension scores predicted the ability of L2 learners to engage in the fluent process. We hypothesize that this finding might be explained by the good-enough parsing account (Ferreira et al., 2002) if we assume that comprehension scores reflect the participants' estimation of their vocabulary size in Russian. The vocabulary knowledge gives these readers an advantage in reading simple child-friendly sentences as the absence of ambiguities and straightforward syntax do not require a potential reanalysis of the sentence. The parser simply 'scans' the words, extracts their meanings, and comes back to interpretation later, resulting in occasional instances of the fluent scanpath reading process while avoiding in-depth processing in reading.

To summarize, the scanpath approach that we adopted in this article draws a general picture of bilingual reading, wherein the difficulties that bilinguals experience both with visual word recognition and morphosyntactic and semantic information integration can be explicitly uncovered and visualized for professionals who work with bilingual speakers. Our findings provide additional support for the theories of bilingual word recognition (*BIA+,* Dijkstra & van Heuven, 2002; *Multilink,* Dijkstra et al., 2019): the language exposure (and, therefore, the proficiency of the readers and the subjective frequency of the words in the language) is critical for the efficiency of lexical access in bilingualism.

Furthermore, our scanpath analysis revealed that bilingual speakers vary in the way they engage in different types of scanpath reading processes. For some, returning and re-reading words, phrases or clauses is the only way to build a whole-sentence representation. Such a process is more in line with retrieval interference theories in which memory-based retrieval operations and individual differences are responsible for difficulties in bilingual language comprehension (Cunnings, 2017; Van Dyke et al., 2014). For others, faster reading without re-reading is the most efficient approach; it is consistent with the theories that place lexical retrieval, semantics and heuristics as the main tools bilinguals use to scaffold sentence meaning (Shallow Structure Hypothesis, Clahsen & Felser, 2006; good-enough parsing account, Ferreira et al., 2002). Crucially, we saw that the choice of the scanpath reading process is not static, as bilingual readers occasionally switched from one process to another. Thus, successful language comprehension can be achieved in multiple ways by an individual, reflecting the importance of individual differences in literacy acquisition in bilingualism.

Limitations and future directions

There are some limitations of the study that should be addressed in future research. First and foremost, the three scanpath reading processes that we identified through cluster analysis are largely descriptive in nature. While they represent concise summaries of the variability in gaze trajectories among different groups of Russian readers, more research is needed to clearly delineate the underlying mechanisms that drive written language comprehension. The fact that many readers in our study produced scanpaths that reflected multiple scanpath reading processes (e.g., the fluent scanpath reading process in Example (1) but the intermediate scanpath reading process in Example (2)), and that there is considerable individual and group-level variation, suggests that scanpath patterns may also be on a continuum.

The scanpath approach used in this investigation is one way to characterize this continuum. One question remains: which underlying factors facilitate the transition of readers, be they monolingual or bilingual, adult learners or children, along the continuum, from the beginner to an intermediate to a fluent scanpath reading process? Specifically, in future research, we should focus on a) linguistic properties, i.e. the lexical and structural effects of the materials on the scanpath reading process (e.g., the surprisal cost of a word in the sentence, word predictability, and sentence complexity) and b) the extra-linguistic properties, i.e. the effect of individual differences between participants that have previously been reported to impact the scanpath reading processes (e.g., working memory capacity, interference factors). Also related to the extra-linguistic properties, future research should additionally consider the possibility of some variability in scanpath reading processes stemming from the cultural and/or instructional differences between the countries of our participants (i.e. the USA and Russia): Will the international context or instructional method of literacy acquisition affect the reading processes?

A second important question for future research concerns the universal nature of scanpath reading processes and how they interact with the orthography for bilingual readers. In our study, bilingual readers read in the Cyrillic alphabet, which is different from the Roman-based alphabet of the dominant English. When Heritage Speakers read in their weaker language that still shares the same script (e.g., Spanish, Italian, French) as the dominant one, would they benefit from the script similarities and move along the reading continuum faster or even skip the beginner process completely? We speculate that while the quality of a given scanpath reading process(es) might stay the same (i.e. *fluent, intermediate* or *beginner*), the number of challenges in grapheme-phoneme conversion will be reduced, thus lowering the cognitive load associated with difficulties in reading in a different script. To speculate even further, the possible differences in the choice of scanpath reading processes might be affected by the typological language proximity in general: The more similar languages are (i.e. the same script, many cognates and syntactic similarity), the easier is the transition to a fluent reading process.

Another limitation of our investigation was the limited number of bilingual participants (60), which might have made it challenging to capture the effects of demographic factors. Thus, future studies with a higher number of Heritage Speakers or L2 participants are needed; for now, the necessity to bring participants physically to the laboratory that houses the eye-tracking facilities remains a barrier to massive online collection of behavioral data that has recently become popular in psycholinguistics. Including a thorough examination of the impact that socio-linguistic and demographic factors have on the development of scanpath reading processes is, however, a promising direction for investigating the similarities and differences between children and bilingual readers.

Finally, we deliberately stayed away from exploring the relationship between scanpath reading processes and comprehension accuracy within and between our groups of readers. Our choice of materials, namely, simple and unambiguous sentences that are appropriate for monolingual children presupposed high accuracy in answering comprehension questions; indeed, accuracy was at ceiling for three of the four groups. Future studies should establish, however, how the ability to engage in a particular scanpath reading process affects the accuracy of comprehension of more complex sentences, or sentences with experimental manipulation; and whether the scanpaths in sentences that result in inaccurate comprehension are qualitatively different from the scanpaths in sentences accurately comprehended. If such a relationship exists (e.g., a higher percentage of word skipping leads to lower accuracy in comprehension in bilingual readers), the results could contribute to the development of targeted literacy instruction in the non-dominant language. We believe that exploring these questions in various populations and in different languages and scripts will provide fruitful lines of future investigation to advance the theories of psycholinguistics and bilingualism.

Acknowledgments

The contribution of Olga Parshina and Irina A. Sekerina was supported by a PSC-CUNY Award # 60616-00-48. The contribution of Anastasiya Lopukhina has been funded by the Center for Language and Brain at HSE University, RF Government Grant, Ag. No. 14.641.31.0004.

References

Barnes, A.E., & Kim, Y.S. (2016). Low-skilled adult readers look like typically developing child readers: A comparison of reading skills and eye movement behavior. *Reading and Writing*, *29,* 1889–1914. https://doi.org/10.1007/s11145-016-9657-5

Benmamoun, E., Montrul, S., & Polinsky, M. (2013). Heritage languages and their speakers: Opportunities and challenges for linguistics. *Theoretical Linguistics*, *39*(3-4), 129–181. https://doi-org.proxy-ub.rug.nl/10.1515/tl-2013-0009

Bicknell, K., & Levy, R. (2011). Why readers regress to previous words: A statistical analysis. In L. Carlson, C. Hölscher, & T.F. Shipley (Eds.), *Proceedings of the 33rd Annual Meeting of the Cognitive Science Society* (pp. 931–936). Boston, MA: Cognitive Science Society.

Blythe, H.I., & Joseph, H.S.S.L. (2011). Children's eye movements during reading. In S.P. Liversedge, I.D. Gilchrist & S. Everling (Eds.), *The Oxford Handbook of Eye Movements* (pp. 643–662). New York, NY: Oxford University Press. doi: 10.1093/oxfordhb/9780199539789.013.0036

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. https://doi.org/10.16910/jemr.2.1.1

Boston, M.F., Hale, J.T., Vasishth, S., & Kliegl, R. (2011). Parallel processing and sentence comprehension difficulty. *Language and Cognitive Processes*, *26*(3), 301–349. https://doi.org/10.1080/01690965.2010.492228

Clahsen, H., & Felser, C. (2006). Grammatical processing in language learners. *Applied Psycholinguistics*, *27*(1), 3–42.

Clifton, C., Staub, A., & Rayner, K. (2007). Eye movements in reading words and sentences. In R.V. Gompel (Ed.), *Eye Movements: A Window on Mind and Brain* (pp. 341–374). Amsterdam, the Netherlands: Elsevier Science Ltd.

Comer, W.J. (2012). Lexical inferencing in reading L2 Russian. *Reading in a Foreign Language*, *24*(2), 209–230. http://hdl.handle.net/10125/66859

Comer, W.J., & Murphy-Lee, M. (2004). Sound correspondence acquisition in first semester Russian. *Canadian Slavonic Papers*, *46*, 23–35. https://doi.org/10.1080/00085006.2004.11092344

Cop, U., Dirix, N., Drieghe, D., & Duyck, W. (2017). Presenting GECO: An eyetracking corpus of monolingual and bilingual sentence reading. *Behavior Research Methods, 49,* 602–615. https://doi.org/10.3758/s13428-016-0734-0

Cop, U., Drieghe, D., & Duyck, W. (2015). Eye movement patterns in natural reading: A comparison of monolingual and bilingual reading of a novel. *PLoS ONE*, *10*, e0134008, https://doi.org/10.1371/journal.pone.0134008.

Cunnings, I. (2017). Parsing and working memory in bilingual sentence processing. *Bilingualism: Language and Cognition*, *20*(4), 659–678. doi:10.1017/S1366728916000675

Demberg, V., & Keller, F. (2008). Data from eye-tracking corpora as evidence for theories of syntactic processing complexity. *Cognition, 109*(2), 193–210. https://doi.org/10.1016/j.cognition.2008.07.008

Dijkstra, T., & Van Heuven, W.J. (2002). The architecture of the bilingual word recognition system: From identification to decision. *Bilingualism: Language and Cognition*, *5*(3), 175–197. https://doi.org/10.1017/S1366728902003012

Dijkstra, T., Wahl, A., Buytenhuijs, F., Van Halem, N., Al-Jibouri, Z., De Korte, M., & Rekké, S. (2019). Multilink: a computational model for bilingual word recognition and word translation. *Bilingualism: Language and Cognition*, *22*(4), 657–679. doi:10.1017/S1366728918000287

Dirix, N., Vander Beken, H., De Bruyne, E., Brysbaert, M., & Duyck, W. (2020). Reading text when studying in a second language: An eye-tracking study. *Reading Research Quarterly*, *55*(3), 371–397. https://doi.org/10.1002/rrq.277

Ferreira, F., Bailey, K.G., & Ferraro, V. (2002). Good-enough representations in language comprehension. *Current Directions in Psychological Science, 11*(1), 11–15. https://doi.org/10.1111/1467-8721.00158

Fraley, C., & Raftery, A.E. (2007). *MCLUST version 3 for R: Normal mixture modeling and model-based clustering* (tech. rep. No. 504). Department of Statistics, University of Washington, Seattle, WA.

Frazier, L., & Rayner, K. (1982). Making and correcting errors during sentence comprehension: Eye movements in the analysis of structurally ambiguous sentences. *Cognitive Psychology, 14*(2), 178–210.

Godfroid, A., Loewen, S., Jung, S., Park, J.H., Gass, S., & Ellis, R. (2015). Timed and untimed grammaticality judgments measure distinct types of knowledge: Evidence from eye-movement patterns. *Studies in Second Language Acquisition*, *37*(2), 269–297. https://doi.org/10.1017/S0272263114000850

Gollan, T.H., Montoya, R.I., Cera, C., & Sandoval, T.C. (2008). More use almost always means a smaller frequency effect: Aging, bilingualism, and the weaker links hypothesis. *Journal of Memory and Language, 58*(3), 787–814. https://doi.org/10.1016/j.jml.2007.07.001

Goodman, K.S. (Ed.). (1979). *Miscue analysis: Applications to reading instruction.* Urbana, IL: National Council of Teachers of English.

Goodman, Y.M. (1996). Revaluing readers while readers revalue themselves: Retrospective miscue analysis. *The Reading Teacher*, *49*(8), 600–609.

Gor, K. (2017). The mental lexicon of L2 Learners of Russian: Phonology and morphology in lexical storage and access. *Journal of Slavic Linguistics, 25*(2), 277–302. doi:10.1353/jsl.2017.0011

Hyönä, J., Lorch, R.F., Jr., & Kaakinen, J.K. (2002). Individual differences in reading to summarize expository text: Evidence from eye fixation patterns. *Journal of Educational Psychology, 94*(1), 44–55. https://doi.org/10.1037/0022-0663.94.1.44

Jiménez, R.T., García, G.E., & Pearson, P.D. (1996). The reading strategies of bilingual Latina/o students who are successful English readers: Opportunities and obstacles. *Reading Research Quarterly*, *31*(1), 90–112. https://doi.org/10.1598/RRQ.31.1.5

Kang, H. (2014). Understanding online reading through the eyes of first and second language readers: An exploratory study. *Computers & Education*, *73*, 1–8. https://doi.org/10.1016/j.compedu.2013.12.005

Ke, C. (1998). Effects of language background on the learning of Chinese characters among foreign language students. *Foreign Language Annals, 31*(1)*,* 91–102. https://doi.org/10.1111/j.1944-9720.1998.tb01335.x

Kerek, E., & Niemi, P. (2012). Grain-size units of phonological awareness among Russian first graders. *Written Language & Literacy*, *15*(1), 80–113. https://doi.org/10.1075/wll.15.1.05ker

Kleiman, E. (2017). *EMAtools: Data Management Tools for Real-Time Monitoring/Ecological Momentary Assessment Data*. R package version 0.1.3. https://CRAN.R-project.org/package=EMAtool

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*, 262–284. https://doi.org/10.1080/09541440340000213

Koda, K. (2005). *Insights into second language reading: A cross-linguistic approach*. New York: Cambridge University Press.

Koda, K. (2007). Reading and language learning: Crosslinguistic constraints on second language reading development. *Language Learning, 57*(1), 1–44. https://doi.org/10.1111/j.1467-9922.2007.00411.x

Kornev, A.N. (1997). *Нарушение чтения и письма у детей: учебно-методическое пособие*. [Disruption in Reading and Writing in Children: A Textbook]. Saint-Petersburg: M & M.

Korneev, A., Akhutina, T., & Matveeva, E. (2017). Silent reading in Russian junior schoolchildren: an eye tracking study. *Psychology: Journal of Higher School of Economics*, *14*(2), 219–235. doi: 10.17323/1813-8918-2017-2-219-235

Kruskal, J. (1964). Multidimensional scaling by optimizing goodness of fit to a nonmetric hypothesis. *Psychometrika*, *29*(1), 1–27. https://doi.org/10.1007/BF02289565

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. https://doi.org/10.1016/j.jml.2011.03.002

Laurinavichyute, A.K., Sekerina, I.A., Alexeeva, S., Bagdasaryan, K., & Kliegl, R. (2019). Russian Sentence Corpus: Benchmark measures of eye movements in reading in

Russian. *Behavior research methods*, *51*(3), 1161–1178. https://doi.org/10.3758/s13428-018-1051-6

Lüdecke, D. (2017). *sjPlot: Data visualization for statistics in social science* (R package version 2.3.3). Retrieved from https://CRAN.Rproject.org/package=sjPlot

Lyashevskaya, O.N., & Sharov, S. A. (2009*). Частотный словарь современного русского языка (на материале Национального Корпуса Русского Языка.* [Frequency Dictionary of Modern Russian (based on the materials of the Russian National Corpus)]. Moscow, Russia: Azbukovnik.

McNamara, D.S. (Ed.). (2007). *Reading comprehension strategies: Theory, interventions, and technologies*. Mahwah, NJ: Erlbaum.

Meseguer, E., Carreiras, M., & Clifton, C. (2002). Overt reanalysis strategies and eye movements during the reading of mild garden path sentences. *Memory & Cognition*, *30*(4), 551–561. https://doi.org/10.3758/BF03194956

Mitchell, D.C., Shen, X., Green, M.J., & Hodgson, T.L. (2008). Accounting for regressive eye-movements in models of sentence processing: A reappraisal of the Selective Reanalysis hypothesis. *Journal of Memory and Language*, *59*(3), 266–293. https://doi.org/10.1016/j.jml.2008.06.002

Montrul, S.A. (2008). *Incomplete acquisition in bilingualism: Re–examining the age factor*. Amsterdam: John Benjamins Publishing.

Parshina, O., Laurinavichyute, A., & Sekerina, I. (2021). Eye-movement benchmarks in Heritage Language reading. *Bilingualism: Language and Cognition, 24*(1), 69–82. https://doi.org/10.1017/S136672892000019X

Polinsky, M., & Scontras, G. (2020). Understanding heritage languages. *Bilingualism: Language and Cognition, 23*(1), 4–20. doi:10.1017/S1366728919000245

R Core Team. (2018). Version 3.5. 1. *R Foundation for Statistical Computing, Vienna, Austria. https://www. r-project. org*.

Rakhlin, N., Kornilov, S.A., & Grigorenko, E.L. (2017). Learning to read Russian. In L. Verhoeven & C. Perfetti (Eds.), *Learning to read across languages and writing systems*, (pp. 371–392). Cambridge, UK: Cambridge University Press.

Raven, J. (2004). Tsvetnye progressivnye matritsy serii A, Ab, B [Raven's progressive matrices]. Moscow: Cogito-Center. (In Russian).

Rayner, K., Chace, K.H., Slattery, T.J., & Ashby, J. (2006). Eye movements as reflections of comprehension processes in reading. *Scientific Studies of Reading*, *10*(3), 241–255. https://doi.org/10.1207/s1532799xssr1003_3

Roberts, L., & Siyanova–Chanturia, A. (2013). Using eye-tracking to investigate topics in L2 acquisition and L2 processing. *Studies in Second Language Acquisition, 35*(2)*,* 213–235. doi:10.1017/S0272263112000861

Schmidtke, D., & Moro, A.L. (in press). Determinants of word-reading development in English Learner university students: A longitudinal eye movement study. *Reading Research Quarterly*, https://doi.org/10.1002/rrq.362

Scontras, G., Fuchs, Z., & Polinsky, M. (2015). Heritage language and linguistic theory. *Frontiers in Psychology*, *6*. https://doi.org/10.3389/fpsyg.2015.01545

Segalowitz, N., & Hulstijn, J.H. (2009). Automaticity in bilingualism and second language learning. In J.F. Kroll & A.M.B. De Groot (Eds.), *Handbook of Bilingualism: Psycholinguistic Approaches* (pp. 371–388). New York: Oxford University Press.

Sheorey, R., & Mokhtari, K. (2001). Differences in the metacognitive awareness of reading strategies among native and non-native readers. *System*, *29*(4), 431–449. https://doi.org/10.1016/S0346-251X(01)00039-2

Stevenson, M., Schoonen, R., & Glopper, K.D. (2003). Inhibition or compensation? A multidimensional comparison of reading processes in Dutch and English. *Language Learning*, *53*(4), 765–815. https://doi.org/10.1046/j.1467-9922.2003.00241.x

Van Dyke, J.A., Johns, C.L., & Kukona, A. (2014). Low working memory capacity is only spuriously related to poor reading comprehension. *Cognition*, *131*(3), 373–403. https://doi.org/10.1016/j.cognition.2014.01.007

Vasishth, S., von der Malsburg, T., & Engelmann, F. (2013). What eye movements can tell us about sentence comprehension. *Wiley Interdisciplinary Reviews: Cognitive Science, 4*(2), 125–134. https://doi.org/10.1002/wcs.1209

Venables, W.N., & Ripley, B.D. (2002). *Modern Applied Statistics with S*. New York: Springer.

von der Malsburg, T., & Vasishth, S. (2011). What is the scanpath signature of syntactic reanalysis? *Journal of Memory and Language, 65*(2), 109–127. https://doi.org/10.1016/j.jml.2011.02.004

von der Malsburg, T., & Vasishth, S. (2013). Scanpaths reveal syntactic underspecification and reanalysis strategies. *Language and Cognitive Processes, 28*(10), 1545–1578. https://doi.org/10.1080/01690965.2012.728232

von der Malsburg, T., Kliegl, R., & Vasishth, S. (2015). Determinants of scanpath regularity in reading. *Cognitive science*, *39*(7), 1675–1703. https://doi.org/10.1111/cogs.12208

Whitford, V., Pivneva, I., & Titone, D. (2016). Eye movement methods to investigate bilingual reading. In R.R. Heredia, J. Altarriba, & A.B. Cie´slicka (Eds.), *Methods in bilingual* *reading comprehension research* (pp. 183–211). New York, NY: Springer. https://doi.org/10.1007/978-1-4939-2993-1_8

Whitford, V., & Titone, D. (2012). Second-language experience modulates first- and second-language word frequency effects: Evidence from eye movement measures of natural paragraph reading. *Psychonomic Bulletin & Review, 19,* 73–80. https://doi.org/10.3758/s13423-011-0179-5

Whitford, V., & Titone, D. (2016). Eye movements and the perceptual span during first- and second-language sentence reading in bilingual older adults. *Psychology and Aging, 31*(1)*,* 58–70. http://dx.doi.org/10.1037/a0039971

Whitford, V., & Titone, D. (2017). The effects of word frequency and word predictability during first- and second-language paragraph reading in bilingual older and younger adults. *Psychology and Aging, 32*(2)*,* 158–177. https://doi.org/10.1037/pag0000151

Woodcock, R.W. (2011). *Wookcock Reading Mastery Tests. 3rd Ed. Manual*. Oxford: Pearson.

Xiao, Y. (2006). Heritage learners in the Chinese language classroom: Home background. *Heritage Language Journal, 4*(1)*,* 47–56.

Yarbus, A.L. (1967). *Eye Movements and Vision*. New York, NY: Plenum Press.

Zhang, H., & Koda, K. (2018). Vocabulary knowledge and morphological awareness in Chinese as a heritage language (CHL) reading comprehension ability. *Reading and Writing, 31,* 53–74. https://doi.org/10.1007/s11145-017-9773-x

Figure captions

1. *Figure 1.* Scanpaths for Example (1) with the least diverse scanpaths. The participant number is at the top, coded by color: *ML1-30* – monolingual, *CH1-30* – child, *HS1-30* – Heritage Speaker, *L21-30* – L2 learner. Scanpath plots referred to in the text are framed. Scanpaths are truncated at 15s for readability of the plots.

2. *Figure 2.* Scanpaths for Example (2) with the most diverse scanpaths. Scanpath plots referred to in the text are framed. Scanpaths are truncated at 15s for readability of the plots.

3. *Figure 3.* Maps of scanpaths showing the two dimensions that explain the most of the scanpath variance: (A) the map for Example (1) that elicited the smallest scanpath variance, (B) the map for Example (2) with the most varied scanpath patterns.

4. *Figure 4. S*ample of the scanpaths coded by clusters for Example (2) that elicited the most diverse eye-movement patterns. The cluster number is at the top, followed by the participant number. Scanpath plots are truncated at 15s for plotting.

5. *Figure 5.* Prototypical scanpath reading processes in Example (2) as identified by the scanpaths closest to the centroids of the clusters.

6. *Figure 6*. Individual differences in the scanpath reading processes: The graph shows how many instances of the scanpath reading process each reader produced. The first 2 top rows are children, the next 2 rows are Heritage Speakers, followed by L2 learners and the last 2 rows are monolingual adults.

Table 1. *Demographic and performance characteristics of the two bilingual groups*

	Heritage Speakers	**L2 learners**
	Mean (*SD)*	Mean (*SD)*
Number of participants	30	30
Age of arrival to the USA (years old)	4.3 (5.4)	0.13 (0.75)
Daily Russian exposure (%)	25.6 (18.9)	7.9 (7.4)
Self-estimated comprehension (1–5)	3.2 (1.1)	2.8 (.87)
Oral Reading Fluency (Russian)	12.3 (6.0)	8.3 (2.7)
Oral Reading Fluency (English)	26.6 (6.3)	28.1 (5.3)

Table 2. *Descriptive characteristics of the child Russian Sentence Corpus*

	Child Russian Sentence Corpus
# of sentences	30
# of words	227
Sentence length (words)	*M* = 8, *Range*: 6–9
Word length (letters)	*M* = 5.6, *Mdn* = 6, *Range*: 1–13
Word frequency (items per million)	*M* = 3088.2, *Mdn* = 2583.3, *Range*: 7.4–7537.7

Table 3. *Means and SDs for the eye-movement measures (top panel) and percentage distribution of the participants comprising each of the scanpath reading processes (SD)*

	Scanpath reading process		
	Fluent	**Intermediate**	**Beginner**
Gaze duration	289.3 (81.8)	689.9 (338.2)	1053.5 (544.4)
Skipping rate (%)	17.1 (13.7)	8.7 (11.9)	5.8 (10.9)
Fixation count/word	1.3 (.390)	2.8 (1.1)	5.1 (2.1)
Regression rate (%)	12.9 (14.7)	25.4 (18.7)	38.2 (22.7)
Count of word readings	1.0 (.270)	1.4 (.492)	2.2 (.923)
Total time reading/sentence (s)	2.1 (.761)	6.4 (2.9)	13.8 (5.9)
Monolinguals	73.6% (4.8)	10.0% (3.4)	0.5% (0.34)
Children	11.0% (2.2)	35.7% (2.8)	22.1% (3.1)
Heritage Speakers	12.2% (2.0)	28.8% (2.9)	35.1% (3.1)
L2 learners	3.3% (0.96)	25.5% (3.8)	42.3% (4.3)

Table 4. *Parameter estimates for GLMMs: Probability of engaging in one of the three scanpath reading processes by group (Bonferroni correction applied).*

	Scanpath reading process											
	Fluent				**Intermediate**				**Beginner**			
	Est	*SE*	*p*	*d*	*Est*	*SE*	*p*	*d*	*Est*	*SE*	*p*	*d*
Monolingual	6.8	.63	**<.001**	3.1	-2.5	.36	**<.001**	-1.3	-6.3	.90	**<.001**	-1.0
Children	-2.4	1.1	.069	-.51	1.6	.39	**<.001**	.74	.15	.67	1.00	-.13
HSs	-3.3	1.1	**.006**	-.47	.54	.41	.579	.22	1.4	.65	.081	.32
L2 learners	-4.6	.96	**<.001**	-.78	.41	.42	.972	.16	2.7	.61	**<.001**	.77

Table 5. *Conceptual comparison of the three scanpath reading processes (to match the descriptive characteristics in Table 3 and statistical analysis in Table A1)*

	Fluent	**Intermediate**	**Beginner**
Fixation duration	289 ms	690 ms	1054 ms
Word skipping	17%	9%	6%
Total reading time	2 s	6 s	14 s
Short leftwards saccades	Few	√	√
Long regressions	Few	√	√
Sentence re-readings	—	Few	√

Author statements

Olga Parshina (corresponding author) is a postdoctoral researcher in the Center for Language and Brain at the HSE University, Russia; e-mail oparshina@hse.ru. Her research interests are bilingual reading and bilingual lexical access.

Irina. A. Sekerina is a professor in the Psychology Department at the College of Staten Island and The Graduate Center of The City University of New York and an affiliated researcher in the Center for Language and Brain at the HSE University, Russia; e-mail irina.sekerina@csi.cuny.edu. Her research interests focus on the investigation of monolingual and bilingual language comprehension using eye-tracking.

Anastasiya Lopukhina is a research fellow in the Center for Language and Brain at the HSE University, Russia; e-mail alopukhina@hse.ru. Her research interests are child reading development and reading disabilities.

Titus von der Malsburg is a junior professor in the Institute of Linguistics at the University of Stuttgart, and a research affiliate in the Department of Brain and Cognitive Sciences at the Massachusetts Institute of Technology (MA, USA); e-mail titus.vondermalsburg@ling.uni-stuttgart.de. His research interests are incremental sentence interpretation and experimental and computational research methods.

Appendix

The following tables present the outcomes for the models and model equations referred to in text.

Table A1. *Comparison of the eye movements characteristics among the three scanpath reading processes. Significant differences are in bold. Bonferroni correction applied.*

	Gaze Duration			Skipping probability			Regression probability			Fixation count			Times word read			Total Time sentence		
Predictors	*Est*	*SE*	*p*	*Est*	*SE*	*p*	*Est*	*SE*	*p*	*Est*	*SE*	*p*	*Est*	*SE*	*p*	*Est*	*SE*	*p*
Reference: Fluent Process																		
(Intercept)	6.1	.05	**<.001**	.15	.01	**<.001**	.12	.01	**<.001**	1.9	.13	**<.001**	.95	.04	**<.001**	8.0	.05	**<.001**
Intermediate	.20	.02	**<.001**	-.05	.01	**<.001**	.14	.01	**<.001**	.90	.06	**<.001**	.46	.03	**<.001**	.55	.02	**<.001**
Beginner	.33	.02	**<.001**	-.08	.01	**<.001**	.27	.01	**<.001**	2.6	.08	**<.001**	1.2	.04	**<.001**	1.1	.02	**<.001**
Reference: Beginner Process																		
(Intercept)	6.5	.05	**<.001**	.07	.01	**<.001**	.38	.01	**<.001**	4.4	.13	**<.001**	2.2	.04	**<.001**	9.1	.05	**<.001**
Intermediate	-.12	.01	**<.001**	.03	.01	**<.001**	-.12	.01	**<.001**	-1.7	.05	**<.001**	-.75	.03	**<.001**	-.47	.02	**<.001**
Fluent	-.33	.02	**<.001**	.08	.01	**<.001**	-.27	.01	**<.001**	-2.6	.08	**<.001**	-1.2	.04	**<.001**	-1.1	.02	**<.001**
Random effects																		
σ^2	.07			.01			.02			.89			.22			.08		
$\tau_{00,\ \text{participants}}$	.21			.00			.01			.86			.11			.18		
$\tau_{00,\ \text{sentence}}$	.02			.00			.00			.19			.02			.02		
Observations	3446			3449			3446			3449			3449			3442		
R^2 / Ω_0^2	.046 / .776			.053 / .309			.204 / .483			.308 / .682			.349 / .580			.323 / .815		

Table A2. *Summary GLMMs for the scanpath reading processes for Heritage Speakers. The cells with estimates in which there is a significant effect are in bold. Bonferroni correction applied.*

	Heritage Speakers											
	Fluent				**Intermediate**				**Beginner**			
	Est	*SE*	*p*	*d*	*Est*	*SE*	*p*	*d*	*Est*	*SE*	*p*	*d*
Fixed Effects												
(Intercept)	-7.1	.95	**<.001**		.49	.36	.516		-.89	.45	.150	
Age of Arrival	.41	.31	.180	.07	.10	.25	1.0	.07	-.40	.34	.717	-.12
Self-assessments:												
Rus. exposure	.21	.45	.1.0	.22	-.32	.30	.876	-.32	.37	.38	.972	.22
Comprehension	.63	.51	.657	.72	-.30	.34	1.0	-.28	-.26	.43	1.0	-.14
Reading pre-tests:												
ORF-Rus	2.6	.57	**<.001**	1.6	.21	.30	1.0	.14	-1.5	.41	**<.001**	-1.2
ORF-Eng	.31	.60	1.0	-.32	.24	.29	1.0	.26	-.09	.35	1.0	-.09
Random Effects:												
$\tau_{00,\ \text{sentence}}$	1.8				.125				.497			
$\tau_{00,\ \text{participant}}$	1.6				2.1				3.1			

Table A3. *Summary GLMMs for the scanpath reading processes for L2 learners. The cells with estimates in which there is a significant effect are in bold. Bonferroni correction applied.*

	L2 learners											
	Fluent				Intermediate				Beginner			
	Est	*SE*	*p*	*d*	*Est*	*SE*	*p*	*d*	*Est*	*SE*	*p*	*d*
Fixed Effects												
(Intercept)	-3.7	.70	**<.001**		.90	.50	.207		-1.5	.58	**.033**	
Self-assessments:												
Rus. exposure	.57	.75	1.0	.23	.47	.53	1.0	-.31	-.47	.62	1.0	.20
Comprehension	2.0	.73	**.018**	.71	.23	.44	1.0	-28	-.72	.51	.465	-.13
Reading pre-tests:												
ORF-Rus	-.64	.89	1.0	1.7	.66	.59	.783	.16	-.65	.69	1.0	-1.2
ORF-Eng	.36	.46	1.0	-.33	-.34	.33	.924	.25	.19	.39	1.0	-.07
Random Effects:												
$\tau_{00, \text{sentence}}$	.291				.495				.725			
$\tau_{00, \text{participant}}$	2.0				2.1				2.8			

Appendix B. *An example for the code used for fitting the generalized linear model for testing the influence of the group membership on the scanpath reading process.*

```
glmer(Fluent process ~ monolingual + (1|sentence) + (1|id), data = strategies, family =
binomial, control = glmerControl(optimizer = "bobyqa"))
```

Appendix C. *An example of the code used for fitting the generalized linear model for testing the influence of the demographic and reading performance factors on the scanpath reading process by Heritage Speaker group.*

```
glmer(Fluent process ~ Age_of_Arrival + Russian_daily_exposure + Comprehension +
ORF_Rus + ORF_Eng + (1|sentence) + (1|id), data = HS, family = binomial, control =
glmerControl(optimizer = "bobyqa"))
```

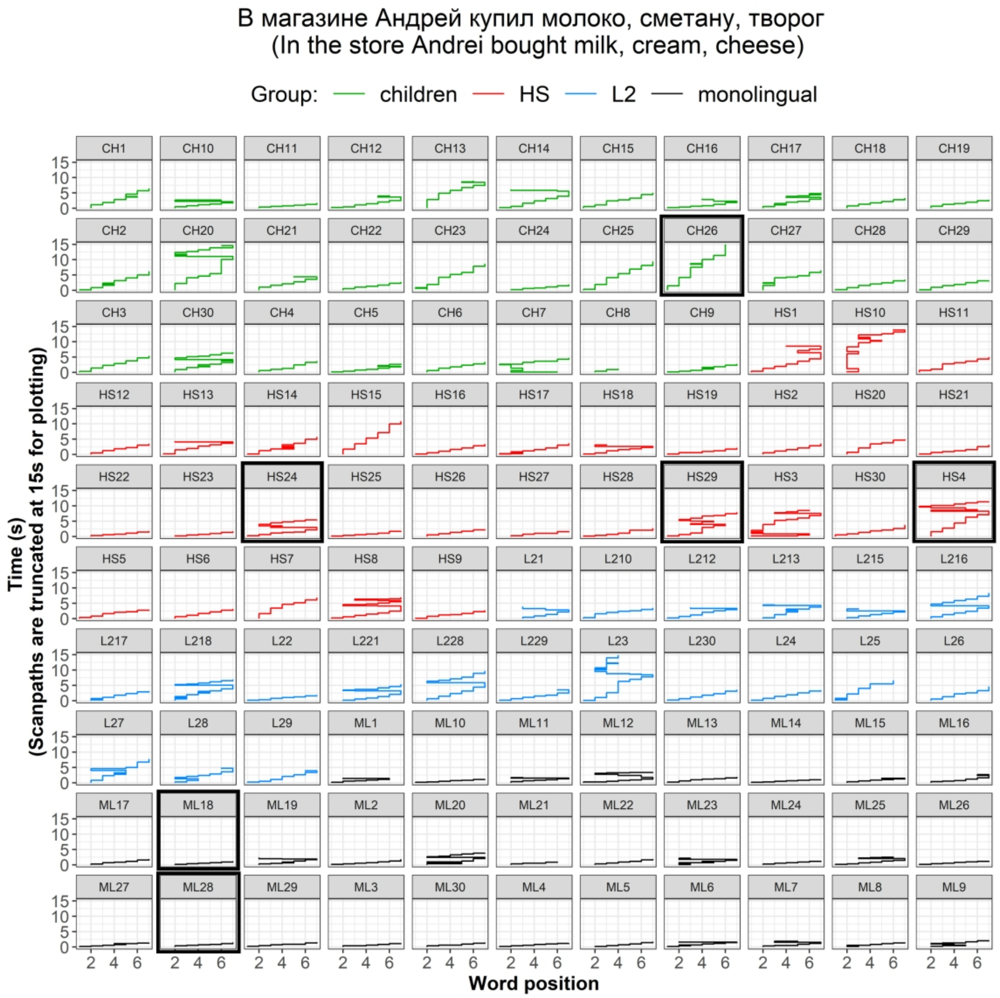

Figure 1. Scanpaths for Example (1) with the least diverse scanpaths. The participant number is at the top, coded by color: ML1-30 – monolingual, CH1-30 – child, HS1-30 – Heritage Speaker, L21-30 – L2 learner. Scanpath plots referred to in the text are framed. Scanpaths are truncated at 15s for readability of the plots.

254x254mm (300 x 300 DPI)

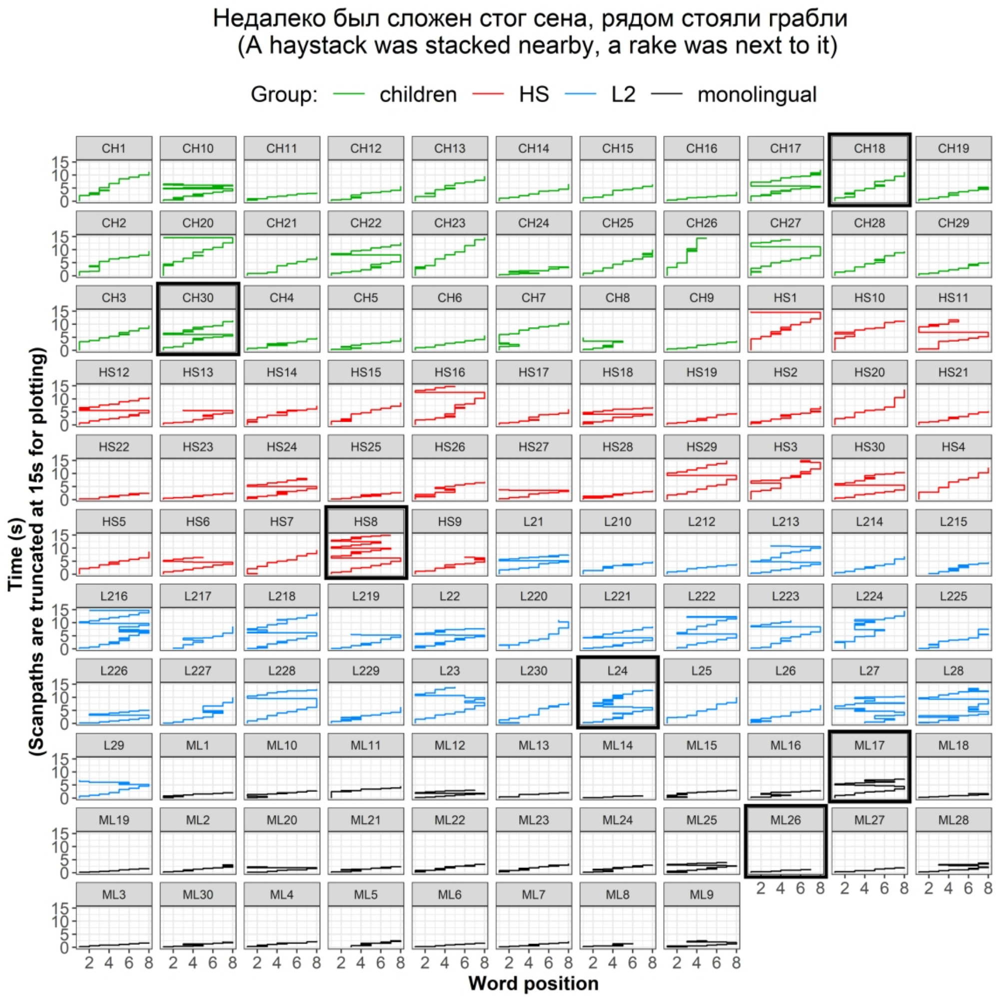

Figure 2. Scanpaths for Example (2) with the most diverse scanpaths. Scanpath plots referred to in the text are framed. Scanpaths are truncated at 15s for readability of the plots.

254x254mm (300 x 300 DPI)

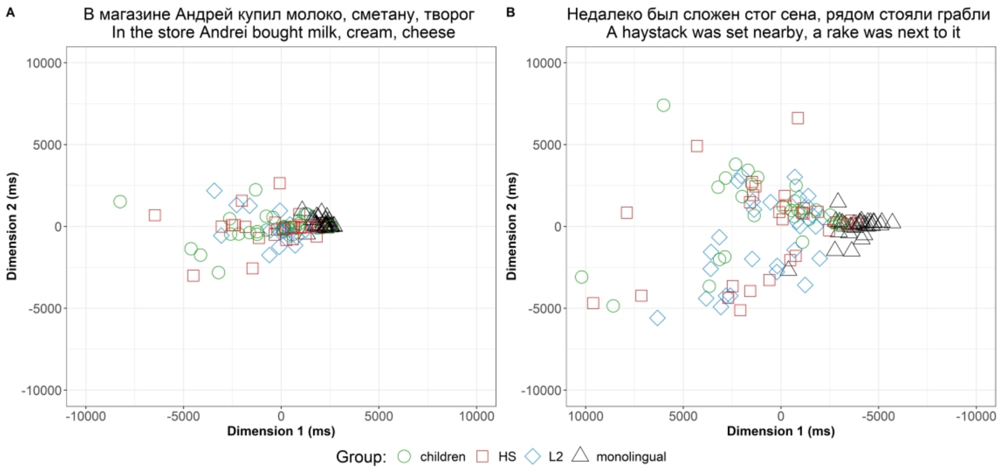

Figure 3. Maps of scanpaths showing the two dimensions that explain the most of the scanpath variance: (A) the map for Example (1) that elicited the smallest scanpath variance, (B) the map for Example (2) with the most varied scanpath patterns.

381x177mm (300 x 300 DPI)

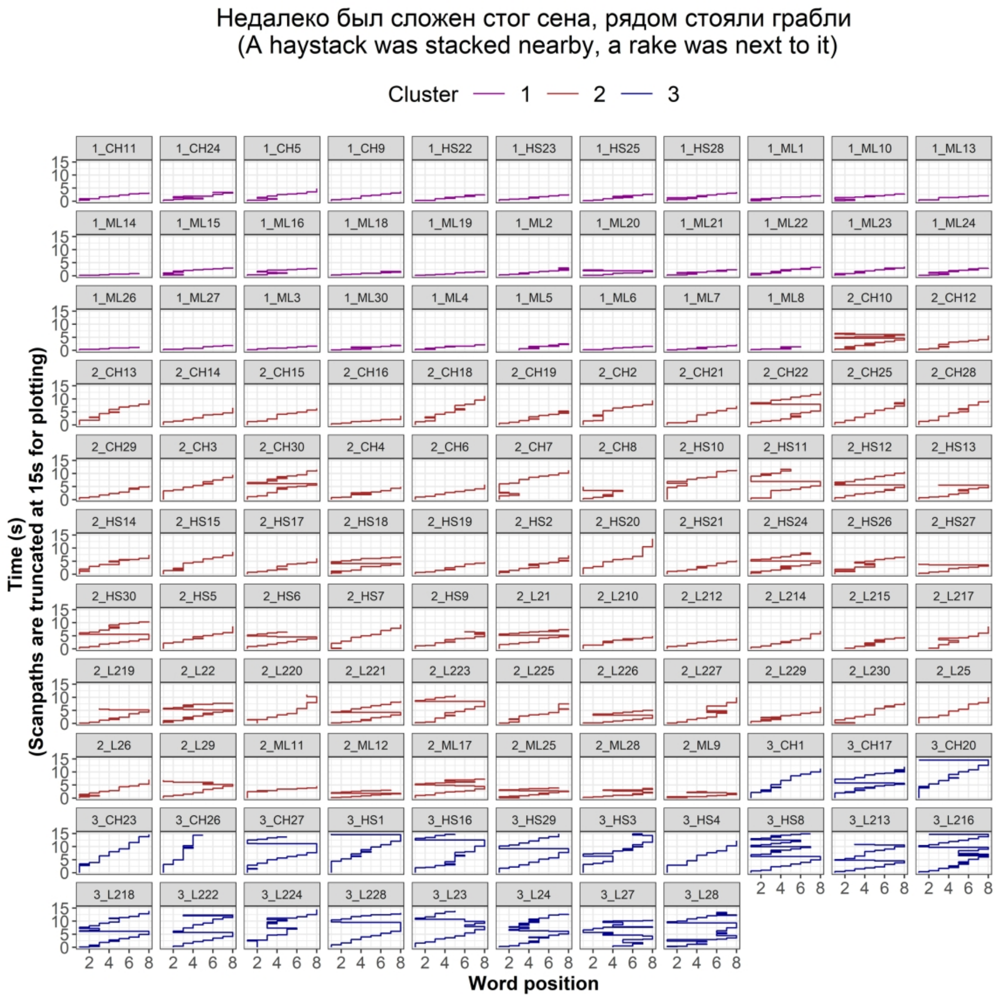

Figure 4. Sample of the scanpaths coded by clusters for Example (2) that elicited the most diverse eye-movement patterns. The cluster number is at the top, followed by the participant number. Scanpath plots are truncated at 15s for plotting.

254x254mm (300 x 300 DPI)

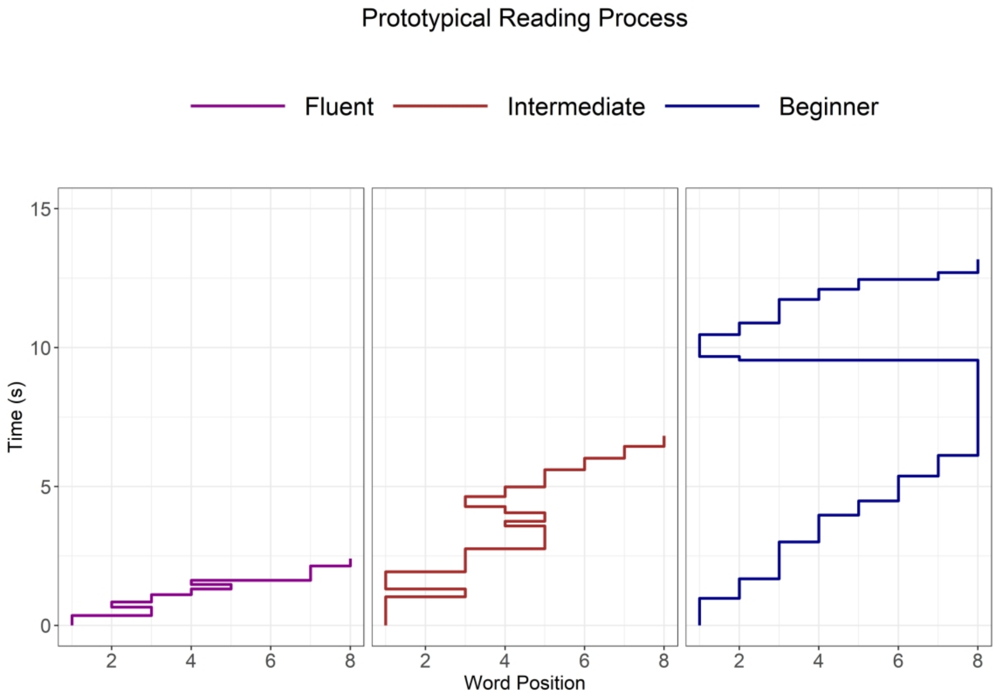

Figure 5. Prototypical scanpath reading processes in Example (2) as identified by the scanpaths closest to the centroids of the clusters.

254x177mm (300 x 300 DPI)

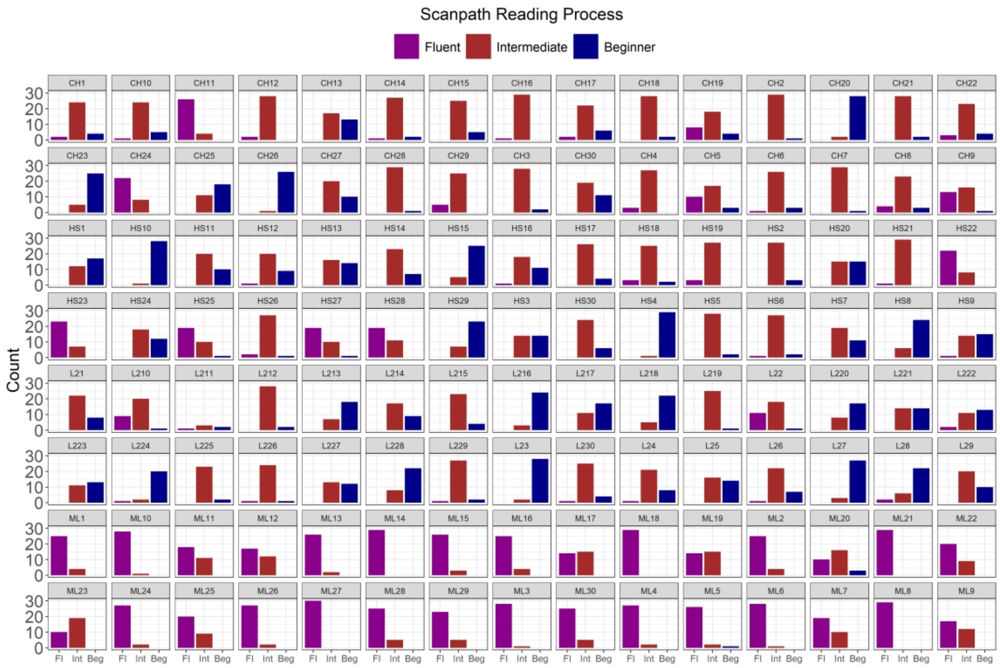

Figure 6. Individual differences in the scanpath reading processes: The graph shows how many instances of the scanpath reading process each reader produced. The first 2 top rows are children, the next 2 rows are Heritage Speakers, followed by L2 learners and the last 2 rows are monolingual adults.

381x254mm (300 x 300 DPI)