Numeral Processing in Second Language Oral Reading: An Eye-Tracking Study

Emrah Dolgunsöz
Bayburt University

ABSTRACT

Numbers (i.e. 45) are symbols that are common in reading texts in various forms such as dates, percentages, and quantity expressions. Each digit corresponds to a word, unlike words in which each letter represents a phoneme. Despite their high frequency in reading texts, eye movement research on numeral processing is still rare. Numerals are easy to process in silent reading due to their standard nature in many languages; however, in oral reading, extra effort is required as numerals are pronounced irregularly among several languages. This study aimed to examine processing differences between words and numerals during second language oral reading by using eye-tracking. 40 intermediate non-native speakers were tested under same three text conditions (familiar words, non-words, and numerals) while their eye movements were recorded. Results showed that the participants spent a considerable amount of time processing numerals when compared to words and non-words. Additionally, numerals caused more regressions and were revisited more causing slower reading performance.

INTRODUCTION

Numbers and letters are basic elements of reading texts. Since the 1970s, word recognition skills during reading have been a popular topic of eye movement research; we have plenty of evidence proposing that readers spend less time on frequent and familiar words both in L1 reading (Williams and Morris, 2004; Chaffin, Morris, and Seely, 2001) and recently in second language (L2) reading (Godfroid, Housen and Boers, 2010, Godfroid, Boers, Housen, 2013; Dolgunsöz, 2015)). Fixation times also differ between silent and oral reading; readers spend more attentional sources during oral reading as eyes get ahead of the words and eye-voice synchronization is required (Kim, Petscher and Vorstius, 2019). Although eye movement literature is abundant in terms of word processing in reading, we still lack empirical evidence on eye movements during numeral processing, especially in the case of L2 oral reading. Numerals and words are both visual symbols of language but they are different in nature. Arabic numerals are logographic which arbitrarily represents quantity while words are alphabetic in which each orthographic symbol (letter) corresponds to phonemes. In both L1 and L2 reading, target language texts are not only comprised of words but may also include various forms of numeric expressions. These expressions are common in the forms of dates, percentages, numerals expressing quantity, and even in
arithmetic expressions. Each of these symbols is called Arabic numerals used in several languages which are similar to the ideographic script since each digit corresponds to one word rather than one phonological unit. Despite their identical form among many languages, the pronunciation, and spelling of Arabic numerals are not regular. The number “15” is pronounced as [fifti:n] in English, [kanz] in French, and [FUUNF-tsayn] in German in which numerals are spelled and articulated irregularly. Traditionally, numbers are taught at the beginner level in the L2 context but learners may be challenged by such irregularities as they are required to name such expressions rapidly along with words for successful oral reading performance. Numbers became even more important for English for specific purposes context (ESP) in which learners were exposed to technical texts loaded with different forms of numerals. The primary aim of this study is to examine eye movements during L2 oral reading by comparing numerals and words by using eye-tracking to examine if words and numbers are processed differently.

**Automaticity, Oral Reading and Eye Movements**

Automaticity as a component of fluency in reading is vital for reading comprehension. The theoretical basis for reading fluency as the predictor of comprehension relies on Automaticity Theory (DeKeyser, 2001; Segalowitz & Segalowitz, 1993; Rasinski & Samuels, 2011) and Verbal Efficiency Theory (Perfetti, 1985). These theories proposed that attention and working memory which are vital for reading processes are limited in capacity. When readers are unable to name a linguistic input rapidly (i.e. an unknown word or a numeral) during reading, they slow down and this unautomatized cognitive process rapidly consumes limited cognitive resources. As a result, reading fluency and rate decrease along with problems in higher levels of reading comprehension processes. Automatic processes are fast, reside in working memory briefly, and require minimal attentional capacity and while unautomatic processes require a higher amount of attentional resources which make them deliberate, slow and effortful (Moors and De Houwer, 2006).

Eye movements during reading are closely linked to automaticity due to these attentional concerns. The eye movement research demonstrated a number of differences in reading between skilled and poor readers; inexperienced readers lack automaticity in word recognition and tend to spend longer fixation durations and make more fixations during silent reading when compared to skilled readers (Rayner, Slattery and Belanger, 2010). On the other hand, skilled readers employ shorter fixations and fewer regressions (Krieber et. al., 2016). Similarly, bilinguals show similar eye movement characteristics in their dominant language (Altarriba et. al., 1996). Longer attentional spans on words signal lexical and phonological difficulties affecting reading performance; skilled readers recognize words faster with less attention span while poor readers spent a higher amount of attention which decreases their fluency and rapid automatic naming. Apart from reading skills, the characteristics of the text also affect eye movements. Several word recognition-oriented eye movement research showed that readers spent more time on less frequent words, on less familiar words (Williams and Morris, 2004; Chaffin, Morris, and Seely, 2001), and longer words (Rayner et. al., 2011). Additionally, skilled readers are fluent and tend to skip 30% of the words in a text (Rayner, 1998); these words are mostly short and familiar (Brysbaert and Vitu, 1998; Drieghe, Brysbaert, Desmet, & De Baecke, 2004).

Although eye movement literature on oral reading is limited, similar effects of reading and word recognition skills have been demonstrated with inflated values. In oral reading, fixation times
are slightly higher than in silent reading as the articulation of the fixated word begins after a saccade is made to the subsequent word as muscles in the speech tract operate more slowly than do cognitive processes (Ashby et. al., 2012) which is also known as eye-voice span. Naturally, this timing difference causes brief inflation in fixation times during oral reading (Inhoff, Solomon, Radach, & Seymour, 2011) when compared to silent reading in which no voice-attention synchronization is required.

Numerals and Words as Two Different Processes

As mentioned above, eye movement literature is rich in terms of word recognition processes during reading but said less about numeral processing. We still lack eye movement related empirical evidence on numeral processing in L2 oral reading and fixation times when compared to eye movements during word recognition processes. However, numbers are common in any reading text as digits, percentages, or as mathematical expressions. In any language, numerals may be presented in a text as word-number (i.e. forty-eight) or in Arabic numerals (i.e. 48). Words and numerals during reading have long been considered as two distinct neural processes (Besner and Coltheart, 1979). Recent related research also confirmed that reading letters and numbers are processed in different neural components (Park et. al., 2012) in the brain and visual recognition of digits and letters are found to be dissociated in some behavioral (Hamilton, Mirkin and Polk, 2006), neuropsychological (Starrfelt, 2007) and neuroimaging studies (James et. al., 2005; Flowers et al., 2004). Letters were found to have activated the left mid fusiform and inferior temporal gyri more than numbers while numerals activated a right lateral occipital area more than letters (Park et. al., 2012). For second language learning, this difference is more complicated when learners are still coping with numerous grammatical, syntactic, and phonological problems.

For L2 context, it is still vague how numerals are processed during L2 oral reading. Depending on the processing differences of numbers and words, the current study hypothesized that numbers required more time to process during L2 oral reading due to their phonological load when compared to familiar words and non-words. In this respect, the primary aim of this study was to compare total time and revisiting times on numerals, non-words, and familiar words during L2 oral reading by utilizing the eye-tracking technique and discuss the challenging nature of numerals in L2 oral reading. The research questions are as follows:

1. How fast numbers were read orally when compared to non-words and familiar words?
2. Was there a significant difference between familiar words, non-words, and numerals regarding total fixation duration, and revisiting times during L2 oral reading?
3. Is there an effect of multiple digits on eye movements during L2 oral reading?

METHODOLOGY

Design

This experimental study adopted a within-subject design in which all participants were tested under the same 3 conditions: Familiar words, non-words, and numerals.
Participants

40 learners of English (20 males and 20 females) with an age range of 18 to 21 voluntarily participated in the study. All learners were intermediate freshmen students with the same L1 background in an ELT department and completed preparatory class a year ago with a proficiency score above 70. All participants had normal or corrected to normal eyesight.

Text Stimuli

A total of 30 short authentic sentences including 10 familiar words, 10 non-words, and 10 numerals were used as textual stimuli. In all three conditions, each sentence was around 7 words long and the total number of words was the same for each condition. To control any effect of proficiency and syntax, beginner-level sentences were used; all sentences were taken from “Oxford Bookworms Stage 1 series” (Love or Money, The Coldest Place on Earth, and The Phantom of the Opera). Since eye movements are sensitive towards word length, the length of non-words and familiar words were controlled. For non-word conditions, a word in each sentence was substituted with a non-word generated from the ARC Non-Word Database (Rastle, Harrington, & Coltheart, 2002). Length of non-words ranged from 4 to 8 letters (M= 6.1, SD=1.28). As the texts were beginner level, familiar words were not modified as they are already highly frequent and easy to pronounce words. Length of these words varied from 4 to 9 letters (M=6.2, SD=2.25). For numerals, 1 to 5 digit numbers were used. Existing numerals in the sentences were slightly modified for multiple digits.

<table>
<thead>
<tr>
<th>Familiar words</th>
<th>Non-words</th>
<th>Numerals</th>
</tr>
</thead>
<tbody>
<tr>
<td>food</td>
<td>cauv</td>
<td>5</td>
</tr>
<tr>
<td>dogs</td>
<td>tarb</td>
<td>3</td>
</tr>
<tr>
<td>room</td>
<td>scwarr</td>
<td>28</td>
</tr>
<tr>
<td>story</td>
<td>shroot</td>
<td>17</td>
</tr>
<tr>
<td>house</td>
<td>clylce</td>
<td>270</td>
</tr>
<tr>
<td>night</td>
<td>swoills</td>
<td>150</td>
</tr>
<tr>
<td>terrible</td>
<td>phighg</td>
<td>1379</td>
</tr>
<tr>
<td>beautiful</td>
<td>krylled</td>
<td>2517</td>
</tr>
<tr>
<td>inspector</td>
<td>squoogs</td>
<td>21465</td>
</tr>
<tr>
<td>snowstorm</td>
<td>groarnte</td>
<td>47852</td>
</tr>
</tbody>
</table>

Apparatus

Eye movements were recorded via a remote GP3 eye tracker with a 60hz speed. GP3 can register a sample every 16 milliseconds with 0.5–1° of visual angle accuracy and a 25 cm (horizontal) x 11 cm (vertical) head movement flexibility. Eye-tracking data was analyzed with GP3 Professional Software.
Procedure

All participants were tested individually under the control of the researcher(s) in a well-lit room. Each sentence was introduced separately on a slide. The participants were instructed to read the sentence on the screen aloud for comprehension purposes and pass to another sentence by pressing space. All sentences were randomized for each participant. Before each experiment, nine-point calibration was applied in front of a 19-inch monitor set up at about 60 cm from the participant’s eye. Eye movements were also observed online in another computer by the researcher(s) to avoid any data loss and data deterioration.

Data Analysis

Time spent (seconds) of each sentence slide was extracted from the software. For the analysis of the eye movement data, AOIs (Areas of interest) for numbers, non-words, familiar words were drawn. For each AOI, values in seconds were extracted from the software and transferred to statistics software. Total fixation duration and revisit counts were analyzed as eye movement measures. Total fixation duration refers to the total time spent on an AOI. Revisits can be defined as the number of rereading attempts on AOIs after exiting the related AOI and regressing it. For all statistical analyses, repeated-measures ANOVA was utilized with a 95% confidence interval adjusted by Bonferroni since all participants were tested with the same 3 conditions.

RESULTS

Finding 1: Reading Speed on 3 conditions

For the first research question, a repeated-measures ANOVA was conducted to see the total time spent on 3 types of sentences and to examine any possible statistically significant difference. The participants finished reading all 10 sentences in the numeral condition in 54.3 seconds (SD=11.8) while they spent 39.1 seconds (SD=9.2) in non-word condition. The participants were observed to have finished familiar word condition in 38.3 seconds (SD=7.31). This difference was found to be statistically significant; F(1.39, 54.5) = 89, p=.000.

Mean time spent for familiar word and non-word conditions were close. Despite the minimal difference, pairwise comparisons revealed a significant difference between familiar word and non-word conditions (p=.001). As expected, the numeral condition differed significantly from both other conditions (p=.000). Hence, the participants showed the slowest oral reading performance in the numeral condition. It was followed by non-word condition and the least time was spent on sentences containing familiar words.

Findings 2: Eye Tracking Findings on Numerals vs non-words and familiar words

Since initial findings revealed slower reading speed in the numeral condition, an effect of numerals was expected in processing times and revisit counts.
**Figure 1.** Sample Gaze Plot for 3 AOIs (familiar words, non-words and numerals, respectively)

The gaze plot in Figure 1 above indicated that fixations were denser on numerals when compared to words and non-words. For a more detailed examination, total fixation duration and revisit times on numerals, non-words, and familiar words as 3 AOIs were analyzed via repeated-measures ANOVA. Mean values were summarized in Table 2 below:

**Table 2. Mean Total Fixation Duration and Revisits**

<table>
<thead>
<tr>
<th>Condition</th>
<th>Total Fixation Duration</th>
<th>Number of Revisits</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean (seconds)</td>
<td>Std. Error</td>
</tr>
<tr>
<td>Familiar words</td>
<td>.49</td>
<td>.02</td>
</tr>
<tr>
<td>Non-words</td>
<td>1.38</td>
<td>.12</td>
</tr>
<tr>
<td>Numerals</td>
<td>2.76</td>
<td>.18</td>
</tr>
</tbody>
</table>

Participants spent only half of a second on familiar words (M=.49, SD=.02) and nearly skipped them which were then followed by non-words (M=1.38, SD=.12). Numerals were processed much slower (M=2.76, SD=.18) when compared to familiar words and non-words. This difference was found to be significant; F(1.62, 63.3) = 121.419, p=.000. The results of the pairwise comparisons indicated a significant difference between numerals and familiar words (p=.000), numerals and non-words (p=.000), and familiar words and non-words (p=.000). These findings indicated that numerals were attended most and consumed a considerable amount of attentional resources in comparison to familiar words and non-words.

For revisit times, learners reread familiar words once (M=.81, SD=.08) and non-words over around 1.5 times (M=1.55, SD=.13). Numerals were revisited twice (M=2.08, SD=.17). This difference
was also found significant; \( F(1.79, 69) = 40.259, p=.000 \). For pairwise comparisons, Numerals were revisited more than familiar words (\( p=.000 \)) and non-words (\( p=.001 \)). Non-words were also reread more than familiar words (\( p=.000 \)). Learners not only spent extra time to process numerals but also regressed and revisited them more which caused slower reading speed. In other words, they were slowed down by paying more attention to numerals, regressing, and rereading them. Hence, sentences containing numerals were processed more slowly.

**Finding 3: The effect of multiple digits on attention during L2 oral reading**

Considering that numeral processing requires more attentional resources and slowed learners down by forcing them to regress and reread during L2 oral reading, an effect of multiple digits on total fixation duration and revisit times was expected. For this aim, a repeated-measures ANOVA was utilized. The results were given in Table 3:

<table>
<thead>
<tr>
<th>Number of Digits</th>
<th>Total Fixation Duration</th>
<th>Number of Revisits</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean (seconds)</td>
<td>Std. Error</td>
</tr>
<tr>
<td>Single-digit</td>
<td>.90</td>
<td>.07</td>
</tr>
<tr>
<td>2 digits</td>
<td>1.21</td>
<td>.10</td>
</tr>
<tr>
<td>3 digits</td>
<td>1.96</td>
<td>.18</td>
</tr>
<tr>
<td>4 digits</td>
<td>3.76</td>
<td>.29</td>
</tr>
<tr>
<td>5 digits</td>
<td>6.59</td>
<td>.46</td>
</tr>
</tbody>
</table>

Single-digit numerals were processed in less than 1 second (\( M=.90, SE=.07 \)) similar to double-digit numerals which were processed slightly above 1 second (\( M=1.21, SE=.10 \)). Total fixation duration values were observed to have dramatically increased for 3 digits and over. Learners spent nearly 2 seconds for 3 digits (\( M=1.96, SE=.18 \)), around 4 seconds for 4 digits (\( M=3.76, SE=.29 \)) and over 6 seconds for 5 digits (\( M=6.59, SE=.46 \)). The number of digits was found to have a significant effect on total fixation duration; \( F(1.89, 73) = 111.134, p=.000 \). According to pairwise comparisons, no significant difference was found between total time spent on single and double-digit numerals (\( p=.058 \)). In all other conditions, a significant difference was observed at \( p<.01 \) level. In sum, learners had hard times in processing 3, 4 and 5 digit numerals while single and double-digit numerals were processed relatively easily.

For revisiting times, single, double, triple, and quadruple-digit numerals were revisited less than twice as presented in Table 6. However, for 5-digit numbers, revisiting times were remarkably high (\( M=3.77, SE=.46 \)). A significant effect of the number of digits was observed on revisit times; \( F(2.06, 80.3) = 19.979, p=.000 \). According to pairwise comparisons, revisit times for 5-digit numerals were found to be significantly different from all other conditions at \( p<.01 \) level. Between 1-, 2-, 3- and 4-digit conditions, no significant difference was observed. In this respect, it can be
concluded that the number of digits had a general effect on revisit times, especially 5-digit numerals cause remarkable regressions and revisits which slowed down L2 oral reading process.

**DISCUSSION**

This study aimed to explore the effect of numbers on eye movements and oral reading speed during L2 oral reading in comparison to familiar words and non-words. The results showed that sentences containing numerals slowed down reading speed; learners spent more attentional resources and employed more revisits while processing numerals when compared to familiar words and non-words. Especially multiple digit numerals required a considerable amount of attention. Although single- and double-digit numerals seemed easier to process, 3 and above digit numbers were found to be challenging. These findings revealed that during L2 oral reading, the learners had difficulty in naming numerals rapidly. Automatization in numeral processing is complex and requires time to acquire. According to the triple-code model of number processing, there are 3 fundamental steps of processing numerals (Dehaene and Cohen, 1995). Readers have 3 representations of numbers in their mental schema and numbers are processed sequentially depending on these 3 forms. The first form is the visual Arabic number form which corresponds to the string of digits on an internal visuospatial sketchpad (i.e. 45). The second form refers to the verbal word frame which is the syntactic representation of visual Arabic number form. In this step, “45” as a visual representation is denoted as “tens(4) and ones(5)”. Finally, “tens(4) and “ones(5) constitute a word lemma and are phonologically processed. This model is L1-based but it can be assumed that similar cognitive procedures are applicable for L2 numeral processing. Depending on this model, it can be assumed that articulating a numeral is not only a linguistic process but also requires basic arithmetic skills, especially in the case of multiple digits. The current findings showed that beginner and intermediate L2 learners were challenged by these cognitive and phonological steps. Lexical access seemed to be less effective on fixation times: Learners already knew what familiar words and numerals referred to; however, there was a considerable attention span difference between them; numerals gained higher fixation time. For non-words, no lexical access was possible but still, numerals were attended more to recognize and articulate. For phonological processing, familiar words were easy to process due to both lexical and phonological access. Non-words were somehow pronounced due to their alphabetic nature and familiarity.

Numerals, especially multiple digit numbers, were processed slowly as they not only required learners to match verbal equivalents of numbers quickly and articulate them but also they needed some arithmetic skills to calculate in the decimal system. As numbers grow, more decimals are involved in the process which caused L2 learners to spend more time. Therefore, processing numerals during L2 oral reading involves both linguistic and arithmetic processes requiring learners to analyze numbers mathematically, retrieve phonological information and articulate it. Hence, these steps were challenging for learners of a second language who were not yet proficient.

Depending on the findings of the current study, it can be proposed that numerals are cognitively demanding and potentially affect the oral reading rate and fluency adversely. Successful readers in L2 are expected to name the textual input without any conscious effort which enables them to read fluently and comprehend. For oral reading, which requires readers to articulate textual input aloud, there is strong empirical evidence referring to the relationship between fluency and reading comprehension concluding that slow reading rates decrease
comprehension (Pinnel et al., 1995; Spear-Swerling, 2006). Although they are limited, several oral reading fluency (ORF) related studies in the L2 context confirms ORF and reading comprehension relationship (Lems, 2003; Crosson and Lesaux, 2010; Jiang, Sawaki and Sabatini, 2012). The consensus is that slow reading rates and lack of automaticity during L2 oral reading decrease reading comprehension performance. Although the current study did not empirically test comprehension issues since simple sentences were used to control syntax and proficiency, higher fixation times and revisits on numerals indicate attempts to construct textual comprehension which refers to spending too much time on low-level processes (i.e. word recognition and pronunciation) which leaves inadequate cognitive sources for higher-level processes (i.e. meaning construction). Hence, it can be proposed that numerals have the potential to adversely affect comprehension and fluency and L2 texts loaded with various forms of numbers need specific attention to ensure comprehension and ORF.

CONCLUSIONS AND PEDAGOGICAL IMPLICATIONS

This study reported eye-tracking results regarding numeral processing during L2 oral reading. L2 texts loaded with numerals were found to have been read slower in L2 oral reading due to regressions and rereads when compared to texts with familiar and unfamiliar words. Numeral instruction and research in L2 reading classrooms are neglected topics since numbers have been regarded as one of the beginner topics of L2 instruction. However, the current results showed that as the number of digits increases, processing times for numerals also increase affecting reading speed.

Although basic oral reading competency of numerals might be sufficient of English for general purposes, the results of the current study are especially important of English for Specific Purposes (ESP) which focuses on different arrays of professions such as engineering, accountancy, aviation, tourism, medicine, logistics and business covering various skills and disciplines (i.e. medical English, scientific English, technical English, English for waiters, English for tourism, English for finance). ESP instruction primarily values technical vocabulary acquisition (Evans and Morrison, 2011; Nation, 2013; Paltridge, Starfield, and Coxhead, 2017) and the estimated size of technical vocabulary in a specific text is around 30% (Chung and Nation, 2003). However, technical texts do not consist of only words. It is quite common to meet a technical text loaded with various forms of numbers such as percentages, arithmetic calculation, equations or figures of calculations, decimal numbers along with math symbols. These numeral forms are widely used in texts in the field of engineering, medicine, finance, banking, statistics, mathematics, and the like. Most of these professions require learners to design and perform oral presentations, reports, and briefings in English in their professional life along with general communication. For instance, the study by Estival and Molesworth (2016) not only examined lexical errors during communication between pilots and flight controllers but also reported some numeral errors or misunderstandings. Such experiences are quite common in international companies pursuing global goals and expecting their employees to produce L2 efficiently with fluency. Therefore, oral competency in reading numerals should be an important interest of L2 reading classrooms, especially in the ESP context. Although this study does not propose a specific instruction technique, it strictly advises enriching reading course materials with more numerical expressions which may help learners to promote their L2 oral reading rate in regular and technical texts.
Limitations and Future Research

During the experiments, the calibration process of some participants was challenging due to the device profile. A faster device would be more suitable and easier to operate for calibration processes.

Further research may be more specific with participants from ESP classrooms such as finance, banking, aviation, or engineering. In this way, the effect of numbers on L2 oral reading would be more specific for each field of ESP and it can be compared to learners of English for General Purposes.

REFERENCES


Rasinski, T. V., & Samuels, S. J. (2011). In S. J. Samuels & A. E. Farstrup (Eds.), *What research has to say about reading instruction* (pp. 94 –114). Newark, DE: IRA.


**APPENDICES**

Appendix 1. Sample Sentences

Was his beautiful Christine mad? (familiar word condition)
She went out of the room and opened the front door. (familiar word condition)
The dogs lived in shroot under the snow. (non-word condition)
It was a hot day and the swoills were open. (non-word condition)
But they travelled 28 kilometres. (numeral condition)
There are more than 2517 doors in the building. (numeral condition)

*Emrah Dolgunsöz* is an Assistant Professor and Head of the Foreign Languages Education Department at Bayburt University. He has been teaching and carrying on his research on reading, eye movements and second/foreign language learning and teaching.

Email: edolgunsoz@gmail.com

Orcid ID: http://orcid.org/0000-0002-1277-2177