Whole Language Method or Phonics Method for Better Reading? An Eye-tracking Study

Mine Yıldız
Department of English Language Education, Ataturk University, Erzurum, Türkiye
ORCID: 0000-0001-6215-4388

Meryem Özdemir Cihan
Department of Primary School Education, Ataturk University, Erzurum, Türkiye
ORCID: 0000-0002-6375-485X

Engin Kurşun*
Department of Computer Education and Instructional Technology, Ataturk University, Türkiye
ORCID: 0000-0002-5649-8595

Selçuk Karaman
Department of Management Information Systems, Ankara Hacı Bayram Veli University, Türkiye
ORCID: 0000-0002-0493-3444

Elif Akan
Department of Primary School Education, Ataturk University, Erzurum, Türkiye
ORCID: 0000-0001-9472-922X

Ömer Yılar
Department of Primary School Education, Ataturk University, Erzurum, Türkiye
ORCID: 0000-0002-3560-0028

Article history
Received: 09.08.2023
Received in revised form: 13.09.2023
Accepted: 25.10.2023

Key words: whole language method, phonics method, silent reading, oral reading, reading comprehension, prosody, eye movement

This study aims to compare the reading skills of individuals who learned reading through the Phonics Method (PM) and those who learned reading through the Whole Language Method (WLM) in terms of silent and oral reading performance and processes. A causal-comparative design was followed to reveal the causal relationship between the methods of teaching reading and individuals’ silent and oral reading processes and performance. The data were collected from 136 undergraduate students in Türkiye in the 2019-2020 academic year through two reading comprehension tests and prosodic assessment of oral reading skills to evaluate their L1 reading performance, and an eye-tracking system to evaluate their silent and oral reading processes. According to the teaching literacy method, no significant difference was obtained in the participants' silent and oral reading comprehension tests. However, the PM group

* Correspondency: ekursun@atauni.edu.tr
demonstrated better prosodic reading skills compared to the WLM group. Similarly, their silent and oral reading processes showed a significant difference in favour of the PM group. Therefore, it can be concluded that the readers who learned reading through the PM had a more favorable reader profile, particularly regarding their reading processes. This study provides direct and cogent support to the discussions in the literature on the effectiveness of these teaching methods.

Introduction

Reading is a skill that plays a vital role in individuals’ lives, from their academic to social lives and to be learned through formal teaching (National Reading Panel, 2000). It is a process of comprehending the text (McIntyre et al., 2011). Therefore, the method of teaching reading to be involved in this process as a basic component has great importance (Şenel, 2004) and also influences the lifelong learning of individuals remarkably. Hence, any mistake encountered in the teaching process of the reading skill may affect all individuals' academic and intellectual lives as much as the learning outcomes regarding this skill (Lonigan et al., 2008). The question of what the best method or approach is to teach reading shows no sign of stopping; instead, it is fraught with controversy (Brooks & Brooks, 2005; Krashen, 2002).

In teaching reading, various methods, including the word method, letter method, phonological method, and phoneme method, have been used and discussed to enhance the effectiveness of this process worldwide. Based on a process of inferring meaning from sentences, the WLM has been involved in the teaching literacy programs of many different countries and still attracts the focus of studies investigating the best method to teach reading. On the other hand, with the National Reading Panel (2000) report, the pendulum was directed to phonics instruction, which depends on teaching children to convert sounds into letters. Although there are several factors, including educational policies, politics, or economics in addition to teachers' decisions based on their own experiences, which have a significant influence on the choice of the method to teach reading, the pendulum goes on swinging forth and back, particularly between these two methods. The disagreement about teaching literacy methods seems to be unresolved. Therefore, the current study investigates the impact of two teaching literacy approaches, the PM and WLM, on individuals’ reading performance and processes using objective metrics such as reading comprehension tests, eye-tracking metrics, and prosodic reading evaluation to determine which teaching literacy method is more effective in enhancing readers' reading skills. Thus, it provides experimental evidence on the efficiency of these two teaching literacy methods and adds to a larger discussion of them.

Methods for teaching reading in Türkiye

Like many countries where the first language (L1) reading is taught through different methods such as the PM, the WLM, the whole word method, and analytic phonics instruction, Türkiye has witnessed similar debates on teaching this skill and thus applied different approaches and methods to teach it effectively. With the current change in the method of teaching reading in 2004 until when the WLM had been used to teach pupils reading, the focus of teaching reading was directed to the sounds rather than the words or sentences.

Based on the reports of the researchers supporting the view that the acquisition of word recognition in languages such as Turkish with a transparent orthography and having consistent phoneme-grapheme mappings is easier (Durgunoğlu & Öney, 1999, 2002; Öney & Durgunoğlu, 1997), the PM was proposed as the method of teaching literacy in the Primary
Education Curriculum (MNE, 2005). The previous method applied in teaching reading, the WLM, ironically, ignores these language characteristics. Although the transparent orthography of the Turkish language facilitates a systematic syllabification of the words and naturally their decoding, literacy instruction used to be based on a procedure in which sentences were first given to memorise, and, only after the whole sentence was memorised, its components such as individual words, syllables, and letters were identified (Durgunoğlu & Öney, 2002). Namely, following such a process in teaching literacy in Turkish seems to be inconsistent with the features of this language. According to the new Primary Education Curriculum (MNE, 2005), children are also required to be equipped with meta-cognitive skills such as analysing, synthesising, questioning, or reasoning in addition to acquiring reading skills (Aydın Yılmaz, 2009). In this sense, the new method of teaching reading, the PM, was believed to enhance learners’ cognitive and metacognitive skills and enable them to be literate as early as possible.

Phonological awareness having a critical role in the acquisition of reading contributes to the development of this metalinguistic skill and the general reading skill (Durgunoğlu & Öney, 1999). According to Öney and Durgunoğlu (1997), before comprehending how orthography reflects spoken language, a child must be familiar with the necessary units of spoken language. Similarly, the report of the National Reading Panel (2000) suggests that when children seek to communicate using written language, they will automatically learn that they need to understand letter-sound relationships and how letters operate in reading and writing and that phonics instruction should at least be integrated into the WLM. Moreover, it has also been reported that literacy instruction in languages with a particular alphabetic writing system should be based on phonics instruction (Hodgson et al., 2013).

Similarly, many studies investigating the new method for literacy instruction in Türkiye pointed out that the PM is more effective in literacy teaching. Whereas the WLM impedes the learning process of reading (Maviş et al., 2014), phonics instruction expedites decoding, helps accurate pronunciation of the words (Şahin et al., 2006), and thus facilitates the acquisition of literacy (Akıncı et al., 2016; Baydık & Kudret, 2012; Bıçak & Susar Kırmızı, 2013). On the other hand, several studies pointed out that the use of the WLM would enhance such skills as focusing on the whole (Maviş et al., 2014; Tok, 2001), fluent reading (Maviş et al., 2014; Şenel, 2004; Ş. Tok, 2001; Ş. Tok et al., 2008), meaningful reading (Akman & Aşkin, 2012), and also permanent learning (Tok, 2001). Nonetheless, the PM has been criticised for miscomprehension (Akıncı et al., 2016; Susar Kırmızı et al., 2012), misreading (Karadağ & Gültekin, 2007), and incorrectly separating syllables with incorrect addition (Bayat, 2014; Baydık & Kudret, 2012).

The choice of the method to teach reading should be given in line with many factors regarding the characteristics of the learners and the language (Aydın Yılmaz, 2009). Therefore, it is of paramount importance to investigate the methods of teaching reading in terms of different variables to identify the weaknesses and strengths of the methods. However, the studies comparing the two methods of teaching reading seem to lack providing experimental findings on the effectiveness of the methods; instead, they mainly depend on the self-reports of the stakeholders, including teachers, students, parents, educational researchers, administrators (Susar Kırmızı et al., 2012; Şenel, 2004). On that account, it is of great necessity to compare these two methods of teaching reading based on empirical studies.
Reading comprehension

Reading is a process of decoding words, but its fundamental purpose is to comprehend the author's message (Wagner et al., 2009). Reading comprehension is defined as the process of obtaining and producing meaning concurrently via interaction and participation with written language (RAND Reading Study Group, 2002). Many studies investigated reading comprehension in terms of different physiological, psychological, sociological, and demographic variables such as gender, anxiety, stress, motivation, reading strategies, reading habits, attitudes towards reading, vocabulary knowledge, working memory, social environment, educational background of the parents, text difficulty, reading mode (oral or silent), planning time (Freeland et al., 2000; Hale et al., 2011; James et al., 2021; Sok et al., 2021; Temizyürek, 2008; Yıldırım et al., 2010; Yıldız & Akyol, 2011).

The method used to teach reading is also believed to affect reading comprehension (Sağırlı, 2020), although those studies seem to disregard the role of literacy instruction in the development of reading comprehension. In this sense, Şahin (2011) compared two methods of teaching reading in terms of reading comprehension and oral reading speed of 5th-grade students and found that the teaching method had no significant effect on these variables. Likewise, Shone (2002), contrasting reading comprehension of the first-grade students learning reading through two different methods, found no significant difference in the students' reading comprehension test results in contrast to the researcher's hypothesis. In contrast, Sağırlı (2020) found a significant difference between the groups who learned reading through two different methods in favour of the students in the WLM group. However, those studies investigated the effects of the method for teaching reading on reading comprehension while teaching reading or after a short while. Therefore, comparing reading comprehension levels of individuals learning reading through two different methods is essential to see whether they have a significant long-lasting effect on their reading performance.

Prosody in reading

Prosody is the central piece of understanding the speech and consists of concepts such as stress and intonation to express the meaning (Breznitz, 2006). Though being a universal concept, it may show differences according to language since they have specific features of stress, intonation, and pause (Dowhower, 1991). Prosody, a reading fluency skill, is suggested to explain reading comprehension best (Baştuğ & Akyol, 2012) and also stated as reading like talking in which punctuation marks, stress and intonation are emphasised (Keskin & Akyol, 2014). Having fluency in oral reading and displaying prosodic features are considered essential skills (National Reading Panel, 2000) and are believed to be gained and improved, especially in childhood, contributing to being a fluent and prosodic reader during adulthood (Dowhower, 1991).

Kentner and Vasishth (2016) revealed a strong interaction between sentence rhythm and understanding the general context in silent and oral reading. Binder et al. (2013) also showed that skilled readers did not use pauses as frequently as less skilled readers. Furthermore, skilled readers decode the text accurately, with relatively few errors, and fluently. In addition, Miller and Schwanenflugel (2006) and Ravid and Mashraki (2007) correlated prosody with reading comprehension and showed that kids reading better had fewer and shorter hesitations in reading. Namely, readers with good reading prosody were seen to have better comprehension than expected (Baştuğ & Akyol, 2012; Benjamin & Schwanenflugel, 2010; Schwanenflugel et al., 2004).
Some studies also examined prosodic reading skills regarding reading fluency, accuracy and rate according to the teaching literacy methods (Ateş & Yıldız, 2011; Şahin, 2011; Şahin et al., 2006), but their findings seem to be confounding. For instance, whereas Ateş and Yıldız (2011) found that prosodic reading skills did not differ according to the method of teaching reading, according to Şahin et al. (2006), the phonics method helps readers recognise and pronounce the sounds accurately. Furthermore, these studies focused on the reading skills of primary school students shortly after they had learned reading. In this sense, it is equally important to investigate the effects of teaching literacy methods on the prosodic reading skills of individuals, particularly in the long run.

**Eye movements in reading**

Objective measurements are required due to the complex nature of reading. At this point, eye tracking systems offer unique and online measurements providing real-time, objective data based on actual behaviour by monitoring eye movements of readers rather than traditional tools primarily based on their self-reports (Bax & Weir, 2012; Rayner, 2009). According to Rayner and Slattery (2009), saccades (ballistic eye movements), fixations (brief pauses), and regressions (back to previous parts of a text) are involved among three significant components of eye movements during reading.

During reading, a reader moves from one place in the text to another, separated by pauses (fixations) that typically last 200 to 250 msec. However, the reader's eyes appear to glide smoothly across the page of text (Rayner & Slattery, 2009). Since vision is surpassed during saccades, the new information is encoded only during fixations. The reader may also make backwards saccades (regressions) between words (interword regressions) or within words (intraword regressions) (Zawoyski et al., 2015). Studies have indicated that good readers have fewer but shorter fixations during reading, whereas they have longer but fewer saccades (Rayner et al., 2006). Furthermore, less frequent regressions are also characterised as indicators of efficient reading (Ashby et al., 2005).

Some studies comparing the methods of teaching literacy proposed findings only based on participants' self-reports rather than monitoring eye movements through eye-tracking systems. For instance, Maviş et al., (2014) suggest that readers acquiring literacy through the WLM make longer and fewer saccades, achieve a better comprehension, and thus have a good reader profile. On the other hand, the readers learning reading through the PM are suggested to ignore the meaning, have shorter saccades but longer fixations, and, hence, be poor readers since they focus on only decoding the words (Akıncı, et al., 2016; Bayat, 2014; Susar Kırımızi et al., 2012; Yurdakal & Susar Kırımızi, 2013). However, such objective measures are unlikely to be considered based on the participants' self-reports; instead, they should be evaluated through eye-tracking systems. In addition, the studies contrasting methods of teaching literacy generally focus on similar variables such as phonics, phonological awareness, phonemic awareness, reading fluency, vocabulary, reading comprehension, writing, spelling, and reading (de Graaff et al., 2009; Sağırlı, 2020; Teale et al., 2018). Besides such variables reflecting the reading performance of readers, eye-tracking measures also provide valuable and objective data about their reading process (Rayner & Slattery, 2009). However, previous literature, particularly in the Turkish context, seems to suffer from those studies comparing the methods of teaching literacy via eye-tracking measurements, although they have a key role in reporting objective findings.

In order to address all these needs outlined above, the current study examines the effects of
the two teaching literacy methods, the PM and WLM, on the reading performance and processes of individuals based on objective measurements through reading comprehension tests, eye-tracking systems, and prosodic reading assessment. It thus provides experimental findings on the effectiveness of the teaching literacy methods and contributes to discussing the methods from a broader perspective. Furthermore, unlike the studies in the literature investigating the effects of teaching literacy methods on the reading skills of young learners, this study focuses on their long-lasting effects.

This study aims to compare the reading skills of individuals who learned reading through the PM and those who learned reading through the WLM in terms of silent and oral reading performance and processes. In line with the purpose of this study, the research questions to be answered are as follows:

- Do the methods of teaching reading have a significant effect on the reading performance of individuals?
- Do the methods of teaching reading have a significant effect on the reading processes of individuals?

Method

Research design

This study follows a causal-comparative design of quantitative methods that provided an opportunity to reveal the causal relationship between the methods of teaching reading and individuals’ silent and oral reading processes and performance. In line with the principles of The Law of Unification of Education (Tevhid-i Tedrisat Kanunu) and the Turkish Education System, the same teaching method is compulsorily used to teach literacy all over the country. Therefore, it does not seem possible to obtain a control or study group of students to teach reading through the WLM that fell into disuse in 2005 to conduct an experimental study.

Participants

The participants of this study consisted of undergraduate students in the Primary Education Department of a state university in Türkiye. The WLM has not been used since 2005 and the PM started to be used to teach reading in that year. Therefore, undergraduate students, third and fourth grades of whom mostly consisted of students having learned reading through the WLM and first and second grades of whom would be those having learned reading through the PM, were the most convenient group for this study to compare the two methods. The participants were selected from the same department where students were placed according to their achievement in a university entrance exam and thus had a similar academic background. Furthermore, before the study, a questionnaire was administered to all students in this department to select the participants elaborately and obtain a homogenous group.

In order to distribute the participants to the right group regarding the method of teaching reading, four different questions (the year of starting the first grade in primary education, the age of starting the first grade in primary education, and the year of birth) were involved in the demographic questionnaire. Their answers for the item asking “the method of teaching reading: (a) the PM, (b) the WLM” were then cross-checked with those for the other three questions since the PM has been obligatorily used to teach literacy in all primary schools.
since 2005. Furthermore, pictures involved in the questionnaire (as illustrated in Figure 1) and oral instructions were utilised to remind the participants, who were expected to know the methods as prospective primary school teachers to teach literacy in the future, the two methods of teaching reading.

15) By which of the following methods did you learn to read?

With "Ali, Look at Horse" reading slips [ ]  "aaaaa" with sounds like [ ]

Figure 1. Pictorial representations of the two methods of teaching literacy.

Accordingly, 356 students completed the questionnaire and those participants who gave various responses to the item asking about the methods of teaching reading (N=10) were first excluded. The consistency of their ages with the years when the methods were used was also examined. The participants under 20 (3) and at 21 (1) were not involved in the WLM and PM groups, respectively. In addition, those aged 27 and over were also excluded from the study to minimize the effect of age and thus their cognitive and academic development. Thus, the ages of participants in the WLM (M= 21.86) ranged from 20 to 25, and those of the PM group (M=19.04) ranged from 18 to 20. The mean score of the WLM group in the reading culture scale was 35.46 (Max. 49, Min. 20), whereas that of the PM group was 34.11 (Max. 48, Min. 20). Descriptive statistics for the ages and reading culture scores of both groups were also displayed in Table 1.

Table 1. Demographic features of participants.

<table>
<thead>
<tr>
<th></th>
<th>Mean</th>
<th>Min.</th>
<th>Max.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age WLM</td>
<td>21.86</td>
<td>20</td>
<td>25</td>
</tr>
<tr>
<td>Age PM</td>
<td>19.04</td>
<td>18</td>
<td>20</td>
</tr>
<tr>
<td>Reading Culture WLM</td>
<td>35.46</td>
<td>20</td>
<td>49</td>
</tr>
<tr>
<td>Reading Culture PM</td>
<td>34.11</td>
<td>20</td>
<td>48</td>
</tr>
</tbody>
</table>

In addition, as illustrated in Figure 2, both groups shared similar reading behaviours in terms of reading frequency.
149 volunteer participants who had appropriate features to participate in eye-tracking experiments (e.g. not wearing contact lenses) were involved in the experiments, but 13 were excluded due to the errors in their eye-tracking data. The Reading Comprehension Tests conducted immediately after eye-tracking experiments were taken by 136 students (24 male, 112 female) signing the informed consent form.

Some participants were not included in the analysis of some dimensions since their results distracted normality distribution. The number of students who were included or excluded in data collection and analysis was illustrated in Figure 3.
356 participants completed the questionnaire  
N=356

10 participants were excluded because of the early literacy method  
N=346

29 participants were excluded because of their age  
N=317

149 participants participated in eye tracking experiment  
N=149

13 participants were excluded due to erroneous eye-tracking data  
N=136

Eye-tracking data in silent reading  
N=136

Reading Comprehension Tests  
N=136

Prosodic Reading  
N=136

Eye-tracking data in oral reading  
N=136

SILENT  
After the normality distribution:  
N=132 for saccade counts  
N=132 for saccade duration  
N=131 saccade length  
N=132 for fixation counts  
N=131 for fixation duration  
N=131 regressions  
N=130 dwell time

ORAL  
After the normality distribution:  
N=134 for saccade counts  
N=131 for saccade duration  
N=116 saccade length  
N=134 for fixation counts  
N=132 for fixation duration  
N=133 regressions  
N=133 dwell time

Figure 3. The number of participants included in data collection and analysis.
Data collection tools

A questionnaire set

A questionnaire set consisting of three different questionnaires was conducted to obtain a homogenous study group regarding their reading skills. The first was the Demographic Information Inventory. The second, the Reading Behaviours Inventory, developed by researchers based on literature, included selective questions to identify participants' reading behaviours. The last was the Reading Culture Scale of Türkel et al., (2017).

Reading texts for eye-tracking experiments

Three different reading texts were used in the eye-tracking experiments and reading comprehension tests. The issues of the Science and Technique Journal by TUBITAK, which appeals to teenagers and adults, were examined, and 13 texts were first chosen considering participants' needs, desires, interests, and levels. These 13 reading texts on different topics were evaluated by nine experts in terms of linguistics, content, structure, and readability principles based on a rubric. Three of the texts were thus chosen according to their scores (398, 338 and 337 points) and comments on the texts. One of the three texts was used as a pilot text to help students understand the process and thus feel relaxed during experiments.

Reading comprehension tests (RCTs)

Three researchers were asked to prepare at least two open-ended questions for each text, one of which would measure the basic reading comprehension of learners whereas the other was intended to be an inferential question. In order to evaluate the reading texts and tests from different perspectives through the lens of participants, a pilot study was carried out with four students learning reading through the PM (N=2) and the WLM (N=2). As a result, six open-ended questions out of 18 prepared by the three researchers were included in RCTs based on the answers and comments of the pilot study participants. After some changes and revisions in the experiment process and texts (page layout, punto, instructions for readers), a new pilot study was conducted with eight students from both groups. Thus, the tests and texts were redesigned for the study.

The reading comprehension levels of participants were assessed through the sum of their scores in silent and oral reading comprehension tests. Although participants answered six questions regarding the two texts they read, the responses to the questions for the pilot test were not included in the assessment. Two raters assessed the RCTs according to an answer key involving keywords to be sought in the participants’ responses. Each question was decided to be evaluated over 10 points, and thus the highest score in each test would equal 20 points. Their scores were analysed for rater reliability through a Pearson correlation test (.82 for silent reading, .88 for oral reading, and .88 for general reading comprehension levels), the mean of the two raters' scores was taken for each question.

Prosodic reading scale

The “Prosodic Reading Scale” developed by Keskin et al. (2013) was used to assess students’ oral reading skills. The scale involves items probably directly or indirectly related to prosodic reading skills, such as punctuation, pitch, intonation, phonics features, reading rhythm, meaningful reading, and automatic word recognition in oral reading. The highest
Participatory Educational Research (PER), 11 (1):121-141, 1 January 2024

score on the scale consisting of 15 items assessed over four points is 60. The Cronbach’s alpha value of the scale was checked and found to be 0.98.

Having discussed unclear points in the scale, three raters first assessed 59 sound records chosen at random. The records in the analysis of which raters had more than a 15-point-difference were reassessed since a 15-point-difference was thought to be probable due to a -1 and +1 difference in scoring of each item. The rest were assessed by the two coders with similar scores in the preceding assessment and higher reliability. The rater reliability level in a Pearson correlation test was .71, and Cronbach’s Alpha value was calculated as 0.98.

Eye-tracking measures

Eye movement measurements were collected via SensoMotorics Instrument (SMI) Eye Tracking Software version 3.7. The sampling rate was 120 Hz, with a high accuracy of 0.5°. The software runs a 5-point calibration cycle with verification, programmed at the beginning of the experiment. A webcam with microphone was placed in front of the monitor screen and approximately 60 cm away from the participant. The texts were presented as PDF files on a monitor (monitor: Samsung) operated at a refresh rate of 60 Hz with a resolution of 1366 × 768 pixels. The Red 250 SMI eye tracking system was mounted on the monitor.

In order for the texts to be read more comfortably and to obtain more reliable data, the texts were set to 14pt, Times New Roman, and 1.5-line spacing. The pilot text used to help the participants adapt to the process consisted of 515 words, and the silent and oral reading texts contained 728 and 808 words, respectively.

The measures regarding fixations, saccades, and dwell time were exported from SMI BeGaze 3.7. version. Since eye-tracking software does not provide regressions for PDF files, regressions and saccade lengths were calculated based on fixation positions with Microsoft Excel 2016.

Procedure

Data were collected at two stages in the 2019-2020 academic years.

Stage 1

After being informed about the project, the participants were requested to complete the questionnaire set. A homogenous study group based on demographic features and reading behaviours in the questionnaire was thus obtained for the study.

Stage 2

Participants were individually invited to the experiment by appointment. They were requested to follow the procedure about head movements, sitting position, and monitoring to obtain accurate and reliable records. Prior to reading, calibration and validation of the eye-tracking system were first monitored and controlled during the whole experiment. Reading comprehension tests were conducted after the eye-tracking experiment.

Pilot experiment

Each participant first read the pilot text and was then asked to answer two questions on a paper-pen test given just after reading it. The process was paused until they completed answering.

Experiment 1- Silent reading
Similarly, they read the following text in silence and then answered the paper test involving two open-ended questions.

**Experiment 2- Oral reading**

Lastly, they read the third text orally and completed its comprehension test. The participants’ voices in oral reading were recorded.

The data collection environment is illustrated in Figure 4.

![Experimental setting](image)

**Figure 4. Experimental setting.**

**Data analysis**

Independent samples t-tests were conducted to see whether there were any differences in the reading performance and reading processes of readers who learned reading through the WLM and PM.

**Results**

The reading performance of participants was considered regarding their silent and oral reading comprehension levels and prosodic reading skills. Their eye movement behaviours in oral and silent reading were examined to reflect their silent and oral reading processes. Thus, they were compared to see whether the method used to teach them literacy led to a significant difference in their silent and oral reading performance and processes.
**Reading performance**

Reading comprehension levels of participants in silent and oral reading and their prosodic reading skills in oral reading were contrasted according to the method via which they learned reading.

**Reading comprehension**

The results of reading comprehension tests were compared to see if there is a significant difference in reading comprehension levels of participants in oral and silent reading, independently and totally, according to the method of teaching literacy, and displayed in Table 2.

Table 2. Reading comprehension levels of participants.

<table>
<thead>
<tr>
<th></th>
<th>N</th>
<th>Mean</th>
<th>SD</th>
<th>t</th>
<th>df</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Silent Reading Comprehension Test</td>
<td>WLM</td>
<td>78</td>
<td>11.42</td>
<td>3.55</td>
<td>1.11</td>
<td>134</td>
</tr>
<tr>
<td></td>
<td>PM</td>
<td>58</td>
<td>10.73</td>
<td>3.52</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Oral Reading Comprehension Test</td>
<td>WLM</td>
<td>78</td>
<td>10.74</td>
<td>4.13</td>
<td>.901</td>
<td>134</td>
</tr>
<tr>
<td></td>
<td>PM</td>
<td>58</td>
<td>10.11</td>
<td>3.92</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Reading Comprehension Levels</td>
<td>WLM</td>
<td>78</td>
<td>22.16</td>
<td>6.30</td>
<td>1.23</td>
<td>134</td>
</tr>
<tr>
<td></td>
<td>PM</td>
<td>58</td>
<td>20.84</td>
<td>6.04</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

According to Table 2, it is clear that the WLM group (M=11.42, SD=3.55) had slightly higher success in the silent RCT than the PM group (M=10.73, SD=3.52). However, such a slight difference between the two groups was not statistically significant, t=1.11, p>.05. Similarly, the method used to teach literacy had no significant effect on the oral reading comprehension levels of participants, t=.901, p>.05. In addition, the reading comprehension levels of participants were defined by the total scores they obtained in the two tests and assessed over 40 points. The difference in their total reading comprehension levels (M= 22.16, SD= 6.30 for the WLM group, M= 20.84, SD= 6.04 for the PM group) was also not at a significant level, t=1.23, p>.05. Both groups acquiring literacy through two different methods had similar success in the oral and silent RCTs. Thus, it can be said that the method of teaching literacy had no significant effects on the reading comprehension levels of individuals.

**Prosodic reading performance**

The results of the tests conducted to identify whether there was a significant difference in the oral reading skills of the participants were displayed in Table 3.

Table 3. The independent t-test results for prosodic reading.

<table>
<thead>
<tr>
<th></th>
<th>N</th>
<th>Mean</th>
<th>SD</th>
<th>t</th>
<th>df</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Prosodic reading</td>
<td>WLM</td>
<td>78</td>
<td>24.97</td>
<td>9.52</td>
<td>-3.24</td>
<td>134</td>
</tr>
<tr>
<td></td>
<td>PM</td>
<td>58</td>
<td>30.30</td>
<td>9.44</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

As shown in Table 3, the prosodic reading levels of the PM group (M=30.30, SD=9.44) was higher than those of the WLM group (M=24.97, SD=9.52) at a significant level, t=-3.24, p<.05. Namely, those who learned reading through the PM could be said to have better prosodic reading performance.

**Reading process**

Participants' silent and oral reading processes were observed through an eye-tracking system and then assessed using eye-tracking measures such as saccade counts, saccade
duration, fixation counts, fixation duration, regressions, saccade length, and dwell time. Eye-movement behaviours of the two groups were compared to see whether the method used to teach them reading led to significant differences in their silent and oral reading processes.

Descriptive results and the independent samples t-tests for silent and oral reading processes were displayed in Table 4.

Table 4. The independent t-test results for silent and oral reading processes.

<table>
<thead>
<tr>
<th></th>
<th>N</th>
<th>Mean</th>
<th>SD</th>
<th>t</th>
<th>df</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Saccade duration (sec)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>WLM</td>
<td>78</td>
<td>20.25</td>
<td>8.10</td>
<td>8.23</td>
<td>96.94</td>
<td>.000</td>
</tr>
<tr>
<td>PM</td>
<td>54</td>
<td>12.20</td>
<td>2.50</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Fixation duration (sec)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>WLM</td>
<td>77</td>
<td>253.94</td>
<td>48.12</td>
<td>0.83</td>
<td>129</td>
<td>.408</td>
</tr>
<tr>
<td>PM</td>
<td>54</td>
<td>246.01</td>
<td>61.02</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Dwell time (sec)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>WLM</td>
<td>78</td>
<td>267.65</td>
<td>51.07</td>
<td>2.62</td>
<td>128</td>
<td>.108</td>
</tr>
<tr>
<td>PM</td>
<td>52</td>
<td>243.71</td>
<td>50.85</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Fixation counts</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>WLM</td>
<td>78</td>
<td>1224.01</td>
<td>284.19</td>
<td>6.37</td>
<td>128.6</td>
<td>.000</td>
</tr>
<tr>
<td>PM</td>
<td>54</td>
<td>970.06</td>
<td>172.45</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Saccade counts</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>WLM</td>
<td>78</td>
<td>1147.71</td>
<td>287.91</td>
<td>5.82</td>
<td>128.2</td>
<td>.000</td>
</tr>
<tr>
<td>PM</td>
<td>55</td>
<td>914.35</td>
<td>173.31</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Regressions</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>WLM</td>
<td>76</td>
<td>44.32</td>
<td>22.94</td>
<td>4.53</td>
<td>128.51</td>
<td>.000</td>
</tr>
<tr>
<td>PM</td>
<td>55</td>
<td>29.07</td>
<td>15.54</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Saccade length</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>WLM</td>
<td>61</td>
<td>80.51</td>
<td>14.57</td>
<td>-6.97</td>
<td>99.04</td>
<td>.000</td>
</tr>
<tr>
<td>PM</td>
<td>56</td>
<td>103.46</td>
<td>20.30</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Saccade duration (sec)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>WLM</td>
<td>76</td>
<td>23.34</td>
<td>8.77</td>
<td>8.867</td>
<td>100.19</td>
<td>.000</td>
</tr>
<tr>
<td>PM</td>
<td>55</td>
<td>13.63</td>
<td>3.19</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Fixation duration (sec)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>WLM</td>
<td>74</td>
<td>398.67</td>
<td>37.44</td>
<td>.843</td>
<td>130</td>
<td>.401</td>
</tr>
<tr>
<td>PM</td>
<td>58</td>
<td>392.96</td>
<td>40.06</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Dwell time (sec)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>WLM</td>
<td>75</td>
<td>414.78</td>
<td>35.83</td>
<td>1.664</td>
<td>131</td>
<td>.098</td>
</tr>
<tr>
<td>PM</td>
<td>58</td>
<td>403.31</td>
<td>43.64</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Fixation counts</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>WLM</td>
<td>78</td>
<td>1456.01</td>
<td>291.31</td>
<td>6.520</td>
<td>132</td>
<td>.000</td>
</tr>
<tr>
<td>PM</td>
<td>56</td>
<td>1153.41</td>
<td>222.97</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Saccade counts</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>WLM</td>
<td>78</td>
<td>1323.06</td>
<td>300.5</td>
<td>7.11</td>
<td>131.83</td>
<td>.000</td>
</tr>
<tr>
<td>PM</td>
<td>56</td>
<td>1011.07</td>
<td>207.38</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Regressions</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>WLM</td>
<td>78</td>
<td>56.37</td>
<td>31.76</td>
<td>5.49</td>
<td>126.57</td>
<td>.000</td>
</tr>
<tr>
<td>PM</td>
<td>55</td>
<td>32.42</td>
<td>18.30</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Saccade length</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>WLM</td>
<td>59</td>
<td>76.21</td>
<td>7.40</td>
<td>-9.388</td>
<td>75.66</td>
<td>.000</td>
</tr>
<tr>
<td>PM</td>
<td>57</td>
<td>99.36</td>
<td>17.13</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Silent reading process

According to Table 4, there is a significant difference in saccade duration in favour of the PM group (M=12.20, SD=2.50), t= 8.23, p<.05. However, the slight difference in fixation duration for the WLM group (M=253.94, SD=48.12) and the PM group (M=246.01, SD=61.02) was statistically non-significant, t=0.83, p>.05. The WLM group (M=267.65, SD=51.07) spent more time reading the text than the PM group (M=243.71, SD=50.85). The dwell time measured through fixation and saccade duration differed significantly in the two groups, t=2.62, p<.05. Furthermore, the counts of fixations made by the WLM group (M=1224.01, SD=284.19) were also found higher than those of the PM group (M=970.06, SD=172.45), at a significant level, t=6.37, p<.05. The WLM group (M=267.65, SD=51.07) spent more time reading the text than the PM group (M=243.71, SD=50.85). The dwell time measured through fixation and saccade duration differed significantly in the two groups, t=2.62, p<.05. Furthermore, the counts of fixations made by the WLM group (M=1224.01, SD=284.19) were also found higher than those of the PM group (M=970.06, SD=172.45), at a significant level, t=6.37, p<.05. Similarly, the WLM group (M=1147.71, SD=287.91) made more saccades than the PM group did (M=914.35, SD=173.31) at a significant level, t=5.82, p<.05. The number of regressions made by the WLM group (M=44.32, SD=22.94) significantly differed from that of the PM group (M=29.07, SD=15.54), t=4.53, p<.05. It is also clear that the difference in saccade length is statistically

Participatory Educational Research (PER)
significant, \( t = -6.97, p < .05 \). The PM group (\( M = 103.46, SD = 20.3 \)) made longer saccades than the WLM group (\( M = 80.51, SD = 14.57 \)). According to these results, it is clear that the PM provided readers with an advantage in the silent reading process and thus helped them have a good reader profile.

**Oral reading process**

According to Table 4, it is also clear that the WLM group (\( M = 23.34, SD = 8.77 \)) spent more time in saccades than the PM group did (\( M = 13.63.67, SD = 3.19 \)) at a significant level, \( t = 8.867, p < .05 \). The difference in the fixation duration between the two groups (\( M = 392.96, SD = 40.06 \) for PM, \( M = 398.67, SD = 37.44 \) for WLM) was slight and non-significant, \( t = .843, p > .05 \). Likewise, the difference in the dwell time of the groups was not significant, \( t = 1.664, p > .05 \), although participants in the WLM completed their oral reading task in a longer time (\( M = 414.78, SD = 35.83 \)). On the other hand, the WLM group (\( M = 1456.01, SD = 291.31 \)) made more fixations than the PM group (\( M = 1153.41, SD = 222.97 \)) did at a significant level, \( t = 7.11, p < .05 \). The WLM group (\( M = 1323.06, SD = 300.5 \)) had similar results for the saccade counts and made more saccades than the PM group did (\( M = 1011.07, SD = 207.38 \)) at a significant level, \( t = 6.520, p < .05 \). Regarding the regressions during oral reading, the WLM group (\( M = 56.37, SD = 31.76 \)) made more regressions than the PM group did (\( M = 32.42, SD = 18.30 \)) at a significant level, \( t = 5.49, p < .05 \). In addition, the PM group (\( M = 99.36, SD = 17.13 \)) had more saccade length at a significant level, \( t = -9.388, p < .05 \). Based on these results, it can be said that participants who learned reading through the PM had more effective oral reading processes and also a good reader profile.

**Discussion**

This study was conducted to contrast the participants who learned reading through the WLM with those who learned reading through the PM in terms of their silent and oral reading performance and processes. Reading performance in this study was evaluated in the dimensions of silent and oral reading comprehension and oral prosodic reading skills. In contrast to the studies pointing out that the WLM helps readers comprehend the text better (Karadağ & Gültekin, 2007; Maviş et al., 2014; Ş. Tok, 2001; Ş. Tok, Tok, & Mazı, 2008) and those suggesting that the phonics method help readers understand what they read better (Arslantaş & Cinoğlu, 2010; Şahin et al., 2006), no significant difference was found in either silent or oral reading comprehension of participants according to the method through which they acquired the reading skill.

The experimental studies had varying findings on the relationship between the methods of teaching reading and reading comprehension. For instance, in contrast to this study, Sağırıl (2020), investigating two methods of teaching reading— the WLM and the PM—in terms of reading skills of 5th-grade students, had a significant difference in their reading skills in favour of the students in the whole language group. Nonetheless, Şahin (2011) and Shone (2002) had consistent findings with the current study and showed no significant difference in participants’ reading performance. This study also supported the findings of Kuşdemir Kayıran and Karabay (2012) pointing out that the reading comprehension levels of the fifth-grade students who learned reading through different methods did not significantly differ.

On the other hand, according to the results of this study regarding prosodic reading skills, those who learned reading through the phonics method used holistic features such as intonation-stress, emotion, phonics features, reading rhythm, grouping, and punctuation features more successfully. This study seems to support the studies showing that the phonics
method helps readers recognise and pronounce the sounds accurately (Şahin et al., 2006). However, Ateş and Yıldız (2011) found contrary findings showing that prosodic reading skills did not differ according to the method of teaching reading.

Poor readers are more likely to pause when encountering commas, whereas good readers do not pause to separate a list of words but pause when the comma separates two of them. Binder et al. (2013) suggest that good readers read with fewer hesitations while reading. Likewise, according to Miller and Schwanenflugel (2006), the students who read better have fewer and shorter hesitations in reading while less successful students generally pause in and between the sentences. Based on those studies, individuals who learned reading through the phonics method in this study can be said to be more successful in terms of prosodic skills and thus have a good reader profile. In line with the findings of this study, the method of teaching literacy can be said to affect the oral reading performance of individuals regarding prosodic reading skills although it does not influence their reading comprehension levels in either oral or silent reading. Namely, the phonics method contributes to the oral reading performance of readers in terms of prosody and thus helps them gain a good reader profile.

Based on these findings of the study and the findings of those studies suggesting that good readers have fewer fixations and saccades but longer saccades and less fixation duration (Rayner et al., 2006; Zawoyski et al., 2015), the readers in the PM group can be said to have a better reading profile since they made fewer saccades, fixations and regressions and naturally spent less time in those metrics while reading a text in silence and orally as well as having shorter dwell time. Furthermore, based on Ashby et al. (2005), the WLM group can also be described as poor readers since they had more frequent regressions that adversely affected their silent and oral reading processes and led to the elongation of dwell time.

However, these findings seem to contradict the findings of Maviş et al., (2014), suggesting that the readers who acquired reading through the WLM make longer but fewer saccades, achieve a better comprehension, and thus have a good reader profile. Furthermore, these findings did not support those studies indicating that the readers learning reading through PM have shorter saccades but longer fixations (Akıncı et al., 2016; Bayat, 2014; Susar Kırmızı et al., 2012; Yurdakal & Susar Kırmızı, 2013). Nonetheless, these findings are suggested to be discussed with caution in the light of those studies since such objective measures considered in those studies are remarkably based on only self-reports of the participants.

This study comparatively investigated the two methods of teaching reading and found differences in some dimensions of reading. It points out that the PM has some advantages over the WLM, particularly in terms of the reading process and prosodic reading skills. Such an advantage can be predicated by the fact that the teaching process of reading begins with the recognition of the sounds, then combining the sounds and thus constructing meaningful syllables, words, and sentences respectively, which helps readers process the knowledge they are exposed to. Since the PM encourages readers to focus on what they see, they may thus make less saccades, fixations, and regressions. On the other hand, the struggle with word identification, recognition, and recalling the schema in the mind, which is the focus of the WLM, may lead readers to make more fixations and regressions.

**Conclusion and Suggestions**

This study provides practitioners and researchers with evidence that the phonics method is more effective in developing a better reading profile. It also provides direct and cogent
support to the discussions in the literature on the effectiveness of these methods. Based on the findings of this study obtained through objective measurements and pointing out that the PM gains the edge over the WLM for better reading performance and silent reading process, the use of the PM is suggested to be sustained as the method to teach Turkish students reading in their language. Furthermore, primary teachers who still have some prejudices towards the effectiveness of the PM in teaching reading are also suggested to consider these findings based on objective measurements rather than being stuck on their self-reports.

According to the findings of this study illustrating that the readers learning reading through the PM had a better reader profile, it can be suggested to maintain the use of the PM to teach students to learn reading in Turkish and to teach reading through the PM in languages with transparent orthographic features like Turkish. In addition, teachers who advocate the WLM and have prejudices against the PM are suggested to consider the objective findings of this study rather than depending on the self-reports. These findings should also be considered in teaching reading in those languages with consistent phoneme-grapheme mappings such as Spanish, Finnish, Latin, and Italian in addition to Turkish. As suggested by Durgunoğlu and Öney (1999), the PM facilitates word recognition, particularly in transparent orthographies in which graphemes and phonemes have one-to-one relationships and words have a consistent spelling and thus the process of teaching reading. A similar conclusion was also drawn in the National Reading Panel (2000) for the languages with relatively deep orthographies such as English that phonics instruction has proved more effective in developing reading skills of children.

This study, however, has some limitations. The data were collected from students at the tertiary level, and there is a significant difference between when they learned reading and when their reading performance and process were assessed. It could not be allowed and ethical to teach children reading through a method different from that in the curriculum; therefore, the undergraduate students sharing similar characteristics, particularly in terms of the reading skill, but having a difference in the method of teaching reading were included in the study. Furthermore, due to the limitations in applying eye-tracking systems in the studies with children, undergraduate students were preferred.

As for the suggestions for researchers, the reading processes can be carried out in the natural environment with mobile eye-tracking tools. The studies can be conducted to determine and evaluate the factors affecting reading comprehension. The instruments providing objective measurement of the prosodic skills can be developed. The results can be compared by using different teaching reading methods for illiterate adults. The teaching methods can be compared in different languages with transparent or non-transparent orthographic features. Although this study points out some differences in the reading processes of readers who learned reading through different teaching methods of reading, further studies investigating the reasons leading to such differences in detail are required. It is remarkably suggested to study the relationship between the cognitive processing of words and the PM.

References

-137-


