The implementation of an intensive second language reading intervention program for Arabic learners of English as a foreign language

Keywords

Reading comprehension, reading instruction, Arabic learners, eye-tracking, vowel blindness, textual enhancement, phonological awareness, spelling, word recognition, reading fluency

The aim of this study was to investigate the effects of focused reading interventions on the eye movement patterns and reading scores of second language (L2) learners. We compared two groups of proficiency matched Arabic L2 learners before and after an intensive reading intervention program. Two classes received a barrage of reading interventions and two classes served as comparison groups. The intervention program consisted of textual enhancement, phonological awareness training, training in word recognition, spelling instruction and oral text reading fluency. Analysis of reading scores showed a significant improvement over time, but no significant impact of the intervention. Both the intervention and comparison groups showed improvement in their eye tracking measures, but no significant difference in gains between the intervention and the comparison groups was found. The study provides a building block for future studies using focused pedagogical interventions with this particular group of learners.

Introduction

There have been numerous studies which indicated that students from an Arabic first language (L1) background often have difficulty learning to read in English as a second language (L2) (see, for example, Abu-Rabia, 1997; Fender, 2003, 2008; Hayes-Harb, 2006; Randall, 2007; Randall & Groom, 2009; Randall & Meara, 1988; Ryan, 1997; Ryan & Meara, 1991; Saigh & Schmitt, 2012; Thompson-Panos & Thomas-Ruzic, 1983). One of the reasons for these reading problems has been assumed to be related to transfer from Arabic L1 reading processes whereby learners focus their attention more on consonant letters and process English vowels inefficiently (Ryan & Meara, 1991). However, word-level decoding difficulties might also be the cause of L2 reading problems for Arabic EFL learners as recently demonstrated by the eye-tracking study of Alhazmi, Milton and Johnston (2019).

Although the literacy-related difficulties of Arabic EFL learners are relatively well documented, very few experimental studies have been conducted to examine how specific pedagogical interventions can facilitate their L2 reading development. More research is also needed in classroom contexts that reflects how teachers use input enhancement in the real world (Winke, 2013). Furthermore, as textual enhancement (TE) alone might result in small learning effects only (cf. Lee & Huang's (2008) meta-analysis), it is also important to examine how TE in combination with more explicit teaching techniques might contribute to L2 reading development. The aim of the present study was to investigate the effects of focused reading interventions on the eye movement patterns and overall reading proficiency of college students from an Arabic L1 background. In this experimental study we examined the impact of an intensive reading intervention programme. The strength of our experimental study which was conducted in an

authentic classroom context is the parallel investigation of measurable reading development and potential changes in eye-movements that reflect cognitive processes involved in L2 reading.

The use of eye-tracking in L2 reading research

Eye-movements can provide a window into reading processes as readers' eyes move along the text in saccades, which are eye movements between pauses (fixations), and skip back to parts of the text (regressions) (for a review see Rayner, 2009). Fixations of eye-movements on a given point in a text indicate where readers' attention is allocated (Rayner, 1998). One can also investigate eye-movements and fixation durations on particular areas of interest to gain insights into whether specific elements of a text require more conscious and effortful processing or re-reading (Holmquist et al., 2011).

Previous eye-tracking research on L2 reading has been scarce. In a pioneering study comparing Dutch-English L2 readers of low-intermediate to advanced proficiency, Cop et al. (2015) found that the bilingual readers demonstrated longer fixations and processed the text in shorter saccades in their L2 than in their L1 but did not show different regression rates across their two languages. Brunfaut and McCray (2015) investigated how L2 readers' eye-movements differ as they engage in reading tasks at different levels of proficiency and how the L2 proficiency of their participants is related to various eye-tracking measures. Their results indicated that as the difficulty of the reading task increased, L2 readers fixated longer on the text and the reading task items, regressed more often and had higher number of forward saccades. In their study, higher proficiency learners demonstrated shorter fixation durations on the text and the items faster. A recent study comparing reading only with reading and listening to a text simultaneously by Conklin

et al. (2020) revealed that L2 readers fixated longer on low-frequency and long words than L1 readers, but fixation durations on high-frequency words were similar for native and non-native speaker participants. However, except for Oakley (2018) (see below), eye-tracking methodology has not been used yet to explore the challenges of Arabic L1 speakers in reading L2 English texts.

Difficulties experienced by Arabic L1 speakers reading in English

Grabe (2009) suggested that among the differences that influence how L2 readers process a text are the various orthographies that visually represent the phonological and morphological systems of each language. Writing systems differ on two dimensions: orthographic representation and depth. Orthographic representation refers to the linguistic unit each graphic symbol denotes. Orthographic Depth (Katz & Frost, 1992) refers to the degree of sound-symbol correspondences. An orthography that closely represents the phonology in a clear one-to–one relationship is called a shallow or transparent orthography. An orthography that does not closely represent the phonology is called deep or opaque. With regards to Arabic, which is the L1 of the EFL participants of this study, vowelized Arabic is considered to be a shallow orthography while unvowelized Arabic is described as a deep orthography (Abu-Rabia & Siegel, 2003). English is often considered as having one of the most opaque orthographies in the world. The distance between English and Arabic in terms of both depth and orthographic representation has been noted by numerous researchers (e.g., Fender, 2003, 2008; Hayes-Harb, 2006; Randall & Meara, 1988; Ryan & Meara, 1991). Skilled Arabic readers have learned to use an orthography that does not include diacritic markers to signal short vowels. Therefore, Arabic readers tend to focus their

attention on the consonants (Hayes-Harb, 2006; Ryan, 1997; Ryan & Meara,1991; Thompson-Panos & Thomas-Ruzic, 1983).

Numerous studies have been published on the difficulties exhibited by native speakers of Arabic when reading in English. In an early study, Ryan and Meara (1991) found that performance by native Arabic speakers was less accurate and slower than performance by EFL learners with non-Arabic native language backgrounds. In particular, the native speakers of Arabic had a higher error rate in judging deleted vowel stimuli than the two other participant groups. Randall and Meara (1988) observed that Arabic L1 learners of English rely heavily on consonants when attempting to recognize an English word and they concentrate on the position of the consonants at the beginning, middle and end of words rather than on the position of vowels. They concluded that "the search functions of Arabic L1 learners were radically different from those readers whose script uses the Roman alphabet" (p.144). This transfer from processing Arabic orthography impeded processing of English script in a phenomenon called "vowel blindness" (Ryan & Meara, 1991). Hayes-Harb (2006) replicated Ryan and Meara's (1991) study and found that native speakers of Arabic were less accurate in detecting vowel letters than consonants and performed lower in the vowel detection task than the two comparison groups from different language background. Based on the results, Hayes-Harb hypothesized that "native Arabic speakers transfer visual word processing strategies from Arabic to reading in English. Fender's (2003, 2008) findings indicated that Arabic ESL students experience 'more difficulty' than other ESL populations in processing English word forms. He maintained that difficulties acquiring English spelling knowledge not only affect word recognition skills but also constrain reading skills (Fender, 2008). In a precursor eye-tracking study, Oakley (2018) found that the reading patterns displayed by an Arabic L1 group of L2 learners were characterized by many long fixations, many

short forward saccades and numerous regressions, which are characteristic of the patterns displayed by beginning and poor readers as well as dyslexics (Ashby, Rayner & Clifton, 2005; Chase, Rayner & Well 2005; Rayner, 1998; Rayner et al., 2006). These word-level decoding problems did not seem to have been caused by 'vowel blindness' as the Arabic L1 EFL participants made a significantly higher proportion of visits on vowels than a comparison group of English L1 participants. The Arabic L1 EFL learners also fixated longer on both consonants and vowels than their L1 speaking counterparts. More recently, Alhazmi et al.'s (2019) study also confirmed these findings and led the authors to speculate that rather than vowel-blindness, incomplete mental representations of vocabulary that primarily rely on phonological rather than orthographic information might cause word level decoding problems (cf. also Perfetti, 1997).

As Hayes-Harb (2006) noted, understanding the specific word recognition problems Arabic L1 learners of English experience has the potential to help teachers develop effective strategies for teaching these learners to read in English. In addition, if the learners are consciously aware of their word identification difficulties, it is possible that conscious strategies may help them to process English words more efficiently.

Reading interventions and instruction

In its comprehensive report, The National Reading Panel in the US (National Reading Panel, 2000) recognized reading fluency as one of the essential elements to be considered in reading instruction. They recommend different types of reading interventions such as computer assisted instruction, phonics instruction, fluency practice, use of text comprehension strategies and numerous others. In this study, we implemented five interventions, the choice of which was based upon current literature and the first author's twenty years' experience teaching Arabic L1 learners

of English. In what follows we will give a brief overview of previous research on the effectiveness of these intervention methods.

Textual Enhancement (TE) is an implicit focus-on-form technique, through which learners' attention is drawn to a language form during an otherwise meaning-focused interaction. TE uses visual enhancement methods such as underlining, bold facing, colour-coding, italicizing, capitalizing or using different fonts as a means to promote the processing of linguistic items (Sharwood Smith, 1991). The perceptual salience created by highlighting the input is intended to draw the learners' attention to the form and, once the first step is successful, learning of the attended form is expected to occur, based on the premise that attention is what mediates input and intake (Izumi, 2002; VanPatten, 2015). In an eye tracking study, Alsadoon and Heift (2015) explored the impact of textual input enhancement on the noticing and intake of vowels by 30 female Saudi Arabian students studying in various ESL schools in Canada. Their results showed that the experimental group spent more time on the target word forms than the comparison group. The authors suggested that the TE drew the experimental group's visual attention to the target forms and their vowels, "implying that the treatment significantly improved the learners' orthographic vowel knowledge and thus reduced their vowel blindness" (p. 69). Their findings, however, need further corroboration because the Areas of Interest (AOIs) were drawn on individual words and not on vowels or consonants.

Another possible way of developing students' word-level reading skills is by phonological awareness (PA) training. PA is the ability to recognize that words are made up of a variety of sound units. A large quantity of research has demonstrated the beneficial effects of PA instruction on reading for English first language children (e.g., Ehri et al., 2001). The National Reading Panel (NRP) (2000) in its meta-analysis examined whether PA instruction was significantly more

effective than alternative forms of training in helping children acquire PA and enabling them to apply this skill in their reading. Results showed that PA instruction is beneficial under different teaching conditions with a variety of learners. Research has indicated that PA instruction has positive effects for ESL and EFL learners as well (e.g., Lesaux & Siegel, 2003).

Training in word recognition and automaticity is another potential method to help develop word-level reading skills. However, findings regarding the benefits of this type of training are contradictory. Fukkink, Hulstijn and Simis (2005)'s pioneering study with Dutch L1 EFL learners did not show a significant improvement in reading comprehension after two days of word recognition training. However, Akamatsu's (2008) results with Japanese EFL learners indicated that the participants benefitted from word-recognition training in speed and accuracy.

A further means of improving students' word-reading skills is training in oral text reading. Fluent reading is the ability to read a text quickly, accurately, and with proper expression. It combines accuracy, automaticity and oral reading prosody and is a factor in both oral and silent reading which can support text comprehension (e.g. Grabe, 2009; Rasinski & Samuels, 2011). The National Reading Panel (2000) concluded that guided oral reading procedures had a significant and positive impact on word recognition, fluency and comprehension across a range of grade levels. In the L2 field, Jiang et al.'s (2012) results demonstrated that oral passage reading fluency was strongly linked to text comprehension.

Word recognition skills can also improve as the quality of the spelling knowledge in the orthographic lexicon develops (e.g. Ehri, 2005). Perfetti (1997) observed that spelling and reading appear to be both sides of the same coin. Incomplete or inaccurate spelling representations or knowledge will result in less efficient and in some cases, less accurate word recognition skills (e.g., Ehri 2005). Ehri (2005) proposed that spelling contributes to L1 reading development by

improving phonemic awareness and strengthening the student's understanding of the alphabetic principle. Furthermore, in the field of L1 reading research, the orthographic knowledge of word forms has also been found to contribute to reading development by enriching the depth of lexical representation (Perfetti, 1997). Spelling is a language skill supported by several linguistic knowledge sources, including phonemic, orthographic and morphological knowledge and requires active consideration of the sounds, patterns and meaning of written language (e.g., Masterson & Apel, 2000).

Rote learning approaches provide little or no instruction for the development of the linguistic sources of knowledge that support spelling. However, research has demonstrated significant improvement in spelling when one or more of these underlying linguistic sources of knowledge are taught (e.g., Apel & Masterson, 2001). A multilinguistic spelling approach provides practice in phonemic and orthographic awareness skills and involves activities that teach the strategy of segmenting words into their individual phonemes and then linking each sound to a letter.

Spelling of words is particularly valuable for ESL and EFL students. When students listen to spoken words while inspecting their spelling, more precise representations are secured in memory (cf. Perfetti, 1997). Unfortunately, very few studies have been conducted on the effects of ESL/EFL spelling interventions, although research with young children has shown that English spelling knowledge and English word-reading skills are closely related (e.g., Wade-Woolley & Siegel, 1997). August (2011) has demonstrated that spelling abilities jointly with vocabulary knowledge contribute to second language reading comprehension. In a recent study O'Connor, Geva and Koh (2019) have also found that bilingual children in Canada experiencing difficulties

with text comprehension are characterized by poor phonological awareness and below average orthographic skills.

The current study

In an eye-tracking study comparing Arabic L1 EFL students and English L1 participants, Oakley (2018) found that the Arabic L1 EFL students displayed eye movements which were similar to poor readers and dyslexic readers. These findings were taken as an indication of processing difficulties with letters and problems with word-level decoding. Therefore, a classroom-based intervention study was designed to improve students' word and text-level reading skills. The study addressed the following research questions:

1. Do focused intervention activities enhance the overall reading comprehension of Arabic L1 EFL students?
2. Do focused intervention activities change the eye movement patterns of Arabic L1 EFL students while reading at the word and sentence level in English?

The specific interventions were activities focusing on rapid word recognition and automaticity, phonemic awareness, oral prosody, spelling and the use of TE (Table 1 lists the eye-tracking measures used and the hypotheses about development).

Table 1. Eye-tracking measures and hypotheses about development

Measures & Definitions	Hypotheses
Fixation The period of time during which the eye remains relatively stationary and reflects when information is being encoded.	
Number of Fixations	As ability increases, the number of fixations decreases (Holmqvist, 2011; Rayner, 1998)

Average Fixation Duration	As ability increases, the average duration of fixations decreases (Holmqvist, 2011; Rayner, 1998)
Saccade A high velocity jump executed by the eye to bring information into foveal vision.	
Number of Forward Saccades	As ability increases, the number of saccades decreases (Hyönä & Olson, 1995; Rayner, 1998; Rayner et al., 2006)
Length of Forward Saccades	As ability increases, the length of saccades increases (Rayner, 1998)
Regressions "In order for a saccade to be a regression, the saccade needs to move in the opposite direction to the text but not necessarily in the opposite direction to the previous saccade". Holmqvist, 2011, p. 263	
Number of Regressions	As ability increases, the number of regressions decreases (Ashby, Rayner & Clifton, 2005)
Length of Regressions	As ability increases, the length of regressions increases (Rayner, 2019; Dussias, 2010)
Mean of all visits on consonants	Mean of all visits on consonants will decrease as a result of the intervention
Mean of all visits on vowels	Mean number of visits on vowels will be higher as a result of the intervention.
Proportion of vowel visits (number of visits on vowels divided by visits on all letters)	The proportion of visits on vowels will be higher after the intervention.

Method

The study had an experimental design with pre-treatment and immediate post-treatment eye movement recordings and a pre-test and immediate post-test of overall reading comprehension (see Figure 1).

Insert Figure 1 around here

Participants

Participants were a sample of 39 Arabic L1 EFL students. Students were enrolled in five sections of an English for Academic Purposes (EAP) course, designed to prepare students for the School of Business. The course was at a B1 level of the Common European Framework of Reference (CEFR) (2001) as specified by the college curriculum. All students followed the same instructional programme in English through their secondary and university years, and hence no significant variation in English language proficiency across groups was expected. There were 28 females and 11 males. Twenty participants were in the 18-20 year range, 15 in the 21-24 range and 4 in the over- 24 range. They were all Qatari nationals who had gone through the same educational system in Qatar. The participants had normal or corrected-to-normal vision.

The study included two intact classes in the treatment group (N=20) and three intact classes in the comparison group (N=19). The interventions were carried out as part of the standard curriculum in students' regular FL classes. The classes for all groups were 17 hours per week over the regular semester which consisted of 12 weeks. Reading instruction, as outlined in the course syllabus, accounted for 40% of the course content and 40% of the final exam grade.

Five teachers participated in the study. All five were EFL instructors, each with over 20 years of EFL teaching experience. All had been teaching in the Middle East for at least 10 years.

Three were male and two were female. The teachers who taught the two treatment groups were aware of the exact nature of the research, while the teachers of the three comparison groups were only told that the research was investigating how Arabic L1 EFL students read in English.

Instruments

Eye-tracking materials

Participants were presented with 16 images which were common everyday signs such as street signs, shop signs, school signs etc. Participants were instructed to look at each sign, and when they were ready, they were to press the space bar, whereupon a new screen would appear with two sentences written on it (see Appendix 1 for examples of the signs and their two corresponding sentences). Participants were asked to silently read the two sentences and tell the researcher which sentence referred to the sign they had just read. Participants could not go back to the previous screen. Sentences were typed in black in a 36 font on a pale yellow background in Courier New. Each pair of sentences was matched as closely as possible for length and word length was held as constant as possible. The sentences were written so that they would not contain words that might be unknown to the Arabic L1 EFL participants. Flesch-Kincaid Grade Level scores were calculated and these provided confirmatory evidence that the pairs of sentences were of comparable readability levels. The same sentences were used as pre- and post-tests.

Reading comprehension tests

The reading comprehension pre-test consisted of a reading text of 750 words on the topic of counterfeit goods being sold in the Gulf. It had a Coh-Metrix L2 readability (Graesser et al., 2004; cohmetrix.memphis.edu) of 7.67, consisted of 35 sentences, with a mean sentence length

of 21.4 and was 6 paragraphs in length. There were five items which required students to identify the main idea in the text, five items which asked students to specify the meaning of vocabulary in the context of the text and five items which assessed understanding of pronoun references. Fifteen items tapped into the understanding of specific information presented in the text. The same assessment was given to participants as a post-test. Participants were not given any feedback or results from the pre-test and as the time lapse between administrations was 12 weeks, we expected little or no carryover of correct response bias or priming. All items were scored as either right or wrong (1 or 0).

Intervention materials

Participants in all groups were given 28 texts to read in class in the same order during the semester. However, the treatment groups received the textually enhanced versions which had all vowels printed in red and in bold (TE). The comparison groups were given texts typed in customary black with no bolding. The comprehension questions following the texts were identical for both groups. The 28 texts, which were part of the in-house materials for the EAP course, were based on the following themes: 1) Recruitment and Retention 2) Managing a Business 3) Information Technology and E-Commerce 4) Marketing Strategies, 5) Conducting Business in a Global Economy 6) Counterfeiting and 7) Employee Compensation. Included in the 28 texts were also several texts on study skills, Qatarisation and Qatari culture.

Tracking activities were also meant to assist students in the treatment group with processing vowels. In this type of activity the students are given a stimulus word at the beginning of a horizontal line of words. Six words follow across the line and differ from the stimulus word by only one letter, which is a vowel (**bat**: bet, but, bit, bat, bit, bat). Students have to underline or

circle the words which are the same as the stimulus word (see Supplementary materials B for examples). In addition to encouraging students to pay attention to the vowels in each word, the activities also served to reinforce the left-to-right directionality while reading in English. The exercises increased in complexity from single words to phrases of three to four words as the semester progressed.

Lists compiled from the first 1,500 words of the New General Service List (NGSL) (Browne, Culligan & Phillips, 2013) were created and given to participants (see Supplementary materials C for an example). As Qian (1999) found that students had a higher rate of retention for decontextualized word lists than for contextualized ones, participants received isolated words in list form. These lists were used as the basis for the Spelling and Rapid Word Reading component of the intervention implemented with the treatment groups and for the Word Meaning component with the comparison groups. The Rapid Word Recognition entailed students taking three columns of 45 words from the NGSL word list and reading them aloud and timing themselves on their mobile phones or watches.

The NGSL list was also used for the spelling intervention. Each week, the teachers of the treatment groups would take one page from the NGSL List. On Day 1, the teachers read aloud the words in the first column on the page. They demonstrated the syllabification and word stress as well as giving spelling rules as applied to specific words in the column (for example, hard and soft 'c', hard and soft 'g', words with a final silent 'e' etc. Students then read one word each going around the class. They were told to study the words for homework. The following day there was a spelling quiz. The same procedure was followed on Days 2-4. On Day 5, there was a quiz on 45 words from the three columns chosen by the teacher. There was no instruction on word meanings as there was in the comparison group.

In order to improve students' text comprehension, teacher read-aloud was combined with relay reading. Relay reading is a type of oral text reading activity in which one student reads one sentence and then the student next to him or her reads the next sentence going around the class. For each of the texts, the teachers first read the text aloud with the students listening and following the words on their papers. Then a relay reading activity follows. As students should be given opportunities to re-read texts, re-reads occurred at the end of each week with students using the pair reading technique.

PA instruction did not have a specific time allotted but arose naturally out of daily spelling, reading and writing activities (see Supplementary materials D for examples of phonemic awareness activities). Activities which are commonly used to teach and improve phonological and phonemic awareness are a) phoneme identification b) phoneme count c) phoneme deletion d) syllable identification e) syllable deletion and f) rhyme.

Procedures

The eye-tracking recordings were made in the first week of the semester before any reading instruction had taken place. Recording sessions, including instructions, took approximately 40 minutes. A 9-point calibration session was carried out before each participant was recorded. Reading materials for eye-tracking were presented on a 23-inch monitor with a screen resolution of 1920 x 1080 pixels. The monitor was attached to an HP Z400 Workstation PC interfaced with a Tobii TX 300 eye tracking system which is equipped with a large head movement box. Each participant sat between 63 and 65 cm from the monitor in a comfortable stationary chair.

The reading pre-test was administered during the first week of classes before any reading instruction had taken place. Tests were administered by the classroom teachers on the same day

during class time. Participants were given one hour to complete the test (see Appendix 2 for the intensity and timing of the intervention).

Participants in the treatment groups were given 20 tracking activities over the course of the semester. These exercises took from five to eight minutes and were usually completed at the beginning of class and served as a warm-up activity before actual reading instruction took place. Participants in the comparison group completed warm-up activities consisting of word searches using words from the NGSL lists, but they did not do the tracking activities. These word searches contained words displayed only in a left-to-right direction and did not contain words displayed vertically, diagonally or in a right-to-left direction as is common in most word search puzzles. This was to ensure that the treatment groups were not potentially disadvantaged because of the directionality of the words given to them.

Both the experimental and comparison groups were given words from the first 1,200 words in the NGSL list at the beginning of each week and were asked to study them as homework. Participants in the treatment groups were given a spelling quiz each week from the 100 words. The comparison groups were given the same NGSL list as the treatment groups each week but were given activities with meaning-focused instruction without a spelling component.

All participants in the treatment groups engaged in rapid word recognition activities. The same word list from the NGSL was used. These activities were completed during class time at the beginning, middle and end of each week to measure any improvement in recognition rate. Participants in the comparison groups were given activities and exercises in the PHRASal Expressions List of 505 of the most frequent non-transparent multiword expressions in English intended especially for receptive use (Martinez & Schmidt, 2012). These activities in the comparison group were added to address the possible ethics concern that the students belonging

to this group might be disadvantaged in terms of general language skills development in comparison with the treatment group.

As mentioned previously, pair reading, relay reading and teacher reading activities were used for the 28 texts read in the treatment groups. The comparison groups read the same 28 texts but participants read individually and silently as was the regular practice at the course at the college.

Analysis

The effect of the treatment on test scores was analysed using a 2-way repeated-measures ANOVA. The within-subject variable was *time*, (pre-and post-tests), and the between-subject variable was the *group* (treatment or comparison). The assumption of normality was not met in the case of some variables. However, ANOVA is fairly robust in terms of violations of the assumptions of normal distribution when sample sizes are equal (Schmider, Ziegler, Danay, Beyer, Bühner, 2010). No corrections for multiple hypotheses were made given the nature of this study which was an exploratory rather than confirmatory study.

After all the recording sessions were completed, the quality of the eye-tracking data was investigated by visual inspection of the gaze plots for the 16 sets of sentences for all participants. Data from three participants which clearly did not represent their eye movements faithfully was removed. The recordings of the participants in this phase of the study were all of high integrity. The general viewing behaviour in the gaze plots provided unexpected patterns. It appeared that the participants made many regressions and forward saccades within the sentences. However, these visual observations needed a more in-depth analysis. This led to a re-examination of the metrics intended for statistical analyses and it was decided that in addition to fixation measures, metrics

pertaining to forward saccades and regressions would be included. The software used, Tobii Studio version 2. X did not calculate forward saccades and regressions; therefore, the raw data for the areas of interest (AOIs) around the sentences were exported into R (2012) to ascertain if there were any significant differences between the two groups of participants in terms of both their fixation and saccadic patterns. Fixation data from the AOIs around the vowels and consonants were calculated in Tobii Studio 2. X. The AOIs for this study were drawn around each of the 32 sentences. Another set of AOIs was drawn around each vowel and each consonant in every sentence so that a separate analysis could be conducted. When capturing letter level data compromise is required, the larger the text the more certain we are in assigning a fixation on a particular letter. However, we also wanted to preserve ecological validity and keep the text small enough such that it elicited the same kind of eye movements and processing that of smaller text, the majority of text we read. The manufacturers of the eye-tracker claim a gaze accuracy (i.e., average distance between the measurement and true location of the fixation) of 0.5º degrees, equivalent to 5.5mm at a distance of 65cm from the screen. Given our vowel level AOIs were 1.5cm by 1.5 cm (e.g., Holmqvist, 2011), we felt that we had sufficiently large AOIs to correctly diagnose the location most of the fixations while maintaining text small enough to elicit ecologically valid eye movement. Holmqvist et al. (2011) note that if the stimulus is simple and the AOIs are so close together that the readers can take in one AOI in peripheral vision while looking at the other, it is questionable to contrast dwell times from the two areas and claim that the visual intake from one AOI is larger than from the other. However, a phenomenon known as "crowding" explains that "as peripheral information becomes more cluttered, it is difficult to distinguish between different elements away from the current point of fixation. Therefore, for complex displays, AOIs which are close together may not cause a problem because crowding

restricts focus to the fovea" (Holmqvist et al, 2011, p. 217). Another issue which needed to be addressed was the space or margin between the AOIs. Adjacent AOIs should have sufficient spacing between them to allow the desired balance of specificity and selectivity by applying the "1 degree" guideline. When designing stimuli and AOIs, the "1 degree" guideline is an easy way to take into account the extent of the foveal field (Holmqvist et al., 2011). Considerable attention was therefore paid to the positioning of the AOIs. Exact positioning is crucial because it can determine whether a significant effect is revealed or not. AOIs "should not overlap because of the danger that single AOI hits and transitions will be counted twice, rendering the statistics difficult if not impossible to interpret" (Holmqvist et al., 2011, p. 221). Appendix 3 includes a list of eye-tracking measures, their explanations and the hypothesized direction of development.

Insert Figure 2 around here

Results

Our first research question asked whether Arabic L1 EFL students' reading comprehension develops as a result of a focused classroom intervention. Table 2 below presents a summary of the pre-and immediate post-test reading scores calculated for the treatment and comparison groups. The results of the repeated measures ANOVA reveal a significant increase with a large effect size in the total reading scores over time. The increase in total score is most likely due to improvement in detecting the main ideas and identifying pronoun reference in the text as these were the scores

that also showed significant changes with large effect size from the pre- to the post-test. There was only one significant difference, with a medium effect size, across groups and this was for the main idea scores, with the treatment group achieving higher means. No interaction effect between time and group was observed suggesting that the treatment did not result in significantly greater gains in the reading comprehension test than regular classroom instruction.

Our second research question aimed to find out whether the focused classroom intervention activities result in any changes in eye-movement patterns. The repeated measures ANOVA tests showed that the number of regressions and the proportion of saccades that are regressions decreased significantly over time, but no significant group effects and no interaction between time and group were detected (see Table 3). Table 4 includes information on eye-fixation measures relating to vowels and consonants. The repeated measures ANOVA test revealed that participants spent significantly longer proportion of time looking at vowels at the end of the study than at the beginning. Overall the comparison group fixated on vowels for a larger proportion of time than the treatment group. The interaction between group and time was near the significance level with a moderate effect size, which indicates that the comparison group maintained its initially higher fixation proportion on vowels, whereas the comparison group increased its initially lower proportion.

Table 2 Descriptive statistics and time, group and interaction effects of reading test scores

Feature	Group	Mean	SD	F time *p* η^2	F group *p* η^2	F interaction *p* η^2
Total Pre	Comparison	11.79	4.00	8.27	0.77	0.981
	Treatment	12.21	3.57	.006	.383	.385
Total Post	Comparison	13.21	3.96	.171	.019	.019
	Treatment	14.58	2.97			
MI Pre	Comparison	1.88	0.95	7.17	4.98	2.78
	Treatment	2.11	1.15	.011	.031	.103
MI Post	Comparison	2.08	1.21	.149	.108	.063
	Treatment	3.00	0.94			

VC Pre	Comparison	1.96	1.08	2.72	2.51	1.13
	Treatment	2.21	1.18	.107	.121	.294
VC Post	Comparison	2.08	1.41	.062	.058	.027
	Treatment	2.79	2.79			
PR Pre	Comparison	2.50	1.29	13.38	0.30	0.86
	Treatment	2.21	1.18	.001	.725	.357
PR Post	Comparison	3.00	1.38	.249	.002	.018
	Treatment	3.05	1.03			
RS Pre	Comparison	5.45	2.41	0.47	0.05	0.33
	Treatment	5.68	2.60	.493	.945	.567
RS Post	Comparison	6.04	2.23	.012	.000	.008
	Treatment	5.73	2.25			

*MI-Main Idea, VC-Vocabulary in Context, PR-Pronoun Reference, RS-Reading for specific information

Table 3. Descriptive statistics and time, group and interaction effects of eye-tracking measures

Feature	Group	Mean	SD	F time *p* η^2	F group *p* η^2	F interaction *p* η^2
NFix Pre	Comparison	473.21	154.60	3.56	0.00	0.32
	Exp.	460.50	146.10	.069	.989	.578
NFix Post	Comparison	420.29	60.62	.106	.001	.000
	Exp.	431.89	120.55			
NSac Pre	Comparison	300.00	93.36	2.85	0.03	0.23
	Exp.	298.22	98.25	.102	.860	.639

NSac Post	Comparison	270.00	35.42	.087	.001	.007
	Exp.	281.39	91.21			
NReg Pre	Comparison	172.57	64.95	6.42	0.03	0.19
	Exp.	166.67	50.24	.017	.869	.675
NReg Post	Comparison	149.93	35.11	.176	.001	.006
	Exp.	150.50	45.27			
	Exp.	0.36	0.07			
FixDur Pre	Comparison	119.54	31.96	1.70	0.09	0.00
	Exp.	112.67	37.05	.203	.765	.983
FixDur Post	Comparison	128.64	39.96	.023	.003	.009
	Exp.	118.00	42.53			
SacLength Pre	Comparison	210.71	37.53	0.25	0.09	0.00
	Exp.	206.81	44.70	.625	.765	.983
SacLength Post	Comparison	212.71	21.86	.008	.003	.009
	Exp.	208.64	42.74			
RegSacLgth Pre	Comparison	-152.07	36.73	1.62	0.24	0.27
	Exp.	-160.61	34.82	.212	.625	.605
RegSacLght Post	Comparison	-148.75	39.88	.024	.008	.008
	Exp.	-152.67	38.79			

* NFix-Number of fixations, NSac-Number of forward saccades, NReg-Number of regressions, FixDur-Fixation duration; SacLength-Length of Saccades, RegSacLgth – length of saccades that are regressions

Table 4. Descriptive statistics and time, group and interaction effects of letter visits.

Feature	Group	Mean	SD	F time *p* η^2	F group *p* η^2	F interaction *p* η^2
Vow Pre	Comparison	37.92	12.38	2.571	0.429	0.240
	Treatment	34.36	13.21	0.119	0.517	0.628
Vow Post	Comparison	33.83	9.83	0.79	0.014	0.008
	Treatment	32.18	13.38			
Con Pre	Comparison	43.21	17.01	8.100	0.193	0.240
	Treatment	46.37	14.47	0.008	0.633	0.628
Con Post	Comparison	37.37	11.21	2.13	0.006	0.008
	Treatment	38.11	13.77			
PropV Pre	Comparison	0.47	0.06	6.04	7.43	3.559
	Treatment	0.42	0.04	0.020	0.011	0.069
PropV Post	Comparison	0.48	0.03	0.168	0.198	0.106
	Treatment	0.45	0.04			

Discussion and conclusion

The present study was an experiment comparing two groups of Arabic L1 EFL students before and after focused reading interventions in the classroom. It assessed changes in reading behaviour and performance by analysing eye movements and using pre-and post-reading tests. The results of the study do not provide evidence that the treatment that included rapid word recognition and automaticity, phonemic awareness, prosody and spelling and the use of TE was more effective than regular classroom instruction either in terms of enhancing participants' L2 reading comprehension (RQ1) or bringing about changes in reading processes that are reflected in eye-tracking measures (RQ2). However, both the treatment as well as the students' regular classroom instruction resulted in relatively large improvement of reading comprehension scores and in the reduction of the number of regressions while reading. Particularly, students' higher order reading skills as assessed by items on the content of main ideas in the text and pronoun reference developed. The eye-tracking measures indicate that students had to reread parts of the text less frequently upon completion of the course. The observed decrease in the number of regressions has

been linked to enhanced reading ability in previous L1 reading research (Ashby, Rayner & Clifton, 2005). Furthermore, the proportion of time students' attention was devoted to vowels also increased substantially over time, indicating a development in the automaticity in the orthographic and phonological processing of vowels.

One of the reasons for the lack of evidence for the differential impact of the treatment might be the relatively low sample size of the research. It is also possible that statistically significant gains in reading proficiency may require interventions of considerable greater intensity than that provided in this study. Another possible reason might be that some of the comparison group's activities i.e. word searches, NGSL list with meanings, and the PHRasal Expressions List (which provided multiple exposures to the vocabulary) may also have had a facilitating effect which was not substantially different from the activities of the treatment group.

TE constituted a key element of the intervention programme, yet it has not been found to be effective in enhancing students' attention to vowels and did not contribute to L2 reading development. It is possible that the lack of significant impact of the intervention may be due to the enhancement itself. In hindsight, highlighting all of the vowels may have been overwhelming for the participants. Enhancing only one vowel letter per reading might have been more effective and might have helped the students to distinguish between the vowels. The results of the study are in line with Leow and Martin's (2018) recent overview which showed that to date almost 80% of studies comparing TE conditions to unenhanced conditions have failed to demonstrate significant learning gains for TE. However, the results of our study are in contrast to those of Alsadoon and Heift (2015) which showed that vowel blindness was significantly reduced for the treatment group due to a longer focus on target words as suggested by their eye tracking data.

PA training was also part of the treatment the treatment group received. Previous studies in the field of L2 learning have found PA training beneficial for L2 speaking children (Lesaux & Siegel, 2003; Yeung et al., 2013). Nonetheless, it has been shown that PA develops only up to a certain age. According to the Simple View of Reading (Gough & Tunmer, 1986) word decoding and word development cease to make a contribution to reading comprehension after a certain age. Therefore, PA training might have been less effective in enhancing L2 reading comprehension for the older participants of our study.

Our results regarding the lack of benefits of word-recognition training do not support Akamatsu's (2008) findings obtained with Japanese university students. Despite previous positive results shown by the National Reading Panel (2000), oral reading fluency training did not contribute to the development reading comprehension or changes in eye-movements during reading in our study either. The relatively small sample size and brevity of the intervention may be possible reasons for these findings.

Our study did not demonstrate a significant contribution of spelling training to reading comprehension or the development of word-level reading skills as assessed by the eye-tracking measures either. There are several possible reasons for this surprising result, which is contradictory to previous findings by Graham et al. (2002) for ESL children. As there were no pre-and post-tests of spelling in any of the groups, it is not possible to ascertain if any significant progress was made in spelling by the treatment groups. Additionally, the brevity of the intervention and the low number of words which were included (the first 2,000 of the NGSL) may have been factors in the results.

The study has several limitations that are inherent in experimental studies conducted in classroom contexts. Firstly, there was a sampling limitation, both in terms of participant numbers

and in terms of the participants being a convenience sample. However, in classroom research such as this, it is not possible to have access to large numbers of participants who are randomly chosen. Another limitation is the length of time that participants are involved in the study. When conducting research in an educational institution, the researcher has no comparison over the length of time students spend in class or the length of terms and semesters. Thus, the time over which the pedagogical interventions could be implemented was short and progress was not detectable when using quantitative measures. Lastly, again because of institutional restraints, it was not possible to conduct delayed post-tests which may have given insight into any long-lasting learning gains. A further limitation is that no pre-treatment or post-treatment assessments were administered to assess rapid word recognition, PA, spelling or oral text reading. Such evaluations would have given an insight into the gains that were made in the individual skill sets that comprise the reading process. Considerable improvement may have taken place in some or all of the skills but this could not be not demonstrated. Future research should be conducted with a larger sample and preferably over a longer period. Follow-up studies should also assess development in lower-order reading skills such as word-level decoding and sentence-level comprehension.

References

Abu-Rabia S. (1997). Reading in Arabic orthography: The effects of vowels and context on reading accuracy of poor and skilled native Arabic speakers. *Reading and Writing*, *9*, 65-78. doi:10.1023/A:1007962408827

Abu-Rabia, S. & Siegel, L. (2003). Reading skills in three orthographies: The case of trilingual Arabic–Hebrew–English-speaking Arab children. *Reading and Writing: An Interdisciplinary Journal, 16*, 611-634. doi: 10.1023/A:1025838029204

Akamatsu, N. (2008). The effects of training on automatization of word recognition in English as a foreign language. *Applied Psycholinguistics*, *29*,175-193. doi:10.1017/s0142716408080089

Alhazmi,K., Milton, J. & Johnston, S .(2019) Examining 'vowel blindness' among native Arabic speakers reading English words from the perspective of eye-tracking. System 80, 235-245. doi:10.1016/j.system.2018.12.005

Alsadoon, R. & and Heift, T. (2015). Textual input enhancement for vowel blindness: A study with Arabic ESL learners. *The Modern Language Journal, 99*, 57-79. doi:10.1111/modl.12188

Apel, K. & Masterson, J. (2001). Theory-guided spelling assessment and intervention: A case study. *Language, Speech, and Hearing Services in the Schools, 32,*182 -195. doi:10.1044/0161-1461(2001/017)

Ashby, J., Rayner, K., & Clifton, C. (2005). Eye movements of highly skilled and average readers: Differential effects of frequency and predictability. *Quarterly Journal of Experimental Psychology, 58A*, 1065-1086. doi:10.1080/02724980443000476

August, G. (2011). Spelling facilitates good ESL reading comprehension. *Journal of Developmental Education*, *35*(1), 14.

Browne, C., Culligan, B. & Phillips, J. (2013). The New General Service List. http://www.newgeneralservicelist.org/

Brunfaut, T., & McCray, G. (2015). *Looking into test-takers' cognitive processes while completing reading tasks* (ARAGs Research Reports Online No. AR/2015/001). British Council.

Chase, K., Rayner, K. &Well, A. (2005). Eye movements and phonological parafoveal preview: Effects of reading skill. *Canadian Journal of Experimental Psychology, 59*, 209-217. doi:10.1037/h0087476

Conklin, K., Alotaibi, S., Pellicer-Sánchez, A., & Vilkaitė-Lozdienė, L. (2020). What eye-tracking tells us about reading-only and reading-while-listening in a first and second language. *Second Language Research*, *36*, 257-276. doi.org:10.1177/0267658320921496

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: 1–38. doi:10.1371/journal.pone.0134008

Council of Europe. (2001). *Common European framework of reference for languages: Learning, teaching, assessmen*t. Cambridge, U.K: Press Syndicate of the University of Cambridge.

Ehri, L. (2005). Learning to read words: Theory, findings, and issues. *Scientific Studies of Reading, 9,* 167–188. doi:10.1207/s1532799xssr0902_4

Ehri, L, Nunes, S., Willows, D., Schuster, B., Yaghoub-Zadeh, Z. & Shanahan, T. (2001). Phonemic awareness instruction helps children learn to read: Evidence from the National Reading Panel's meta-analysis. *Reading Research Quarterly, 36,* 250–287. doi:10.1598/rrq.36.3.2 =

Fender, M. (2003). English word recognition and word integration skills of native Arabic and Japanese-speaking learners of English. *Applied Psycholinguistics*, *24*, 289-315. doi:10.1017/s014271640300016x

Fender, M. (2008). Spelling knowledge and reading development: Insights from Arab ESL learners. *Reading in a Foreign Language*, *20*, 19-42.

Fukkink, R., Hulstijn, J. & Simis, A. (2005). Does training in second language word recognition skills affect reading comprehension? An experimental study. *The Modern Language Journal* 89(1), 54–75. doi:10.1111/j.0026-7902.2005.00265.x

Gough, P. & Tunmer, W. (1986). Decoding, reading and reading disability. *Remedial and Special Education, 7*, 6-10. doi:10.1177/074193258600700104

Grabe, W. (2009*). Reading in a second language: Moving from theory to practice.* New York, NY: Cambridge University Press.

Graham, S., Harris, K. & Chorzempa, B. (2002). Contribution of spelling instruction to the spelling, writing, and reading of poor spellers. *Journal of Educational Psychology, 94*, 669-686. doi:10.1037/0022-0663.94.4.669

Hayes-Harb, R. (2006). Native speakers of Arabic and ESL texts: Evidence for the transfer of written word identification processes. *TESOL Quarterly, 40*, 321-339. doi:10.2307/40264525

Holmqvist, K., Nyström, M., Andersson, R., Dewhurst, R., Jarodzka, H., & Van de Weijer, J. (2011). *Eye tracking: A comprehensive guide to methods and measures*. Oxford: Oxford University Press.

Izumi, S. (2002). Output enhancement and the noticing hypothesis. *Studies in Second Language Acquisition, 24*, 541-577. doi:10.1017/s0272263102004023

Jiang, X., Sawaki, Y. & Sabatini, J. (2012). Word reading efficiency and oral reading fluency in ESL reading comprehension. *Reading Psychology, 33*, 323–349. doi:10.1080/02702711.2010.526051

Katz, L. & Frost, R. (1992). The reading process is different for different orthographies: The orthographic depth hypothesis. In R. Frost & L. Katz (Eds.), *Orthography, Phonology, Morphology, and Meaning* (pp. 67-84). Amsterdam, NL: Elsevier North Holland Press.

Lee, S. & Huang, H. (2008). Visual input enhancement and grammar learning: A meta-analytic review. *Studies in Second Language Acquisition, 30*, 307-331. doi:10.1017/s0272263108080479

Leow, R., & Martin, A. (2018). Enhancing the input to promote salience of the L2. In P. Spinner, S. M. Gass, & J. Behney (Eds.), *Salience in second language acquisition* (pp. 167–186). New York: Routledge.

Lesaux, N. & Siegel, L. (2003). The development of reading in children who speak English as a second language. *Developmental Psychology, 39,* 1005-1019. doi:10.1037/0012-1649.39.6.1005

Martinez, R. & Schmitt, N. (2012). A phrasal expressions list. *Applied Linguistics, 33,* 299-320. doi:10.1093/applin/ams010

Masterson, J. & Apel, K. (2000). Spelling assessment: Charting a path to optimal instruction. *Topics in Language Disorders, 20*, 50-6. doi:10.1097/00011363-200020030-00007

National Reading Panel. (2000). *Teaching children to read: An evidence-based assessment of the scientific research literature on reading and its implications for reading instruction.* Retrieved from http://www.nichd.nih.gov/publications/nrp/report.cfm

Oakley, J. M. (2018). *Coming from different directions: A comparison of the eye movements of English L1 and Arabic L1 speakers reading in English and the implementation of an intensive reading intervention programme* [Unpublished doctoral dissertation]. Lancaster University.

O'Connor, M., Geva, E., & Koh, P. W. (2019). Examining reading comprehension profiles of grade 5 monolinguals and English language learners through the lexical quality hypothesis lens. *Journal of Learning Disabilities*, *52*(3), 232-246. doi:10.1177/0022219418815646

Perfetti, C. (1997). The psycholinguistics of spelling and reading. In C. Perfetti, L. Rieben, & M. Fayol (Eds.), *Learning to spell: Research, theory, and practice* (pp. 21-38). Mahwah, NJ: Lawrence Erlbaum Associates.

Plonsky, L. (2015). Statistical power, p values, descriptive statistics, and effect sizes: A "back-to-basics" approach to advancing quantitative methods in L2 research. In: Plonsky, L, (Ed.) *Advancing quantitative methods in second language research***.** (pp. 23-45). Routledge: New York.

Qian, D. (1999). Assessing the role of depth and breadth of vocabulary knowledge in reading comprehension. *The Canadian Modern Language Review, 56*, 282–307. doi:10.3138/cmlr.56.2.282

Randall, M. (2007). *Memory, psychology and second language learning*. Amsterdam, NL: John Benjamins.

Randall, M., & Groom, N. (2009). The BUiD Arab learner corpus: A resource for studying the acquisition of spelling. Retrieved from http://ucrel.lancs.ac.uk/publications/CL 2009/54_FullPaper.doc.

Randall, M. & Meara, P. (1988). How Arabs read Roman letters. *Reading in a Foreign Language*, *4,* 133-145.

Rasinski, T. & Samuels, S. (2011). Reading fluency: What it is and what it is not. In S. Samuels & A. Farstrup (Eds.), *What research has to say about reading instruction* (pp. 94-113). Newark, DE: International Reading Association.

Rayner, K. (1998). Eye movements in reading and information processing: 20 years of research. *Psycholinguistic Bulletin, 124*, 372-422. doi:10.1037/0033-2909.124.3.372

Rayner K (2009). Eye movements and attention in reading, scene perception, and visual search. *Quarterly Journal of Experimental Psychology, 62,* 1457–1506. doi:10.1080/17470210902816461

Ryan, A. (1997). Learning the orthographical form of L2 vocabulary-a receptive and a productive process. In N. Schmitt & M. McCarthy (Eds.), *Vocabulary: Description, acquisition and pedagogy* (pp. 181-198). Cambridge, UK: Cambridge University Press.

Rayner K., Reichle, E., Stroud, M., Williams, C. & Pollatsek, A. (2006). The effect of word frequency, word predictability, and font difficulty on the eye movements of young and elderly readers. *Psychology and Aging, 21*, 448-465. doi:10.1037/0882-7974.21.3.448

Ryan, A., & Meara P. (1991). The case of the invisible vowels: Arabic speakers reading English words. *Reading in a Foreign Language*, *7,* 531-540.

Saigh, K., & Schmitt, N. (2012). Difficulties with vocabulary word form: The case of Arabic ESL learners. *System*, *40*, 24-36. doi:10.1016/j.system.2012.01.005

Schmider, E., Ziegler, M., Danay, E., Beyer, L., & Bühner, M. (2010). Is it really robust? Reinvestigating the robustness of ANOVA against violations of the normal distribution. *European Research Journal of Methods for the Behavioral and Social Sciences*, *6*, 147-151. doi:10.1027/1614-2241/a000016

Sharwood Smith, M (1991). Speaking to many minds: On the relevance of different types of language information for the L2 learner. *Second Language Research, 7*, 118-132. doi:10.1177/026765839100700204

Thompson-Panos, K., &Thomas-Ružič, M. (1983). The least you should know about Arabic: Implications for the ESL writing instructor. *TESOL Quarterly*, *17,* 609-623. doi:10.2307/3586616

VanPatten, B. (2015). Foundations of processing instruction. *International Review of Applied Linguistics, 53*(2), 91–109. doi: 10.1515/iral-2015-00

Wade-Woolley, L., & Siegel, L. (1997). The spelling performance of ESL and native speakers of English as a function of reading skill. *Reading and Writing: An Interdisciplinary Journal, 9,* 387-406. doi:10.1007/978-94-017-3054-9_5

Winke, P. (2013). The effects of input enhancement on grammar learning and comprehension. *Studies in Second Language Acquisition*. *35*, 323-352. doi:10.1017/s0272263112000903

Yeung, S., Siegel, L., & Chan, C. (2013). Effects of a phonological awareness program on English reading and spelling among Hong Kong Chinese ESL children**.** *Reading and Writing, 26,* 681-704. doi:10.1007/s11145-012-9383-6

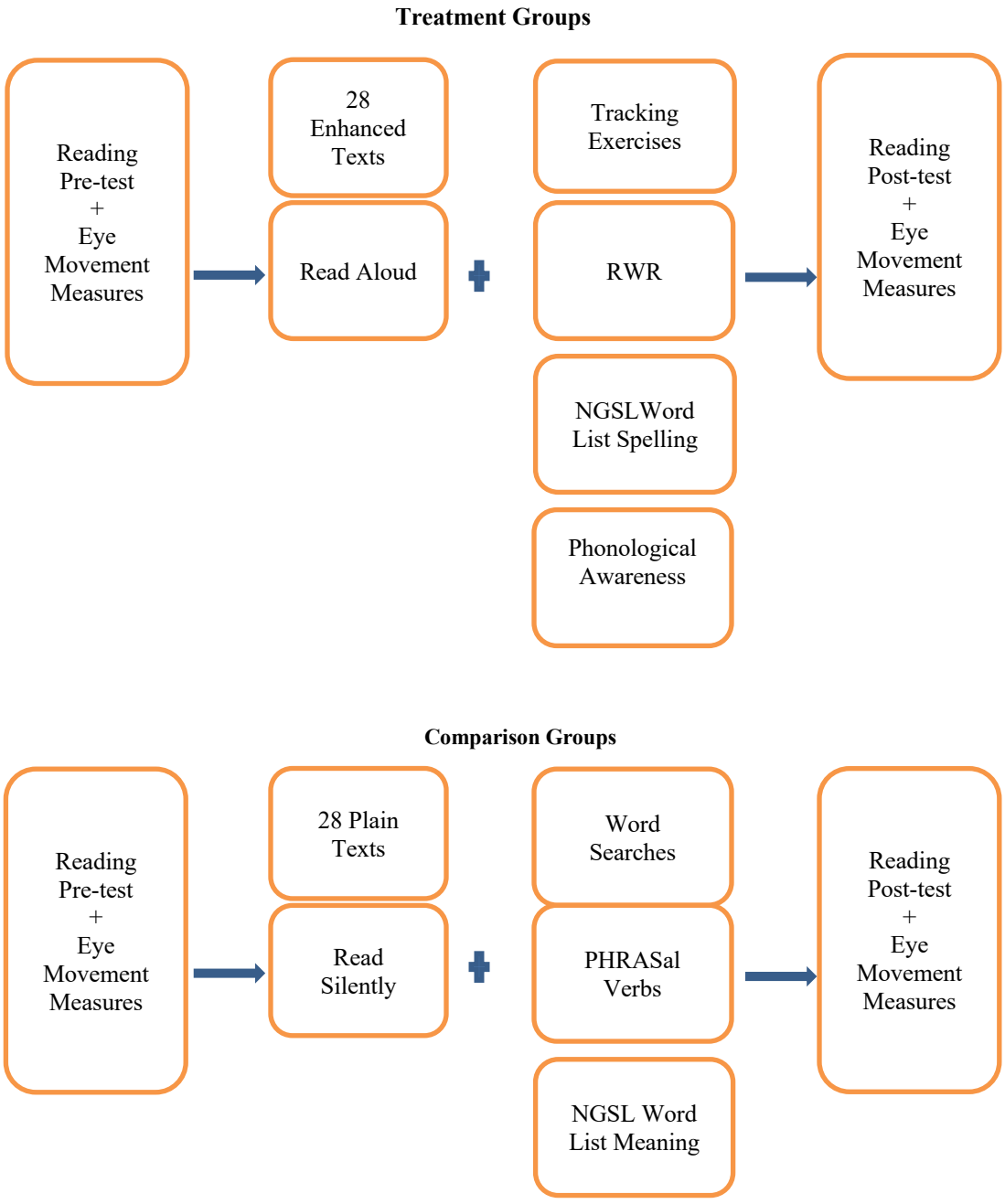

Figure 1. Research design of the study

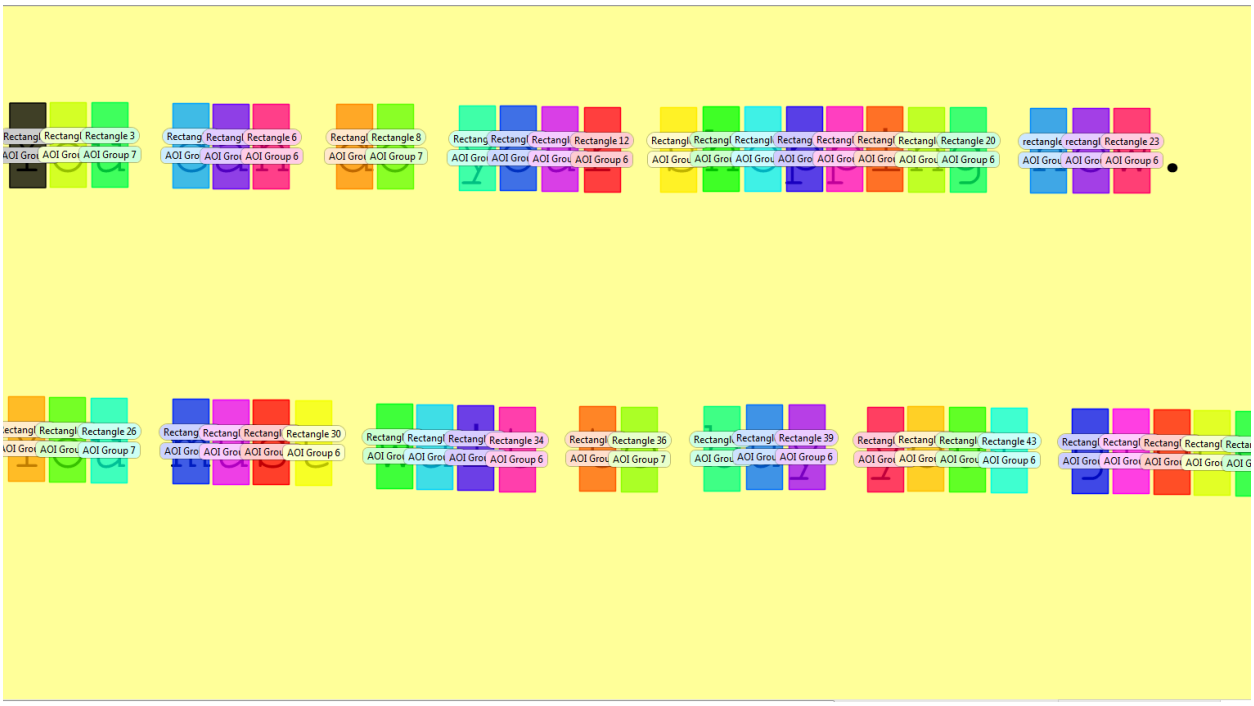

Figure 2. AOIs around vowels and consonants

Appendix 1. Example of the signs and sentences

Teachers and workers can park here.

Students cannot leave their cars here.

Appendix 2 Overview of the timing of the intervention tasks

Interventions	Approx. number per week	Approx. time per week
Enhanced texts read orally	5 texts	5 hours
Spelling	3 columns of 15 words each	20 minutes
Rapid word recognition	45 words a day x 5 = 225	10 minutes x 5 = 50 minutes
Tracking exercises	2-5 sheets	60 minutes
Phonemic awareness activities	daily but not scheduled	50 minutes

Appendix 3. Eye-tracking measures and hypotheses about development

Measures & Definitions	Hypotheses
Fixation The period of time during which the eye remains relatively stationary and reflects when information is being encoded.	
Number of Fixations	As ability increases, the number of fixations decreases (Holmqvist, 2011; Rayner, 1998)
Average Fixation Duration	As ability increases, the average duration of fixations decreases (Holmqvist, 2011; Rayner, 1998)
Saccade A high velocity jump executed by the eye to bring information into foveal vision.	
Number of Forward Saccades	As ability increases, the number of saccades decreases (Hyönä & Olson, 1995; Rayner, 1998; Rayner et al., 2006)
Length of Forward Saccades	As ability increases, the length of saccades increases (Rayner, 1998)
Regressions "In order for a saccade to be a regression, the saccade needs to move in the opposite direction to the text but not necessarily in the opposite direction to the previous saccade". Holmqvist, 2011, p. 263	
Number of Regressions	As ability increases, the number of regressions decreases (Ashby, Rayner & Clifton, 2005)
Length of Regressions	As ability increases, the length of regressions increases (Rayner, 2009; Dussias, 2010)
Mean of all visits on consonants	Mean of all visits on consonants will decrease as a result of the intervention
Mean of all visits on vowels	Mean number of visits on vowels will be higher as a result of the intervention.
Proportion of vowel visits (number of visits on vowels divided by visits on all letters)	The proportion of visits on vowels will be higher after the intervention.

Authors and affiliations

Dr. Joan Molly Oakley, College of the North Atlantic-Qatar, Doha

Professor Judit Kormos, Department of English Language and Linguistics, Lancaster University

Dr. Gareth Gareth McCray, School of Medicine, Keele University

Bio-data

Dr. Joan Oakley taught ESL and was involved in teacher education in Canada at Université du Québec à Chicoutimi, Concordia University and Memorial University. After a move to the Middle East in 2000, she taught EFL at UAE University. In 2006, she moved to Qatar where she has been teaching EFL at the College of the North Atlantic-Qatar, researching how best to help Arabic L1 speakers learn to successfully read in English.

Judit Kormos is a Professor in Second Language Acquisition at Lancaster University. She is the co-author of the book Teaching Languages to Students with Specific Learning Differences with Anne Margaret Smith. She has published widely on the effect of dyslexia on learning additional languages including a book entitled The second language learning processes of students with specific learning difficulties. She is the author of multiple research papers that investigate the role of cognitive factors in second language acquisition.

Dr Gareth McCray is currently working as a lecturer in biostatistics in the School of Medicine at Keele University. His current research includes modelling child development trajectories in developing countries, test linking, IRT, and the diagnosis of Neurodevelopmental delay in 0-3 year-olds. He has expertise in psychometrics, simulation methods and expert judgement collection and analysis. He worked extensively on the construction of the IYCD (Infant and Young Child Development Indicators) for the WHO. He has also worked as a consultant for Oxford University Press, and The British Council language testing divisions.