Expanding English Vocabulary Knowledge through Reading: Insights from Eye-tracking Studies

Ana Pellicer-Sánchez
UCL Institute of Education, UK

Abstract
Students in academic contexts are expected to engage with large amounts of reading and they frequently meet unknown words and phrases in those reading materials. Previous research has shown that second and foreign language learners can acquire some of the unknown vocabulary that they encounter during reading. However, these previous findings were mainly based on scores in off-line, post-reading tests and thus, our understanding of the cognitive processes involved during learning from reading has been rather limited. Technological advancements have made it easier for researchers to explore learners’ online processing behaviour. One of such advancements is eye-tracking, which provides a rich record of online reading behaviour. The last decade has witnessed an unprecedented increase in the number of eye-tracking studies conducted in second/foreign language learning research, with a particular focus on vocabulary learning from reading. This article illustrates how the use of eye-tracking has helped researchers gain a better understanding of the process of vocabulary learning from reading and of the relationship between eye-movements and performance measures. This article discusses recent research findings and identifies directions for future research.

Keywords
Vocabulary, second language, reading, eye-tracking, incidental learning

Introduction
English for Academic Purposes (EAP) students need to read extensively and intensively. EAP researchers have been concerned with developing strategies to help learners cope with the reading demands placed upon them. The large amounts of reading that EAP
students engage with become an excellent source for vocabulary growth. Moreover, EAP students usually want to increase their vocabulary size, ‘regarding it as a yardstick of their language improvement’ (Jordan, 1997: 149). Thus, gaining a better understanding of the reading process and of how learners develop their lexical knowledge through reading is an important concern for EAP researchers and practitioners.

Advancements in technology have brought important methodological innovations in second and foreign language acquisition research. These advances have had a significant impact not only on the use of English in academic and professional contexts but also on the way we study the acquisition and use of English (as well as other second/foreign languages). One of these important advancements is eye tracking. Despite being an established technique in other fields and being extensively used in psychology and psycholinguistic research, only in the last decade has it started to be used in second and foreign language learning research. Technological advancements have made eye-tracking systems more affordable and readily available to second language researchers, explaining the surge in eye-tracking studies conducted in this area in the last decade.

Eye-tracking allows researchers to record participants’ eye movements while processing different types of stimuli on a computer screen (i.e. text, static images, dynamic video), providing a precise record of where participants are looking and for how long. It provides ‘a real-time, direct measure of cognitive processing and processing effort’ (Conklin et al., 2018: 13). Eye-tracking has been a common tool in reading research. Much of what we know about first language (L1) and second language (L2) reading comes from eye-tracking studies. However, this research has traditionally focussed on describing the processing of words and sequences of different characteristics and on examining the role of different factors on that processing. Very few studies in the psycholinguistic literature were concerned with the learning of words and sequences while reading. At the same time, second language researchers interested in learning processes were mainly using off-line, post-reading measures to examine vocabulary learning from reading. Recent interdisciplinary research has used eye-tracking, in combination with the more traditional off-line tests, to investigate the processing and learning of new words from reading, providing useful insights about the cognitive processes involved in vocabulary learning from reading. The considerable amount of eye-tracking studies conducted in this area in the last few years allows us to draw tentative conclusions and to suggest future research directions that would help to move the field forward.

**Eye-tracking in Reading Research**

Eye-tracking has been an essential tool in examining the processing of written stimuli, leading to an extensive literature on eye movements during reading. This section summarizes the main types of eye movements during reading and their main features (for a comprehensive review of eye-tracking research in reading see Conklin et al., 2018). As argued earlier, a lot of what we know about L1 and L2 reading comes from eye-tracking studies. One of the main advantages of eye-tracking is that it can be performed without a secondary task (Conklin and Pellicer-Sánchez, 2016). Importantly, eye-tracking allows researchers to measure the processing of single words or phrases in sentence or story contexts, providing a more natural reading experience (Pellicer-Sánchez
Eye-tracking measures different types of eye movements while reading. It measures the rapid movements of the eyes, i.e. *saccades*; the moments where the eyes stop, i.e. *fixations*; and the movements back to parts of the text that were already processed, i.e. *regressions* (see Figure 1). These different types of eye movements usually occur automatically, without conscious control (Conklin et al., 2018), and are measures of overt attention in reading and other complex cognitive tasks (Rayner, 2009). Eye-tracking is based on the eye-mind link (Just and Carpenter, 1980), which assumes that eye movements are a measure of cognitive processing. Thus, we can make inferences about the cognitive effort readers expend in processing specific parts of the written stimuli (Conklin et al., 2018). Eye-movements are recorded for a particular area of interest (AOI). In reading, AOIs are commonly single words or sequences of words. Eye-movement measures to a particular AOI are usually classified into early and late measures. Early measures (e.g. first fixation duration, first pass reading time) reflect automatic processing and initial stages of lexical access, while late measures (e.g. fixation count, total reading time) are said to reflect more controlled processing and lexical integration (see Conklin et al., 2018, for a detailed review of most commonly used measures).

When we read, we do not fixate on every word in a text. Instead, many of the words in a text are skipped, with a higher skipping rate for shorter words (e.g. Brysbaert et al., 2005; Rayner, 1998) and more predictable words (e.g. Rayner and Well, 1996; Rayner et al., 2005). It is estimated that 35% of function words and 15% of content words are skipped in reading (Carpenter and Just, 1983). The general assumption is that more and longer fixations on a word reflect higher cognitive effort. In silent L1 reading, fixation durations are on average 225–250 milliseconds, and the mean saccade size 7–9 letter spaces (Rayner, 1998). Different factors have been shown to affect eye movements in reading, including length, frequency, and predictability. In the L2 context, evidence has also been provided for the facilitation effects of cognate status and interlingual homographs (e.g. Lemhöfer et al., 2004) (see Rayner, 2009 for a review of all factors affecting the processing of written stimuli). Although most processing studies traditionally focused on single words, the last decade has witnessed a considerable increase in the number of studies examining the processing of formulaic sequences, such as idioms, e.g.

---

**Figure 1.** An example sentence with an illustrative eye movement pattern depicted below the sentence. Fixations are indicated by circles and the number indicates their order (1–8); saccades are indicated by forward curves (a–f) and regressions by backward curves (g).
beat around the bush, (e.g. Carrol and Conklin, 2015; Siyanova-Chanturia et al., 2011), and binomial expressions, e.g. bride and groom (e.g. Siyanova-Chanturia et al., 2011) (see Siyanova-Chanturia and Van Lancker Sidtis, 2019, for an overview of online processing of formulaic language).

**Vocabulary Learning from Reading**

Evidence for the effectiveness of reading as a source of L1 and L2 vocabulary growth abounds. Many studies have shown that different components of lexical knowledge can be acquired incidentally from reading (e.g. Pellicer-Sánchez, 2016; Pellicer-Sánchez and Schmitt, 2010; Brown et al., 2008; Horst et al., 1998; Pigada and Schmitt, 2006; Waring and Takaki, 2003; Webb, 2005; Webb, 2007a; Webb, 2007b) and that lexical knowledge is acquired with different levels of certainty and confidence (e.g. Pellicer-Sánchez, 2016; Wesche and Paribakht, 2000). The positive role of reading for vocabulary learning expands to the learning of formulaic sequences, with recent studies showing that collocations can also be learnt incidentally from reading (e.g. Pellicer-Sánchez, 2017; Webb et al., 2013).

A common feature of these studies is that findings were based on scores in off-line measures, i.e. post-reading vocabulary tests that measured learners’ knowledge of different components of lexical mastery. These findings provided useful insights into the amount and type of vocabulary that learners remembered after a reading activity. However, they do not contribute to our understanding of the online processing of new lexical items during reading and how online behaviour could be related to lexical gains. In order to overcome this limitation, vocabulary researchers started to use eye-tracking, in combination with the more traditional post-reading vocabulary tests, to explore both the online processing of new lexical items and the knowledge demonstrated in subsequent vocabulary tests. These studies have provided the most comprehensive picture to date of the process of incidental learning from reading and have opened a new direction in incidental vocabulary learning research.

In the L1 context, a few eye-tracking studies explored the processing of novel words in sentence contexts and found that in general, L1 readers spend more time reading novel words than familiar words (e.g. Chaffin et al., 2001; William and Morris, 2004) and that processing time decreases with subsequent exposures to the novel items (e.g. Joseph et al., 2014). In the L2 context, the first study to use eye-tracking to examine vocabulary learning from reading was conducted by Godfroid, Boers, and Housen (2013). L2 learners were asked to read a set of sentences that contained either a pseudoword (e.g. paniplines), a known control word (e.g. boundaries), or both. After the reading, learners’ knowledge of the pseudowords was assessed in a multiple-choice gap-filling test. Results showed that a single exposure to novel words during reading improved learners’ ability to recognize those words, with participants recognizing on average 23% of the target items. Fixation durations on novel pseudowords were significantly longer than those on known controls. A similar pattern was found by Pellicer-Sánchez (2016). In this study learners were asked to read a longer story text that contained eight repetitions of six pseudowords and to complete a set of vocabulary tests assessing knowledge of those pseudowords. Results showed that repeated exposures to novel lexical items led to a
significant increase in learners’ ability to recognize the form and meaning of the novel items. Learners’ ability to recall the meaning of the novel pseudowords was the most difficult aspect to acquire and learners also seemed to be less certain about their knowledge of this lexical component. Results of the analyses of the eye-movement data showed that the number of fixations and fixation durations on pseudowords were longer than on known control words but that reading times decreased with subsequent encounters to the target items. By the eighth repetition, L2 learners were reading known control items and pseudowords in a similar way. This finding suggested that learners not only developed declarative knowledge of new vocabulary from reading, as measured by the post-reading tests, but they also improved the fluency with which new items were read, in line with results of studies conducted with L1 readers (Joseph et al., 2014).

These two studies were the first to show L2 learners’ acquisition and online processing of novel words while reading and constituted an important contribution to our understanding of incidental vocabulary learning from reading. However, these two investigations used reading materials that had been specifically designed for the purposes of the studies. While allowing researchers to control for a range of factors that have been shown to affect eye-movement patterns (e.g. position of target items in the sentence, length of text, text difficulty, etc.), it could be argued that their findings are not representative of authentic reading contexts. This limitation was overcome in other recent studies that employed authentic materials as the reading stimuli for their eye-tracking experiments. Mohamed (2018) examined L2 learners’ eye movements while reading a graded reader containing twenty pseudowords that were repeated several times (frequency range= 1–30). Learners’ knowledge of the pseudowords was assessed in three post-reading tests (form recognition, meaning recognition, and meaning recall). Results showed that learners acquired a considerable percentage of the target items, with form recognition being the easiest component to acquire, followed by meaning recognition (30%), and meaning recall (13%). Analysis of eye movements confirmed that in general more time was spent reading the pseudowords than known control words, and that processing time decreased significantly as a function of frequency of exposure. Learners spent more time processing the pseudowords in the initial exposures.

Similarly, Godfroid et al., (2018) examined learners’ acquisition of vocabulary from reading a long, authentic text. Participants in this study read five chapters of an authentic novel while their eye movements were recorded. After the reading participants completed three vocabulary tests that measured their knowledge of the foreign words that appeared in the text. Results showed that participants were able to recognize the form and the meaning of around 30% of the items measured, while they could only recall the meaning of around 13% of the items, further confirming the challenging nature of recall knowledge. The analysis of the eye-movement data showed that reading times on novel words decreased as an effect of exposure in a non-linear fashion. Results of the growth curve modelling analysis showed that the decrease in reading times followed an S shaped curve, with a rapid decrease in reading times from the first to the fourth exposures, followed by a slight increase, and a final gradual decrease.

These findings were confirmed in a study by Elgort et al., (2018). L2 learners in this study were asked to read a continuous expository text containing low-frequency target words and high-frequency control words. Analyses of learners’ eye movements showed
that, as expected, learners spent more time processing the low-frequency target words than the high-frequency control words. Importantly, they found that these differences in the processing of target and control items reduced significantly by the eighth exposure, in line with the findings by Pellicer-Sánchez (2016). Elgort et al., (2018) also wanted to examine if the lexical representations that had been created from the reading activity could be accessed in other semantically neutral contexts, or whether they would be restricted to the specific context of the experimental text. In order to achieve this aim, participants were asked to complete a second reading activity consisting of reading a set of semantically-neutral sentences that contained the target items. They then compared the reading of target items in these neutral sentences with the reading of the same target items in the last exposure in the initial expository text. Interestingly, results of the eye-movement analyses showed that there were differences in the reading of target items in these two contexts, with more fixations and longer reading times in the neutral sentences (second reading activity) than in the last occurrence of the continuous text (first reading activity). This was interpreted by the authors as evidence that the lexical representations that learners had acquired from the reading activity were not strong enough to be translated to a different, semantically-neutral context.

Overall, the available eye-tracking studies on incidental vocabulary learning from reading have shown that reading times to novel lexical items are in general longer than those on matched controls. More processing time and hence, higher processing effort, is required when processing novel words at the initial encounters in a text, but processing speed increases with subsequent exposures, leading to more fluent and automatic reading of the newly learned words. L2 learners seem to need around eight encounters for novel items to be read in a similar way to previously known words. Crucially, these patterns have been consistently found in studies using a range of reading materials, i.e. including both experimental materials designed for the purposes of the studies (e.g. Godfroid et al., 2013; Pellicer-Sánchez, 2016) and authentic materials (e.g. Elgort et al., 2018; Godfroid et al., 2018), as well as using target items of different characteristics, i.e. pseudowords (e.g. Godfroid et al., 2013; Mohamed, 2018; Pellicer-Sánchez, 2016), foreign language words (e.g. Godfroid et al., 2018), and low-frequency, real words (Elgort et al., 2018), attesting to the robustness of the effects found.

Enhancing Vocabulary Gains from Reading

One of the most common findings from the incidental learning literature is that vocabulary gains from reading tend to be relatively modest, and that acquiring large numbers of words is only possible if learners are engaged in large amounts of reading. This has led vocabulary researchers to explore ways to maximize learning gains from reading. One of such techniques is input enhancement, which usually involves making target words more salient to learners with typographical enhancement techniques, such as bolding, underlining, upper case, colour letters, and highlighting. This could be particularly beneficial in the academic contexts where reading introduces new specialized vocabulary. Input enhancement was initially put forward by Sharwood Smith (1991; Sharwood Smith,1993) and since then, most studies on the effect of input enhancement have been conducted in grammar learning. Empirical evidence has suggested that input enhancement can also be
beneficial for the deliberate acquisition of vocabulary (e.g. Barcroft, 2003), although for their effects to be significant it might need to be combined with lexical elaboration (Kim, 2006). The positive effect of input enhancement is supported by the argument that the amount of attention and engagement with lexical items predict their learning (Laufer, 2017; Schmitt, 2008) and is based on the assumption that typographical enhancement leads to increased attention that is then reflected in higher vocabulary gains. However, previous claims about the effect of input enhancement were mainly based on learners’ performance on post-reading tests, and no direct evidence of the assumed increased attention was provided. Eye-tracking provides the means to directly measure attention to enhanced lexical items while reading. Choi (2017) examined the effect of input enhancement, i.e. bolding, on the acquisition of collocations. L2 learners in this study were asked to read a passage containing 14 target lexical and grammatical collocations. Participants were divided into two experimental groups: the enhanced group read the passage with the target items boldfaced, whereas the un-enhanced group read the same text with no input enhancement. Results of the eye-movement analysis showed that input enhancement led to more attention on target collocations, with longer total fixation durations and more fixations on the enhanced items. Results of a post-reading recall test showed that learners in both conditions significantly improved their knowledge of the target collocations, but the learners in the enhanced conditions outperformed those in the un-enhanced group.

Another approach used to increase learners’ attention to target items while reading is in-text glosses, which drive students’ attention to the target vocabulary and provide the meaning of target items, usually by an L1 translation or definition or a combination of both. Warren et al., (2018) examined the effectiveness of different types of glosses, i.e. text only, picture only, and multimodal glosses, and used eye-tracking to explore learners’ attention to the different gloss types. Learners read a text containing six glossed pseudowords repeated three times while their eye movements were recorded and then completed a form recall and a meaning recognition test. Knowledge of word meaning was best retained when pseudowords were presented with picture-only glosses, but this was indeed the condition that attracted less attention. Cumulative reading times on the pseudowords in the text predicted performance on the post-reading vocabulary tests but amount of attention to the glosses was not a clear predictor of vocabulary gains.

**Relationship between Eye Movements and Performance Measures**

A main concern of the majority of studies reviewed in the previous sections was to explore the relationship between eye movements and performance measures. The ultimate goal of vocabulary researchers interested in eye-movement behaviour is to gain a better understanding of what those measures might tell us about the type of knowledge that is acquired from reading. The findings reported in the previous section are an important contribution to our understanding of how learners process unknown words while reading. However, I would argue that the real drive to use eye-tracking in incidental vocabulary learning research is to find out whether eye-movement patterns can predict
vocabulary learning gains. Results about the relationship between eye-movement measures and vocabulary gains have yielded conflicting findings.

In the L1 context, the first study to explore the connection between eye movements and vocabulary learning gains was conducted by William and Morris (2004). To the best of my knowledge, this was the first study to show a systematic relationship between online processing and memory for novel words. This connection was later confirmed in the L2 context by Godfroid, Boers and Housen (2013), who demonstrated that total reading times on the target items significantly predicted scores in a multiple-choice, gap-filling test. Longer total fixation durations on the pseudowords were related to higher recognition scores. Similarly, Mohamed (2018) found that the amount of attention on pseudowords positively predicted learning outcomes in all the lexical measures included. Cumulative reading times on the pseudowords (i.e. the sum of fixation duration of all encounters of the same target item) embedded in the graded reader significantly predicted learners’ ability to recognize the form and the meaning of the pseudowords, as well as their ability to recall their meaning. The finding that increased attention to target items leads to higher vocabulary gains supports the role of noticing and attention for successful language learning to take place (Schmidt, 1994; Schmidt, 2001).

A weaker relationship between reading times and vocabulary gains was found by Pellicer-Sánchez (2016). Results of her study showed that cumulative reading times were positively related to participants’ ability to recall the meaning of the pseudowords, supporting previous research findings. However, this connection was only found for one of the lexical measures employed in the study, i.e. meaning recall. No significant connection was found between cumulative reading times and form or meaning recognition, which questions the strength of such relationship. Similarly, Warren et al., (2018) found a significant relationship between cumulative reading times on target words in the text and vocabulary scores, but no connection was found between time spent processing the glosses and vocabulary scores. In fact, other recent studies have failed to find any significant and reliable connection between reading times and vocabulary scores (e.g. Elgort et al., 2018).

These contradictory findings are indeed in line with the empirical evidence available from other areas of vocabulary learning research. For instance, in the context of vocabulary learning from multimodal exposure, Bisson et al., (2015) examined the acquisition of words presented auditorily one at a time with either the written L1 translation or translation and picture. They found that longer processing times on the images were related to higher vocabulary gains, but no relationship was found between reading times on words and vocabulary gains. Montero Perez, Peters, and Desmet (2015) examined vocabulary learning from captioned videos in incidental and intentional conditions, operationalized as presence or absence of test announcement. Interestingly, results showed that longer reading times on target words were related to higher scores in the post-viewing vocabulary test only in the intentional condition. Longer reading times on target words in the incidental conditions were negatively related to vocabulary scores, which the authors interpreted as a sign of processing difficulties. This resembles findings of studies examining reading comprehension, with evidence suggesting a negative relationship between reading times and levels of comprehension (e.g. Chang and Choi, 2014; Serrano and Pellicer-Sánchez, 2019). When this relationship has been examined in the context of
grammar learning, studies have also yielded inconsistent findings. A positive relationship between reading times and gains in grammatical knowledge was observed only in some of the conditions explored (e.g. Indrarathne and Kormos, 2017) and only with some of the target structures examined (e.g. Godfroid and Uggen, 2013).

These conflicting and inconsistent findings demonstrate the complexity of the relationship between reading times and performance measures and question the direct link between amount of attention and learning gains, at least when attention is measured by total fixation durations. As Pellicer-Sánchez and Siyanova-Chanturia (2018: 24) claimed, ‘this relationship is presently tentative at best and more research is needed before we can make stronger claims about the link between processing times and vocabulary learning’. The empirical evidence available suggests that the relationship between eye movements and performance measures could potentially go in two directions. Longer processing times on target lexical items could reflect deeper engagement that is later positively reflected in larger gains. This deeper engagement could reflect readers’ attempts to guess the meaning from context and encode it in memory. However, longer reading times on novel items can also be a reflection of processing difficulties that are later reflected in lower test scores. The exploration of eye movements on their own might not provide enough information to disambiguate these distinctive effects. We clearly need more fine-grained analysis of eye-movement data. As Godfroid et al., claimed, we need to ‘isolate and identify these different subprocesses in readers’ eye-movement data more clearly’ (Godfroid et al., 2018: 35). Researchers have recently argued for the need to combine eye-movement data with performance measures and stimulated recall data. Some initial attempts to such triangulation approach have been conducted in the context of L2 writing (e.g. Revesz et al., 2019; Yu et al., 2017). In vocabulary learning research, Godfroid and Schmidkt (2013) were the first to triangulate data from stimulated recalls, eye movements, and vocabulary test scores. Results of this initial investigation showed that words that participants could remember having read them in context were fixated longer and better learned, showing an interesting relationship between eye movements, level of awareness, and learning.

**Conclusion and Future Directions**

Technological advancements have had an important impact on the research techniques available to second language acquisition researchers. Eye-tracking is one of the techniques recently adopted in second language learning research and its use has rapidly increased to the point of becoming a common technique in research in this area. A large number of these recent eye-tracking studies have focussed on vocabulary learning from reading.

Eye-tracking studies in this context have shown that learners do not only acquire declarative vocabulary knowledge incidentally from reading but that their procedural knowledge is also improved, as reflected in a consistent decrease in reading times with repeated exposures. However, as Elgort et al., (2018) argued, this increase in processing times does not seem to be translated to other new neutral contexts, suggesting that the lexical representations created from reading might be weak and contextually-bound. These findings call for more research in the area to gain a better understanding of the
type and amount of exposure that leads to strong lexical representations that are then able to be transferred to other contexts.

All available eye-tracking studies on incidental vocabulary learning from reading have examined learning from reading-only conditions. However, written texts are often combined with auditory and non-verbal input that supports the reading comprehension process and that are believed to maximize its learning gains. Recent eye-tracking studies have shown that the presence of auditory input affects L2 learners processing of text and pictures in multimodal reading conditions without a negative effect on reading comprehension (e.g. Pellicer-Sánchez, et al., 2018, in press; Serrano and Pellicer-Sánchez, 2019). The effect that these processing differences may have on vocabulary learning gains is yet to be examined. Future eye-tracking studies should explore vocabulary learning from different reading conditions. While eye-tracking research is usually conducted in laboratory settings with careful control of the factors affecting eye movements and might be difficult to be conducted in a classroom environment, language teachers and practitioners can make use of results from eye-tracking experiments to gain a better understanding of the type of knowledge that learners develop during different types of reading activities and reading modalities, which can in turn support their pedagogical decisions.

Finally, while one of the main aims of the use of eye-tracking in vocabulary learning research is to examine how eye-movement measures predict vocabulary gains, the findings to date are inconclusive. As argued in this article, future studies should perform more detailed analysis of eye-movement data in order to clarify the different subprocesses that are reflected in eye movements during reading. Only through these more precise investigations of eye movements in reading, will we be able to obtain an accurate understanding of the cognitive processes involved in incidental vocabulary learning.

Funding

The author received no financial support for the research, authorship, and/or publication of this article.

ORCID iD
Ana Pellicer-Sánchez https://orcid.org/0000-0002-9401-7027

References


