

Lookups as an e-reading resource for vocabulary learning: Readers' experiences with complexified text

Avery Little, Averil Coxhead, and Anna Siyanova-Chanturia

School of Linguistics and Applied Language Studies, Te Herenga Waka – Victoria University of Wellington

Abstract

In this small-scale qualitative investigation, we explored second language readers' use of an e-reading programme wherein readers primarily focus on meaning-making. We further examined learners' reflections on their use of the programme. Four English as a foreign language adult learners (L1 Arabic) read a text for comprehension with the help of Calibre, an open-source e-reading software. We found that while participants did make use of the lookup tool, the use was infrequent; participants also differed in the ways in which the tool was used. Lack of familiarity with the tool was also found to discourage learners from looking up words, often leading participants to guess meaning from context and skipping unknown words.

Keywords: vocabulary learning, reading, second language, lookups, definitions

Introduction

A linguistic paradox at the heart of reading in a second language is that learners need to read to learn words, but they need to learn words to read (Cobb, 2007). Graded readers are useful to many readers to improve fluency and achieve an authentic reading experience (see Nation & Bonesteel, 2010, *Real reading* textbook series; Nation & Macalister, 2020); however, they can be difficult or expensive to access. In developing countries such as Morocco, interesting and accessible texts in English are a rarity for many learners, tutors, and even schools (Boutieri, 2012). This lack of accessible materials contributes to underdevelopment of learners' skills in reading and writing compared to speaking and listening (Abouabdelkader & Bouziane, 2016). However, reliance on printed text is waning. Much reading in English in Morocco is already online and motivated by desire to access knowledge or to participate in online discourse (Soussi, 2021). A promising compromise that could allow learners to access written material on the internet and learn vocabulary independently is reading on a digital device (e-reading), which can include integrated access to an online dictionary or lookup tool. These tools, or resources, can support reading more quickly and

This work is licensed under a [CC BY-NC-ND 4.0 license](https://creativecommons.org/licenses/by-nc-nd/4.0/).

Correspondence concerning this article should be addressed to Averil Coxhead. Email: averil.coxhead@vuw.ac.nz

spontaneously than dictionaries, and more independently than glosses; however, they remain under-researched as a language-learning tool (Boers, 2022).

The present small-scale exploratory study took place during COVID when data gathering in schools was difficult and a larger face-to-face study was not deemed possible. We thus set out to explore a small group of second language (L2) readers' experiences of use and reflections on an e-reading programme where readers primarily focused on meaning-making, although word-learning still happened.

Literature review

Vocabulary and reading language

Vocabulary can pose a challenge to second- or foreign-language readers of English, particularly when access to materials written for learners is limited and independent readers are searching for interesting materials. Learners might not have sufficient vocabulary to be able to cope with reading (Sun & Dang, 2020). Around 80% of daily English language use comes from the first (most frequent) 1,000-word families (Nation, 2022, p. 98), but texts usually use greater variety of lexis. Using Nation's (2006) BNC lists, Schmitt and Schmitt (2014) revisited the findings of a wide range of corpora studies to redefine high frequency as the most common 1st to 3rd 1,000-word families, mid-frequency as the 4th to 8th 1,000, e.g., *mediation* and *pedagogy*, and low frequency as items above the 9th 1,000-word families. High and mid frequency vocabulary is essential for independently understanding and enjoying text and media intended for English speakers (Schmitt & Schmitt, 2014). Given that the many of the texts readers will encounter independently on the internet are intended for first language (L1) speakers, we estimate that even learners who know the 1st 3,000-word families in Nation's (2017) BNC/COCA word lists can face a daunting gap of 6,000 or so word families to read authentic texts (Schmitt & Schmitt, 2014).

A further vocabulary-related challenge with reading is that vocabulary uptake can be slow (Nation, 2015). Reading at 150 words per minute, Nation (2015) calculates eight hours per week over 40 weeks are needed to encounter the 1st to 9th 1,000-word families at least 12 times. To maximize vocabulary learning, Nation (2015) argues that readers can deliberately notice unknown words, guess meaning, and use a dictionary to confirm their understanding of new words. Cobb (2007) is pessimistic about the amount of word learning possible from extensive reading. He calculates the likelihood of learning vocabulary from free reading using a generous threshold of six repetitions for the 3rd 1,000 words and concludes that "even the largest plausible amounts of free reading will not take the learner very far" (2007, p. 44) without resources to support learning. He then proposes that computer-based reading provides resources to "preserve the free in free reading" by placing "encounters with new words at any level

within the act of reading ... for any type of text, and for lots of texts” (Cobb, 2007, p. 45). As of 2024, Cobb’s website, the *Compleat Lexical Tutor* (<https://www.lextutor.ca/>), hosts a reading resource called *HyperText*, which links each word to a concordance and dictionary definition; in other e-reading programs, this function is called ‘look-up’, the term we use in this paper. Despite being designed for L2 readers, *HyperText* is not, to our knowledge, in widespread use in language teaching/learning communities.

Reading with resources (e.g. look-up functions) permits learners to access more word meanings if they know how to use the resource and approach reading with an interest in word learning. Research to date supports the idea of a threshold around 95-98% of words which must be known for a reader or listener to comprehend a message (Hu & Nation, 2000; Laufer, 1989; Schmitt et al., 2011) (c.f. the coverage comprehension model, see Laufer, 2020; McLean, 2021). How best to measure participants’ vocabulary prior to reading (McLean, 2021; Schmitt et al., 2011) and what constitutes “adequate” comprehension (Nation, 2015) is under debate, but research to date generally supports the conclusion that knowing below 90% of running words leads to very little comprehension. A study by Schmitt et al. (2011) suggests that above 90%, each additional percentage point of vocabulary coverage leads to a 2.3% comprehension increase. This threshold model helps us define prerequisite knowledge for comprehending reading.

Challenges in measuring word-learning from reading

Measuring word learning from independent reading is difficult. One solution is self-rating, while another is dimensional testing, and each has advantages and disadvantages which must be weighed in research design. Self-rating is often justified in terms of practicality. Horst (2005) investigated adult immigrant ESL learners in Montreal with access to a classroom library of 37 books to read over six weeks. She used three tests: a self-reported ‘yes / not sure / no’ word recognition pre-test of items (n=100) from the entire library, a post-test limited to words from books students had read, and finally a delayed meaning recall test on words that elicited a pre-test “no” and post-test “yes”. The third test only included one to three items per learner, reflecting the difficulties of item sampling. This study found atypically high measurements of word learning, with an average pre- to post-test change in ‘yes’ responses of 6.59 additional high-frequency words and 10.29 additional mid- and low-frequency words (Horst, 2005). In another study opting to use self-rating, Schmitt et al. (2011) had readers check “yes/no” of 120 target words representing a 50% sample of the mid- and low-frequency words from a reading text plus 30 nonwords to improve validity, to estimate their participants’ vocabulary coverage prior to reading. Self-rating allows for a wider range of words to be tested, but does not indicate what has been learned, and validity is a concern.

Another approach, which we have adopted in this study, is to test different dimensions of word knowledge, such as form (e.g., spelling), meaning, and use (e.g., register or collocational patterns) (Nation, 2013). This method offers more detailed information about word learning, but only works with shorter word lists. An early study of incidental vocabulary acquisition through extensive reading had first language (L1) English participants read Burgess' 1962 novel *A Clockwork Orange*, complete a comprehension and literary criticism task, as well as a multiple-choice meaning recognition test on its Russian-based slang terms (Saragi et al., 1978). Subsequently, Pellicer-Sánchez and Schmitt (2010) had 20 L1 Spanish, L2 English participants read Chinua Achebe's 1958 novel *Things Fall Apart*, which includes 34 words of Ibo, a Nigerian language. Their findings from this dimensional test suggest that independent reading supports meaning recognition and spelling recognition far more than meaning recall or word class recall, and ten or more repetitions in the text led to better learning (Pellicer-Sánchez & Schmitt, 2010). These dimensional tests can offer more detailed measurements of word-learning through reading but require a much shorter list of items and potentially fatigue participants.

Glossing and e-reading with resources

Glossing can offer significant lexical learning opportunities in intensive reading (Nation, 2013). The more attention a gloss draws, the better the language learning results (Boers et al., 2017). Creating glosses takes time and effort and limits vocabulary-learning attention to specific words and overall reading to texts chosen by the materials developer or teacher (Nation, 2004; Cobb, 2007). E-reading offers many potential vocabulary resources for independent readers. In one early study, Laufer and Hill (2000) asked Israeli and Hong Kong learners of English to read a 120-word academic passage including 12 highlighted target words using a computer program ("Words in your ear", Laufer and Hill (2000)) which allowed access to audio recordings, L2 definitions, L1 translations (Hebrew and Chinese), and the root of target words. Analysis of clicks showed that most of the Israeli learners looked up L1 translations and most of the Hong Kong learners looked up L2 definitions, a difference likely originating in learned reading strategies. In both groups, participants were receptive to the convenience of the lookup tool. Most importantly, target look-up words were better retained than previous dictionary-supported reading experiments (Laufer & Hill, 2000). Unfortunately, their program is not publicly available. The Compleat Lexical Tutor (<https://www.lextutor.ca/>) hosts two browser-based language learning tools for resourced e-reading, HYPERTEXT and RA-READING (Cobb, 2023a, 2023b). However, they are not mainstream. Beyond broad claims about "ease of use" (Nation, 2015), there is still much to be learned about how today's readers use online resources. This study sought to evaluate one such tool – Calibre (Goyal, 2022),

an open-source e-reading software with a lookup function, as a resource for reading and learning English. Our research questions were as follows:

1. How do the L2 learners of English in this study use the demonstrated lookup tool when reading on a computer?
2. What motivates L2 learners to look up a word when reading?

Methodology

Participants

Four young adult volunteer participants were recruited from English language classes in Rabat, Morocco. Table 1 shows their goals for learning English, age and gender.

Table 1

Descriptions of Participants and Their Aims

Pseudonyms	English learning aims	Age	Gender
Mohamed	Studying for PhD, read academic papers in English	37	M
Saad	Self-improvement	28	M
Jihane	Preparing for TOEFL*, applying for English MA programs	20	F
Asmae	Using English social media, reading memes	18	F

Note. *Test of English as a Foreign Language

Pre- and post-reading measures

The Language Experience and Proficiency Questionnaire (LEAP-Q) (Kaushanskaya et al., 2020) was adapted and delivered through Qualtrics. LexTALE (Lemhöfer & Broersma, 2012), a real/non-real word recognition test was used to estimate proficiency scores. LexTALE scores were produced using this formula: $((\text{number of words correct}/40 \times 100) + (\text{number of nonwords correct}/20 \times 100)) / 2$ (see <https://www.lextale.com/scoring.html>). This test was validated using the Test of English for International Communication (TOEIC) and the Quick Placement Test (QPT) (Lemhöfer & Broersma, 2012). TOEIC defines 785 points as the threshold score of a B2 level in the Common European Framework (Tannenbaum & Wylie, 2021). L1 Korean participants with an average TOEIC score of 887 had a corrected average LexTALE score of 65.3% (Lemhöfer & Broersma, 2012). Our participants scored between 56% to 80% on the LexTALE (see Table 2), which is similar to the Lemhöfer and Broersma (2012) B2 participants (mean score 75.5%, SD 12.5% for L1 Dutch; 65.3%, SD 10.3% for L1 Korean).

Table 2

Participants' Linguistic Background and English Proficiency, Ordered by Overall Proficiency

Names	Languages known, ordered by preference of use				Lex-TALE (%)	Self-rated English proficiency		
	1 st	2 nd	3 rd	4 th		Speak	Listen	Read
Asmae	French	English	Arabic		80	6/10	6/10	5/10
Jihane	Darija	Classic Arabic	English	French	79	9/10	9/10	8/10
Mohamed	Arabic	French	English		62.5	5/10	7/10	7/10
Saad	Arabic	English	French		56	6/10	7/10	7/10

The participants' activities, including lookups and time, were recorded on Zoom while reading. In a follow-up stimulated recall interview, participants watched the recording of their reading and narrated their thinking to offer insights into their reading experience. Table 3 summarizes the measures used.

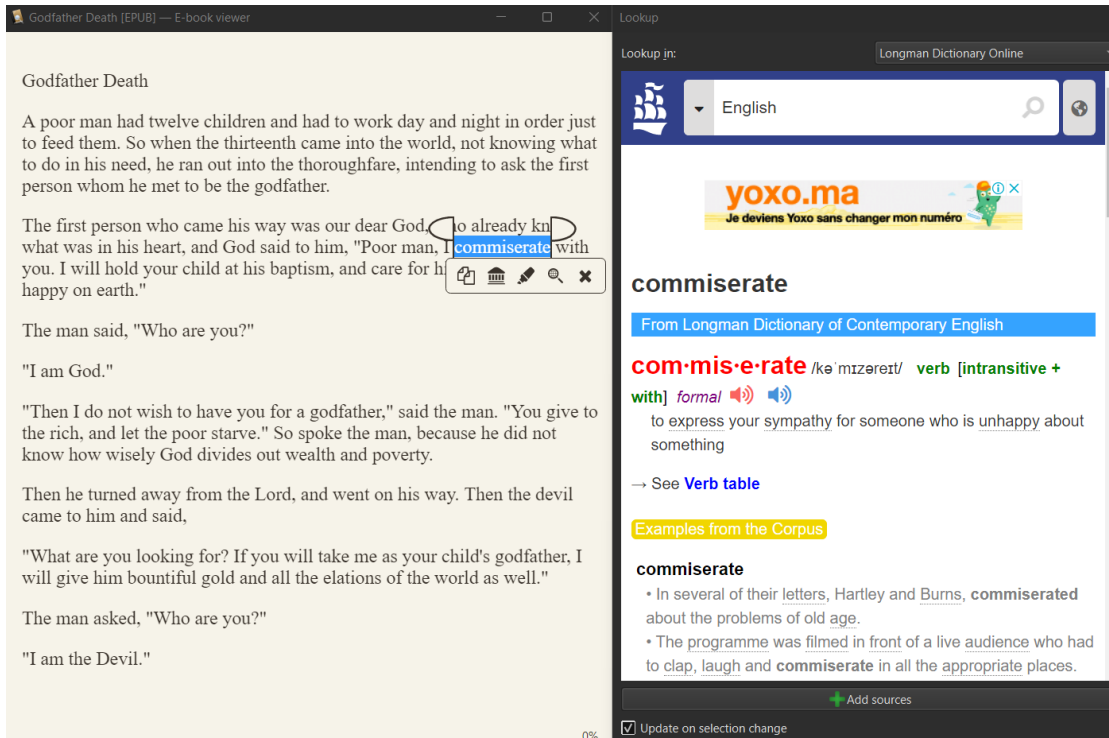
Table 3

Summary of Measures

Measure	Data	Purpose
Entrance questionnaire based on LEAP-Q (Kaushanskaya et al., 2020)	Self-reported language background & reading habits (Likert)	Describe the group
LexTALE (Lemhöfer & Broersma, 2012)	English word recognition ability (%)	Describe the group
Screen-record of reading (Zoom)	Count lookups (#), time on task	RQ 1
Stimulated recall interview with volunteers (Zoom)	Verbal recollections stimulated by video recording	RQ 2

Materials

Calibre, an open-source e-reading software, was chosen for its accessibility and adaptability. It offers a library function and an e-reader. Although it was not designed for language learning (c.f. Cobb's HyperText and Laufer's Words in your Ear), it can easily be modified for language learning. Users can search any online dictionary or translator in the e-reader itself (Goyal, 2022). Figure 1 is an image of the lookup tool querying Longman Dictionary of Contemporary English (LDOCE) (Pearson Longman, 2022).

Figure 1*The Calibre Lookup Tool Querying ‘Commiserate’ from the Modified Text “Godfather Death”*

A 1,200-word Grimm folktale named “Godfather Death” was modified by means of word-replacement, such that learners would encounter unknown target words within the text which they could look up. The principles of frequency and textual proportionality guided the selection of target words. First, the vocabulary in the text was profiled on the Compleat Lexical Tutor (Cobb, 2023c) using Nation’s (2017) British National Corpus (BNC) (Aston & Burnard, 1998) and the Corpus of Contemporary American English (COCA) (Davies, 2008) (BNC/COCA) word lists. A total of 41 words in the 2nd to 4th 1,000-word bands of Nation’s BNC/COCA frequency lists (2018) were identified. Words occurring more than three times in the text, indicating centrality to the story, were left unchanged. Words repeated two or three times in the text had only the final occurrence modified. Only up to 4% of words in a paragraph were modified, and never two words in sequence. The highest frequency targets were replaced in paragraphs where more than 4% of the words were replacement candidates. Approximately 2% of high-to-mid frequency words were replaced with low frequency words to encourage participants to notice unknown words without disrupting overall comprehension (Laufer, 2020, 2021; Nation, 2013).

Second, up to 12 single-word synonyms from a thesaurus for each target candidate word were gathered and checked. Any words in the 5th 1,000-word band or lower

frequency were eliminated. French cognates were excluded by checking the Cambridge Online English-French Dictionary (<https://dictionary.cambridge.org/dictionary/english-french>) and Linguee (<https://www.linguee.com/english-french>); items closely resembling French equivalents were eliminated.

The final list of potential replacement words (n=36), plus three original low-frequency words (*astray*, *infatuated*, and *whosoever*), was normed using a Qualtrics survey with eight Moroccans at a similar English level to the participants. Items which two or more norming participants recognized and were able to explain were eliminated (n=12). The final 24 low-frequency target words were inserted into the text with minor syntactical adjustments. The final text was checked to ensure that Longman Dictionary of Contemporary English (LDOCE) lookups would access suitable information (Pearson Longman, 2022).

Table 4

Target words (n=24)

thoroughfare	commiserate with	bountiful	discrepancy
mettle	predicament	hoodwinked	condone
be bequeathed	vise	beguiled	clutched
wheedled	purported	yearning for	elations
forthwith	thence	puny	reap
thwart	portending	astray	infatuated

For example, *abundance* could have been replaced by *bounty* (7th 1,000), *plethora* (9th), or *plenitude* (15th). *Plethora* and *plenitude* have French cognates. *Bounty* was not recognized by the norming participants and was added to the sentence by changing “give him an abundance of gold” to “give him bountiful gold.” The LDOCE definition in the lookup says, “if something is bountiful, there is more than enough of it” (Pearson Longman, 2022). Finally, the text was checked by three L1 English speakers for clarity and cohesion, to ensure replacements did not negatively affect the authenticity of the reading experience.

Procedure

Consent was signed and the entrance questionnaire was completed by participants. The reading and stimulated recall (Figure 2) were conducted via Zoom by the first author, with each participant individually. Total data collection took approximately 90 minutes per person.

Figure 2*Research Procedure*

Zoom meetings were screen-recorded and began with a brief tour of Calibre, including the lookup function. Participants then received the target text by email, added the text to Calibre and opened it in the e-reader. Participants were told they would have tasks to complete post-reading and then were left to read independently. Finally, they participated in a recorded stimulated recall interview using the screen-recording of their reading to reflect on the experience.

Data analysis and results

Descriptive statistics and individualized analysis of each participant's reading strategies and stimulated recall interview were carried out to explore their individual approaches to e-reading with a lookup tool.

Question 1: Do the L2 learners in this study use the demonstrated lookup tool when reading on a computer? If so, how?

Each participant read differently according to their habits and purpose for reading. Three participants completed the reading task as intended (Table 5). They each used the lookup tool between three and six times during the reading activity, spending between 10 and 13 minutes to read the text. The participants read at an average of 104 words per minute. Jihane had the highest self-rating of English reading proficiency and the lowest number of lookups. She read for comprehension, then re-read paragraphs for unknown words. She was preparing to take the TOEFL exam, which may have impressed on her the need for strategic reading. Mohamed and Saad both took a practical, beginning-to-end approach, like leisurely extensive reading. They read more slowly than Jihane and looked up more words. Mohamed said that he tried to read "like I read normally in my house."

Table 5*Participants by Self-Rated Reading Proficiency, and Reading Time and Lookup Use*

Participant (pseudonym)	Self-rating of English reading	Lookup use: Number of words	Time spent on task	Words per minute
Jihane	8 very good	3	10:06	118
Mohamed	7 good	6	13:07	92
Saad	7 good	6	11:54	101
Asmae	5 adequate	0	N/A – read aloud	N/A

Asmae read the text aloud and did not use the lookup tool. She said when she reads, “most of time ... I read it loudly [sic] to practice my English skills. Although, when I am reading something in French or Arabic, I read it silently.” A total of 19 words caused a hesitation or mispronunciation in her reading, indicating unfamiliarity. Of these, 15 were low-frequency target words (e.g., *infatuated*).

Question 2: What motivates L2 learners to look up a word when reading?

Overall, participants checked 11 words which included five target items (asterisked in Table 6). *Thoroughfare* appeared in the very beginning of the story and was looked up by all participants. Five mid- and four low-frequency items were looked up, along with two high-frequency words.

Table 6*Words Looked Up by Each Participants and Time Elapsed from Start of Reading*

Jihane		Mohamed		Saad	
Word and frequency by BNC/COCA lists	Time	Word and frequency by BNC/COCA lists	Time	Word and frequency by BNC/COCA lists	Time
*thoroughfare (10 th)	0:31	*thoroughfare (10 th)	0:30	*thoroughfare (10 th)	0:38
*reap (6 th)	3:36	baptism (7 th)	1:19	withered (6 th)	2:53
*commiserate (12 th)	4:03	withered (6 th)	2:38	herb (4 th)	4:28
		*reap (6 th)	4:54	*thence (11 th)	6:19
		stump (5 th)	11:08	threatened (2 nd)	7:04
		slip (1 st)	12:38	*bequeathed (9 th)	8:25

The interviews revealed that Mohamed, Jihane, and Saad took the reading as a meaning-focused activity and prioritized story comprehension. Missing meaning was the prompt in 11 out of 15 lookups. Jihane said, “I didn’t know the word

[*thoroughfare*]... I read the word [*definition*], it says ‘main road,’ and I understand what the word means. The man, he ran out into the road.” It is not surprising that not knowing the meaning caused a look-up.

Another theme was a focus on reading speed; reduced fluency and unfamiliarity with the tool discouraged the participants from looking up words. Although the lookup was quicker than switching mediums and looking for a word in a print dictionary or on a phone, it still took about five seconds to find a suitable meaning, and lookups were quickly abandoned when the participants focused on comprehending the story. Mohamed and Jihane looked up words in the first five minutes of their reading and then stopped (Table 6). Jihane set a goal of reading the text in 10 minutes and Mohamed said that “I tried to use my habit in this document ... When I read a book, I try to read it so quickly.” It is clear that their focus was on fluently making meaning, not learning individual words.

Participants also guessed meaning from context and skipped unknown words when reading. Mohamed did not look up words between 4:54 and 11:08, and explained that skipping unknown words was satisfactory because, “...my feeling is [that] I enjoy the story” but “with literature I find a lot of words I did not know because it is not my area.” Tolerating some ambiguity is to be encouraged in independent and extensive reading (Bamford & Day, 2002).

Interestingly, after reading the story, Mohamed and Saad looked up high-frequency words that they partially knew, thereby using the tool to accumulate deeper word knowledge after being satisfied with their comprehension of the text. About his lookup of *threatened*, Saad said, “I was thinking, um, I knew the meaning but ... I want to know if I’m thinking in the right way, or maybe I heard it before and I forget [sic] the meaning.” These lookups at the end of reading could suggest a shift from meaning-making to word-learning (Nation & Macalister, 2020). The overall timing pattern of lookups in this small-scale pilot study supports a distinction between language-focused learning and meaning-focused input (Nation & Macalister, 2020). These participants appear to have been doing one activity or the other, rather than doing both at the same time.

Discussion and implications

The present study was originally designed to gather both qualitative and quantitative data, but challenges in recruitment forced our focus towards a more qualitative view. An area of interest for discussion is the way learners’ skills and experiences, as well

as motivations for reading, may affect e-reading behaviours and experiences. This has implications for teaching using extensive e-reading tools, as well as future research and possibly product design.

First, familiarity and ease of use was found to be key for language learners. The lookup tool used in this project was new to our participants, who were not explicitly trained in using lookups, and as a result, the use was limited. Dictionary-familiarization activities can show learners the kinds of information available in a dictionary and help them make effective use of this tool (Nation, 2013). The LDOCE lookups present frequency, pronunciation, and usage information, but we observed that advertisements in LDOCE, like those present in other free educational resources, covered most of the dictionary page. It was therefore a challenge for the participants to find the information they sought. The overall average reading speed was 104 words per minute with only a few lookups. With reference to Nation's (2015) estimate of 150 words per minute for L2 extensive reading, our participants' relatively slow reading could indicate how challenging they found it to use the tool.

Further, although our participants were confident in using computers, learners with less developed computer skills would require much more support. Teachers considering using Calibre, or other similar tools, need to be aware of potential technical challenges. For example, participants sometimes selected the word including a space which caused the lookup function in Calibre/LDOCE to present a list titled "Did you mean," making lookups less efficient. Avoiding technical errors and giving learners practice with using strategies will help to ensure successful integration of resources into a reading programme.

Our participants were asked to read a short literary text, modelled after studies such as Pellicer-Sánchez and Schmitt (2010). Yet, none of our participants was particularly interested in English literature. A more naturalistic study might have allowed them to choose their own text. After all, one of the goals of e-reading is to allow learners to access texts that they find of interest. This would be motivational for learners, and observations of this more agentive free reading could offer further insight into the ways L2 readers learn words and material through reading.

This research has further uncovered an opportunity to develop a language learning tool for an aspiring programme developer. Using Application Programming Interfaces, it would be possible to create a Calibre-based (or similar) open-source e-reader that could query audio, definition, examples, or translations similar to "Words in your ear" (Laufer & Hill, 2000). Cobb's "Hypertext" (2023a) offers a similar function for short texts that can be copy-pasted into the browser-based tool; while the "Resource Assisted Reading" tool (Cobb, 2023b) contains a small library of 13 texts with lookup capabilities. However, these tools do not allow readers to collect texts, as Calibre can.

It is clear there is a long way to go in developing suitable e-reading tools that comprehensively serve learners' interests in free reading.

Limitations and future considerations

This small-scale pilot study has several limitations. Only a small number of participants was used, meaning no generalizable conclusions can be drawn to wider learner populations. Although we modified the text by adding low-frequency words, we cannot be certain participants were indeed encountering fully unknown words. Also, over half of the looked-up words were not target words. The learners also looked up partially known mid-frequency words, suggesting the need for a broader sample. One possible solution is to pre-test and/or post-test a much larger sample of mid- and low-frequency words from the target text and use self-rating test with validation controls to reduce fatigue (as in Schmitt et al., 2011). Another option is to purposefully design a post-test of word learning to include any items that were looked up by participants, mimicking Horst (2005). Future studies should also allow participants to choose their readings.

Conclusion

Reading in a second language can be challenging at best, especially for learners whose L1 writing system differs from their L2, or who come from relatively under-developed literacy education programmes (Koda & Zehler, 2008). Increasing literacy through reading is a goal for many language education programs, so we must analyse the tools available to teachers and learners to improve literacy practice. With strategic instruction in dictionary and lookup tool usage, readers may be able to effectively accelerate their vocabulary learning and increase the range of texts available to them. Now that computers are more easily accessible and more powerful than ever, online tools can tear down obstacles to language learning – but only with resources designed to meet learners' needs in place.

Acknowledgements

The authors wish to thank the anonymous reviewers and editor for their insightful comments and suggestions on this paper.

References

- Abouabdelkader, H., & Bouziane, A. (2016). The teaching of EFL writing in Morocco: Realities and challenges. In A. Ahmed & H. Abouabdelkader (Eds.), *Teaching EFL writing in the 21st Century Arab world* (pp. 35–68). Palgrave Macmillan.
- Achebe, C. (2009). *Things fall apart*. Doubleday.
- Aston, G., & Burnard, L. (1998). *The BNC handbook: Exploring the British National Corpus with SARA*. Edinburgh University Press.
- Bamford, J., & Day, R. (2002). Top ten principles for teaching Extensive Reading. *Reading in a Foreign Language, 14*(2).
- Boers, F. (2022). Glossing and vocabulary learning. *Language Teaching, 55*(1), 1–23.
- Boers, F., Warren, P., Grimshaw, G., & Siyanova-Chanturia, A. (2017). On the benefits of multimodal annotations for vocabulary uptake from reading. *Computer Assisted Language Learning, 30*(7), 709–725.
- Cambridge University Press and Assessment. (2022). Cambridge English–French Dictionary. <https://dictionary.cambridge.org/dictionary/english-french/> (2 September, 2024).
- Boutieri, C. (2012). In two speeds (À deux vitesses): Linguistic pluralism and educational anxiety in contemporary Morocco. *International Journal of Middle East Studies, 44*(3), 443–464.
- Burgess, A. (1962). *A clockwork orange*. Heinemann.
- Cobb, T. (2007). Computing the vocabulary demands of L2 reading. *Language Learning & Technology, 11*(3), 38–63.
- Cobb, T. (2023a). *HYPertext 2*. Hypertext Builder 2. <https://www.lextutor.ca/hyp/2/>
- Cobb, T. (2023b). *RA-READING HOME*. Resource-Assisted Reading. https://www.lextutor.ca/ra_read/
- Cobb, T. (2023c). *VocabProfile Compleat* [Website]. LexTutor. <https://www.lextutor.ca/cgi-bin/vp/comp/output.pl#editor>
- Davies, M. (2008). *English Corpora: The Corpus of Contemporary American English (COCA)*. English Corpora. <https://www.english-corpora.org/coca/>
- DeepL SE. (2022). Linguee English-French Dictionary. <https://www.linguee.com/english-french> (2 September 2024).
- Goyal, K. (2022). *Calibre E-book Management* (6.6.1) [Windows]. <https://calibre-ebook.com/>
- Horst, M. (2005). Learning L2 vocabulary through extensive reading: A measurement study. *The Canadian Modern Language Review / La Revue Canadienne Des Langues Vivantes, 61*(3), 355–382.

- Hu, M. H., & Nation, P. (2000). Unknown vocabulary density and reading comprehension. *Reading in a Foreign Language*, 13(1), 403–430.
- Kaushanskaya, M., Blumenfeld, H. K., & Marian, V. (2020). The Language Experience and Proficiency Questionnaire (LEAP-Q): Ten years later. *Bilingualism: Language and Cognition*, 23(5), 945–950.
- Koda, K., & Zehler, A. M. (Eds.). (2008). *Learning to read across languages: Cross-linguistic relationships in first- and second-language literacy development*. Routledge.
- Laufer, B. (1989). What percentage of text lexis is essential for comprehension? In C. Lauren & M. Nordman (Eds.), *Special language: From humans thinking to thinking machines* (pp. 316–323). Multilingual Matters.
- Laufer, B. (2020). Lexical coverages, inferencing unknown words and reading comprehension: How are they related? *TESOL Quarterly*, 54(4), 1076–1085.
- Laufer, B. (2021). Lexical thresholds and alleged threats to validity: A storm in a teacup? *Reading in a Foreign Language*, 33(2), 238–246.
- Laufer, B., & Hill, M. (2000). What lexical information do L2 learners select in a CALL dictionary and how does it affect word retention? *Language, Learning and Technology*, 3, 58–76.
- Lemhöfer, K., & Broersma, M. (2012). Introducing LexTALE: A quick and valid Lexical Test for Advanced Learners of English. *Behavior Research Methods*, 44(2), 325–343.
- McLean, S. (2021). The coverage comprehension model, its importance to pedagogy and research, and threats to the validity with which it is operationalized. *Reading in a Foreign Language*, 33(1), 126–140.
- Nation, I. S. P. (2004). Vocabulary learning and intensive reading. *English Australia Journal*, 21(2), 20–29.
- Nation, I. S. P. (2006). How large a vocabulary is needed for reading and listening? *Canadian Modern Language Review* 63.1, 59–82.
- Nation, I. S. P. (2013). *Learning vocabulary in another language* (2nd ed.). Cambridge University Press.
- Nation, I. S. P. (2022). *Learning vocabulary in another language* (3rd ed.). Cambridge University Press.
- Nation, I. S. P. (2014). How much input do you need to learn the most frequent 9,000 words? *Reading in a Foreign Language*, 26(2), 1–16.
- Nation, I. S. P. (2015). Principles guiding vocabulary learning through extensive reading. *Reading in a Foreign Language*, 27(1), 136–145.
- Nation, I. S. P. (2017). *Information on the BNC/COCA word family lists*. Victoria University of Wellington: School of Linguistics and Applied Language Studies.
https://www.victoria.ac.nz/_data/assets/pdf_file/0004/1689349/Information-on-the-BNC_COCA-word-family-lists-20180705.pdf

- Nation, I. S. P., & Bonesteel, L. (2010). *Real reading 1*. Pearson Longman ELT.
- Nation, I. S. P., & Macalister, J. (2020). *Teaching ESL/EFL reading and writing*. Taylor & Francis Group.
- Pearson Longman. (2022). *About the Longman Dictionary of Contemporary English*. <https://www.ldoceonline.com/about.html>
- Pellicer-Sánchez, A., & Schmitt, N. (2010). Incidental vocabulary acquisition from an authentic novel: Do “Things Fall Apart”? *Reading in a Foreign Language*, 22(1), 31–55.
- Qualtrics. (September 2022) [Computer software]. <https://www.qualtrics.com>
- Saragi, T., Nation, I. S. P., & Meister, G. F. (1978). Vocabulary learning and reading. *System*, 6(2), 7.
- Schmitt, N., Jiang, X., & Grabe, W. (2011). The percentage of words known in a text and reading comprehension. *The Modern Language Journal*, 95(1), 26–43.
- Schmitt, N., & Schmitt, D. (2014). A reassessment of frequency and vocabulary size in L2 vocabulary teaching. *Language Teaching*, 47(4), 484–503.
- Soussi, H. (2021). World Englishes in multilingual Morocco. *World Englishes*, 40(2), 259–267.
- Sun, Y., & Dang, T. N. Y. (2020). Vocabulary in high-school EFL textbooks: Texts and learner knowledge. *System*, 93, Article 102279.
- Tannenbaum, R. and E. C. Wylie. (2021). *Mapping the TOEIC ® Tests on the CEFR*. ETS. https://etswebsiteprod.cdn.prismic.io/etswebsiteprod/7acf5750-eceb-4a5f-8271-cf4db75d4432_TOEIC+Tests+-+CEFR+Correlation+Table.pdf