Reading & Writing - Journal of the Literacy Association of South Africa
ISSN: (Online) 2308-1422, (Print) 2079-8245

Words matter – but how? A study of high school learners' vocabulary and reading comprehension

Authors:
Belden L. Liswaniso[1]
Elizabeth Pretorius[2]

Affiliations:
[1]Department of Intermediate and Vocational Education, University of Namibia, Katima Mulilo, Namibia

[2]Department of Linguistics and Modern Languages, University of South Africa, Pretoria, South Africa

Corresponding author:
Belden Liswaniso,
belden1333@gmail.com

Dates:
Received: 23 Sept. 2024
Accepted: 17 Jan. 2025
Published: 31 Oct. 2025

How to cite this article:
Liswaniso, B.L. & Pretorius, E., 2025, 'Words matter – but how? A study of high school learners' vocabulary and reading comprehension', *Reading & Writing* 16(1), a529. https://doi.org/10.4102/rw.v16i1.529

Background: Learners' vocabulary knowledge in English additional language (EAL) can serve as an indicator of their reading comprehension as well as their chances of academic success.

Objectives: This article describes and examines patterns and relationships between the vocabulary knowledge of Grade 11 learners at different word frequency levels (2000, 3000, 5000, university word list and 10 000 word level) and their reading comprehension performance.

Method: In this cross-sectional study quantitative data were collected from 67 senior high school EAL learners in two low-performing schools from a low socio-economic context in the Zambezi region in Namibia. A vocabulary levels test and a reading comprehension test were used to assess the EAL learners.

Results: The results showed that the learners had limited vocabulary knowledge and low comprehension levels. The study found a significant modest correlation between knowledge of vocabulary and reading comprehension. The study further revealed that while all word levels correlated significantly with reading comprehension, the 5000 word level emerged as the strongest predicator of reading comprehension.

Conclusion: The findings suggest that EAL learners need to be supported more explicitly from the start of their EAL journey to increase their vocabulary knowledge and to boost their chances of comprehending their academic texts.

Contribution: This study adds to the understanding of how learning and academic achievement in high-poverty contexts might be slowed down by poor schooling. The study also contributes knowledge of which vocabulary high school learners need in aiding their reading comprehension.

Keywords: vocabulary; word frequency levels; vocabulary size; reading comprehension; English additional language.

Introduction

The whole purpose of reading is to make sense of the texts we read. Reading comprehension has been described as the process of extracting and constructing meaning through engagement with written language (RAND Reading Study Group 2002). This raises the question as to what knowledge and skills are needed for readers to 'extract' and 'construct' meaning from texts. Research has shown that vocabulary knowledge is a critical component in supporting reading comprehension (Nation 2015; National Reading Panel 2000; Pretorius & Murray 2019; Stoffelsma & Spooren 2019). This applies equally to reading in a home language or an additional language. Learners' success in English as an additional language (EAL) and as the language of learning and teaching (LoLT) requires them to comprehend what they read and have a large vocabulary that enables them to read efficiently for pleasure and for study purposes (Brooks, Clenton & Fraser 2021; Pretorius & Murray 2019; Schmitt 2008). However, very little research has been done in the Namibian context on learners' vocabulary knowledge in relation to their reading skills or academic success.

Namibian learners' scores in reading assessments have been consistently low (e.g. Liswaniso 2021; The Southern and Eastern Africa Consortium for Monitoring Educational Quality [SACMEQ] 2005, 2010). Tötemeyer (2013) points out that because learners have poor reading skills, which makes it difficult for them to read for meaning, they tend to rely on rote learning

of their content subject and their academic performance is poor. Only a small number of students qualify to enrol in tertiary institutions (Education Management Information System [EMIS] 2022, 2023) and the majority of students who gain access to tertiary institutions are not competent readers equipped to study independently (Liswaniso & Mubanga 2019; Tötemeyer 2010). Harris (2011:14) points out that 69% of college and university students in Namibia fail to understand their course materials because of poor English proficiency. Similarly, Liswaniso and Mubanga (2019) found that students find comprehending English texts difficult throughout their schooling, and rely on rote learning rather than reading to learn.

The Namibian basic education curriculum covers pre-primary (ages 5–6) and Grades 1–12 (ages 7–18). As of 2019, Grade 11 has become the first school exit point from basic education when learners can apply to study in vocational training centres, seek or create a job, or begin their young adult life. Learners who perform well in Grade 11 final examinations have an option to continue to Grade 12 to the Namibia Senior Secondary Certificate Advanced Subsidiary (NSSCA). Some universities in the country such as the University of Namibia only admit students who passed at least two subjects at NSSCA level. The NSSCA is an internationally recognised certificate and gives learners access to higher institutions in Namibia and beyond. Learners who find it difficult to understand reading materials in schools and are unable to succeed academically tend to drop out of school.

Despite the widely studied relationship between reading comprehension and vocabulary knowledge in more affluent and literate English education contexts, the research context of EAL and vocabulary in low- and middle-income countries (LMIC) such as Namibia has not been well explored, and the relationship between vocabulary and reading comprehension in academic achievement necessitates ongoing investigation of these two aspects. Although information is available about Namibian learners' poor reading abilities and their academic challenges, research on their EAL vocabulary knowledge is sparse. What kinds of words should they know, and how large should their vocabulary be for them to better understand their texts and to use reading as an effective means of extracting and constructing meaning? This article reports on a study that examined the relationship between reading comprehension and vocabulary knowledge among Grade 11[1] EAL learners with low academic performance in a print-poor environment in the Zambezi region of Namibia. It assessed the learners' reading comprehension as well as their vocabulary and analysed the interplay between these two variables to better understand how vocabulary profiles affected reading ability.

1.Currently in Namibia, Grade 11 is the first exit point from secondary school. At the time of the study, data were collected from Grade 11 and Grade 12 learners. The previous Grade 12 (ordinary level) is roughly equivalent to the current Grade 11; due to these changes, Grade 11 is used to refer to the cohort reported here.

Theoretical framework and literature review

This section provides a theoretical lens for conceptualising both reading comprehension and vocabulary and how they might interact. It also reviews prior research in these domains and highlights main trends that have emerged thus far.

Reading: A complex process

Reading is a complex process whereby learners identify words and sentences in row after row of printed text, integrating their linguistic knowledge of the printed words with their cognitive thinking, and reasoning skills and background knowledge of the world in order to extract and construct meaning from a text (Castles, Rastle & Nation 2018; Cunningham & Stanovich 2001; Day & Bamford 1998; Lee & Spratley 2010). This complex process is reliant on numerous code, linguistic, cognitive, and background knowledge factors. The simple view of reading (Gough & Tunmer 1986) describes reading comprehension as the product of two main components, viz. decoding and linguistic comprehension. Both components are necessary for reading comprehension, but on their own are not sufficient. Decoding is the process of using code or alphabetic knowledge to identify words in a text, while linguistic comprehension refers to the (oral) language proficiency needed to make sense of the words. Poor decoding skills compromise reading comprehension; if decoding skills are weak and slow, readers cannot access the text, as the process of identifying words consumes too much cognitive energy and attention, depleting memory resources needed to extract information and construct meaning during reading. Similarly, poor language proficiency compromises reading comprehension. Even if readers can decode a text, inadequate knowledge of linguistic structures and vocabulary can stymie their ability to make sense of the text, make inferences, and utilise background knowledge to construct meaning at a deeper level. The simple view of reading is not simplistic; rather, it condenses the complexity of reading to two core components. It has stood the test of time and is empirically supported across alphabetic orthographies and languages (Castles et al 2018; Kim 2017). In turn, the two core components comprise several subcomponents, as illustrated in Scarborough's rope model (2001) and also form an integral part of more recent models of reading such as the Direct and Indirect Effect Model of Reading (DIER) (Kim 2017, 2020).

Comprehension refers to the understanding process that occurs when meaning is 'extracted' from the explicitly stated information in the text, which forms the basis for 'constructing' meaning by integrating, interpreting and evaluating information during reading (Pikulski & Chard 2005; Pretorius 2002; RAND Reading Study Group 2002). Here the reader draws on linguistic and vocabulary knowledge, general background knowledge, knowledge of text genres, structures and conventions, and thinking and reasoning skills such as inferencing, perceiving temporal, causal or contrastive relations, and detecting discrepancies.

Skilled reading is regarded as interactive reading. The term *interactive* has a dual meaning. A learner interacts with a text, extracting information from it and making sense of it as they read in terms of available linguistic, cognitive, and background knowledge (Nassaji 2003). Skilled reading is also seen as an interactive process because, as described above, 'many skills work together simultaneously in the process' (Grabe 1991:378). However, a reasonable level of fluent decoding is necessary for interaction to occur during reading. Fluency refers to the accuracy and speed at which reading occurs and has long been associated with skilled reading and with reading comprehension in both first language reading (Grabe 2010) and EAL reading (Mutema & Pretorius 2024; Pretorius & Spaull 2016; Wills et al. 2022).

Categories of reading ability

Padgett (1997:42) describes reading ability as the combination of accuracy, speed, and comprehension. The relative status of decoding fluency vis-à-vis comprehension provides a nuanced way to identify where learners lie on a cline of reading ability. For example, McCormick (1995:100) groups readers into four categories of reading ability: independent, instructional, borderline and frustration level. Readers at the independent level have 98% or more decoding accuracy and a comprehension level of at least 95% of a text at their age or Grade level. Independent readers do not find reading challenging. They are highly skilled readers who utilise a number of skills in reading and can learn effectively from a text at their maturational level (Pretorius 2002). Instructional level readers have 95% decoding accuracy and about 75% comprehension level. These readers benefit from reading instruction but do not have major decoding or reading problems. At the borderline level, readers have 90% – 94% decoding accuracy, with comprehension ranging anywhere between 55% and 74%. Their decoding skills are not yet accurate enough to properly support reading comprehension. Borderline readers need additional reading exposure and practice to improve their accuracy and in order to benefit from reading experience. Learners at the frustration level read with 90% or lower decoding accuracy and have about 50% or lower accuracy in comprehension. These are learners with major reading problems, where their poor decoding skills inhibit text comprehension.

Scientific studies of reading have long recognised that skilled reading is accurate, precise, and rapid, allowing for automaticity in text processing which, in turn frees up cognitive effort and attention for reading comprehension (DeFord 1991; Grabe 1991; Smith 2004), even for EAL readers (Wills et al. 2022).

Levels of reading comprehension

Well-designed assessments of reading comprehension focus on learners' ability to read a text on their own and answer a range of easy to more difficult questions on it. Their responses to the questions are considered to reflect how well they have extracted and constructed meaning from the text, that is, comprehended it. A reading comprehension test does not assess decoding, fluency or language use and grammar, nor is it a memory test, as the text is always available for the reader to re-read or search for answers. In other words, a reading comprehension test reflects readers' comprehension of written texts. If readers perform poorly on a reading comprehension test, possible reasons for their poor comprehension can be surmised, but unless separate data on their decoding, fluency, or language proficiency are available, we cannot say for sure what it is hampering reading comprehension.

There have been several frameworks or taxonomies for assessing reading comprehension (see Liswaniso 2021 for a summary). For example, the framework used in the Progress in International Reading Literacy Study (PIRLS) that assesses the reading comprehension abilities of 9–10 year olds (typically Grade 4s) worldwide includes four types of comprehension processes to assess reading comprehension. These processes of comprehension are comprehension of explicitly stated information (i.e. literal information), making straightforward inference, integrating information and ideas, and critical or evaluation comprehension (eds. Mullis & Martin 2019).

Literal comprehension is described as a basic level of reading where a learner recognises information stated explicitly in a text (Howie et al. 2008; Liu 2010). Literal comprehension does reflect some reading comprehension ability, but only at a superficial level. Although this is regarded as the easiest level of text comprehension, it is nevertheless important for extracting explicitly stated information from a text on which further meaning is constructed.

At the straightforward inferential level, a learner needs to infer information not explicitly stated in a text or see links between information in a text. These involve fairly straightforward inferences within a sentence or across adjacent sentences. At this level, readers are beginning to comprehend the author's purpose and are developing a deeper comprehension of the overall text (eds. Mullis & Martin 2019).

In integrating information and ideas, readers attempt to construct more specific understanding of a text by fusing their own knowledge and experience with the meaning in the text (eds. Mullis & Martin 2019). This involves perceiving links between bits of information across larger chunks of text. For example, readers may determine the overall meaning of a text and the tone or mood of a story, or they may perceive relations at a deeper level. The last level is critical or evaluative comprehension, where a reader may evaluate and analyse the text's meaning or content objectively, or from a personal point of view. The second, third and fourth levels of text processing typically involve making inferences and perceiving links between items of information or ideas within and across paragraphs in the text, taking the reader beyond explicitly stated information.

This study focused on two broad categories of reading comprehension, namely literal and inferential reading, to distinguish between ability to understand explicit information as opposed to constructing deeper levels of meaning from what is implied or inferred in a text. Given that linguistic knowledge is necessary for reading comprehension, as indicated by the simple view of reading, the study used vocabulary as an index of language proficiency. Research has long shown strong associations between vocabulary knowledge and other aspects of verbal proficiency such as oral language proficiency and reading comprehension (Pretorius & Stoffelsma 2017; Stæhr 2008). We turn now to a review of vocabulary research.

Vocabulary knowledge

Vocabulary refers to knowledge of words and their meanings. It includes recognising a word from its aural or written form, knowing its meaning and possible connotations, recognising its grammatical function (noun, verb, adjective, etc.), and knowing whether it can be used in formal or informal contexts. The more one knows about words the more comprehensive is one's vocabulary knowledge (referred to as vocabulary depth). Vocabulary is generally categorised into receptive and productive knowledge. Words that we can recognise or comprehend when reading or listening are referred to as receptive vocabulary knowledge (Zhou 2010). Productive vocabulary knowledge refers to words that we actively use to communicate when speaking or writing (Laufer 1998). Receptive vocabulary is usually much larger than productive vocabulary; over time, words that were formerly known receptively can become part of our productive vocabulary repertoire (Meara 2009).

Vocabulary knowledge can be assessed in terms of breadth or size (the number of words that learners generally know 'at a particular level of competence') and depth (the degree to which a word is understood, e.g. its derivations and connotations) (Akbarian 2010; Hatami & Tavakoli 2012; Li & Kirby 2014). One factor that affects vocabulary development is the frequency with which words are used in a language – words that occur frequently in our everyday spoken and written interactions tend to be acquired earlier and more readily than words less commonly used. The frequency with which words occur in a language thus provides a systematic framework for examining what kinds of words learners know and at what frequency level.

Vocabulary frequency levels and their role in shaping vocabulary knowledge

The notion of 'frequency level' refers to the arrangement of words based on their frequency of occurrence in use. Data on the frequency of words originate from electronic corpora containing samples of written and spoken language use and which are used, inter alia, to analyse the occurrence of words in a particular language (Bennett 2010). The work of Kučera and Francis (1967) served as the foundation for the Brown University Standard Corpus of Present-Day American English, or Brown Corpus. It was one of the earliest linguistic databanks (created in 1961) and contained about 1 million words (Bennett 2010). Nowadays, linguistic corpora are much larger. For example, the British National Corpus contains over 100 million words. Corpora are subjected to statistical analyses of the frequency distribution of words. This provides an empirically reliable way of determining how often words are used across contexts and modality (spoken or written language). Frequency levels are usually determined according to bands of a 1000 words, ranging from high frequency (the 3000 most frequent words in a language), mid-frequency (words between the 4000 and 9000 frequency ranges) and low frequency words (words at the 10 000 word level or beyond) (Nation & Anthony 2013; Schmitt & Schmitt 2014). Core academic words (words that occur in formal learning or scientific contexts across disciplines, e.g. *contrast*, *establish*, and *significant*) have also been identified as a separate category: the Academic Word List (AWL). Academic words tend to occur in the mid-frequency range.

Vocabulary exposure refers to how often a new word is encountered in oral or written language and the chances of learning it (uptake) so that it becomes part of a learner's receptive or productive word knowledge. The different frequency levels of words described above play a critical role in language acquisition, language proficiency and reading comprehension. High frequency words are words that appear most often in both oral language and in texts. Coady (1993) refers to high frequency words as the core vocabulary. For example, high frequency words at the 2000 word level form the largest part of running words[2] in generic English texts, ranging from 86% to 90% of words in texts (Nation 2006). Schmitt and Schmitt (2014) suggest that words occurring at the 1000–3000 word levels be categorised as high frequency words, and they also introduced the mid-frequency category (4000–9000 word levels). Mid-frequency words occur more in written language than oral language. Everyday spoken conversations tend to recycle a relatively small number of common high frequency words, even among adult speakers. Low frequency words occur mostly in written language, especially of a more technical nature and discipline-specific, for example geology or medical texts. Words from the AWL occur in the mid-frequency range and are used mainly in more formal learning contexts (oral and written) related to different content subjects and disciplines. They account for roughly 10% of running words in academic texts (Coxhead 2000). They also cover 4% of words in newspapers, 4% in popular magazines and 1.7% in fiction texts (Nation & Waring 1997). Although these figures appear low, knowledge of these words can affect the ease with which learners read and understand the textbooks related to their different content subjects.

Children tend to learn high frequency words before they learn words of mid and low frequency. Learners are regularly exposed to high frequency words throughout their early

2.The term 'running words' (i.e. tokens) 'refers to the total number of word forms in a text' (Coxhead 2000:216; Nation & Anthony 2013:5).

school years (Li & MacGregor 2010). Books intended for young EAL learners are likely to contain vocabulary from the higher word frequency levels than books intended for native speakers (Nation 2015). Research suggests that the basic high frequency vocabulary at the 2000 word level needs to be mastered by EAL learners because of their frequency in everyday conversations and in texts. However, the 2000 word level seems to be necessary but not sufficient for success in reading. For effective inferencing and learning in EAL, for example, a learner needs mastery of the 3000 word level (Laufer 1997; Nation 2015). Exposure to books and print material increases the chances of learning words from the mid and low frequency ranges.

Although the proportion of academic words might seem low, they play a critical role in understanding academic texts. Wood, Schatschneider and VelDink (2021) found a significant relationship between EAL learners' use of academic words and their reading comprehension. Because socio-economic factors affect access and exposure to books and literate activities, learners from a low socio-economic background tend to show lower levels of knowledge of academic words (Wood et al. 2021). A study by Stoffelsma, Antwi and Hanson (2024) among university students in Ghana, found academic word knowledge and knowledge of words at the 3000 word level were significant predictors of the students' academic achievement. Low frequency words are more difficult to learn because of their relative infrequent occurrence in texts. However, because many of them are technical words (e.g. *fiscal austerity*, *solar radiation*, *thermodynamics*) their meanings are often explicitly defined in textbooks.

As more words are acquired, the coverage of words known to a learner increases. Ideally, at least 95% of words should be known to a reader for reading comprehension to be manageable. Nation (2006) suggests that knowledge of mid-frequency words at the 8000–9000 word level is sufficient for a learner to read a variety of authentic texts.

In sum, once children have learned to read on their own, reading becomes a powerful tool for building vocabulary knowledge. Reading exposes learners to vast lexical resources that are not encountered in daily oral interactions. Knowledge of mid-frequency words, academic words and low frequency words is acquired mainly through regular exposure to a range of reading materials and increases learners' chances of academic success (Nation & Waring 1997). It is no surprise therefore that skilled readers tend to have much larger vocabularies than unskilled readers or people who can read but who chose not to.

Measures of knowledge of words at different frequency levels can serve as a proxy for the proportion of times a learner may have encountered a certain word in English (Lawrence et al. 2021). Information about adequate vocabulary reading thresholds is important for EAL education. As teachers and course designers can utilise such information in 'setting vocabulary goals and designing lexical syllabi' (Laufer & Ravenhorst-Kalovski 2010:6). In order to help learners increase their vocabulary, Nation (2012) suggests that teachers should determine which of the frequency levels their learners have mastered and which levels still need to be mastered. Without understanding the role of vocabulary size, chances are that vocabulary learning and teaching will not receive the required attention in schools. Although far more words are explicitly taught in an additional language compared to a home language, there are too many words in any additional language for all of them to be taught. Many thousands of words in both a home language and EAL are acquired through incidental learning, through exposure to both oral and written language, especially the latter.

Measures of vocabulary knowledge

Vocabulary size or breadth is measured with the Vocabulary Size Test (e.g. Nation & Beglar 2007; Nguyen & Nation 2011; Webb & Nation 2008; Webb & Sasao 2013). Another way to measure vocabulary knowledge is to use word frequency levels as an organising device, as in the Vocabulary Levels Test (VLT) (Read 2000). According to Li and MacGregor (2010), the VLT is based on a vocabulary learning model in which words are learned according to their frequency in a text. The VLT test was originally developed by Nation in 1983 as a test of receptive vocabulary knowledge at the 2000, 3000, 5000, academic and 10 000 word levels, with a sample of 18 words from each level included in each section. This was later replaced by Schmitt, Schmitt and Clapham's (2001) VLT, where the items at each level increased from 18 to 30. Nation later revised his VLT as a test of productive knowledge at the above-mentioned five different levels (Nation 2006; Laufer & Nation 1999). Webb, Sasao and Ballance (2017) used more recent electronic databanks to validate possible changes in word frequency levels over time and introduced the Updated VLT at the 1000, 2000, 3000, 4000 and 5000 word levels, sampled from the up-to-date British National Corpus and Corpus of Contemporary American English at the time.

In the productive VLT a testee's word knowledge is assessed by prompting a learner to produce a targeted word by filling in a word in a specific context (Meara 2009; Nation 2006; Zhou 2010). For example:

> I sent him an invitation so I ex..... him to come to my party tonight.

The first two letters of the missing word are provided to prompt a learner for the target word (*expect*), while the sentence context prevents them from producing a non-target word that might be similar (e.g. *exhort)* but would be inappropriate for the context or from a different frequency level.

Mastery of a level in the VLT is assumed if a testee scores at least 83% in a section (i.e. 15 out of 18) (Laufer & Nation 1999). However, mastery levels have differed somewhat over

the years. In Schmitt et al.'s (2001) receptive VLT, mastery at each level is placed at 87% (26 out of 30). In subsequent studies, a lower mastery level of 80% was suggested as sufficient for mastery of a word level (Xing & Fulcher 2007), while Webb et al. (2017) proposed a more nuanced system for different frequency levels, recommending a mastery level of 97% (29 out of 30) on the 1000, 2000, and 3000 word frequency levels, and 80% (24 out of 30) on the 4000 and 5000 word frequency levels. Elmasry (2012) points out that a vocabulary test based on frequency levels serves a diagnostic purpose, as it indicates at which level attention needs to be given to boost learners' vocabulary development to meet their age or Grade needs.

Another way to measure productive vocabulary size is via word analysis from student essays. For example, leaners can be asked to write a 300-word essay, and then the occurrence of high-, mid- and low frequency words that they use in their writing can be tallied (Laufer 1998). However, this is a more time-consuming method, and an essay topic may elicit particular word usage so that analyses of several essays on various topics would be needed to get a more representative word frequency profile.

Vocabulary knowledge and reading comprehension

Research associates vocabulary knowledge with world knowledge (Lawrence et al. 2021): the more a learner knows about words the more they know about the world, and the better their reading comprehension becomes (Castles et al. 2018). For example, if a learner knows words called connectives (e.g. *consequently*, *therefore*, *by contrast*), they are more likely to make causal or contrastive connections in academic texts (cf. Crosson, Lesaux & Martiniello 2008; Lawrence et al. 2021). Since knowledge of connectives shows that a learner knows something about the world (e.g. how arguments or information are integrated in a text), the learner may find it easier to understand a text in which such a relation is suggested, even if the text does not use a known word (such as *consequently*) to indicate the nature of the relationship (Lawrence et al. 2021). Learners for whom English is not their home language but is the LoLT need a large vocabulary size to comprehend what they read. Stæhr (2008:144) found vocabulary size to correlate significantly with reading comprehension ($r = 0.83$), listening comprehension and writing among junior secondary school EAL learners in Denmark. Zano and Phatudi (2019) also found a strong correlation between vocabulary size and reading comprehension ($r = 0.90$) among Grade 11 EAL learners in South Africa. In some studies, vocabulary size has been found to be a predictor of reading success in EAL (Bintz 2011; Pikulski & Templeton 2004; Webb 2009).

It must be emphasised that although vocabulary size has an influence on reading comprehension, it does not necessarily predict reading comprehension or academic performance. Vocabulary richness per se is not sufficient for reading comprehension (Schmitt, Jiang & Grabe 2011). As pointed out earlier, fluency is also needed for reading comprehension. Language proficiency, prior topic knowledge, familiarity with genre conventions in different texts, motivation, amount of exposure to print, the amount of information in a reading text and the overlap between the known meanings of words and the context meaning are all factors that may affect reading comprehension (Castles et al. 2018; Schmitt et al. 2011; Webb 2009). This means that vocabulary interacts with other factors in increasing reading comprehension.

Although studies examining the link between reading comprehension and vocabulary size in southern Africa have been undertaken (Cooper 1999; Zano & Phatudi 2019), it seems not many have been conducted in Namibia. Considering the low socio-economic status of many Namibian learners, it is important to establish whether or not a similar pattern as for the literature review emerges. This study therefore attempts to add knowledge to the Namibian knowledge base about the link between reading comprehension and vocabulary size. In this study, vocabulary knowledge was measured in terms of vocabulary levels. The research questions that guide the study are the following:

- What are the vocabulary levels and reading comprehension of Namibian learners?
- What is the relationship between vocabulary levels and reading comprehension performance among Namibian learners?
- Which vocabulary frequency level best predicts performance in reading comprehension of Namibian learners?

Research methodology

The research context, instruments and procedures for this study are briefly described below.

The research context of this study

The data presented here derive from a study that investigated issues related to EAL vocabulary and reading comprehension in Namibian high schools (Liswaniso 2015). Data reported here were gathered in two Namibian high schools (School A and School B) in Katima Mulilo, a town in the Zambezi region with low academic performance and limited reading materials. They were the only two schools that had classes from Grade 8 to Grade 12 in the town. Since then, Grade 11 has become the first exit point for school leaving. Nationally, statistics shows that only 35% and 34% of Grade 11 learners obtained A–C results in EAL in 2021 and 2022, while learners from the Zambezi region performed even more poorly in the Grade 11 final examination, with only about 20% obtaining A–C results across school subjects (EMIS 2022, 2023). These results are normally required for admission to universities.

The participants and sampling

The participants in this study were 70 Grade 11 and Grade 12 learners drawn from School A and School B. School A and School B had several streams (A–F classes) of Grade

11 learners, with 35 learners in each class. These classes were streamed according to learners' academic performance, with best-performing learners placed in Class A. Three learners were withdrawn from the study as they did not complete both assessments. This brought the total number of learners who participated in the study to 67, comprising 31 female learners and 36 male learners. All learners who participated in the study had English as their LoLT and Silozi as their first language (L1). Most of the learners also speak additional mother tongue languages at home which are neither their L1 nor LoLT at school. The majority of the learners are from a low socio-economic background.

Even though there were several Grade 11 and Grade 12 classes in each school, only two classes were selected from the Class A–C stream in both schools, which represent the stronger learners. At School A, Class A and Class C studied Natural Sciences and Class B studied Social Sciences, whereas at School B only Class A studied Natural Sciences and the other classes took Social Sciences. The strong classes were purposively selected for being the best-performing learners in the schools, and only learners who signed informed consent forms participated in the study. Learners' ages ranged between 15 and 23 years, with a mean age of 17.6.

An independent samples *t*-test showed no significant differences in the mean reading comprehension scores of the Grade 11 and Grade 12 learners ($t = –1.31$ (65), $p = 0.19$). Likewise, a repeated measure analysis of variance for the vocabulary levels was applied but no significant differences emerged between Grade 11 and Grade 12: $F = 0.550$ (1), $p = 0.46$. In this article, the learners are treated as a single cohort and referred to as Grade 11 learners, to align with changes in the phases of Namibian senior secondary school, where the previous Grade 12 is roughly equivalent to current Grade 11, which is now the first exit point from secondary school.

Data collection instruments

Two instruments were used to collect data for this study: a productive VLT and a reading comprehension test.

The productive vocabulary levels test

The productive VLT designed by Laufer and Nation (1999) was used to test learners' vocabulary knowledge at the five levels of word frequency (the 2000, 3000, 5000, UWL and 10 000 word levels). Each word level had 18 items. The mastery level of the test, according to Laufer and Nation (1999:41), is seen as 'a matter of judgement' and depends on the vocabulary level being considered. It falls between 83% and 85% (15 out of 18). Cronbach's reliability index of the VLT in this study was 0.90.

Reading comprehension test

The reading comprehension test was taken from the EAL specimen papers that were used when the Namibian English syllabus was adjusted to suit the Namibian context in 2010. It consisted of two texts: the first text (about sharks) comprised 373 words and the second text (about forest primates) comprised 380 words. There were 11 questions for the first text and 12 questions for the second one, with each text comprising literal and inferential comprehension questions. In total, there were 17 literal questions and 18 inferential questions. The inferential questions in the reading comprehension test comprised vocabulary inferencing and text-semantic inferences. The reading comprehension test totalled 35 marks. Information about the official reliability score was not available. Its reliability index in this study was 0.66. Even though this test was designed by the Ministry of Education for high school learners and had already been used before, the low reliability index is a limitation and the results should be interpreted with caution.

Data collection

Data were collected in the second school semester over 2 days. After permission from school principals had been granted, the data collection started with the reading comprehension test on the first day, followed by the productive VLT on the second day. The first author administered the instruments in person. The reading comprehension test was set to be written in 1 h 30 min and the productive VLT was set for 40 min. Both assessments were completed within the allotted time.

Data capture and analysis

Using the marking scheme provided by the test designers, the reading comprehension test was marked twice. In the first marking, the scores of learners were noted on a separate sheet and after 1 week the test was marked again. Grammatical and spelling errors were not penalised during marking. This double scoring was done to ensure scoring reliability, as suggested by Hughes (2003). No scoring differences emerged in both marking. The subscores and the total test score of 35 were converted into percentages for convenient analysis.

Similarly, the productive VLT was also marked twice. Spelling errors were not penalised, for example '*injared*' for '*injured*' was accepted as correct. The vocabulary subscores (out of 18 for each level) and the total raw score (out of 90) were converted into percentages. Although performance at each level is of interest in the VLT, the mean total score helps to indicate the vocabulary profile of a cohort overall.

This study made use of descriptive and inferential statistics to analyse the coded data by using the SPSS program (version 29, IBM Corporation 2020), as described in the next section.

Results

Descriptive statistics were used to address the first research question, viz: *What are the vocabulary levels and reading comprehension of Namibian learners?*

Table 1 provides the descriptive statistics for performance of the Grade 11 learners on the VLT, across all five frequency levels. The mean scores are reflected as percentages. A total percentage score based on the averages obtained for the five levels is shown in the bottom row.

Considering that mastery for a word level is 83%, it is evident from the means in Table 1 that this cohort of learners had not yet mastered any of the vocabulary levels. Mean performance on the most frequent words, the 2000 word level, was 69.9%. Only learners at the 75th percentile showed mastery of the 2000 most frequent words. There is a big drop in mean learner performance from 69.9% to 36.9% at the next most frequent word level, the 3000 word level. Performance at 5000 word level and UWL is similar (26.3% and 27%). Not unexpectedly, the lowest performance was at the 10 000 word level, at 14.9%. The low total vocabulary mean of 35% does not augur well for the reading comprehension competencies of these learners.

Table 2 shows the descriptive statistics for the reading comprehension test. Performance on literal and inferential questions and for total reading comprehension is reflected as mean percentages for the group as a whole and for performance at the 25th, 50th and 75th percentiles.

As Table 2 shows, the overall comprehension of the learners was very low, with a group mean of 28.5%. Learners found the inferential questions particularly challenging (17.7%), and three learners scored zero for these. Performance doubled on the literal questions (39.9%), although performance even on these easier literal questions was still low.

TABLE 1: Vocabulary performance on the productive vocabulary levels test (N = 67).

Word frequency level	Mean %	Standard deviation	Minimum – maximum	Percentiles		
				25th	50th	75th
2000	69.9	14.18	33 – 100	61.1	72.2	83.3
3000	36.9	14.68	6 – 67	27.7	33.3	44.4
5000	26.3	12.01	6 – 61	16.6	27.7	33.3
University word list	27	16.32	0 – 78	16.6	27.7	38.8
10 000	14.9	10.19	0 – 44	5.5	11.1	22.2
Total vocabulary	35	11.21	16 – 66	26.6	35.5	41.1

TABLE 2: Reading comprehension performance.

Reading comprehension	Mean %	Standard deviation	Minimum – maximum score	Percentiles		
				25th	50th	75th
Literal score	39.9	12.89	18 – 71	29.4	41.1	47.0
Inferential score	17.7	14.05	0 – 61	5.56	11.1	27.7
Total score	28.5	11.49	11 – 60	20	25.7	34.2

In sum, the descriptive profiles of the learners show very poor performance on both vocabulary knowledge and reading comprehension.

The second research question examines the relationship between these variables, viz: *What is the relationship between vocabulary levels and reading comprehension performance among Namibian learners?*

The Pearson product-moment correlation was used to determine correlations between vocabulary performance at the different word levels and overall, and literal, inferential, and total reading comprehension scores. The results are presented in Table 3.

As is apparent from Table 3, all the vocabulary levels correlate significantly with literal comprehension, inferential comprehension and total reading comprehension, but some were slightly more robust than others. A moderate correlation (r = 0.532, p = 0.000) was obtained between overall vocabulary and overall reading comprehension. Knowledge of words at the 5000 word frequency level also showed a moderate correlation with reading comprehension (r = 0.539, p = 0.000). The findings indicate that the more words learners knew, the better they performed on reading comprehension, relatively speaking. Knowledge of words at the 5000 word level in particular correlated with reading comprehension.

Given the overall poor performance on both tests, the relationship between vocabulary knowledge and reading comprehension was further explored by coding the learners into three categories, based on their range of performance on the VLT: 'very weak', 'weak' and 'emerging' vocabulary groups (Table 4). These three groups were suggested by the data as in each case there was a gap of 3% and 4% separating the highest performer in the previous cluster from the lowest performer in the next cluster.

The 'very weak' vocabulary group had mean vocabulary scores ranging between 16% and 24%, comprising 16

TABLE 4: Vocabulary mean percentages across the three vocabulary groups.

Word frequency level	Very weak (14% – 24%) (n = 16)	Weak (27% – 44%) (n = 42)	Emerging (48% – 66%) (n = 9)
2000	53.4	72.3	87.6
3000	21.1	38.2	59.2
5000	14.9	26.5	45.6
University word list	10.4	27.7	53.09
10 000	4.8	15.7	29.01
Total mean	20.9	36.1	54.9

TABLE 3: Correlation between vocabulary knowledge and reading comprehension (N = 67).

Word level	2000	3000	5000	University word list	10 000	Total
Literal comprehension	0.328**	0.291*	0.439**	0.351**	0.292*	0.408**
Inferential comprehension	0.479**	0.368**	0.477**	0.368**	0.356**	0.492**
Total reading comprehension	0.480**	0.390**	0.539**	0.423**	0.383**	0.532**

*, Correlation is significant at the 0.05 level (two-tailed).

**, Correlation is highly significant at the 0.001 level (two-tailed).

learners (24%). The mean scores of the 'weak' vocabulary group ranged between 27% and 44% and consisted of 42 learners (62%). The smallest group of 9 learners (13%) comprising the 'emerging' vocabulary group, had mean vocabulary scores between 48% and 66%. Table 4 and Table 5 show the mean performance profiles in vocabulary knowledge and reading comprehension across the 'very weak', 'weak' and 'emerging' vocabulary groups.

It is interesting to note how vocabulary knowledge consistently increases by as much as 10% or more at each frequency level across the groups. For example, the 'very weak' mean of 21% at the 3000 word level increases by 7% – 38% in the 'weak' group, with an even bigger jump of over 20% – 59% in the 'emerging' vocabulary group. Table 5 shows similar increases in mean reading comprehension across the three vocabulary groups.

The box and whisker plots in Figure 1 display the relationship between vocabulary and reading comprehension visually. The bottom and top lines in each box show first quartile (25th percentile) and third quartile (75th percentile) means, while the black lines in the boxes show median reading comprehension scores for each vocabulary group. The whiskers indicate minimum and maximum comprehension scores within each vocabulary group. While there is still considerable variance within and across the groups, Figure 1 shows the depressive effect that weak vocabulary knowledge has on reading comprehension: the 'very weak' vocabulary group clearly struggled with comprehension, but comprehension remained sluggish in the 'weak' group too. The 'emerging' vocabulary group was the only one that started pulling ahead of the other groups in terms of comprehension. Although this small group of learners knew many more words than their peers and performed the best in reading comprehension compared to their peers, their comprehension was still very low.

TABLE 5: Reading comprehension across the three vocabulary groups.

Vocabulary groups	Vocabulary		
	Very weak (n = 16)	Weak (n = 42)	Emerging (n = 9)
Literal mean	34.1	39.9	50.3
Inferential mean	10.4	17.2	33.3
Total reading comprehension mean	21.9	28.2	41.5

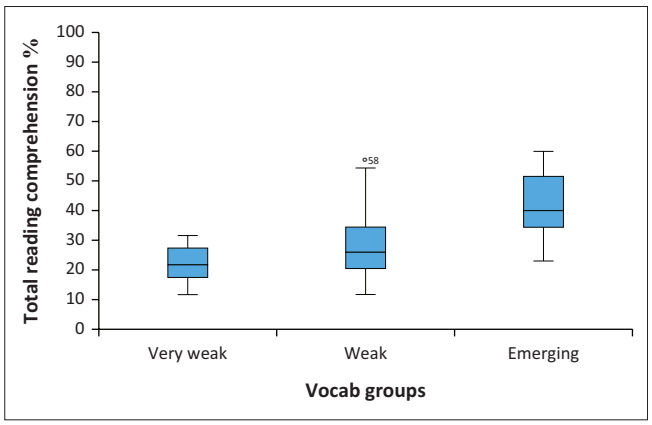

FIGURE 1: Box and whisker plots – Reading comprehension across vocabulary groups.

Finally, the third research question: *Which vocabulary frequency level best predicts performance in reading comprehension of Namibian learners?* A multiple regression was applied with the respective word levels as the predictor variables and reading comprehension as the dependent variable. The scores for the five vocabulary levels were entered as independent variables and the overall reading comprehension score was entered as the dependent variable. A significant model emerged for reading comprehension; F (5, 61) = 6.038, p = 0.000, R^2 = 0.331. Table 6 below shows the vocabulary levels and the coefficients.

Only the 5000 word level emerged as a significant predictor of reading comprehension, accounting for 35% of the variance in reading comprehension (b = 0.370, p = 0.019), with a positive coefficient, indicating that an increase in knowledge at the 5000 word level is associated with an increase in reading comprehension.

Discussion

Although various research studies have shown a strong association between vocabulary knowledge and reading comprehension, very little attention has been given to this topic in the Namibian context. The aim of this study was to assess Grade 11 learners' vocabulary knowledge at five different word frequency levels (2000, 3000, 5000, university word list and 10 000 word level), using the productive VLT, and to examine and describe possible patterns and relationships between their vocabulary knowledge at different word levels and their reading comprehension performance. In this section, we discuss the results according to the three research questions, consider the pedagogical implications for EAL learners and identify the limitations of the study.

In terms of RQ1 (*What are the vocabulary levels and reading comprehension of Namibian learners?),* the results of the vocabulary and the reading comprehension tests show that the Grade 11 learners had weak vocabulary knowledge (total mean of 35%) and they had not yet mastered the most common English words at the 2000 and 3000 word frequency levels. Reading comprehension performance was likewise poor (total mean of 28.5%), even at the easier literal level (39.9%), dropping to 17% in response to inferential questions. As indicated in the literature review, a reading comprehension

TABLE 6: Vocabulary levels and regression coefficients.

Word level	Unstandardised coefficients		Standardised coefficients	t	Significance
	B	Standard error	Beta		
2000	0.192	0.121	0.237	1.587	0.118
3000	-0.041	0.123	-0.053	-0.337	0.737
5000	0.355	0.147	0.370	2.405	0.019
University word list	0.061	0.114	0.087	0.533	0.596
10 000	0.005	0.173	0.005	0.030	0.977

Note: Dependent variable: Total reading comprehension.

level of 50% or less indicates that learners are reading at frustration level with inaccurate decoding skills, which means that they will find it difficult to read and make sense of texts appropriate to their grade level (McCormick 1995). This whole cohort of high school learners was reading at frustration level despite 11 years of schooling.

It is concerning that the majority of these high school learners had not yet mastered basic vocabulary at even the most frequent 2000 word level. Only 17 learners in the cohort (25%) achieved mastery at this basic level, and the highest score at the 3000 word level was only 67%, achieved by four learners in the Emerging vocabulary group. Mastery of the 2000–3000 word level should be attained early in primary school to enable learners to use an additional language effectively, especially if it is a LoLT, and acquire other word levels as they advance from primary to high school. Their limited knowledge of mid-frequency and academic words that occur frequently in written language and are required to adequately comprehend texts in high school is also cause for concern considering that these high school learners use English as the LoLT and that a limited vocabulary will make the reading of high school textbooks extremely challenging (Nation 2001; Nation & Waring 1997). Knowledge of the 10 000 word level typically is usually associated with skilled readers who read independently, so it is not surprising that these EAL learners knew very few words at this frequency range, but their poor knowledge of high frequency words is concerning.

The research literature points to several factors that underpin differential vocabulary growth among both L1 and EAL leaners, such as socio-economic factors, print material in learners' home and school environments, low literacy levels among parents (and teachers), under-resourced schools, and poor instructional practice (Aebersold & Field 1997; Carbonaro & Gamoran 2002; Day & Bamford 1998; Mullis et al. 2012; Rapp & Van den Broek 2005; Webb 2009). While this study was not designed to identify contributory factors to vocabulary development, the poor performance in the vocabulary and comprehension tests suggests that neither the vocabulary development nor the reading abilities of these EAL learners have been given adequate attention and support since primary school for building language or reading proficiency.

Research question 2 examined the *relationship between vocabulary levels and reading comprehension*. The obtained significant correlation of 0.53 between vocabulary knowledge and reading comprehension confirms what many other studies have found, namely that increased vocabulary knowledge is associated with better reading comprehension, for both L1 and EAL readers (Stæhr 2008; Stoffelsma et al. 2024; Zhang 2012). Although some studies have found stronger correlations between vocabulary and reading (e.g. in Staehr's 2008 study a correlation of 0.83 was reported between vocabulary and reading comprehension), consistent low to modest significant associations were obtained between reading comprehension across all the vocabulary levels in this Namibian study, despite variation in performance across the cohort. Some of this variance may have been due to poor decoding skills, as suggested by the low comprehension performance even in easy literal questions. Decoding below a threshold level has been found to 'flatten' the relationship between decoding and reading comprehension (Wang et al. 2019). In other words, decoding is so poor that reading comprehension cannot happen, as predicted by the simple view of reading (Gough & Tunmer 1986). It is also possible that vocabulary knowledge below a threshold level can flatten its association with reading comprehension. This is an area that calls for further research.

Research question 3 examined *which of the vocabulary levels best predicted reading comprehension*. The regression analysis revealed that in this cohort of learners the 5000 word level was a strong predictor of reading comprehension. This was somewhat surprising, as mastery of the 3000 word level has been shown to be important for basic reading comprehension, since these words occur frequently not only in spoken language but also in texts (Nation 2015). Stoffelsma et al. (2024) found that both the 3000 word level and academic words predicted reading comprehension among Ghanaian students. Although all the Grade 11 learners performed poorly at the 5000 word level, those who knew more words at this level also performed far better in reading comprehension. Thus, it seems that for senior high school learners, knowledge of words in the mid-frequency range gives them the edge in reading comprehension.

Pedagogical implications

What is disconcerting about the findings from this study is that despite having been in school for 11 years, these EAL learners had still not gained mastery of the 3000 most frequent words in English, words that are commonly used in daily conversations and also in texts.

The poor vocabulary knowledge of the Grade 11 learners suggests that vocabulary development needs to be given far more attention in Namibian schools. Mastery of the 3000 most frequent words should be a pedagogic goal in early primary school, especially if English is used as the LoLT. For example, the curriculum in South Africa for EAL in Foundation Phase (Grades 1–3) recommends that EAL learners should know between 1500 and 2500 words by the end of Grade 3 (Department of Basic Education 2011:22). As research has shown, knowledge of words at the high frequency level not only enables conversations in English at a basic level, it also makes it easier for learners to make inferences and learn new things while reading (Laufer 1997; Nation 2015). Strong vocabulary growth also increases young learners' motivation and confidence in EAL learning. The fact that the Grade 11 learners only had 70% knowledge of the 2000 word regarded as core vocabulary highlights how wide the vocabulary gap becomes for EAL learners when pedagogy fails them in early primary school.

As children get older, much of their input for learning new words shifts to the written mode and exposure to new words in texts strongly supports word acquisition. The new word and its form are more easily noticed and recognised in written language, and unlike the ephemeral nature of speech, written text is always readily available for rereading and rechecking what a new word looks like and what it might mean in its context. In one study it was found that at least 30 encounters with a new word were required for uptake in oral classroom discourse (Brown, Waring & Donkwaebua 2008), whereas research indicates that in the written mode between 6 and 16 encounters with a new word are required for a new word to be acquired (Rott 1999; Van Hees & Nation 2017; Zahar, Cobb & Spada 2001). It is thus vitally important for teachers to maximise exposure to words in senior primary school and throughout high school, both through explicit instruction and through incidental exposure in the written mode. Schools in disadvantaged areas tend to have fewer book resources, thus putting EAL learners at greater risk of developing inadequate reading skills and constraining vocabulary knowledge.

Limitations and future directions

This was a small study with a sample size of 67 learners from two schools only, so generalising to larger regional or national groups requires caution. However, given the schooling conditions of these Grade 11 learners and the EMIS Education statistics (EMIS 2022, 2023), the outcomes may be a fairly credible reflection of the vocabulary knowledge and reading comprehension of Grade 11 learners in the Zambezi region generally. Despite size limitations, this study provides data on vocabulary knowledge at different frequency levels of high school learners in the Namibian context and contributes to our understanding of the relationship between reading comprehension and vocabulary knowledge, especially for EAL learners in LMIC educational contexts.

The unsatisfactory reliability score of the reading comprehension test used in this study is a limitation. Standardised reading comprehension tests for EAL high school learners in LMIC educational contexts are not readily available or accessible, and even the use of tests designed by a ministry need to be carefully validated before they can be used for research purposes. In future vocabulary studies, the use of more updated productive VLTs that are readily available is also recommended to improve the quality of scholarship in this area.

The decoding fluency skills of the learners were not included in this study. Given the poor reading comprehension levels of the learners, it is possible that inaccurate and slow decoding impeded their reading fluency, making reading slower and more effortful; thus leaving far less working memory and attention available for comprehension. The fact that so many of their responses to the easier literal questions were incorrect suggests that extracting even explicitly stated information in a text is a challenge. Studies of the reading abilities of EAL learners in high school seldom assess fluency, so this is an area that merits attention as reading fluency opens up access to the curriculum. Reading fluency should already be developed in primary school as early as from Grade 3 onwards. Examining the relationship between fluency and knowledge of core vocabulary can also shed light on how inferencing abilities and reading to learn can be supported in EAL learners for whom English is the LoLT.

Conclusion

The relationship between vocabulary development and reading comprehension has been well documented in both home language and EAL readers (Lawrence et al. 2021; Stæhr 2008; Stoffelsma et al. 2024; Webb 2009; Zano & Phatudi 2019). This study confirms this relationship in the Namibian schooling context, where the size of Grade 11 learners' EAL vocabulary was moderately associated with performance in reading comprehension. Knowledge at the 5000 word level emerged as the best predictor of reading comprehension in this context, confirming the value of knowing mid-frequency words in aiding reading comprehension in high school.

Having a large vocabulary may increase one's chances of better understanding a text and, in turn, skilled reading increases one's chances of reading more and increasing one's vocabulary through exposure to the vast range of words that occur in written language. Being able to read with comprehension is associated with the ability to read to learn, becoming an independent learner, expanding one's knowledge, academic success, and gaining access to studies at higher institutions.

Learners of EAL from disadvantaged backgrounds and schools are at greater risk of not having access to quality teaching or ready access to reading resources. The findings in this study that by Grade 11 many learners had still not mastered core words at the basic 2000 word frequency level – a level which typically should be mastered early in primary school – and struggled to answer even easy literal questions in a comprehension test highlights the braking effect that poor schooling can have on reading, learning and academic success. Stakeholders need to be aware of the barriers to learning in high-poverty contexts and make sure that schools become strong sites of language and literacy development. More focused attention needs to be given to developing EAL vocabulary in primary school to support reading comprehension so that learners can cope more effectively with curriculum demands by the time they reach high school.

Acknowledgements

The authors sincerely thank the schools and learners in this study, whose willing participation contributed to useful data on EAL vocabulary in Namibian high schools. The authors also thank the anonymous reviewers of the article for their helpful comments.

Competing interests

The authors declare that they have no financial or personal relationships that may have inappropriately influenced them in writing this article.

Authors' contributions

Both B.L. and E.P. were involved in the design of the methodology and in the conceptualisation of the study. B.L. collected the data and B.L. and E.P. analysed the data. B.L. wrote the original draft. Both authors, B.L. and E.P., revised and edited the final version.

Ethical considerations

An application for full ethical approval was made to the Higher Degrees Committee and Ethics subcommittee, Department of Linguistics and ethics consent was received on 23 January 2013. The ethics approval number is AL_BLL1_2013. The principals of the two schools gave their consent for the schools to participate in the study. The Grade 11 learners' participation was voluntary, and the anonymity and confidentiality of the participating schools and learners was respected.

Funding information

This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

Data availability

The data that support the findings of this study are not openly available due to confidentiality and are available from the corresponding author, B.L., upon reasonable request.

Disclaimer

The views and opinions expressed in this article are those of the authors and are the product of professional research. It does not necessarily reflect the official policy or position of any affiliated institution, funder, agency, or that of the publisher. The authors are responsible for this article's results, findings, and content.

References

Aebersold, J.A. & Field, M.L., 1997, *From reading to reading teacher: Issues and strategies for second language classrooms*, Cambridge University Press, Cambridge.

Akbarian, I., 2010, 'The relationship between vocabulary size and depth for ESP/EAP learners', *System* 38(3), 391–401. https://doi.org/10.1016/j.system.2010.06.013

Bennett, G.R., 2010, *Using corpora in the language learning classroom: Corpus linguistics for teachers*, University of Michigan Press, Ann Arbor, MI.

Bintz, W.P., 2011, 'Teaching vocabulary across the curriculum', *Middle School Journal* 42(4), 44–53. https://doi.org/10.1080/00940771.2011.11461773

Brooks, G., Clenton, J. & Fraser, S., 2021, 'Exploring the importance of vocabulary for English as an additional language learners' reading comprehension', *Studies in Second Language Learning and Teaching* 11(3), 351–376. https://doi.org/10.14746/ssllt.2021.11.3.3

Brown, R., Waring, B. & Donkaewbua, S., 2008, 'Incidental vocabulary acquisition from reading, reading-while-listening, and listening to stories', *Reading in a Foreign Language* 20(2), 136–163. https://doi.org/10.64152/10125/66816

Carbonaro, W.J. & Gamoran, A., 2002, 'The production of achievement inequality in high school English', *American Educational Research Journal* 39(4), 801–827. https://doi.org/10.3102/00028312039004801

Castles, A., Rastle, K. & Nation, K., 2018, 'Ending the reading wars: Reading acquisition from novice to expert', *Psychological Science in the Public Interest* 19(1), 5–51. https://doi.org/10.1177/1529100618772271

Coady, J., 1993, 'Research on ESL/EFL vocabulary acquisition: Putting it all in context', in T. Huckin, M. Haynes & J. Coady (eds.), *Second language reading and vocabulary learning*, pp. 3–23, Ablex, Norwood, NJ.

Cooper, P.A., 1999, 'Lexis and the undergraduate: Analysing vocabulary needs, proficiencies and problems', Unpublished MA dissertation, University of South Africa, Pretoria.

Coxhead, A., 2000, 'A new academic word list', *TESOL Quarterly* 34(2), 213–238. https://doi.org/10.2307/3587951

Crosson, A.C., Lesaux, N.K. & Martiniello, M., 2008, 'Factors that influence comprehension of connectives among language minority children from Spanish-speaking backgrounds', *Applied Psycholinguistics* 29(4), 603–625. https://doi.org/10.1017/S0142716408080260

Cunningham, A.E. & Stanovich, K.E., 2001, 'What reading does for the mind', *Journal of Direct Instructions* 1(2), 137–149.

Day, R.R. & Bamford, J., 1998, *Extensive reading in the second language classroom*, Cambridge University Press, Cambridge.

DeFord, D.E., 1991, 'Fluency in initial reading instruction: A reading recovery lesson', *Theory into Practice* 30(3), 201–210. https://doi.org/10.1080/00405849109543501

Department of Basic Education, 2011, *Curriculum and Assessment Policy Statement (CAPS) Foundation Phase grades 1–3, English first additional language*, Department of Basic Education, Republic of South Africa, Pretoria.

Education Management Information System (EMIS), 2022, *EMIS Education statistics 2021*, Ministry of Education, Arts and Culture, Windhoek.

Education Management Information System (EMIS), 2023, *EMIS Education statistics 2022*, Ministry of Education, Arts and Culture, Windhoek.

Elmasry, H.I., 2012, 'Depth and breadth of vocabulary knowledge: Assessing their roles in reading comprehension of high-school EFL learners in the UAE', MEd thesis, The British University in Dubai.

Gough, P.B. & Tunmer, W.E., 1986, 'Decoding reading and reading disability', *Remedial and Special Education* 7(1), 6–10. https://doi.org/10.1177/074193258600700104

Grabe, W., 1991, 'Current development in second language reading research', *TESOL Quarterly* 25(3), 375–396. https://doi.org/10.2307/3586977

Grabe, W., 2010, 'Fluency in reading: Thirty-five years later', *Reading in a Foreign Language* 22(1), 71–83. https://doi.org/10.1093/oxfordhb/9780195384253.013.0006

Harris, P.G., 2011, *Language in schools in Namibia: The missing link in educational achievement?* The Urban Trust of Namibia monograph no. 1, The Urban Trust of Namibia, Windhoek.

Hatami, S. & Tavakoli, M., 2012, 'The role of depth versus breadth of vocabulary knowledge in success and ease in L2 lexical inferencing', *TESL Canada Journal* 30(1), 1–21. https://doi.org/10.18806/tesl.v30i1.1123

Howie, S., Venter, E., Van Staden, S., Zimmerman, L., Long, C., Du Toit, C. et al., 2008, *PIRLS 2006 summary report: South African children's reading achievement*, Centre for Evaluation and Assessment, Pretoria.

Hughes, A., 2003, *Testing for language teachers*, Cambridge University Press, Cambridge.

IBM Corp., 2020, *IBM SPSS statistics for windows (Version 27.0)*, [Computer software], IBM Corp., Armonk, New York.

Kim, Y.-S.G., 2017, 'Why the simple view of reading is not simplistic: Unpacking the simple view of reading using a direct and indirect effect model of reading (DIER)', *Scientific Studies of Reading* 21(4), 310–333. https://doi.org/10.1080/10888438.2017.1291643

Kim, Y.S.G., 2020, 'Toward integrative reading science: The direct and indirect effects model of reading', *Journal of Learning Disabilities* 53(6), 469–491. https://doi.org/10.1177/0022219420908239

Kučera, H. & Francis, W.N., 1967, *Computational analysis of present day American English*, Brown University Press, Providence, RI.

Laufer, B. & Nation, P., 1999, 'A vocabulary size test of controlled productive ability', *Language Testing* 16(1), 33–51. https://doi.org/10.1177/026553229901600103

Laufer, B. & Ravenhorst-Kalovski, G.C., 2010, 'Lexical threshold revisited: Lexical text coverage, learners' vocabulary size and reading comprehension', *Reading in a Foreign Language* 22(1), 15–30.

Laufer, B., 1997, 'The lexical plight in second acquisition: Words you don't know, words you think you know, and words you can't guess', in J. Coady & T. Huckin (eds.), *Second language vocabulary acquisition*, pp. 20–34, Cambridge University Press, New York, NY.

Laufer, B., 1998, 'The development of passive and active vocabulary in a second language: Same or different?', *Applied Linguistics* 19(2), 255–271. https://doi.org/10.1093/applin/19.2.255

Lawrence, J.F., Knoph, R., McIlraith, A., Kulesz, P.A. & Francis, D.J., 2021, 'Reading comprehension and academic vocabulary: Exploring relations of item features and reading proficiency', *Reading Research Quarterly* 57(2), 669–690. https://doi.org/10.1002/rrq.434

Lee, C.D. & Spratley, A., 2010, *Reading in the disciplines: The challenges of adolescent literacy*, Carnegie Corporation of New York, New York, NY.

Li, L. & MacGregor, L.J., 2010, 'Investigating the receptive vocabulary size of university-level Chinese learners of English: How suitable is the vocabulary levels test?', *Language and Education* 24(3), 239–249. https://doi.org/10.1080/09500781003642478

Li, M. & Kirby, J.R., 2014, 'The effects of vocabulary breadth and depth on English reading', *Applied Linguistics* 36(5), 1–25. https://doi.org/10.1093/applin/amu007

Liswaniso, B.L. & Mubanga, G.N., 2019, 'Examining the reading habits of university students: A study of students at the Katima Mulilo Campus of the University of Namibia', *Electronic Journal of Foreign Language Teaching* 16(1), 140–152. https://doi.org/10.56040/bdlg1619

Liswaniso, B.L., 2015, 'Exploring the relationship between English additional language reading comprehension and vocabulary size', Unpublished master's dissertation, Dept. of Linguistics and Modern Languages, University of South Africa.

Liswaniso, B.L., 2021, 'The design and effects of a catch-up reading intervention for Grade 5 teachers and learners in Namibia', PhD thesis, Dept. of Linguistics and Modern Languages, University of South Africa.

Liu, F., 2010, 'Reading abilities and strategies: A short introduction', *International Education Studies* 3(3), 153–157. https://doi.org/10.5539/ies.v3n3p153

McCormick, S., 1995, *Instructing students who have literacy problems*, Merrill, Englewood Cliffs, NJ.

Meara, P.M., 2009, *Connected words: Words associations and second language vocabulary acquisition*, John Benjamins Publishing, Amsterdam.

Mullis, I.V.S. & Martin, M.O. (eds.), 2019, *PIRLS 2021 assessment frameworks*, TIMSS & PIRLS International Study Centre, Boston College, Chestnut Hill, MA.

Mullis, I.V.S., Martin, M.O., Foy, P. & Drucker, K.T., 2012, *PIRLS 2011 international results in reading*, TIMSS & PIRLS International Study Centre, Boston College, Chestnut Hill, MA.

Mutema, F. & Pretorius, E.J., 2024, 'Does it matter if some learners read slowly? Exploring relationships between reading comprehension and oral reading fluency', *South African Journal of Childhood Education* 14(1), a1518. https://doi.org/10.4102/sajce.v14i1.1518

Nassaji, H., 2003, 'Higher-level and lower level text processing in advanced ESL reading comprehension', *The Modern Language Journal* 87(2), 261–276. https://doi.org/10.1111/1540-4781.00189

Nation, I.S.P. & Beglar, D., 2007, 'A vocabulary size test', *Language Teacher* 31(7), 9–13.

Nation, I.S.P. & Waring, P., 1997, 'Vocabulary size, text coverage and word lists', in N. Schmitt & M. McCarthy (eds.), *Vocabulary: Description, acquisition and pedagogy*, pp. 9–19, Cambridge University Press, Cambridge.

Nation, I.S.P., 2001, 'How good is your vocabulary program?', *ESL Magazine* 4(3), 22–24.

Nation, I.S.P., 2006, 'How large a vocabulary is needed for reading and listening', *The Canadian Modern Language Review* 63(1), 59–82. https://doi.org/10.3138/cmlr.63.1.59

Nation, P. & Anthony, L., 2013, 'Mid-frequency readers', *Journal of Extensive Reading* 1, 5–16.

Nation, P., 2012, *The vocabulary size test*, viewed 02 July 2024, from https://www.wgtn.ac.nz/lals/about/staff/Publications/paul-nation/Vocabulary-Size-Test-information-and-specifications.pdf.

Nation, P., 2015, 'Principles guiding vocabulary learning through extensive reading', *Reading in a Foreign Language* 27(1), 136–145.

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, reports of the subgroups*, National Institute of Child Health and Human Development, Rockville, MD.

Nguyen, L.T.C. & Nation, I.S.P., 2011, 'A bilingual vocabulary size test of English for Vietnamese learners', *RELC Journal* 42(1), 86–99. https://doi.org/10.1177/0033688210390264

Padgett, R., 1997, *Creative reading: What it is, how to do it, and why*, National Council of Teachers, Urbana, IL.

Pikulski, J.J. & Chard, D.J., 2005, 'Fluency: Bridge between decoding and reading comprehension', *The Reading Teacher* 58(6), 510–519. https://doi.org/10.1598/RT.58.6.2

Pikulski, J.J. & Templeton, S., 2004, *Teaching and developing vocabulary: Key to long-term reading success*, Houghton Mifflin Company, Boston, MA.

Pretorius, E.J. & Murray, S., 2019, *Teaching reading comprehension*, Oxford University Press Southern Africa, Cape Town.

Pretorius, E.J. & Spaull, N., 2016, 'Exploring relationships between oral reading fluency and reading comprehension amongst English second language readers in South Africa', *Reading and Writing* 29(7), 1449–1471. https://doi.org/10.1007/s11145-016-9645-9

Pretorius, E.J. & Stoffelsma L., 2017, 'How is their word knowledge growing? Exploring Grade 3 vocabulary in South African township schools', *South African Journal of Childhood Education* 7(1), 1–13. https://doi.org/10.4102/sajce.v7i1.553

Pretorius, E.J., 2002, 'Reading ability and academic performance in South Africa: Are we fiddling while Rome is burning?', *Language Matters* 33(1), 169–196. https://doi.org/10.1080/10228190208566183

RAND Reading Study Group (RRSG), 2002, *Toward an R&D program in reading comprehension*, RAND, Santa Monica, CA.

Rapp, D.N. & Van den Broek, P., 2005, 'Dynamic text comprehension: An integrative view of reading', *Current Directions in Psychological Science* 14(5), 276–279. https://doi.org/10.1111/j.0963-7214.2005.00380.x

Read, J., 2000, *Assessing vocabulary*, Cambridge University Press, Cambridge.

Rott, S., 1999, 'The effect of exposure frequency on intermediate language learners' incidental vocabulary acquisition through reading', *Studies in Second Language Acquisition* 21(4), 589–619. https://doi.org/10.1017/S0272263199004039

Scarborough, H.S., 2001, 'Connecting early language and literacy to later reading (dis) abilities: Evidence, theory, and practice', in S. Neuman & D. Dickinson (eds.), *Handbook for research in early literacy*, pp. 97–110, Guilford Press, New York, NY.

Schmitt, D. & Schmitt, N., 2014, 'A reassessment of frequency and vocabulary size in L2 vocabulary teaching', *Language Teaching* 47(4), 484–503. https://doi.org/10.1017/S0261444812000018

Schmitt, N., 2008, 'Review article: Instructed second language vocabulary learning', *Language Teaching Research* 12(3), 329–363. https://doi.org/10.1177/1362168808089921

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. https://doi.org/10.1111/j.1540-4781.2011.01146.x

Schmitt, N., Schmitt, D. & Clapham, C., 2001, 'Developing and exploring the behaviour of two new versions of the vocabulary levels test', *Language Testing* 18(1), 55–88. https://doi.org/10.1177/026553220101800103

Smith, F., 2004, *Understanding reading: A psycholinguistic analysis of reading and learning to read*, Lawrence Erlbaum Associates, London.

Southern and Eastern Africa Consortium for Monitoring Educational Quality, 2005, *The SACMEQ II project in Namibia: A study of the conditions of schooling and quality of education*, [Compiled by] D. Makuwa, SACMEQ, Harare.

Southern and Eastern Africa Consortium for Monitoring Educational Quality, 2010, *SACMEQ III project results: Pupil achievement levels in reading and Mathematics*, Working document no. 1., Paris, SACMEQ

Stæhr, L.S., 2008, 'Vocabulary size and the skills of listening, reading and writing', *Language Learning Journal* 33(2), 139–152. https://doi.org/10.1080/09571730802389975

Stæhr, L.S., 2008, 'Vocabulary size and the skills of listening, reading and writing', *Language Learning Journal* 33(2), 139–152. https://doi.org/10.1080/09571730802389975

Stoffelsma, L. & Spooren, W., 2019, 'The relationship between English reading proficiency and academic achievement of first-year science and mathematics students in a multilingual context', *International Journal of Science and Mathematics Education* 17(5), 905–922. https://doi.org/10.1007/s10763-018-9905-z

Stoffelsma, L., Antwi, V. & Hanson, R., 2024, 'Incidental vocabulary learning in a postcolonial English second language context', *Language Matters* 55(1–2), 92–113. https://doi.org/10.1080/10228195.2024.2341010

Tötemeyer, A.-J., 2010, *Multilingualism and the language policy for Namibian schools*, Project for alternative education in South Africa PRAESA occasional papers no. 37, University of Cape Town, Cape Town.

Tötemeyer, A.-J., 2013, 'Few books, many languages: Namibian publishing for children', *Mousaion* 31(1), 3–26.

Van Hees, J. & Nation, P., 2017, *What every primary school teacher should know about vocabulary*, NZCER, Wellington.

Wang, Z., Sabatini, J., O'Reilly, T. & Weeks, J., 2019, 'Decoding and reading comprehension: A test of the decoding threshold hypothesis', *Journal of Educational Psychology* 111(3), 387–401. https://doi.org/10.1037/edu0000302

Webb, S. & Nation, I.S.P., 2008, 'Evaluating the vocabulary load of written text', *TESOLANZ Journal* 16, 1–10.

Webb, S., Sasao, Y. & Ballance, O., 2017, 'The updated vocabulary levels test: Developing and validating two new forms of VLT', *International Journal of Applied Linguistics* 168(1), 33–69. https://doi.org/10.1075/itl.168.1.02web

Webb, S.A. & Sasao, Y., 2013, 'New directions in vocabulary testing', *RELC Journal* 44(3), 263–277. https://doi.org/10.1177/0033688213500582

Webb, S.A., 2009, 'The effects of pre-learning vocabulary vocabulary on reading comprehension and writing', *The Canadian Modern Language Review* 65(3), 441–470. https://doi.org/10.3138/cmlr.65.3.441

Wills, G., Ardington, C., Pretorius, E.J. & Sebaeng, L., 2022, *Benchmarking early grade reading skills: English first additional language. Summary report*, Khulisa Management Services, Johannesburg.

Wood, C.L., Schatschneider, C. & VelDink, A., 2021, 'The relation between academic word use and reading comprehension for students from diverse backgrounds', *Language, Speech, and Hearing Services in Schools* 52(1), 273–287. https://doi.org/10.1044/2020_LSHSS-19-00099

Xing, P. & Fulcher, G., 2007, 'Reliability assessment for two versions of vocabulary levels tests', *System* 35(2), 182–191. https://doi.org/10.1016/j.system.2006.12.009

Zahar, R., Cobb, T. & Spada, N., 2001, 'Acquiring vocabulary through reading: Effects of frequency and contextual richness', *Canadian Modern Language Review* 57(4), 541–572. https://doi.org/10.3138/cmlr.57.4.541

Zano, K. & Phatudi, N., 2019, 'Relationship between vocabulary knowledge and reading comprehension of South African EFAL high school learners', *Per Linguam* 35(3), 16–28. https://doi.org/10.5785/35-3-830

Zhang, D., 2012, 'Vocabulary and grammar knowledge in second language reading comprehension: A structural equation modelling study', *The Modern Language Journal* 96(4) 558–575. https://doi.org/10.1111/j.1540-4781.2012.01398.x

Zhou, S., 2010, 'Comparing receptive and productive academic vocabulary knowledge of Chinese EFL learners', *Asian Social Science* 6(10), 14–19. https://doi.org/10.5539/ass.v6n10p14