


English vocabulary exposure in South African township schools: Pitfalls and opportunities

**Author:**Lieke Stoffelsma^{1,2} **Affiliations:**

¹Department of Linguistics and Modern Languages, University of South Africa, South Africa

²Centre for Language Studies, Faculty of Arts, Radboud University Nijmegen, The Netherlands

Corresponding author:

Lieke Stoffelsma,
l.stoffelsma@let.ru.nl

Dates:

Received: 25 July 2018

Accepted: 19 Nov. 2018

Published: 21 Feb. 2019

How to cite this article:

Stoffelsma, L., 2019, 'English vocabulary exposure in South African township schools: Pitfalls and opportunities', *Reading & Writing* 10(1), a209. <https://doi.org/10.4102/rw.v10i1.209>

Copyright:

© 2019. The Authors.
Licensee: AOSIS. This work is licensed under the Creative Commons Attribution License.

Background: This small-scale study investigated English vocabulary exposure from graded readers and teacher talk in Grade 3 classrooms in poorly resourced township schools in South Africa. Vocabulary is one of the key building blocks for becoming a fluent reader. Most words are learnt through incidental exposure to oral or written language.

Objectives: This study is a first attempt to investigate opportunities for incidental vocabulary exposure in poorly resourced classrooms in South Africa.

Method: A corpus linguistics approach was used to analyse a written corpus of 57 143 tokens and a spoken corpus of 12 242 tokens.

Results: The study showed that there are vast differences between levels of written and spoken vocabulary in the classrooms and that the role for oral vocabulary exposure in classrooms is restricted. Spoken vocabulary registered above the K-3 word frequency level largely came from teachers' read alouds of print materials.

Conclusion: The study findings show that even in contexts where print exposure is limited, oral language cannot compensate for the richness of written vocabulary. Situational constraints, such as lack of books, negatively influenced the effective use of graded readers. Opportunities for incidental vocabulary learning, as well as implications for policy and further research, are discussed.

Introduction

Despite South Africa's economic growth, literacy levels in the country remain problematic. South Africa's participation in three cycles of Progress in International Reading Literacy Study (PIRLS) (2006, 2011 and 2016) shows consistently low reading comprehension levels at the fourth grade. The most recent data indicate that 78% of South African Grade 4 children cannot read for meaning or retrieve basic information from the text to answer simplistic questions, compared to 4% of students internationally (Mullis et al. 2017:55). Although South Africa's reading performance in the African languages has been consistently low over the past three PIRLS cycles, English reading results have also not improved over the past 10 years. English is one of the 11 official languages in South Africa and is used as the most common language of learning from Grade 4 onwards.

The inability to read is problematic because it restricts children from participating in formal education. It prevents them from successfully learning textbook content, participating in classroom discourse or responding reliably to tests (Abadzi 2006). Children who do not master reading skills from an early age will find it difficult to catch up to their required grade level and run the risk of falling further behind when they move on to higher grades (Spaull 2017).

What are the causes of these low reading levels in South Africa? Spaull (2017) argues that the most prominent reason for low reading levels in South African primary schools is that school teachers have not acquired the knowledge and skills in their professional development to teach reading. Lack of resources in the system is another major prominent cause. The indicator on access to libraries and multimedia centres shows that only 37% of learners in primary schools in South Africa have access to a library in their school (Department of Basic Education 2014). Scarcity of educational resources is particularly prevalent in historically African township schools and rural schools because of inequalities that originated in the apartheid era, and which are often complicated by poor home environments (Amin & Ramathan 2009; Department of Basic Education 2014). The percentage of learners that have their own reading textbook in schools serving poor communities is low: from 35.6% in Quintile 1 schools¹ to 43.4% in Quintile 3 schools

1.State aid to public schools in South Africa is determined by socio-economic (SE) factors. Schools serving poor communities receive the most funding. Schools are categorised from quintiles 1 to 5, with quintiles 1–3 being the poorer schools.

Read online:

Scan this QR code with your smart phone or mobile device to read online.

(Spaull 2011). Evidence shows that under-resourced schools produce lower results in the Grade 4 PIRLS reading test (Mullis et al. 2017). Based on the analysis of the Southern Africa Consortium for Monitoring Educational Quality (SACMEQ) results for South Africa, Van der Berg (2008) argues that school resources in the country matter only conditionally and there are vast differences between schools in their ability to convert these resources into outcomes. In summary, simply adding more resources does not necessarily improve school performance.

A better understanding of literacy practices in poorly resourced classrooms in South Africa is important for two reasons. Firstly, the social, political and economic benefits that literacy brings to nations and individuals have been acknowledged by the international community at large (Mullis et al. 2007; UNESCO 2005). Improving literacy levels of future generations is needed to improve the development of countries that struggle with their literacy levels. Secondly, most of the knowledge that we have about written and spoken literacy development comes from the so-called Western contexts, where language exposure in classrooms tends to be rich and abundant. It is important to derive theories from local empirical research, rather than only infer from research contexts that do not share the typical characteristics of poorly resourced classrooms.

In an attempt to contribute to a better understanding of literacy practices in poorly resourced classrooms, this study investigates the English exposure in Grade 3 township school classrooms that serve poor communities in South Africa. By taking a linguistics perspective on school resources, the study aims to gain a better understanding of the potential contribution of graded readers and teacher talk to incidental vocabulary development. The emphasis on English exposure in Grade 3 is relevant because more than 70% of learners in South Africa shift from an African Home Language (HL) in Grade 3 to English as the medium of instruction in Grade 4 (Fleisch 2008) and many of these learners are not adequately prepared to meet the English requirements in the intermediate phase (grades 4–6) (Pretorius & Stoffelsma 2017; MacDonald & Burroughs 1991). Before we proceed further, a brief overview of oral and written language exposure in relation to vocabulary development in the early school years is presented.

Theoretical framework

Vocabulary development

The amount of reading that students do for school and for personal enjoyment has a positive effect on their reading achievement and comprehension skills and contributes to their academic knowledge development (Mol & Bus 2011; Mullis et al. 2007). One of the key building blocks for becoming a fluent reader is vocabulary knowledge. Research in both first-language (L1) and second-language (L2) contexts shows that vocabulary knowledge correlates strongly with reading proficiency and is an important prerequisite for the development of reading skills (Helman & Burns 2008;

Read 2004) and academic performance (Nagy & Townsend 2012). Research further shows that vocabulary is a strong predictor of reading achievement in Grade 4 (Scarborough 2001). Moreover, vocabulary size at the end of Grade 1 is a significant predictor of reading comprehension 10 years later (Cunningham & Stanovich 2001). Print exposure is essential in the development of vocabulary (Stanovich 2000). People who are highly exposed to print are more likely to encounter rare words, which enhances their vocabulary growth; they will encounter more complex syntactic structures and will acquire more knowledge about the world (Long, Johns & Morris 2006).

There are large differences in vocabulary development of children, depending on their home background and school context. In general, children from low socio-economic backgrounds have a poorer understanding of words. Research has identified that a socio-economic gap in vocabulary can be established by the age of 3 years (Hart & Risley 1995) and even as early as the age of 18 months (Fernald, Marchman & Weisleder 2013). This gap is difficult to bridge in the years that follow.

Lexical acquisition from reading and listening

The majority of vocabulary growth occurs through incidental exposure to oral or written language (Cunningham 2005). Research is somewhat divided on how many reading exposures are necessary to obtain an initial receptive knowledge of words through reading and listening. Webb (2007) discovered that if L2 learners of English encounter unknown words 10 times while reading, this led to sizeable vocabulary learning gains. However, he argues that more than 10 repetitions may be needed to develop a full knowledge of a word. Waring and Tataka (2003) found that L2 learners of English, who engage in reading graded readers, need to meet a word at least eight times in order to have about a 50% chance of recognising a word's form 3 months later. However, for word *learning* to occur, their data suggest that it would take at least 20 or even 30 reading encounters with a word for it to be learnt. There is less knowledge in the research field about incidental vocabulary development through listening. How vocabulary is learnt from spoken discourse and how many repetitions are needed for a word to be learnt are questions that still require further investigation (Schmitt 2010). There are, however, a few studies that have investigated the effect of oral language exposure in educational contexts. Horst (2010) shows that acquiring vocabulary knowledge through teacher talk is not an efficient method for adult ESL learners, mainly because many words that are frequently used in texts are unlikely to be encountered in speech. On the other hand, studies that have investigated the effects of storybook reading to children have found that extensive listening to easy texts for enjoyment is beneficial for incidental vocabulary learning of children in L1 and L2 contexts (Elley 1991; Nation 2001).

Research into the effect of storybook reading on vocabulary development shows that many word repetitions are needed

before a word can be learnt. In their study amongst 18–21-year-olds, Brown, Waring and Donkaewbua (2008) showed that storybook reading only resulted in word learning if the word was encountered more than 20 times. However, they suggested that probably as much as 50 or even 100 word meetings may be needed to acquire a word's meaning from listening-only. Clearly, factors such as quality of the reading and/or listening engagement (intensive or extensive) and the proficiency level of the learners will influence the learning of a new word.

It is generally accepted that written language offers a more diverse and abundant vocabulary than oral language. Mol and Bus (2011) argue that reading for enjoyment exposes students to a rich vocabulary and words that they are unlikely to encounter in speech. Cunningham (2005) argues that the lexical density of oral language, compared to written language, is 'substantially degraded or impoverished' (p. 50). She argues that people who are engaged in conversation rely greatly on the use of common words because of time pressure. However, when people produce written language, they are usually allowed more time to compose a more refined communication. Consequently, written language is a more effective way of building a child's vocabulary than oral language (Cunningham 2005). This does not mean that oral exposure is not important. On the contrary, verbal input is also considered to be a precondition for good vocabulary growth. Teacher talk as a resource for vocabulary learning can be particularly important when the L1 in the learner's home environment is different from the Language of Teaching and Learning (LoTL). Research shows that schools in high-poverty contexts play an important role in compensating for the limited literacy opportunities in children's homes (Howie & van Staden 2012).

The difference between oral and written exposure is a given and does not need further corroboration. However, how this difference plays out in low-resourced L2 classrooms has not been investigated. Information from the field, albeit at a small scale, presented in this article, will provide a better understanding of these differences in practice.

Vocabulary in Grade 3

The transition from Grade 3 to Grade 4 implies a change in the reading demands of learners. Rather than learning to read, learners will focus more on reading to learn. This transition often implies a sudden decrease in reading scores (Hirsch 2003). One way to prepare learners for this linguistic challenge is to ensure that they have a sufficient level of English vocabulary by the end of Grade 3. A recent study found that English HL learners in Western Cape township schools had a receptive vocabulary knowledge of only 61% of the most frequent words by the end of Grade 3, whereas their English Foreign Additional Language (FAL) counterparts only knew 27% of these words (Pretorius & Stoffelsma 2017). This not only shows that there are vast differences between English HL and FAL learners, but it also shows that they are not adequately prepared for the vocabulary level of Grade 4.

Research has produced clear indications as to how many words are needed for adequate reading comprehension. Hu and Nation (2000) calculated that most non-native speakers of English would need to know 98% of the words in a text to gain adequate reading comprehension. Based on a word coverage of 98%, Nation (2006) calculated that for L2 learners of English to be able to read and understand graded readers, they would need knowledge of the 3000 most frequent words of the British National Corpus (BNC). This idea is in line with Nation and Waring's (1997) recommendation for second-language learners to learn at least 3000 high-frequency (HF) words.

Schmitt (2010) argues that the frequency at which a word occurs is 'arguably the single most important characteristic of lexis that researchers must address' (p. 63). High-frequency words give an impression of important vocabulary in a language and HF word lists provide a valuable indication of which words should be prioritised in vocabulary development. There is a wealth of HF word lists for various contexts. For Grade 3 learners in South Africa, the following word lists are some of the useful lists based on which a learner's word levels can be measured: the Dolch list (Dolch 1948), the Sibanda list (Sibanda 2014) and the Academic Word List (AWL) (Coxhead 2000). The Dolch list contains 220 HF English sight words based on children's books from pre-primary to third grade, which children should master by the end of Grade 3. Although it was developed long ago, it is still a commonly used HF word list in primary schools (cf. Miles, Rubin & Gonzalez-Frey 2017). The Sibanda list was developed for all South African Grade 3 learners and it contains English HF words that learners should master before their transition to Grade 4. The AWL contains 570 word families and is especially suitable for university students. However, 83.5% of the AWL contains words from 1000–3000 levels (words such as *area*, *evaluate*, *normal*, *respond* and *team*) that students are likely to encounter in Grade 4 expository texts.

Graded readers

Assisting learners in building their vocabulary repertoire while controlling for HF words can be done by offering graded reader schemes. These readers are written for learners of English, controlling for lexis and syntax, with increasing difficulty in terms of language, length and format, as the reader moves on to a higher level (Hill 2008). Graded readers are a very important resource for developing reading skills, language consolidation and vocabulary building, as they provide ample opportunity for extensive reading at the appropriate vocabulary level (Waring & Nation 2004).

In general, reading through all the levels of a graded reader scheme can provide favourable conditions for incidental vocabulary learning. Corpus research by Nation and Wang (1999) shows that the conditions for incidental vocabulary learning are good at higher levels of graded readers. Their study shows that graded readers can be an effective way of ensuring that learners are exposed to HF words with sufficient repetition. In order for this to be effective, they argue that

learners need to work their way through the levels, read at least five books per level and at an intensive rate of about one book per week.

The South African Curriculum and Assessment Policy Statement (CAPS 2011) stipulates that Grade R–Grade 3 learners use graded readers as part of group-guided reading. This is a group reading session where the teacher scaffolds learning (Department of Basic Education 2011).

Purpose of the study

There is a lack of published evidence on print exposure in high-poverty schools in South Africa. The aim of this article was to localise international research on the nature of oral and written language exposure in order to get a better understanding of what is happening inside South African classrooms. The aim of this small-scale corpus linguistics study was to investigate vocabulary exposure from graded readers and teacher talk in Grade 3 classrooms to determine what kind of incidental vocabulary learning opportunities this exposure provides. The focus on Grade 3 is intentionally chosen because reading levels by the end of Grade 3 are a significant predictor for later achievement in secondary school (Lesnick et al. 2010). The following research questions guided this article:

- To what written English vocabulary from graded readers are Grade 3 learners in South African township schools exposed and what opportunities do these exposures offer for incidental word learning?
- How are English-graded readers being used in poorly resourced South African classrooms?
- To what English oral vocabulary from teacher talk are Grade 3 learners in South African township schools exposed and what opportunities do these exposures offer for incidental word learning?

Research methods and design

Schooling context

Four Grade 3 classes from two different urban township schools in the Western Cape province participated in the study. Both schools participated in a 3-year literacy project, the Zenex Literacy Project (2015–2017). The aim of the project was to improve literacy levels of Foundation Phase learners through improved classroom literacy practices at quintiles 1–3 schools across three provinces. The two schools were English HL Quintile 3 primary schools, with Afrikaans as their FAL. The schools served mainly low-income communities and were situated in the township areas of Cape Town. The learners consisted of both Afrikaans and English mother tongue speakers.

Sample

Convenience sampling was applied for the selection of the schools. The teachers were asked to provide a list of all the available graded readers used in their classrooms. The 24 graded readers that they listed included 19 readers from the

TABLE 1: Overview of recorded lessons.

Teacher	School	Children in class	Recorded class time	Topic of lesson
Teacher A	School A	38	27 min 51 s	<ul style="list-style-type: none"> • Mathematics • Short story: 'The grasshopper and the ant'
Teacher B	School A	35	47 min 9 s	<ul style="list-style-type: none"> • News of the day • Life skills: insects
Teacher C	School B	41	56 min 53 s	<ul style="list-style-type: none"> • Short story: 'Another mystery' • Short story: 'The first blue jeans' • Language test
Teacher D	School B	40	1 h 7 min 8 s	<ul style="list-style-type: none"> • Short story 'The fall' • Poem 'The spider and the fly' • Reader: <i>Jungle shorts</i>
Total	-	-	3 h 19 min 1 s	-

Oxford Reading Tree scheme (stages 10–12), published between 1995 and 2011. These readers accounted for 60% of the corpus. The remaining 40% of the corpus consisted of Kathy and Mark basic readers, which were published between 1966 and 1971. All readers were fiction-graded readers, with reading levels clearly indicated for each book, indicating increasing difficulty in terms of language, length and format. The target audience of the readers are children aged 7–11 years. Permission was granted by the same four teachers to video-record a randomly selected lesson (for details, see Table 1).

Data analysis

Data analysis followed the argument by McCarthy and Carter (1997) that spoken and written discourses are essentially different and, therefore, written and spoken corpora should be separately constructed when examined on word distribution. Readers were transcribed into digital formats and recorded lessons were transcribed verbatim, leading to a separate written and spoken corpus. Learners' responses were not included in the transcriptions. Some teachers administered tests or assignments during class time, during which there was no oral exposure. These 'silent' parts were not included in the analysis. Consequently, the length of the recordings varied per class.

Both files were converted into plain output texts, in which all punctuation was eliminated and all numerals (1, 20, etc.) were replaced by words (one, twenty, etc.). Contractions were replaced by constituent words (*won't* => *will not*), and all proper nouns were excluded from the files (adding up to a total of 3052 nouns for the written corpus and 320 nouns for the spoken corpus). Single letters were eliminated as words except for 'a' and 'I'. Vocabprofile (<http://www.lexutor.ca/vp/comp/>) was used to compare the written and spoken corpora with the BNC-COCA 1–25 framework² and determine word frequency levels. Generated output included the total number of words (word tokens), the number of distinct words (word types) and the total number of word families. Emphasis in the results section will be on analysis of the word families. Using the family as a unit of comparison means that if the root form *stimulate* is in corpus 1 and the regular derivation *stimulation* in corpus 2, then this is

²The BNC-COCA is an integration of the British National Corpus (BNC) containing 100 million English words and the Corpus of Contemporary American English (COCA) containing 450 million American-English words.

considered a repetition of the same family, that is, these are equivalent tokens.

Because of a lack of agreement on the number of written and spoken exposures necessary for incidental word learning to occur, it was difficult to determine a cut-off point for the generation of HF words. It was decided to follow the local study by Sibanda (2014) who used 30 word occurrences as a cut-off point. Two HF word lists were composed: one for the written and one for the spoken corpus. High-frequency words for both corpora were calculated with WordSmith. These lists were compared with the Dolch list (1948), the Sibanda list (2014) and the AWL (Coxhead 2000) using Tex Lex Compare (<http://www.lex tutor.ca/vp/comp/>). Lexical diversity of the corpora was measured by using a standardised type-token ratio (STTR) in WordSmith, with a basis of 100. Using an STTR compensates for the problem that differences in text length cannot influence the outcome (cf. Schmitt 2010). Teacher surveys were used to investigate how many readers were available in each class and how they were being used in practice (see Appendix 1).

Results

General characteristics of the corpora

The written corpus contained far more words (tokens) than the spoken corpus (see Table 2). This is because of the difference in sample sizes (written words vs. teacher talk) and is in line with corpus linguistics studies where spoken corpora are generally much smaller than written corpora (Nation 2006). In spite of the difference in size, it was possible to measure the lexical diversity for each corpus by using the STTR. Table 2 shows that the STTR is higher for the written corpus than for the spoken corpus, indicating a greater lexical variety for the written corpus than the spoken corpus.

Frequency levels

The outcomes of the word frequency analyses show different distributions of word frequency levels for each corpus (Table 3). A total of 72.9% of the written corpus comprised words from the K-1 to K-3 levels, which are the most important vocabulary levels for our target group (cf. Nation 2006; Nation & Waring 1997). The spoken corpus consisted for 88.5% of words from the K-1 to K-3 levels.

From the 3000 word frequency level and up, words occurred more often in the written than in the spoken corpus: 11.5% of the spoken corpus consisted of words from the 4000 frequency and higher, whereas for the written corpus, this was 27.1%.

The spoken corpus shows a vast drop of word families after the K-2 level (Table 3). Further analysis of these higher word levels in the spoken corpus³ revealed that many words originated from read alouds by the teachers (see Figure 1).

3. Analysis performed from level K-3 and up. The number of words at the K-1 and K-2 levels at both teacher talk and textbook talk made it impossible to make a distinction between the two. The K-6-K-20 were combined because of the low number of words in each category.

TABLE 2: Characteristics of the written and spoken corpora.

Generated output	Written corpus	Spoken corpus
Tokens	57143.00	12242.00
Types	2926.00	1174.00
Families	1813.00	825.00
STTR	65.4	56.0

STTR, standardised type-token ratio.

TABLE 3: Number of words per word frequency level in written and spoken corpora.

Word frequency level [†]	Written corpus: Word families		Spoken corpus: Word families	
	<i>N</i>	%	<i>N</i>	%
K-1	843	46.5	570	69.2
K-2	400	22.1	122	14.8
K-3	77	4.3	37	4.5
K1-3	1320	72.9	729	88.5
K-4	136	7.5	33	4.0
K-5	121	6.7	19	2.3
K-6-10	180	9.9	32	3.9
K-11-15	38	2.1	6	0.7
K-16-20	13	0.7	5	0.6
K-21-25 [‡]	5	0.3	0	0.0
Total	1813	100.0	825	100.0

[†], BNC-COCA K1-25 Word Frequency levels; [‡], The highest frequency level reached in the written corpus was K-22, and K-19 in the spoken corpus.

At the K-5 level, almost 50% of the spoken language was not simply 'teacher talk' but vocabulary from textbooks or stories that were read aloud by the teachers in the classroom (for practical purposes this is referred to as 'textbook talk'). These read aloud excerpts came from story books, poems, short stories or textbook assignments, as the following examples illustrate (underlined words are at the 3000 word level and higher):

'Okay, everybody read. She then puts on his leg, it is *bruised* but it will soon *heal*. Maybe it was not such a good idea to race along the *gravel* road.' (Teacher C, female, Grade 3 teacher)

'Will you walk into my *parlour* said the *spider* to the fly? It's the prettiest little *parlour* that ever you did? Spy. The way into my *parlour* is up the winding? *Stair*.' (Teacher D, female, Grade 3 teacher)

High-frequency word lists

Besides knowing the word frequency levels within the corpora, it is important to know how often learners encounter these words. The 30-word frequency cut-off point applied to the corpora yielded a total of 281 words from 228 word families in the written corpus, and 84 words from 69 word families in the spoken corpus. This indicates that if learners were to read through all the readers, they would encounter 281 words at least 30 times, which means there is a fair chance that these words would be added to their vocabulary. If they were to listen to all the words from the spoken corpus, they would hear 84 words at least 30 times, which means there is a fair chance that these words would be added to their vocabulary. The follow-up question is whether these are relevant words for South African Grade 3 learners to know. A comparison with three relevant word lists shows that the written corpus HF words covered 73.7% of the Dolch list, 34.3% of the Sibanda list and 0.9% of the AWL. The spoken

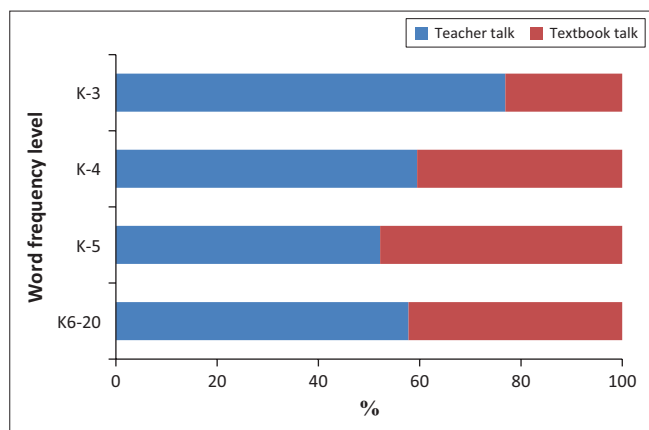


FIGURE 1: Distribution of teacher talk and textbook talk in percentage per word frequency level.

TABLE 4: Corpora coverage of Sibanda list, Dolch list and Academic Word List.

Type of corpus	Coverage (%) of Dolch list (179 families)	Coverage (%) of Sibanda list (487 families)	Coverage (%) of AWL (570 families)
Written corpus (1854 families)	99.4	80.1	5.8
Written corpus HF words (228 families) 30x exposure	73.7	34.3	0.9
Spoken corpus (843 families)	93.9	69.6	6.7
Spoken corpus HF words (69 families) 30x exposure	33.5	12.1	0.6

AWL, academic word list; HF, high-frequency.

corpus HF words covered 33.5% of the Dolch list, 12.1% of the Sibanda list and 0.6% of the AWL (Table 4).

Individual teacher talk

To find out whether there were any differences between the teachers' talk, a comparison at the individual level was made. Table 5 shows that the number of spoken words differed per teacher, and the STTR ranged from 54.9 (Teacher D) to 59.8 (Teacher C). The data also show that a higher number of word families used by a teacher did not automatically result in a higher lexical diversity of speech. Although the topics that were discussed differed per class, and the number of word families used per teacher varied, the majority of teacher talk was based on vocabulary from the K-1 to K-3 word levels, which varied from 91.2% (Teacher D) to 95.3% (Teacher C) (see Figure 2).

Use of readers in practice

In addition to knowing the language exposure that learners could get from reading graded readers, it is important to determine how often they actually read these texts. Results from the teacher survey showed that the use of the readers in practice varied per classroom, ranging from twice a week to daily use. The readers were reported to be used for different reading tasks, including group-guided reading, individual reading, paired reading (aloud) and reading with the teacher (one on one).

Only one teacher was of the opinion that the readers are a valuable source for vocabulary learning. Some teachers noted that the readers were based on American English and not on

TABLE 5: Overview of word frequencies in teacher talk.

Generated output	Teacher A	Teacher B	Teacher C	Teacher D
Tokens	1961	2925	2491	4865
Types	421	500	433	629
Word families	337	385	339	466
Standardised TTR	58.4	56.4	59.8	54.9
Words per minute	70	62	44	73

TTR, type-token ratio.

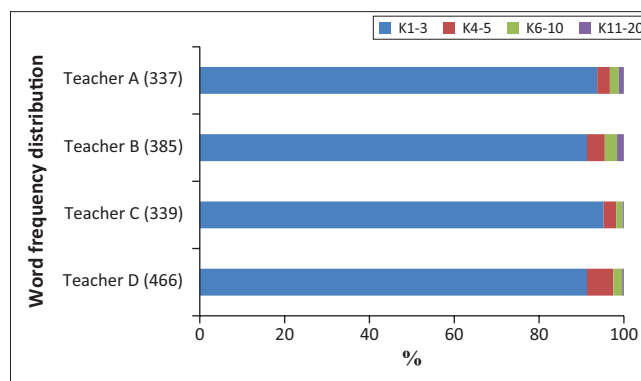


FIGURE 2: Word frequency distribution per teacher (total word families).

South African English. It was also argued that the Kathy and Mark series was more appropriate for reading practice (repetition of words) than for vocabulary building. Although it could be argued that repetition of words leads to increased exposure of the same word(s), and therefore an increase in vocabulary, the teachers did not make this connection. It was observed that although the Kathy and Mark readers were published between 1966 and 1971, none of the teachers mentioned that these were outdated. The availability of the readers varied greatly per classroom, ranging from 1 to 11 books per title, and between one and nine titles per stage. Not all teachers reported the number of books that were available per title (for details, see Table 6).

None of the teachers reported strictly adhering to the sequential order of the levels or stages for individual learners. Arguments for switching between levels included that good or exceptional readers should be allowed to skip a level, and sometimes teachers reported looking for specific vocabulary which necessitated the switch to another reader. Lack of readers was not reported as a reason to skip levels. All teachers reported that their learners were allowed to take the readers or copies of the readers home. Classes B, C and D did not possess the minimum number of five titles per level, which is the recommended minimum for the reading scheme to be effective (Nation & Wang 1999).

Discussion and conclusion

This article set off to investigate the potential of written (RQ1 + RQ2) and oral (RQ3) vocabulary exposure in poorly resourced South African township schools. The study was based on the knowledge that graded readers are an important resource for vocabulary building (cf. Waring & Nation 2004) and the assumption that teacher talk can be an important source of incidental vocabulary learning in low

TABLE 6: Availability of graded readers (titles and books) per classroom.

Graded reader series	Class A (39 learners)		Class B (40 learners)		Class C (37 learners)		Class D (37 learners)	
	Titles	Books	Titles	Books	Titles	Books	Titles	Books
Oxford Treetops Stage 10	9	99	6	-	3	23	3	22
Oxford Treetops Stage 11	9	-	5	-	4	26	4	28
Oxford Treetops Stage 12	9	-	4	-	5	16	1	14
Mark and Kathy Readers	0	-	3	-	5	30	5	40
Other	2†	-	1‡	-	0	-	0	-

†, Wide range readers; Grade 2 Oxford readers, not reported how many titles; ‡, Oxford reading tree, not reported how many titles.

socio-economic environments, especially when the LoTL is different from the language used at home.

While interpreting the data, a few limitations should be taken into account. This was a small-scale study and the yielded corpora were limited by the number of readers available and the volume of teacher talk that was recorded. Only four teachers participated in the study; however, their word frequency levels were similar and mainly included words from the K-1 to K-3 word level, indicating a similar level of vocabulary. The yielded HF word lists were not equal in their design; the written corpus produced a more elaborate HF word list than the spoken corpus. This was not only because of difference in input, but corpus linguistics studies, in general, show that spoken corpora are much smaller than written corpora (Nation 2006). A final limitation of this study is that, as in all educational research data collection, the collected data were snapshots of real-life classroom interactions transferred into a laboratory setting and should be interpreted as such. In spite of the limitations, the study provides an insight into what kind of language exposure takes place in poorly resourced classrooms, and should be considered a modest attempt to better understand vocabulary exposure in these particular settings.

The study localised research findings from more affluent Western contexts by showing that there are vast differences between levels of written and spoken vocabulary in poorly resourced classrooms. Although this seems evident, it is important to understand how these differences play out in practice. If we accept that schools in South Africa play a compensatory role for the limited literacy opportunities in homes of children from low socio-economic backgrounds (Howie & van Staden 2012; Snow et al. 1991), it is important to understand the different literacy components that contribute to that role. In spite of the schools being Quintile 3 and poorly resourced, the written texts from classroom readers that were available showed a language that was more lexical diverse and with a higher number of words from lower word frequency levels (K-4-K-22) than the teacher talk. The vast drop of word families after the K-2 level in the spoken corpus confirms findings that spoken language makes greater use of HF words than written language (Nation 2006). It also shows that the role of oral exposure in classrooms is restricted, which is in line with the study by Horst (2010) amongst adult ESL learners in Canada. The fact that it is difficult to compensate for the richness of written vocabulary through teacher talk was further demonstrated

by the finding that the richer spoken vocabulary largely came from teachers' read alouds of print materials. This supports the observation by Abadzi (2006) that access to print material is important for the academic achievement of poor students. Moreover, it confirms that reading aloud to children is important for their literacy development, as has been shown in numerous studies (cf. Wan 2000).

The compilation of HF word lists based on both corpora gives an indication of the type of vocabulary that Grade 3 learners could acquire incidentally through written or spoken language exposure. Naturally, learners will listen to many more hours of teacher talk throughout the academic year, and will also be exposed to other reading materials (e.g. textbooks) than the graded readers that were included in this study. Based on the collected data, we can carefully state that the learners are likely to frequently listen to at least 33.5% of the words from the Dolch list, 12.1% of the Sibanda list and 0.6% of the AWL. Moreover, it can be expected that they are likely to frequently read at least 73.72% of the Dolch list, 34.3% of the Sibanda list and least 0.9% of the AWL.

These figures show that children are likely to encounter more HF words from reading books than from listening to teacher talk.

Whether or not children will actually learn these words depends on contextual quality of the encounters, but based on the studies reported earlier (cf. Brown et al. 2008; Waring & Takaki 2003; Webb 2007), it can be argued that they have sufficient opportunities to learn these words and expand on their vocabulary. It is not surprising that the easier Dolch list reaches the highest coverage of the corpora, whereas the more difficult AWL has the lowest coverage. It is important to keep in mind that incidental vocabulary learning from reading should not be considered the primary source of learning new words (Schmitt 2010) and that it is essential that learners receive direct vocabulary instruction as well. Future research should investigate to what extent direct vocabulary instruction can be useful to build the vocabulary levels of these learners.

Implications for policy and further research

One of the major pitfalls discovered in the study was that the teachers did not adhere to Nation's (1999) argument that learners must work their way sequentially through the

graded reader levels and that they should read at least five books per level. This was because of the following situational constraints: a lack of titles per level and teachers' apparent lack of knowledge about the importance of following the sequence of the reading schemes. Skipping levels limits the amount of practice to decode HF words and learn sight words and clearly impacts learners' exposure to written vocabulary. The CAPS 2011 document does not prescribe that learners must pass through a graded reading scheme entirely before moving on to the next. It also does not mention the number of readers that students should read on a weekly basis. Given the CAPS instruction that graded readers are used for group sessions, there is little opportunity for learners to do individual reading and follow their own pace. It is recommended that future CAPS documents incorporate more accurate and research-based instructions about graded readers' routines to reach the maximum effect of their use.

Although none of the teachers mentioned that the readers were outdated, it is important to stress that especially the Kathy and Mark reading schemes were published more than 50 years ago and use stereotypes and bias that is no longer appropriate or politically correct in the 21st century. Reading egalitarian books to children over a sustained period of time shapes their attitudes and beliefs towards racial diversity and gender equality. Furthermore, children's motivation for reading, which is currently believed to be as important as their skills development (Guthrie 2013), is more likely to develop if they have access to interesting modern stories to which they can relate. It is therefore highly recommended that schools use more contemporary reading schemes that reflect life as we know it in the 21st century.

In order for educational resources to have an impact, schools need to understand under which conditions the resources can be translated into effective outcomes (Van der Berg 2008). The current study showed that learners can have access to important HF words through graded readers, but that additional materials, such as textbooks or storybooks, need to be used to expand their exposure to HF words. In this study, the following important conditions for incidental vocabulary learning to occur were identified: emphasis on written English exposure in Grade 3 classrooms, rather than on oral exposure, and compliance with the sequential order of graded reading schemes. We know that learners in South African township schools are capable of increasing their active word knowledge by about 9% per year (Pretorius & Stoffelsma 2017). Through the current study, we are now beginning to understand what the opportunities are for incidental vocabulary exposure from graded readers and teacher talk. Further research should investigate vocabulary exposure at a larger scale and include a broader range of written resources used in classrooms to determine the exact impact.

Acknowledgements

The author would like to extend sincere thanks to the schools, Grade 3 teachers and learners for their participation in the

study. The author would also like to express sincere thanks to the Zenex Foundation for accommodating this small study within the larger reading project, the Zenex Literacy Project (2015–2017), which was aimed at Foundation Phase teachers. This project was funded by the Van Coeverden Adriani Stichting (CAS), VU University Amsterdam, the Netherlands, and supported by the Zenex Foundation, South Africa.

Competing interests

The author declares that she has no financial or personal relationships that may have inappropriately influenced her in writing this article.

References

- Abadzi, H., 2006, *Efficient learning for the poor. Insights from the frontier of cognitive neuroscience directions in development (36619)*, World Bank, Washington, DC.
- Amin, N. & Ramrathan, P., 2009, 'Preparing students to teach in and for diverse contexts: A learning to teach approach', *Perspectives in Education* 27(1), 69–77.
- Brown, R., Waring, R. & Donkaewbua, S., 2008, 'Incidental vocabulary acquisition from reading, reading-while-listening, and listening to stories', *Reading in a Foreign Language* 2(0), 136–163.
- Coxhead, A., 2000, 'A new academic word list', *TESOL Quarterly* 34(2), 213–238. <https://doi.org/10.2307/3587951>
- Cunningham, A., 2005, 'Chapter 3 vocabulary growth through independent reading and reading aloud to children', in E.H. Hiebert & M.L. Kamil (eds.), *Teaching and learning vocabulary, bringing research to practice*, pp. 45–68, Lawrence Erlbaum, Mahwah, NJ.
- Cunningham, A. & Stanovich, K., 2001, 'What reading does for the mind', *Journal of Direct Instruction, Summer*, 1(2), 137–149.
- 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.
- Department of Basic Education, 2014, *Second detailed indicator report for basic education sector*, Department of Basic Education, Pretoria, Republic of South Africa.
- Dolch, E.W., 1948, *Problems in reading*, Garrard Press, Champaign, Ill.
- Elley, W.B., 1991, 'Acquiring literacy in a second language: The effect of book-based programs language learning', *Language Learning* 41(3), 375–411. <https://doi.org/10.1111/j.1467-1770.1991.tb00611.x>
- Fernald, A., Marchman, V.A. & Weisleder, A., 2013, 'SES differences in language processing skill and vocabulary are evident at 18 months', *Developmental Science* 16(2), 234–248. <https://doi.org/10.1111/desc.12019>
- Fleisch, B., 2008, *Primary education in crisis: Why South African schoolchildren underachieve in reading and mathematics*, Juta, Cape Town.
- Guthrie, J.T., 2013, 'Engagement and motivational processes in reading (key note)', paper presented at the 8th Annual Conference, Reading Association South Africa Johannesburg, n.d., n.p. South Africa.
- Hart, B. & Risley, T., 1995, *Meaningful differences in the everyday experience of young American children*, Paul Brookes, Baltimore, MD.
- Helman, L.A. & Burns, M.K., 2008, 'What does oral language have to do with it? Helping young English-language learners acquire a sight word vocabulary', *The Reading Teacher* 62(1), 14–19. <https://doi.org/10.1598/RT.62.1.2>
- Hill, D.R., 2008, 'Survey review: Graded readers in English', *ELT Journal* 62(2), 184–204. <https://doi.org/10.1093/elt/ccn006>
- Hirsch, E.D., 2003, 'Reading comprehension requires knowledge-of words and the world: Scientific insights into the fourth-grade slump and the nation's stagnant comprehension scores', *American Educator*, Spring, 10–22, 28–29, 44.
- Horst, M., 2010, 'How well does teacher talk support incidental vocabulary acquisition?', *Reading in a Foreign Language* 22(1), 161–180.
- Howie, S. & van Staden, S., 2012, *South African children's reading literacy achievement – PIRLS and prePIRLS 2011*, Centre for Evaluation and Assessment, Pretoria.
- Hu, M. & Nation, I.S.P., 2000, 'Vocabulary density and reading comprehension', *Reading in a Foreign Language* 13(1), 403–430.
- Lesnick, J., Goerge, R., Smithgall, C. & Gwynne, J., 2010, *Reading on grade level in third grade: How is it related to high school performance and college enrollment?*, Chapin Hall University of Chicago, Chicago, IL.
- Long, D.L., Johns, C.L. & Morris, P.E., 2006, 'Comprehension ability in mature readers (chapter 20)', in M. Traxler & M. Gernsbacher (eds.), *Handbook of psycholinguistics*, pp. 801–834, 2nd edn., Elsevier/Academic Press, Amsterdam.
- MacDonald, C. & Burroughs, E., 1991, *Eager to talk and learn and think (Consolidated report of The Threshold Project)*, Maskew Miller Longman, Cape Town.
- McCarthy, M. & Carter, R., (eds.), 1997, *Written and spoken vocabulary*, Cambridge University Press, Cambridge.

- Miles, K.P., Rubin, G.B. & Gonzalez-Frey, S., 2017, 'Rethinking sight words', *The Reading Teacher* 0(0), 1–12. <https://doi.org/10.1002/trtr.1658>
- Mol, S.E. & Bus, A.G., 2011, 'To read or not to read: A meta-analysis of print exposure from infancy to early adulthood', *Psychological Bulletin* 137(2), 267–296. <https://doi.org/10.1037/a0021890>
- Mullis, I.V.S., Martin, M.O., Foy, P. & Hooper, M., 2017, *PIRLS 2016 International Results in Reading: International Association for the Evaluation of Educational Achievement (IEA)*, Boston College, Chestnut Hill, MA.
- Mullis, I.V.S., Martin, M.O., Kennedy, A.M. & Foy, P., 2007, *PIRLS 2006 international report: IEA's progress in international reading literacy study in primary schools in 40 countries*, International Association for the Evaluation of Educational Achievement (IEA), Boston, MA.
- Nagy, W. & Townsend, D., 2012, 'Words as tools: Learning academic vocabulary as language acquisition', *Reading Research Quarterly* 47(1), 91–108. <https://doi.org/10.1002/RRQ.011>
- Nation, I.S.P., 2001, *Learning vocabulary in another language*, Cambridge University Press, New York.
- Nation, I.S.P., 2006, 'How large a vocabulary is needed for reading and listening?', *The Canadian Modern Language Review* 63(1), 59–81. <https://doi.org/10.1353/cml.2006.0049>
- Nation, I.S.P. & Wang, K., 1999, 'Graded readers and vocabulary', *Reading in a Foreign Language* 12(2), 355–380.
- Nation, I.S.P. & Waring, R. (eds.), 1997, *Vocabulary size, text coverage and word lists*, Cambridge University Press, Cambridge.
- 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 Early Childhood Development* 7(1), 1–13. <https://doi.org/10.4102/sajce.v7i1.553>
- Read, J., 2004, 'Research in teaching vocabulary', *Annual Review of Applied Linguistics* 24, 146–161. <https://doi.org/10.1017/S0267190504000078>
- Scarborough, H.S., (ed.), 2001, *Connecting early language and literacy to reading (dis)abilities: Evidence, theory and practice*, Guilford Press, New York.
- Schmitt, N., 2010, *Researching vocabulary*, Palgrave MacMillan, New York.
- Sibanda, J., 2014, 'Investigating the English vocabulary needs, exposure, and knowledge of isiXhosa speaking learners for transition from learning to read in the foundation phase to reading to learn in the intermediate phase: A case study', PhD, Rhodes University, Grahamstown.
- Snow, C.E., Barnes, W.S., Chandler, J., Goodman, I.F. & Hempill, L., 1991, *Unfulfilled expectations: Home and school influences on literacy*, Harvard University Press, Cambridge, MA.
- Spaull, N., 2011, *A preliminary analysis of SACMEQ III South Africa*, Stellenbosch Economic Working Papers: 11/11, University of Stellenbosch, Stellenbosch.
- Spaull, N., 2017, *Teaching reading for meaning: The Funda Wandé project*, viewed 08 February 2018, from <http://www.allangrayorbis.org/entrepreneurship-blog/cultivation/teaching-reading-meaning-funda-wande-project-dr-nic-spaull/>
- Stanovich, K., 2000, *Progress in understanding reading: Scientific foundations and new frontiers*, Guilford Press, New York.
- UNESCO, 2005, *EFA Global Monitoring Report 2006: Literacy for life*, UNESCO, Paris.
- Van der Berg, S., 2008, 'How effective are poor schools? Poverty and educational outcomes in South Africa', *Studies in Educational Evaluation* 34(3), 145–154. <https://doi.org/10.1016/j.stueduc.2008.07.005>
- Wan, G., 2000, 'Reading aloud to children: The past, the present, and the future', *Reading Improvement* 37, 148–160.
- Waring, R. & Nation, I.S.P., 2004, 'Second language reading and incidental vocabulary learning', *Angles on the English-Speaking World* 4, 11–23.
- Waring, R. & Takaki, M., 2003, 'At what rate do learners learn and retain new vocabulary from reading a graded reader?', *Reading in a Foreign Language* 15(2), 130–163.
- Webb, S., 2007, 'The effects of repetition on vocabulary knowledge', *Applied Linguistics* 28(1), 46–65. <https://doi.org/10.1093/applin/aml048>

Appendix starts on the next page →

Appendix 1

Teacher survey

1. How many different titles of graded readers are available in your classroom? Please indicate if there is more than one copy per title available.
2. How often do the learners in your classroom read these readers? For example, every day, once a week, once a month? Please be as specific as possible and also estimate the amount of time they spend on reading these books (in minutes/hours per week).
3. How are the readers used in your classroom? Do students read them individually (silent reading, read alouds)? Or do you use them for group-guided reading? Other?
4. Do you think the readers are a valuable resource for vocabulary learning? If yes, why? If no, why?
5. The Oxford Treetops and Mark and Cathy series use different levels. Do you keep records of the levels or stages that learners have read? If yes, how?
6. Do you follow the sequential order of the levels or stages for individual learners? For example, can a learner only continue with Stage 11 readers if he or she has finished all the Stage 10 readers in the Oxford Treetops scheme?
7. Are children allowed to take the books home?
8. Is your school considered a low-resourced school? Do you agree or disagree with your school status? Please explain.
9. Do you have any comments or suggestions about the use of graded readers or use of other reading resources in your classroom?