AOSIS

How is their word knowledge growing? Exploring Grade 3 vocabulary in South African township schools

CrossMark
click for updates

Authors:
Elizabeth J. Pretorius[1]
Lieke Stoffelsma[1,2]

Affiliations:
[1]Department of Linguistics and Modern Languages, University of South Africa, South Africa

[2]Centre for Language Studies, Radboud University Nijmegen, Netherlands

Corresponding author:
Lieke Stoffelsma,
l.stoffelsma@let.ru.nl

In this article, we report on a study that examined the active and receptive English vocabulary of two different groups of Grade 3 learners in South African township schools. The groups consisted of English Home Language (HL) learners in the Western Cape and Xhosa HL and English First Additional Language (FAL) learners in the Eastern Cape. The purpose was to document their different vocabulary trajectories during Grade 3. The Woodcock-Muñoz Language Survey was used to measure the active vocabulary levels of 118 learners at the beginning and the end of the school year. Another 284 learners from the same eight Grade 3 classes participated in a receptive vocabulary test at the end of the year. This test assessed their knowledge of the 60 most frequent words that occur in South Africa Grade 4 English textbooks. Results showed that although the HL learners knew almost double the number of words their English FAL peers did, both groups of learners increased their active word knowledge through the year by about 9%. Regarding their receptive vocabulary, the English FAL learners on average only knew 27% of the most frequent words at the end of their Grade 3. No significant gender differences were found. Learners in both language groups who were above their grade age had significantly lower scores than their younger peers. This confirms findings that children who start school with weak language skills tend to stay weak. Finally, initial active vocabulary knowledge was found to be a strong predictor of vocabulary development during the school year.

Introduction

In 2006, South Africa was one of the 45 countries that participated in the Progress in International Reading Literacy Study (PIRLS), which assesses reading comprehension. Grade 4 South African learners wrote the test in the language which had been the language of learning and teaching (LoLT) during the Foundation Phase. In PIRLS, the Low International Benchmark of 400 reflects basic, literal understanding of explicitly stated information in a text. Not being able to achieve this basic level of reading puts learners at risk of not being able to read. Only 13% of Grade 4 learners in South Africa reached this benchmark; in other words, 87% of Grade 4 learners in South Africa were at serious risk of not being able to read. Reading performance in the African languages was particularly low. When South Africa participated in the next round of PIRLS in 2011, the results were not much better. This time, Grade 4 learners participated in pre-PIRLS, an easier reading assessment with shorter texts (Howie & Van Staden 2012). The South African mean of 461 was considerably lower than the international mean of 500, and 29% of learners could not read at all. The results showed large inequalities in reading literacy according to language, again with reading comprehension poorest in the African languages, these learners being 2–3 years' worth of learning behind their peers (Spaull 2016). As in many other countries, the gender gap in South Africa was considerable, with Grade 4 girls outperforming their male counterparts significantly (Howie & Van Staden 2012).

Given that PIRLS assesses reading comprehension, poor performance in PIRLS obviously points to comprehension challenges that need to be addressed in South African classrooms. But what other lessons, besides the obvious comprehension ones, can be drawn from PIRLS? Reading research has consistently shown that poor reading comprehension is also strongly associated with two other factors: poor decoding skills and inadequate vocabulary. Decoding enables comprehension. If learners cannot yet decode written language fluently, their comprehension is severely compromised (Hoover & Gough 1990; Pretorius 2014). According to Hoover and Gough's (1990) 'simple view of reading', reading ability reflects an individual's decoding skill (the ability to recognise letter-sounds and their combination into words) as well as oral language ability (including receptive and

expressive vocabulary skills). In line with this model, other research has shown that vocabulary is strongly correlated with reading comprehension – children who know fewer words typically comprehend at lower levels (Chall, Jacobs & Baldwin 1990; Hart & Risley 1995; Scarborough 2001).

From this, it can be argued that South Africa's poor performance in PIRLS also points to vocabulary challenges that need to be addressed in the classrooms. Spaull (2016) points to the correlations between language, socio-economic status (SES), geographic location and school functionality within the South African schooling system and suggests that inequalities in learning will also be reflected in other dimensions. One of the 'other dimensions' that we wish to explore in this article is that of vocabulary, particularly the vocabulary knowledge of Grade 3 learners.

Research has shown that vocabulary correlates strongly with all aspects of language proficiency, such as oral comprehension (Alderson 2005), reading comprehension (Cain, Bryant & Oakhill 2004; Helman & Burns 2008; Stæhr 2008) and writing (Alderson 2005). It also correlates with academic performance (Dickinson & Tabors 2001; Scheepers 2016), general knowledge (Cunningham & Stanovich 2001; Dougherty Stahl & Stahl 2012) and even intelligence quotient (Bell et al. 2001; Marchman & Fernald 2008; Smith et al. 2003). The more refined and extensive a child's vocabulary is, the more it reflects the child's degree of knowledge about his or her world.

Vocabulary also shows strong associations with socio-economic status, and there are large disparities in word knowledge amongst children when they start formal schooling (Corson 1997; Hart & Risley 2003). Children from poorer homes tend to know fewer words – not because there is anything wrong with them but because they are not exposed to as rich oral and written verbal input as their peers from middle-income homes. Furthermore, research indicates that unless intervention occurs, it is difficult to change children's vocabulary trajectory once it is established in early childhood (Biemiller 2012; Hart & Risley 1995, 2003).

Despite the importance of vocabulary in language and reading achievement, there is surprisingly little formal vocabulary research being undertaken in Home Language (HL) or First Additional Language (FAL)[1] within the South African education context. The complex linguistic landscape in South Africa results in the majority of learners doing their first three years of schooling in an African HL, with English becoming the LoLT in Grade 4. This requires English FAL learners to have a fairly solid vocabulary by the time they enter the Intermediate Phase in order to understand teacher discourse on content subjects and to understand their textbooks in order to 'read to learn'. Furthermore, for some English HL learners, the English variety they speak may deviate from the formal standard English that occurs in

written texts. Added to this is the fact that the schools that these learners attend occur in low socioeconomic communities, which in turn puts them at risk of reading difficulties and poor school performance.

In this article, we report on a small-scale study that examined the active and receptive English vocabulary of two different groups of Grade 3 learners in schools that serve poor communities. The first Grade 3 group comprised English HL learners in township schools in the Cape Flats in the Western Cape, while the other group were Xhosa HL and English FAL learners in township schools in Port Elizabeth. While the learners differed in terms of whether they were HL or FAL users of English, what they had in common was that they attended schools in disadvantaged areas. Given the large differences in vocabulary knowledge that naturally exist between HL and FAL learners, the purpose was not to compare the two groups, but rather to document their different vocabulary trajectories during Grade 3 in non-affluent schooling conditions. Before we turn to the study itself, we first provide a brief review of vocabulary development in the early school years and the effects that poverty can have on vocabulary.

Vocabulary development in the early primary school years

Children learn their first words from social interaction with people in their lives. This interaction is grounded in the concrete, here-and-now of daily events and actions of the young child's life, and provides a potentially rich oral communicative context within which words are acquired.

Although many words are explicitly 'taught' to young children during social interaction, it is generally agreed that there is a large amount of vocabulary that, for all practical purposes, could not have been explicitly taught and must therefore have been incidentally acquired. In other words, it is generally held that children infer the meanings of a large number of unfamiliar words from oral contexts before they start school (Nagy & Herman 1987; Sternberg 1987). By the time 6-year-olds start school, they have a 'basic' vocabulary that consists mostly of high frequency words that occur in oral contexts. But how large is a 'basic' vocabulary? There are no easy answers to the thorny issue of estimating vocabulary size. With regard to English, Chall et al. (1990) put it at about 4000 words at age 6, while Sénéchal and Cornell (1993) posit 8000 words.

Estimates about vocabulary size and rate of development vary, depending on how the concept of 'word' is defined, how word counts in a particular language are undertaken, how word knowledge is defined and assessed, and whether one is referring to HL or FAL vocabulary development. Some scholars define a word as a graphically distinct type (e.g. White, Graves & Slater 1990). However, several scholars such as Laufer and Nation, and Anderson, Nagy and colleagues work with the notion of word families, where a word is defined 'as a base form with its inflected and some derived

---

1. Terms such as first language (L1) or mother tongue and second languguae (L2 or ESL) or additional language (AL) are also used in the literature. In this article, we adopt the terms 'Home Language' (HL) and 'First Additional Language' (FAL) as used in the South African curriculum documents.

word forms' (Laufer & Nation 1995:312). Thus, for example, *happy, unhappy, happily* and *happiness* belong to the same word family. Knowing the meaning of one of these words is assumed to increase the chances of being able to infer the meaning of the other members of the word family. There are usually three to four words per word family.

Furthermore, vocabulary development does not only involve the quantitative accumulation of new words to the lexicon; qualitative development is also an integral part of vocabulary acquisition. Many words in a language's lexicon are polysemous, conveying different meanings in different contexts (e.g. light comes from the sun; a light wind; light blue). Vocabulary development should therefore also take into account the degree of knowledge, which progresses from receptive knowledge (recognising words when heard or seen) to active or productive knowledge (using words in speech or writing). We always know more words than we use. There are a large number of words that never progress to free active use but remain part of a learner's receptive vocabulary repertoire.

As the discussion so far indicates, vocabulary development is quantitatively and qualitatively cumulative, and for HL English learners, fairly rapid too, with about 1000 word families (equivalent to about 3000 words) added each year. Increasing age typically reflects increasing word knowledge, unless children have language or learning difficulties, in which case their vocabulary is often below that of their age peers. Proposing a developmental sequence for HL vocabulary, Nation (1990) suggests that by age 5 at least 1528 word families are known (more than 4000 words), by age 10 a child knows about 7020 word families, by age 15 approximately 12 000 word families are known and by age 18 there are at least 17 600 known word families. Researchers (Nagy & Herman 1987; White et al. 1990) estimate that the average school child learns more than 3000 new words each year during the school years, of which at least half are estimated to be acquired from written contexts and the other half from oral contexts. However, there is considerable variation within this average, with some children learning 5000 new words a year and others only about 1000.

Variations in vocabulary development

Research indicates that variables such as home and school background as well as individual ability affect vocabulary development considerably. As already indicated, socioeconomic status has been found to have an early and persistent effect on children's vocabulary. In their groundbreaking research, Hart and Risley (1995, 2003) found that children from middle-class homes had vocabularies five times larger than children from very disadvantaged homes. In their comparative study of children's vocabulary from Grade 1 to Grade 5, White et al. (1990) found that by the end of Grade 4, middle-class children knew an estimated 15 000 words, while the same grade SES children knew an estimated 10 000 words (these estimates were based on graphically distinct words, not word families). In other words, middle-

class children's vocabulary skill grew at a rate of 5200 words per annum, while low-SE status children had a growth rate of around 3500 words.

Further evidence of vocabulary differences between socioeconomic groups comes from Corson's British research (1983, 1997). Corson found not only quantitative but also qualitative vocabulary differences between working-class and middle-class children. Corson suggested that speakers of some social dialects in English, especially working-class dialects, are hindered from accessing knowledge categories in their school curriculum because of differences in the types of words they acquire. Research evidence in the USA points to SE status differences in vocabulary growing larger with age. In other words, unless there is active intervention to narrow the vocabulary gap, differences between SES groups increase rather than level off (Hart & Risley 1995, 2003).

What about FAL vocabulary development, and what are the estimates with regard to the size of English FAL learners' vocabularies? The notion of incidental word learning in context is obviously less problematic when applied to the HL because, by the time children start school, they have already acquired basic competence in the language, which provides a scaffold for inferring word meaning from context. Furthermore, they are extensively exposed to the language in print or oral form on a daily basis. The underlying linguistic competence of FAL students is not commensurate with that of HL students, and they typically have less exposure to the language. Thus, not only do they have fewer bootstrapping mechanisms to infer word meaning but they also have fewer opportunities in which to apply word inferencing processes. Laufer (1992) estimates that for English FAL students a threshold of 3000 word families (about 5000 lexical items) is needed for 'minimal comprehension' in English. She argues that for learners below this threshold, no amount of general academic ability or reading skills in the HL will enable them to read in a satisfactory manner (Laufer 1992:100).

Turning to South Africa, in the Threshold report, Macdonald (1990) pointed to the immense vocabulary gap that exists between the words that Grade 4 FAL learners know at the end of their school year and the words they were then expected to know the following year in order to understand their textbooks in Grade 5. The learners in her study had a vocabulary of about 800 words, and they did not seem to have encountered more than half of the words used in their science textbooks. Twenty-four years later, Sibanda's study (2014) had similar findings, where Grade 4 learners tested on 60 high-frequency words that occur in South African textbooks performed poorly. These findings point to the critical need for special emphasis to be placed on vocabulary interventions in schools.

Vocabulary and reading

Although there is a strong association between vocabulary and reading, the nature of the relationship changes at different stages of schooling, with vocabulary becoming an

increasingly stronger determinant of reading ability as children move up the schooling ladder. The following trends have been observed:

- Some studies have shown vocabulary size to be a weak predictor of Grade 1 reading ability (e.g. Scarborough 2001). This may be because much of the effort in learning to read in the first year of schooling is focused on developing accurate decoding skills. However, locally, Wilsenach (2015) found receptive vocabulary to be a robust predictor of early literacy development in Northern Sotho-English emergent bilingual children. Although receptive vocabulary knowledge was found to be low in both Northern Sotho and English groups of learners, it had a significant effect on early literacy skills in both languages.
- Vocabulary is a strong predictor of reading achievement in Grade 4 (Scarborough 2001), and the main predictor of reading achievement in Grades 7 and 8 (Lescaux & Kieffer 2010).
- Vocabulary size at the end of Grade 1 is a significant predictor of reading comprehension 10 years later (Cunningham & Stanovich 2001).
- Children with restricted vocabularies at the end of Grade 3 show declining reading comprehension scores in later primary school (Chall et al. 1990; Lescaux & Kieffer 2010).

Vocabulary and the curriculum

Given the multilingual nature of South African schooling, what guidelines does the curriculum provide regarding vocabulary development? The Foundation Phase[2] Curriculum and Assessment Policy Statement document (referred to as CAPS) for English FAL states that by the end of Grade 3 learners should understand 1500–2500 English words in context (Department of Basic Education 2011:22). No comparable recommendations about vocabulary size are found in the English or African Home Language CAPS document. The somewhat inconsistent guidelines for vocabulary development across HL and FAL documents suggest that further research is called for in this domain.

Mindful of two factors – the expected differences in vocabulary knowledge in HL and FAL, and the barriers that low SE status can impose on learning and vocabulary knowledge across languages – the main aim of this study was to investigate the vocabulary levels of Grade 3 learners in schools that serve disadvantaged communities. This is a small-scale exploratory study of Grade 3 learners in four schools in the Western and Eastern Cape provinces. We captured the aim in the following research questions:

1. How does active word knowledge of Grade 3 learners in English HL and English FAL schools develop during the course of the Grade 3 year?
2. What is the receptive word knowledge of high-frequency words of Grade 3 learners at the end of the Foundation Phase?
3. Are there age and gender differences in active and receptive word knowledge of Grade 3 learners?

4. Is there a relationship between the learners' active and receptive word knowledge?
5. What is the role of initial active vocabulary knowledge on vocabulary development during the year?

Research methods and design

A pre-test post-test design using intact classroom groups was used in which Grade 3 learners' active vocabulary knowledge in English was measured at the beginning and at the end of the school year. The learners' receptive knowledge of high-frequency words was also measured at the end of the school year. Because no vocabulary assessment instruments and development norms have yet been established in the African languages, and because very little attention has been given to vocabulary research in the South African context, it was decided to make do with two existing instruments: the Woodcock-Muñoz picture vocabulary test and Sibanda's high-frequency word tests (Sibanda 2014). A detailed description of the instruments is given below. The former assesses active vocabulary knowledge, whereas the latter assesses receptive vocabulary knowledge.

Schooling context

Two different low-SE urban primary schools in the Eastern Cape and two in the Western Cape participated in the study. The two schools in the Western Cape were English HL primary schools, situated in the Cape Flats and serving mainly low-income mixed race communities. The schools had Afrikaans as their First Additional Language in Foundation Phase. These will be referred to as schools WC1 and WC2.

The two schools in the Eastern Cape had Xhosa Home Language as their LoLT and English as their FAL in Foundation Phase. These are Quintile 3 schools situated in township areas of Port Elizabeth. These will be referred to as schools EC1 and EC2. From each school, two classes participated in the vocabulary research, amounting to a total of eight classes, four from each province.

All the schools were part of a larger three year project, the Zenex Literacy Project (2015-2017), that worked with Foundation Phase teachers in Quintile 1–3 schools[3] across three provinces, the aim of which was to improve learner literacy levels through improved classroom literacy practices. The central focus of this intervention was on improving teacher reading instruction and increasing reading opportunities in the Foundation Phase, primarily in the HL but also in the FAL. The vocabulary study formed a small part of this larger intervention. The larger study is not discussed further in this article and serves merely to contextualise the vocabulary study.

Participants

In total, 118 Grade 3 learners from eight different classes from the four selected schools participated in an active

2.In South Africa, the Foundation Phase includes Grade R plus Grades 1–3.

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

TABLE 1: Home languages spoken by Grade 3 learners' parents (percentages) (*n* = 118).

Language	Western Cape		Eastern Cape	
	HL Mother	HL Father	HL Mother	HL Father
English	79.3	74.1	0.0	1.7
Xhosa	13.8	12.1	98.3	95.0
Afrikaans	5.2	10.3	0.0	3.3
Shona	1.7	1.7	0.0	0.0
Zulu	0.0	0.0	1.7	0.0
Unknown	0.0	1.7	0.0	0.0

HL, Home Language.

vocabulary test. Their active vocabulary was assessed twice: at the start of the year (January 2016) and at the end of the year (November 2016). Their average age at the start of the year was 8.4 years and towards the end of the year it was 9.1 years. The genders were fairly evenly distributed across the classes, except for school EC1 where there were more girls than boys in the class (for details, see data collection procedures). The dominant HL spoken by the learners' parents was Xhosa in the Eastern Cape and English in the Western Cape (see Table 1). Another 284 learners from the same eight Grade 3 classes participated in a receptive vocabulary test at the end of the year.

Test instruments

Woodcock-Muñoz picture vocabulary test

No standardised or normed vocabulary tests have yet been developed in South Africa. The Woodcock-Muñoz Language Survey was used to measure the active vocabulary levels of 15 Grade 3 learners from each of the eight classes in our sample. It was developed in the USA to assess the language development of bilingual children (mainly Hispanic). Because it was developed specifically for bilingual children, it was decided to use this test in the South African context. The Woodcock-Muñoz Language Survey (English) is a standardised set of four subtests that measure proficiency in oral language, reading and writing in speakers of English as a second language, with a minimum starting age of at least 2 years. The vocabulary subtest tests active vocabulary where testees must name 59 picture objects. The picture items start with familiar everyday items (e.g. *ball* and *house*) and progress to more unfamiliar items (e.g. *trombone* and, *chimney*). The pictures depict objects (i.e. concrete nouns) and reflect the extent to which children have lexicalised an object in the language and can name it. The picture vocabulary test does not tap into knowledge of abstract nouns or actions or attributes (as lexicalised in verbs, adjectives or adverbs).

The test is administered one-on-one, and provides scores for individual vocabulary knowledge which can be used to classify a learner's English vocabulary level according to age and grade level. To ensure the test items were appropriate for the sample, an analysis of the items was performed to assess suitability for primary school children in the South African context. As a result, 15 test items that were considered to be potentially unfamiliar to South African primary school students were replaced by items from a similar instrument

TABLE 2: Vocabulary tests used.

Instruments	Dimension of word knowledge tested	Type of words tested
1. Woodcock-Muñoz test	Active vocabulary knowledge	59 words (concrete nouns)
2. Definition matching test	Receptive vocabulary Knowledge	13 words from Sibanda's HFW list
3. Cloze (gap-filling) test	Receptive vocabulary Knowledge	23 words from Sibanda's HFW list

HFW, high frequency word.

(the Renfrew Vocabulary Test[4]). Because the Woodcock-Muñoz vocabulary test was modified, the vocabulary age norms are not applicable. The items in the Woodcock-Muñoz vocabulary test that were replaced and the alternative items that were used are presented in Appendix 1. The full list of items is presented in Appendix 2. The test had a strong Cronbach's alpha reliability coefficient (0.95 in January and 0.94 in November), which means that this is a robustly reliable test.

Sibanda's (2014) high-frequency word tests

Two high-frequency word tests were also used based on the work of Sibanda (2014), who developed a list of 60 high-frequency words generated from a corpus analysis of Grade 4 subject textbooks from the South African curriculum (see Appendix 3). These are words that Grade 3 learners should master before their transition to Grade 4. Sibanda developed a range of nine subtests that assess the 60 high-frequency words in various ways. Because of time constraints, we used only two of these instruments, administered to the whole class of Grade 3 learners at each of the selected schools:

- A definition matching test (Sibanda's Test 4): here a list of definitions is given on the left side of the page, and the testee is asked to match them up with a list of relevant words given in an oval on the right of the page (e.g. *someone who teaches – teacher).* This test assesses receptive vocabulary knowledge as prompts are provided in the form of matching definitions to relevant words.

- A modified cloze (gap-filling) test (Sibanda's Test 6): a list of sentences is provided on the left side of the page with one word missing from the sentence, and the testee fills in the gap with words chosen from an oval containing words on the right of the page. This test assesses receptive knowledge, as prompts to the gaps are provided. Several of the words in this subtest were function words (e.g. prepositions, auxiliaries, etc.)

A total of 36 items were included in the tests. The Cronbach's alpha reliability coefficient for the combined tests was high (0.95), which means that this is a robustly reliable test. Details of the tests used in the current research are presented in Table 2.

Data collection procedures

The tests were administered by two external researchers during school hours. Because of time and budgetary

4.The Renfrew vocabulary test is also a picture vocabulary test developed in the UK. It has been standardised on socio-economically diverse groups of children in Britain, Australia and South Africa.

constraints, for the one-on-one active vocabulary assessment, a purposive stratified sample of 15 learners from each of the eight Grade 3 classes was used. The classroom teachers were asked to select participants based on their class performance: five top learners, five average learners and five weak learners, respectively. Demographic characteristics of the test participants are presented in Table 3.

Sibanda's receptive word knowledge test of high-frequency words was administered in November to all 284 learners from the same eight Grade 3 classes. The pen-and-paper test was administered to the whole class, in the classroom, by two external researchers during regular class hours. Demographic characteristics of the test participants are presented in Table 4.

Data analysis

For both samples, the demographic items on gender, age and HL were averaged per school. The average score on the vocabulary test was calculated for all participants. The Kolmogorov–Smirnov test was used to determine normal distributions within the two separate groups (Western Cape [WC] and Eastern Cape [EC]). Scores were normally distributed within each sample: the EC, $D(60) = 0.09$, $p > 0.05$; and the WC, $D(58) = 0.09$, $p > 0.05$. For these normally distributed data, independent t-tests (independent samples) and one-way analysis of variance (ANOVA) were used to compare subgroups, that is, schools within each province, followed by post hoc tests in case of significance. Dependent t-tests were used to compare pre- and post-test results within groups, and an effect size analysis on the pre-post results for each group was calculated, using Cohen's d (e.g. Becker 2000). Pearson's correlation coefficient was used to investigate the relationship between active and receptive word knowledge. Finally, a simple regression analysis was used to determine the role of initial active vocabulary knowledge on vocabulary development during the year.

Ethical considerations

All the schools were part of a larger three year project that worked with Foundation Phase teachers, the aim of which

TABLE 3: Active word knowledge test: Participants per school, gender distribution and average age (January 2016).

School	N	Female (%)	Male (%)	Average age
EC School 1	30	63.3	36.7	8.7
EC School 2	30	50.0	50.0	8.8
WC School 1	28	53.6	46.4	8.2
WC School 2	30	43.3	56.7	8.0
Overall	118	52.5	47.5	8.4

TABLE 4: Receptive word knowledge test of high-frequency words: Participants per school, gender distribution and average age (in November 2016).

School	No. of students	Female (%)	Male (%)	Average age
EC School 1	71	60.6	39.4	9.3
EC School 2	70	50.0	50.0	9.4
WC School 1	72	51.4	48.6	9.0
WC School 2	71	46.5	53.5	8.8
Overall	284	52.1	47.9	9.1

was to improve learner literacy levels through improved classroom literacy practices. Permission to work in these schools was obtained at national and provincial levels. The focus of this larger intervention was on improving teacher reading instruction and increasing reading opportunities, primarily in the HL and also in the FAL.

All the schools were informed about the vocabulary study at the beginning of the year and two schools from each province were asked to volunteer to participate. Permission to use the Sibanda's high-frequency word tests was obtained in March 2016.

Results

The first research question is related to how active vocabulary knowledge of Grade 3 learners in English HL and English FAL schools developed during the course of the Grade 3 year, using the Woodcock-Muñoz vocabulary test.

Woodcock-Muñoz vocabulary test (active vocabulary knowledge)

Results at province level are presented in Table 5. The scores have been converted to percentages and the standard deviation is given in brackets. On average, participants in the Western Cape scored significantly higher on the post-test (mean [M] = 68.8, standard error [SE] = 1.7) than on the pre-test ($M = 60.3$, $SE = 1.8$, $t[57] = -12.01$, $p = 0.001$). Cohen's effect size value suggested a medium to large vocabulary learning effect ($d = 0.65$). Participants in the Eastern Cape also scored significantly higher on the post-test ($M = 41.2$, $SE = 1.1$) than on the pre-test ($M = 31.4$, $SE = 1.2$, $t[59] = -13.23$, $p = 0.001$), and the effect size shows a large vocabulary learning effect ($d = 1.09$). As to be expected, Grade 3 English HL learners knew almost double the number of words the English FAL learners did. What is of interest, however, is the fact that on the whole the learners from both English HL and English FAL schools increased by about 9% from January (pre-test) to November (post-test). At the end of the year, the active word knowledge gains of learners from English FAL schools in the Eastern Cape ($M = 9.7$, $SE = 0.7$) had increased slightly more

TABLE 5: Woodcock-Muñoz vocabulary test scores per province (%).

Province			Pre-test	Post-test	Gains
Western Cape English HL ($n = 58$)	Mean		60.3	68.8	8.5
	SD		13.4	12.6	5.4
	SE		1.8	1.7	0.7
	Minimum		36.2	46.6	10.4
	Maximum		86.2	94.8	8.6
	Percentiles	25	51.7	58.6	4.7
		50	58.6	67.2	9.5
		75	69.8	77.6	12.1
Eastern Cape English FAL ($n = 60$)	Mean		31.4	41.2	9.7
	SD		9.2	8.6	5.7
	SE		1.2	1.1	0.7
	Minimum		10.3	20.7	10.4
	Maximum		60.3	70.7	10.4
	Percentiles	25	24.1	36.2	5.2
		50	31.0	41.4	8.6
		75	37.9	46.6	13.8

HL, home language; FAL, first additional language; SD, standard deviation.

than the active word knowledge score of learners from the Western Cape schools ($M = 8.5$, $SE = 0.7$), but this difference is not statistically significant (t [22] = -1.21, $p > 0.05$).

By the end of the year, this gap remained more or less the same, with the mean score of English FAL learners in the Eastern Cape at 41.2%, whereas the mean score of the English HL learners in the Western Cape was 68.8%. What is also interesting to note is that across all the schools, the vocabulary growth of children at the 25th percentile (i.e. those with lower scores) was lower, at around 5%, than the better performing learners at the 75th percentile, who showed a vocabulary growth of 12% – 13%.

The second research question focuses on the receptive vocabulary knowledge of high-frequency words of Grade 3 learners at the end of the Foundation Phase, using two of Sibanda's receptive knowledge tests.

Sibanda's high-frequency word tests (receptive word knowledge)

Results of the high-frequency word test are presented in Table 6, which includes the separate test results (word meaning and cloze test), as well as the combined results (average receptive vocabulary knowledge of high-frequency words). Here too, as expected, the data show that the English FAL schools in the Eastern Cape obtained lower scores on both receptive tests than the English HL schools in the Western Cape. The average scores of the learners in the Western Cape ($M = 60.6$, $SE = 2.5$) are twice as high as the scores of learners in the Eastern Cape ($M = 27.4$, $SE = 1.8$). This difference is statistically significant (t [282] = 10.94, $p = 0.001$). Not unexpectedly, the English HL learners knew twice as many words as their English FAL peers. Furthermore, the data show that all learners found the word meaning test challenging and obtained a lower score on the word meaning test than on the gap-filling cloze test.

The results illustrate two important things. Firstly, there are large differences in terms of distribution of scores between the HL and FAL learners. Secondly, the score differences within the language groups are equally large, suggesting that the vocabulary levels of the learners are highly unequal.

Age and gender differences

We turn now to age and gender differences in active and receptive word knowledge of Grade 3 learners. Age distribution at the province level is presented in Figure 1. With an average age of 9.4 years, the Grade 3 learners in the Eastern Cape were older than the learners in the Western Cape (8.9 years).

Results of the receptive vocabulary knowledge average (word meaning and cloze tests) across the age groups are shown in Table 7.

There were significant differences between the age groups in the Western Cape (F [3,138] = 3.50, $p < 0.05$, $\omega = 0.22$). Post hoc tests indicated that both 8- and 9-year-olds had a significantly higher score on receptive knowledge of high-frequency words than 10-year-olds. No other significant differences were found. In the Eastern Cape, there were also significant differences between the age groups (F [4,136] = 2.54, $p < 0.05$, $\omega = 0.20$). Post hoc tests indicated that the 9-year-olds scored significantly higher on receptive vocabulary knowledge than their 10-year-old peers.

There is typically an increase in vocabulary knowledge of high-frequency words with age during the school years; exceptions do of course occur, and children with learning problems often present with lower vocabulary knowledge. The majority of Grade 3 children in this cohort were 9 years old. The older learners are likely to be learners who repeated a grade in Foundation Phase. The results seem to confirm this: the 10-, 11- and 12-year olds in particular seemed to
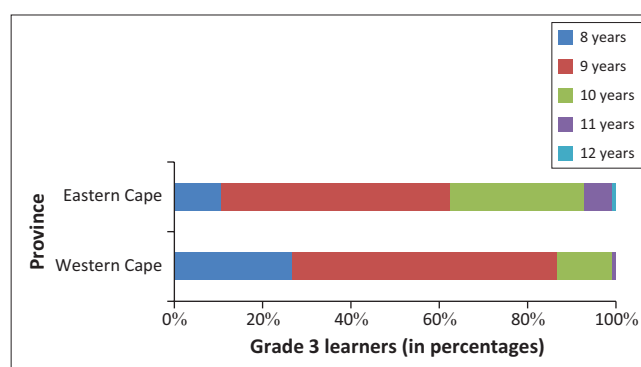
TABLE 6: Receptive knowledge of high-frequency words per province (%).

Province			Word meaning test	Cloze test	Receptive vocabulary knowledge average
Western Cape English HL ($n = 143$)	Mean		46.7	69.9	60.6
	SD		33.2	31.9	29.4
	SE		2.8	2.7	2.5
	Minimum		0.0	0.0	3.0
	Maximum		100.0	100.0	100.0
	Percentiles	25	16.7	50.0	36.7
		50	41.7	83.3	66.7
		75	75.0	100.0	86.7
Eastern Cape English FAL ($n = 141$)	Mean		18.9	33.0	27.4
	SD		18.8	26.2	21.1
	SE		1.6	2.2	1.8
	Minimum		0.0	0.0	0.0
	Maximum		100.0	100.0	100.0
	Percentiles	25	8.3	11.1	10.0
		50	16.7	22.2	20.0
		75	25.0	52.8	40.0

SD, standard deviation.

FIGURE 1: Age distribution of Grade 3 learners per province.

TABLE 7: Receptive knowledge of high-frequency words per province and age group (%).

Province		8 years	9 years	10 years	11 years	12 years
Eastern Cape	Mean	23.6	32.5	21.9	20.0	6.7
	SE	5.8	2.6	2.8	6.3	-
	SD	22.3	21.9	18.2	19.0	-
Western Cape	Mean	66.5	62.7	41.1	62.7	-
	SE	4.6	3.3	4.5	3.3	-
	SD	28.6	30.1	19.3	30.1	-

SD, standard deviation.

display much lower vocabulary knowledge, suggesting that they were weak learners.

Table 8 reflects gender differences in active and receptive vocabulary knowledge across English HL and English FAL Grade 3 learners.

Although female FAL learners in the Eastern Cape scored slightly higher on active vocabulary ($M = 41.8$, $SE = 7.1$) than their male peers ($M = 40.4$, $SE = 10.3$), this difference was not significant ($t[58] = 0.62$, $p > 0.05$). The difference in active vocabulary knowledge between HL boys ($M = 68.9$, $SE = 12.8$) and girls ($M = 68.7$, $SE = 12.5$) in the Western Cape was also not significant ($t[56] = -0.04$, $p > 0.05$).

Girls in the Eastern Cape scored slightly higher in receptive word knowledge ($M = 28.6$, SE $= 21.2$) than their male peers ($M = 25.8$, $SE = 21.1$), but this was not significant ($t[139] = 0.77$, $p > 0.05$). The difference in receptive vocabulary knowledge between boys ($M = 57.4$, $SE = 30.5$) and girls ($M = 64.0$, $SE = 28.0$) in the Western Cape was also not significant ($t[141] = 1.35$, $p > 0.05$).

Relationship between active and receptive word knowledge

The fourth research question investigated the relationship between the learners' active and receptive word knowledge. Pearson's correlation coefficient was calculated for active and receptive vocabulary knowledge, and the results are presented in Table 9. No significant correlations were found between both active vocabulary test results (January and November) and the receptive vocabulary results. The correlation between the active vocabulary tests of January and November was highly significant.

The fact that no correlation was found between learners' active and receptive vocabulary knowledge might be explained by the different levels of vocabulary that are measured by both tests. In the Sibanda tests, 97.1% of the words are from the 1000 word family list (i.e. the most common words in English). In contrast, the Woodcock-Muñoz test only has 39.5% words at the 1000 word family level and includes far more difficult vocabulary (up to the

TABLE 8: Active and receptive vocabulary knowledge (mean % and SE) as measured in November.

Province	Active vocabulary knowledge		Receptive vocabulary knowledge	
	Male	**Female**	**Male**	**Female**
Eastern Cape	40.4 (10.3)	41.8 (7.1)	25.8 (21.1)	28.6 (21.2)
Western Cape	68.9 (12.8)	68.7 (12.5)	57.4 (30.5)	64.0 (28.0)

TABLE 9: Pearson's correlation coefficients of active and receptive word knowledge.

Active word knowledge (Jan)	Active word knowledge (Nov)	Receptive word knowledge (Nov)
1.00	0.95†	ns
	1.00	ns
		1.00

ns, not significant

†Correlation is significant at the 0.01 level (two tailed).

14 000 word family list). Details of the word family levels per test are presented in Table 10.

Role of initial vocabulary knowledge

The final research question looked at the role of initial vocabulary knowledge on vocabulary development during the year. To what extent did the learners' active vocabulary knowledge at the beginning of the year predict their vocabulary growth during the year, as reflected in their final active vocabulary knowledge scores and their scores on the high-frequency words?

In both English HL and English FAL, learners who knew more words at the beginning of the year seemed to acquire more words by the end of the year. In order to investigate the role of initial active vocabulary in vocabulary development during the year, a simple linear regression was calculated to predict post-test active vocabulary knowledge based on pre-test active vocabulary knowledge. A significant regression equation was found in the Western Cape: F (1, 56) $= 288.647$, $p < 0.001$, with an R^2 of 0.838. Similarly, in the Eastern Cape, a significant regression equation was found: F (1, 58) $= 100.697$, $p < 0.001$, with an R^2 of 0.797.

Another simple linear regression was calculated to predict receptive vocabulary knowledge at the end of the year based on active vocabulary knowledge at the beginning of the year. No significant regression equation was found either in the Western Cape (F [1, 56] $= 1.157$, $p > 0.05$) or in the Eastern Cape (F [1, 58] $= 0.476$, $p > 0.05$).

Discussion

The paucity of local research into vocabulary development in the early primary school years by English HL and English FAL learners prompted this study, the aim of which was to explore the receptive and active vocabulary knowledge of Grade 3 learners who attend schools that serve low-SE communities. Although the four schools in this vocabulary study were involved in an intervention aimed at improving reading levels in the schools, the purpose of this study was simply to record the learners' vocabulary knowledge within this context, and not to evaluate the impact of the intervention on the learners' vocabulary development. In this study the different classroom contexts were not compared, nor were there measures to determine how much reading was happening in HL or FAL across the classrooms and how this

TABLE 10: Word family representations in the active and receptive vocabulary tests.

BNC/COCA word family list	Woodcock-Muñoz test (token %)	Sibanda word list (token %)
K-1	39.5	97.1
K-2	18.4	2.9
K-3–5	18.4	
K 6–10	11.9	
K 11–14	9.2	
Proper nouns	2.6	

BNC/COCA, British National Corpus/Corpus of Contemporary American English.

might affect vocabulary acquisition. The main findings are summarised and discussed below.

Firstly, as to be expected, the HL learners knew almost double the number of words their English FAL peers did, and this difference remained fairly constant throughout the year. Although both groups of learners increased their active word knowledge through the year, their growth was similar, at about 9%. An important finding was that the learning effect in practical terms was substantial for both groups of learners. A limitation of the current study is that it does not have a control group; therefore, whether this reflects a 'normal' growth pattern or whether the learners' vocabulary was benefitting from the larger reading intervention are issues that require further research. Home language learners have far greater exposure to their language at home and at school than FAL learners, whom for many, school is the only place where they get exposed to the FAL. According to CAPS, learners should have 7–8 HL instructional hours a week, and FAL learners 3–4 instructional hours. Despite these differences in exposure to language, it is surprising that the HL learners' rate of acquisition in terms of percentage increase in scores from pre- to posttest time was the same as that of the English FAL learners. A further area of interest would be to compare this growth pattern with that of learners in high performing Quintile 5 schools, and to see how SES impacts vocabulary growth patterns in our schools, and whether vocabulary growth can be accelerated in low income schools, for both HL and FAL learners.

Research on vocabulary development in the English-speaking world suggests that children acquire about 3000 words on average every year, with children from lower SE backgrounds learning fewer words than their middle-class peers because of differences in exposure to opportunities for vocabulary acquisition. Given that approximately 70% of schools in South Africa serve low-SE communities (Fleisch 2008), far more local research is needed to see how SE home and associated school factors affect vocabulary development in HL and FAL groups of learners across the economic spectrum, in all languages, and how SE barriers to learning can be overcome.

Vocabulary development is a complex phenomenon that can be assessed in terms of size and breadth as well as receptive or active knowledge. Receptive knowledge is usually in advance of active knowledge; we know more words than we actively use in speaking and writing. Assessing vocabulary is always a challenging task, and it is made more so when there is a lack of culture-appropriate tests that have been normed and standardised on a large group of children across SE and regional differences within the same country. The receptive vocabulary test used in this study assessed knowledge of the 60 most frequent words that occur in South African Grade 4 English textbooks (Sibanda 2014; Sibanda & Baxen 2016). However, task effects clearly emerged even though the same group of high-frequency words were being assessed: the word meaning tasks were more challenging than the gap-filing cloze task for both the HL and the English

FAL learners. Consistent with Sibanda's (2014) study, the learners found it more difficult to match up words with definitions than to find an appropriate word to complete a sentence. The former is a more abstract task and is thus cognitively more challenging.

Given that the receptive vocabulary test in this study assessed knowledge of the 60 most frequent words that occur in South African Grade 4 English textbooks, it is disconcerting to note that the English FAL learners knew on average only 27% of these most common words at the end of their Grade 3 year. This means that these learners were entering the Intermediate Phase with barely a third of the basic vocabulary knowledge needed to cope with the demands of the Intermediate Phase and the need to 'read to learn' from content subject textbooks. Even the strongest English FAL learners (at the 75th percentile) knew only 40% of these words, while the weakest learners (at the 25th percentile) knew merely 10% of the most common English words. However, the fact that both the HL and FAL learners were unfamiliar with these kinds of vocabulary tasks speaks to instructional practices in the classrooms. Writing and various forms of written assessment tend to be neglected in South African classrooms and children on the whole lack familiarity with more literate tasks and activities (Hoadley 2012; Pretorius 2014; Spaull 2016). This not only deprives children of opportunities to develop their language and literacy proficiency but also makes vocabulary assessment more difficult.

Surprisingly, no significant gender differences were found in this cohort of Grade 3 learners, although in early primary school girls typically perform better than boys on literacy measures (Howie & Van Staden 2012). This is an aspect that requires further research. However, interesting age differences emerged. While maturationally, older learners tend to know more words than younger learners, in this study, learners in both language groups who were above their grade age fared worst of all. Most of the Grade 3 learners in this study were 9 years old. No significant differences in vocabulary scores were found between the 8- and 9-year-olds in the study, but there were significant differences between children who were older than the average grade age; the 10-year-old learners knew fewer words than their 8- and 9-year-old Grade 3 peers. These older children had probably been retained a year at some point in the Foundation Phase. Clearly, a year's retention had not benefitted them in terms of vocabulary. Vocabulary research consistently shows that children who start school with weak language skills tend to stay weak and they also have greater difficulty learning to read and write (Chall et al. 1990; Scarborough 2001). The results of this study confirm this trend, with the 10-year-old learners knowing fewer words than their grade peers who were one or two years younger. The pedagogical implications are clear: simply retaining learners does not help them catch up. Learners who are retained need additional attention and enriched interventions.

Finally, there was considerable variation in the learners' vocabulary knowledge, both within and across English HL

and English FAL learners. This is in line with Sibanda's (2014) study in which learners' performance showed great divergence. There were some English HL learners who performed well on the active vocabulary test, but there were also some whose word knowledge was poor. The same applied to the English FAL learners: even though their scores were on average 30% lower than that of the HL learners, some FAL learners performed relatively well while others knew hardly any words. What was noticeable was that the learners' vocabulary trajectory remained relatively constant across the year: the weaker learners, those at the 25th percentile, in both language groups learned fewer words (about 5% gains) than their peers at the 75th percentile, whose word gains were double (about 12%). As the regression analysis showed, initial vocabulary knowledge was a strong predictor of vocabulary development through the year. Learners with higher initial vocabulary scores acquired more new words during the year. This finding is consistent with other vocabulary studies that show that it is difficult to change vocabulary trajectories established in early childhood (Collins 2005; Hart & Risley 1995, 2003).

Concluding thoughts and the way forward

This was an exploratory study that adapted a standardised active picture word test to be more in line with South African English vocabulary conventions. Although the reliability of the adapted test was acceptable, further refinements of such a test are still needed. It was also a small-scale study; therefore, generalisations to the larger population are necessarily curtailed. Vocabulary acquisition is largely a function of exposure to rich oral and print-based input in homes and schools. International vocabulary research has shown that children from lower SE backgrounds tend to come to school with smaller vocabularies than their more affluent peers. Vocabulary is also central to additional language learning, and adequate vocabularies are especially important for children for whom the FAL is also their LoLT. As prior local studies have suggested, young South African English FAL learners are struggling with vocabulary (Macdonald 1990; Sibanda 2014; Wilsenach 2015), but performance in the African HL is also a challenge (Pretorius & Mokhwesana 2009; Wilsenach 2015). Given the centrality of vocabulary in language proficiency in general and reading in particular, and given too its links to school success, vocabulary development, whether in the HL or FAL, should not be left to chance in classrooms. There is a clear need for further research in this domain, in African HLs as well as in English FAL:

- There is an urgent need to develop linguistically appropriate and culturally sensitive standardised vocabulary tests suitable for the South African multilingual educational context.
- Reliable tests that tap into both receptive and active vocabulary knowledge at different frequency levels are needed.

- It is important to establish a broad framework of normative vocabulary acquisition at different grade levels in South African schools, both in the HL and FAL, especially if the FAL also becomes the LoLT from Grade 4 onwards. This should derive from local empirical research and not be extrapolated from research contexts that are different from our own. How much vocabulary growth (or stasis) learners display when left to their own devices (i.e. when schools do not pay particular attention to improving vocabulary in classrooms) and how vocabulary growth can be accelerated are issues that need attention.
- Learners at risk of learning and reading difficulty should be identified as early as possible. Well-developed vocabulary tests that have been standardised and normed on South African children can help in this regard. To what extent can initial vocabulary gaps be narrowed as learners move up the education ladder?
- Rich verbal input is a prerequisite for good vocabulary growth. Interventions that increase exposure to incidental word learning through print-rich classrooms and extended teacher talk are needed. Research has also consistently documented the effects that regular storybook reading in and outside the classroom have on children's vocabulary development, both in the HL and FAL (Agustín-Llach & Alonso 2016; Collins 2005; Mol & Bus 2011).
- At the same time, explicit vocabulary instruction that occurs within the prescribed instructional hours should target high-frequency words, especially in English FAL, so that the 200 words that make up 90% of the running words of English texts in Grade 3 can be acquired as quickly as possible.

The importance of vocabulary for reading achievement and for school success cannot be overstated. As this small-scale study shows, there is still much that we do not know about vocabulary development amongst South African learners. The sooner we embark on a systematic and comprehensive research programme to address gaps in the understanding of this phenomenon, the sooner our learners can benefit from instruction that closes gaps in their vocabulary acquisition.

Acknowledgements

Competing interests

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

Authors' contributions

E.J.P. was the project leader of the vocabulary study and responsible for the testing of the learners, while L.S. was responsible for the testing of the learners, data analysis and reporting. Both authors contributed jointly to the writing of the article.

References

Agustín-Llach, M.P. & Alonso, A.C., 2016, 'Vocabulary growth in young CLIL and traditional EFL learners: Evidence from research and implications for education', *International Journal of Applied Linguistics* 26(2), 211–227. https://doi.org/10.1111/ijal.12090

Alderson, J.C., 2005, *Diagnosing foreign language proficiency*, Continuum, London.

Becker, L.A., 2000, *Effect Size (ES)*, viewed 13 September 2017, from http://www.bwgriffin.com/gsu/courses/edur9131/content/EffectSizeBecker.pdf.

Bell, N.L., Lassiter, K.S., Matthews, T.D. & Hitchinson, M.B., 2001, 'Comparison of the Peabody Picture Vocabulary test-third edition and Wechsler adult intelligence scale-third edition with university students', *Journal of Clinical Psychology* 57(3), 417–422. https://doi.org/10.1002/jclp.1024

Biemiller, A., 2012, 'Teaching vocabulary in the primary grades: Vocabulary instruction needed', in E.K. Kame'enui & J.F. Baumann (eds.), *Vocabulary instruction: Research to practice*, pp. 34–50, The Guilford Press, New York.

Cain, K.E., Bryant, P.E. & Oakhill, J., 2004, 'Children's reading comprehension ability: Concurrent prediction by working memory, verbal ability, and component skills', *Journal of Educational Psychology* 96(1), 31–42. https://doi.org/10.1037/0022-0663.96.1.31

Chall, J.S., Jacobs, V.A. & Baldwin, L.E., 1990, *The reading crisis: Why poor children fall behind*, Harvard University Press, Cambridge, MA.

Collins, M.F., 2005, 'ESL preschoolers' English vocabulary acquisition from storybook reading', *Reading Research Quarterly* 40(4), 406–408. https://doi.org/10.1598/RRQ.40.4.1

Corson, D., 1983, 'Social dialect, the semantic barrier and access to curricular knowledge', *Language in Society* 12, 213–223. https://doi.org/10.1017/S0047404500009817

Corson, D., 1997, 'The learning and use of academic English words', *Language Learning* 47(4), 671–718. https://doi.org/10.1111/0023-8333.00025

Cunningham, A.E. & Stanovich, K.E., 2001, 'What reading does for the mind', *Journal of Direct Instruction* 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.

Dickinson, D.K. & Tabors, P.O., 2001, *Beginning literacy with language: Young children learning at home and school*, Brookes, Baltimore, MD.

Dougherty Stahl, K.A. & Stahl, S.A., 2012, 'Young word wizards! Fostering vocabulary development in preschool and primary education', in E.J. Kame'enui & J.F. Baumann (eds.), *Vocabulary instruction: Research to practice*, pp. 72–92, The Guilford Press, New York.

Fleisch, B., 2008, *Primary education in crisis: Why South African schoolchildren underachieve in reading and mathematics*, Juta, Cape Town.

Hart, B. & Risley, T., 1995, *Meaningful differences in the everyday experience of young American children*, Brookes, Baltimore, MD.

Hart, B. & Risley, T., 2003, 'The early catastrophe: The 30 million word gap by age 3', *American Educator* 22, 4–9.

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

Hoadley, U., 2012, 'What do we know about teaching and learning in South African primary schools?', *Education as Change* 16(2), 37–41. https://doi.org/10.1080/16823206.2012.745725

Hoover, W.A. & Gough, P.B., 1990, 'The simple view of reading', *Reading and Writing: An Interdisciplinary Journal* 2, 127–160. https://doi.org/10.1007/BF00401799

Howie, S. & Van Staden, S., 2012, *South African children's reading literacy achievement – PIRLS and prePIRLS 2011*, Centre for Evaluation and Assessment, Pretoria.

Laufer, B., 1992, 'How much lexis is necessary for reading comprehension?', in P.J.L. Arnoud & H. Béjoint (eds.), *Vocabulary and applied linguistics*, pp. 126–132, Macmillan, London.

Laufer, B., & Nation, P., 1995, 'Vocabulary size and use: Lexical richness in L2 written production', *Applied Linguistics* 16(3), 307–322.

Lescaux, N.K. & Kieffer, M.J., 2010, 'Exploring sources of reading comprehension difficulties among language minority learners and their classmates in early adolescence', *American Educational Research Journal* 47(3), 596–632. https://doi.org/10.3102/0002831209355469

Macdonald, C.A., 1990, *Crossing the threshold into standard three*, HSRC, Pretoria.

Marchman, V.A. & Fernald, A., 2008, 'Speed of word recognition and vocabulary knowledge in infancy predict cognitive and language outcomes in later childhood', *Developmental Science* 11(3), F9–F16. https://doi.org/10.1111/j.1467-7687.2008.00671.x

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

Nation, I.S.P., 1990, *Teaching and learning vocabulary*, Newbury House, New York.

Nagy, W. & Herman, P., 1987, 'Breadth and depth of vocabulary knowledge: Implications for acquisition and instruction', in M.G. McKeown & M.E. Curtis (eds.), *The nature of vocabulary acquisition*, pp. 19–35, Erlbaum, Hillsdale, MI.

Pretorius, E.J., 2014, 'Supporting transition or playing catch-up in Grade 4? Implications for standards in education and training', *Perspectives in Education* 32(1), 51–76.

Pretorius, E.J. & Mokhwesana, M.M., 2009, 'Putting reading in Northern Sotho on track in the early years: Changing resources, expectations and practices in a high poverty school', *South African Journal of African Languages* 29(1), 54–73.

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

Scheepers, R.A., 2016, 'The importance of vocabulary at tertiary level', *Journal for Language Teaching* 50(1), 53–77. https://doi.org/10.4314/jlt.v50i1.3

Sénéchal, M. & Cornell, E.H., 1993, 'Vocabulary acquisition through shared reading experiences', *Reading Research Quarterly* 28, 360–374. https://doi.org/10.2307/747933

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*, Rhodes University, Grahamstown.

Sibanda, J. & Baxen, J., 2016, 'Determining ESL learners' vocabulary needs from a textbook corpus: Challenges and prospects', *Southern African Linguistics and Applied Language Studies* 34(1), 57–70. https://doi.org/10.2989/16073614.2016.1159521

Smith, B.L., Smith, T.D., Taylor, L. & Hobby, M., 2003, 'Relationship between intelligence and vocabulary', *Perceptual and Motor Skills* 100(1), 101–108. https://doi.org/10.2466/pms.100.1.101-108

Spaull, N., 2016, 'Disentangling the language effect in South African schools: Measuring the impact of language of assessment in grade 3 literacy and numeracy', *South African Journal of Childhood Education* 6(1), a475. https://doi.org/10.4102/sajce.v6i1.475

Stæhr, L.S., 2008, 'Vocabulary size and the skills of listening, reading and writing', *Language Learning Journal* 36(2), 139–152. https://doi.org/10.1080/09571730802389975

Sternberg, R.J., 1987, 'Most vocabulary is learned from context', in M.G. McKeown & M.E. Curtis (eds.), *The nature of vocabulary acquisition*, pp. 89–105, Psychology Press, New York.

White, T., Graves, M. & Slater, W., 1990, 'Growth of reading vocabulary in diverse elementary schools: Decoding and word meaning', *Journal of Educational Psychology* 82, 281–289. https://doi.org/10.1037/0022-0663.82.2.281

Wilsenach, C., 2015, 'Receptive vocabulary and early literacy skills in bilingual Northern Sotho-English bilinguals', *Reading and Writing* 6(1), Art. #77, 1–11.

Appendix 1

Overview of Woodcock-Muñoz test items that were replaced.

Original item	Alternative item	Original item	Alternative item
21 Piggy bank	Elbow	53 Mortarboard	Parachute
34 Zeppelin	Helicopter	54 Radar dome	Binoculars
39 Platypus	Penguin	55 Chevron	Saddle
41 Thermostat	Thermometer	56 Transom	Spanner
42 Churn	Igloo	57 Baluster	Chimney
48 Tuning fork	Violin	58 Amphora	Racket
50 T-square	Calendar	59 Mortise	Compass
51 Cornucopia	Oval		

Appendix 2

Woodcock-Muñoz test items.[†]

Item #	Eastern Cape mean score	Western Cape mean score	Item #	Eastern Cape mean score	Western Cape mean score
1 (Ball)	0.97	1.00	18 (Truck)	0.08	0.95
8 (Apple)	0.95	1.00	**21 (Elbow)**	0.07	0.82
4 (Balloons)	0.90	1.00	31 (Jellyfish)	0.07	0.50
12 (Ice cream)	0.90	1.00	25 (Eagle)	0.03	0.55
5 (Cat)	0.89	1.00	**48 (Violin)**	0.03	0.38
11 (Sailboat)	0.89	0.98	**51 (Oval)**	0.03	0.53
7 (House)	0.87	1.00	**56 (Spanner)**	0.03	0.22
16 (Flashlight)	0.85	0.82	28 (Astronaut)	0.02	0.63
14 (Toothbrush)	0.77	1.00	29 (Notebook)	0.02	0.20
2 (Mouth)	0.75	0.93	37 (Eyedropper)	0.02	0.02
24 (Stapler)	0.75	0.93	**42 (Igloo)**	0.02	0.35
9 (Sock)	0.70	1.00	44 (India)	0.02	0.00
10 (Glasses)	0.70	1.00	46 (Waterwheel)	0.02	0.12
13 (Cow)	0.70	1.00	52 (Trap)	0.02	0.00
50 (Calendar)	0.70	0.77	27 (Lawn mower)	0.00	0.27
6 (Chair)	0.67	0.98	32 (Weights)	0.00	0.15
20 (Hanger)	0.64	0.90	33 (Crutches)	0.00	0.45
3 (Window)	0.57	1.00	35 (Knots)	0.00	0.45
23 (Microphone)	0.49	0.90	36 (File cabinet)	0.00	0.12
47 (Biplane)	0.43	0.87	38 (Pocket watch)	0.00	0.08
34 (Helicopter)	0.41	0.92	40 (Trombone)	0.00	0.05
22 (Toaster)	0.36	0.92	**41 (Thermometer)**	0.00	0.05
17 (Rocking chair)	0.33	0.98	43 (Beaker)	0.00	0.12
39 (Penguin)	0.33	0.87	49 (Weathervane)	0.00	0.00
15 (Drum)	0.31	0.93	**54 (Binoculars)**	0.00	0.22
45 (Mallet)	0.31	0.83	**55 (Saddle)**	0.00	0.05
30 (High chair)	0.25	0.70	**57 (Chimney)**	0.00	0.45
26 (Bookcase)	0.20	0.27	**58 (Racket)**	0.00	0.20
19 (Spider web)	0.13	0.92	**59 (Compass)**	0.00	0.15
53 (Parachute)	0.10	0.28			

NB: The mean scores for some of the replaced items are rather high, especially for the higher item numbers (more difficult vocabulary) that were replaced. Examples include 50 (Calendar), 34 (Helicopter), 39 (Penguin), 53 (Parachute) and 21 (Elbow). If a more difficult vocabulary would have been selected to replace the items, the scores for the Western Cape group might have been much lower.

[†]Scores from the pre-test; Eastern Cape mean scores determine the chronological order of the items, from high scores to low scores. Replaced items are shown in bold.

APPENDIX 3

Sibanda's 60 high-frequency words.

Words 1–22	Words 23–55	Words 56–98	Words 99–151	Words 152–212
and	*have*	*if*	only	*every*
you	**need**/*need*	own	between	stop
use	*like*	but	*colour*	**both**
can	*from*	explain	because	**understand**
do	*other*	why	*example*	*enough*
what	*out*	some	*important*	until
are	each	time	**together**	**always**
for	**different**	*day*	**after**	again
it	*answer*	**same**	**piece**	before
or	*about*	**also**/*also*	*most*	easy
on	*not*	more	*must*	**once**
that	look	**school**	*another*	family

Source: Sibanda 2014
Words in bold were used in Test 2: definition matching test (12 words, the word 'teacher' is not in the list).
Words *in italic* were used in Test 3: cloze (gap-filling) test (23 words).
The words 'need' and 'also' were repeated in both tests.