

The Osmosis Plateau: The Limits of Incidental Vocabulary Acquisition in South African Education Students

Ilse Fouché and Maria Prozesky

The University of Witwatersrand, Johannesburg

Abstract

This study investigates the vocabulary proficiency of undergraduate Bachelor of Education (BEd) students at a South African university. Using a quantitative longitudinal design, students ($n = 643$) were assessed on high-frequency, mid-frequency, and low-frequency vocabulary levels. Results reveal students perform well below the thresholds required for accessing unsimplified texts. A sharp proficiency decline was observed first at the 3000 word level, and then again at the 5000 word level, with mastery at these levels remaining relatively stagnant across two years of university study. The findings suggest that incidental learning is insufficient for this demographic. Without explicit intervention aimed at building students' vocabulary knowledge, future teachers risk graduating without the vocabulary knowledge required to mediate the curriculum effectively, perpetuating the country's literacy crisis.

Keywords:

vocabulary levels; higher education; incidental vocabulary acquisition; English additional language; pre-service teachers

Introduction

English is one of 12 official languages in South Africa. It is primarily taught as an additional language from Grades 1 to 3 and becomes the language of learning and teaching from Grade 4 onwards for the vast majority of learners, despite it being the home language of only 8.7% of the population (Statistics South Africa, 2023). It is widely acknowledged that the country faces a reading crisis (Carter et al., 2024; Spaull & Pretorius, 2019). This is evident in the Progress in International Reading Literacy Study (PIRLS) assessments, where South African Grade 4 learners consistently achieve alarmingly low results (Department of Basic Education, 2023). There is a scarcity of research on reading comprehension in the language of learning and teaching at higher education level within multilingual contexts such as that of South Africa. Yet increasingly, emerging studies show the predictive validity of poor reading skills in primary school extending into higher education (Andrianatos, 2018; Cliff, 2014; Cronje, 2022; Weder, 2022).

While the ability to read with understanding for academic purposes is multifaceted, arguably one of the most important variables connected to reading ability is vocabulary. Vocabulary refers to the knowledge that speakers have about words, their meanings, and how they are used in language (Pretorius & Stoffelsma, 2021). Vocabulary levels have been researched extensively at all levels of education. However, there is a dearth of research on the vocabulary levels of multilingual additional language students in higher education contexts such as South Africa, where the majority of students embarking on their academic studies are English additional language speakers. Furthermore, little attention is paid to vocabulary acquisition as a stand-alone construct in South African schooling curricula across all subjects. Research demonstrates that a substantial amount of vocabulary is typically developed incidentally through exposure, particularly through extensive reading (Teng & Reynolds,



Except where otherwise noted, content in this journal is licensed under a [Creative Commons Attribution 4.0 International Licence](https://creativecommons.org/licenses/by/4.0/). As an open access journal, articles are free to use with proper attribution. ISSN: 2205-0795

2024). However, relying solely on this “osmosis” is problematic in contexts where the requisite conditions, such as high volumes of reading and rich linguistic input, are lacking (cf. Webb, 2019).

This article focuses on investigating undergraduate Bachelor of Education (BEd) students’ vocabulary levels, using three different data points: pre-test (beginning of first year), post-test 1 (end of first year), and post-test 2 (end of second year). An enhanced understanding of the vocabulary levels with which non-traditional and academically vulnerable students enter universities, and how these change over time through exposure to a higher education context, will better position institutions to support students by means of evidence-based, targeted interventions.

Drawing on Laufer and Nation’s (1995) productive vocabulary levels test, this study addresses the following research questions:

1. To what extent do university students’ vocabulary levels increase a) from the beginning to the end of their first year and b) from the end of their first year to the end of their second year at university?
2. At which vocabulary levels do these students achieve mastery at pre-test, post-test 1, and post-test 2 stages?

Literature Review

The Relationship Between Vocabulary Knowledge and Reading Comprehension

There are several ways in which the relationship between variables can be measured. These include determining correlations between these variables, measuring effect sizes, and doing regression analyses. The relationship between reading comprehension and vocabulary size has been found to be robust in several meta-analyses, with statistically significant moderate to strong correlations ranging from $r = .57$ to $r = .998$ (Jeon & Yamashita, 2014; Zhang & Zhang, 2022). Studies by Jeon and Yamashita (2014) and Zhang and Zhang (2022) indicate that correlations between these variables are stronger for undergraduate university students than they are for school learners. Importantly though, vocabulary knowledge is not the only variable that correlates strongly to reading comprehension. Jeon and Yamashita (2014), for example, found that grammar knowledge ($r = .85$) is particularly strongly correlated to reading comprehension. There is also emerging research indicating that if readers have not met certain reading fluency thresholds, the impact of factors such as vocabulary size on reading comprehension remains dormant (Stoffelsma et al., 2026; Wang et al., 2019). While the current study focuses on vocabulary knowledge, it is important to keep in mind the effects of other variables in students’ reading comprehension journeys.

While it is important to keep in mind that correlation is not causation, the current article argues that it is important to understand the vocabulary levels of students entering higher education as a predictor of reading performance and, ultimately, academic performance (cf. Scheepers, 2016). Moving beyond correlations, Dong et al. (2020) found an effect size of .55 ($p < .0001$) and a Q -value of 39.97 in the relationship between the constructs of vocabulary size and reading comprehension for undergraduate students. In a regression analysis by Jeon and Yamashita (2014), vocabulary knowledge accounted for a 62% shared variance in reading comprehension scores; thus, 62% of the differences in students’ reading comprehension scores could be explained by differences in their vocabulary knowledge. Within the African context, with undergraduate Tanzanian students, Biseko (2023) also found vocabulary to be a strong predictor for reading comprehension, with a variance of 56.5%. In a regression model, Biseko (2023) found that each additional raw score point on the vocabulary size test (equivalent to knowledge of 100 word families) corresponded with an increase of 1.13 percentage points on the reading comprehension test.

Understanding the vocabulary levels with which students enter universities would seem to provide effective predictions as to their reading levels and might assist in the development of empirically-based, targeted interventions that include a focus on vocabulary acquisition. Research on the vocabulary levels of teachers in training is particularly important as these future teachers are likely to impact the vocabulary levels of their future learners, yet research in this area remains sparse (Creighton, 2022; Pretorius & Murray, 2025). Given the strong predictive relationship between vocabulary knowledge and reading comprehension, it becomes important to understand how vocabulary is measured.

Assessing Vocabulary Knowledge

Vocabulary can be categorised based on the frequency with which words are used in spoken and written discourse across various domains. In English, this is measured through the British National Corpus (BNC) and the Corpus of Contemporary

American English (COCA); drawing on these corpora, Nation (2016) compiled ordered lists of 1000 word families (words sharing the same root, including all its inflections and derivations), ranked by frequency of occurrence. The Vocabulary Levels Test (VLT; Laufer & Nation, 1995; Nation, 2022) assesses students' vocabulary knowledge at the 2000, 3000, 5000, and 10000 word levels (henceforth referred to as L2000, L3000, L5000 and L10000). An additional category is the University Word List (UWL), which consists of 836 word families which contain non-technical vocabulary that is not contained in the first 2000 most frequently used words in English but which is still used frequently in academic writing (Nation, 2022).

Words found in the 1000 to 3000 word frequency levels are considered high-frequency words (Nation, 2006). Studies show that mastery of these high-frequency words is not sufficient to understand typical texts. Laufer and Ravenhorst-Kalovski (2010) suggest two lexical thresholds for reading with understanding: one optimal and one minimal. The optimal threshold predicts that students can read with functional independence and requires students to have knowledge of 8000 word families (thus, having achieved mastery at L8000 in the BNC/COCA lists). This translates to students understanding 98% of the words in a text. The minimal threshold requires students to have a knowledge of between 4000 and 5000 words (thus, having mastery at L5000), which translates to them understanding 95% of the words in typical texts. Students who have mastered L5000 are likely to read with understanding with support. This is in line with other studies like Nation (2006) and Hu and Nation (2000).

When considering vocabulary size (estimates of how many words are known in a target language; Van Hees & Nation, 2017), it is useful to distinguish between receptive (words understood when reading or listening) and productive vocabulary (words that students can produce while writing or speaking) (Cheng & Matthews, 2018; Pretorius & Murray, 2025). Receptive vocabulary size is always larger than that of productive vocabulary and precedes it; thus, students are likely to score better on receptive vocabulary tasks than on productive vocabulary tasks (Pretorius & Murray, 2025). Productive vocabulary has shown stronger predictive associations to academic performance than receptive vocabulary (Scheepers, 2016); in line with this, Nation (2022) specifically recommends the monitoring of students' productive vocabulary knowledge. For these reasons, the productive version of the VLT was used for the current study.

Research on Vocabulary in Low- and Middle-income Countries

Research on the vocabulary levels of university students in low- and middle-income countries (LMICs) remains scarce. LMICs are often characterised by societies that are multilingual and multicultural, with high poverty levels, high wealth disparities, and limited access to books and other print-based materials in most households (Pretorius & Murray, 2025). This section discusses some of the research that has been conducted within this context.

Some LMIC studies have measured the English vocabulary levels of high school and university students. A study by Dubravac et al. (2024) considered the vocabulary levels of first-year English as a foreign language (EFL) university students in Bosnia and Herzegovina. Although these students scored relatively well at L3000 (close to the 85–90% mastery level proposed by Laufer and Nation (1999)¹, very few students indicated mastery at L5000, indicating that they would likely face difficulties understanding authentic materials. Biseko's (2023) Tanzanian study showed that first-year EFL students, on average, had knowledge of approximately 4300 word families. However, 85% mastery level was only achieved at L2000, with a 68% mean score at L3000, and a 44% mean at L5000.

The little research that has been conducted in South Africa shows that vocabulary levels remain low despite 12 years of exposure to English in school. For example, a study on the productive vocabulary levels of South African foundation phase teachers across three provinces found that mastery or near-mastery scores were only achieved at L2000 (with means of between 81.8–95.3%; Pretorius & Murray, 2025). At L3000, means ranged from 54.6% to mastery at this level. No cohort of teachers in this study achieved mastery at L5000 or UWL level. Scheepers (2016) used the productive VLT to assess the vocabulary levels of South African literature and law first-year students at the University of South Africa. While both literature and law students achieved mastery at L2000 (with means of 91.4% and 86.2%, respectively), neither group achieved mastery at L3000 (with means of 79.5% and 69.8%, respectively). This study also shows medium to strong correlations between students' vocabulary levels and academic performance, particularly at L3000 and the UWL level. Unlike the studies cited above, Nizonkiza and Van Dyk (2015) used the receptive version of the VLT. Very few students were reported to be below the

¹ Other studies have suggested mastery levels of between 80% and 87% (Schmitt et al., 2001; Xing & Fulcher, 2007). The current study defines mastery at 85%, in line with Laufer and Nation (1999), while acknowledging that this score is somewhat arbitrary. Results are discussed with this in mind.

mastery level of the UWL and L3000, with approximately one third of students not reaching mastery at the L5000 word level. The differences between this study and the Scheepers (2016) and Pretorius and Murray (2025) studies point to marked differences in receptive and productive vocabulary outcomes, with the former relying on multiple choice questions and the latter likely being a more direct test of ready word knowledge.

Research on vocabulary levels in LMICs reveals a consistent pattern of students entering higher education with vocabulary knowledge below the thresholds required for independent reading. This pattern is closely tied to the limitations of incidental vocabulary acquisition in these contexts, as discussed below.

Incidental Vocabulary Learning

Incidental vocabulary learning refers to the acquisition of new words as a byproduct of engaging in communicative activities, including reading and listening, rather than through deliberate study (Nation, 2022). The concept is rooted in theories like Krashen's (1989) input hypothesis and suggests that learners naturally "pick up" vocabulary when exposed to comprehensible input over time. Incidental vocabulary acquisition, however, heavily relies on a rich linguistic environment. South African literature demonstrates that many English additional language learners come from under-resourced schools with a lack of access to rich print environments and have limited English exposure outside the classroom (Pretorius, 2014). The efficacy of incidental learning is therefore likely to be context dependent.

Methodology

Research Context and Design

This study employed a quantitative longitudinal research design using both paired and full samples at three time points to investigate the vocabulary levels of first-year BEd students at a South African university. South African BEd degrees are four-year degrees. The students reported on in this study were full-time students, enrolled in seven subjects that entail between 28 and 34 45-minute on-campus classes and tutorial sessions per week in the first year. This is slightly reduced in the second year, but students have full timetables throughout their BEd degrees. Most subjects prescribe both textbook chapters and academic articles. Students are therefore exposed to both informal English discussions in tutorials (which typically recycle high-frequency words) and more formal spoken and written academic English through lectures and course material, which should draw heavily on mid- and low-frequency words. There is thus a range of sources from which new words should be acquired incidentally.

In this longitudinal study, participants from two student cohorts (2023 and 2024) were invited to write the VLT as a pre-test (PRE) at the beginning of their first year, and as a post-test (POST1) at the end of their first year. The 2023 cohort was invited to write a second post-test (POST2) at the end of their second year to further extend the longitudinal lens. The research design allowed for cross-sectional comparisons between cohorts to determine whether the cohorts could be considered comparable, in order to ultimately use combined numbers for stronger inferential statistical pre- and post-test analyses. Further, within-subject analysis over time evaluated vocabulary development using paired samples, in which each student's score at one time point was matched to their own score at the subsequent time point (PRE to POST1, $n = 180$; POST1 to POST2, $n = 49$) analysed using paired samples t -tests for continuous scores and McNemar tests for categorised scores. To determine whether the findings for POST1 and POST2 could be generalised to the broader cohort, participants (i.e., students who returned for the POST tests) were compared to non-participants (thus, students who only wrote the PRE test) on age, gender, ethnicity, home language, and L5000 PRE score using chi-squared tests for categorical variables and the Wilcoxon Rank Sum test for continuous variables. No significant differences could be found. The results of the POST tests are therefore generalisable to the PRE group.

Ethical clearance was obtained, data were anonymised, and all participants provided informed consent. While data collection following paired samples of the original cohorts extended into 2025, these data have not yet been analysed and are not included in this article. The research project will also continue into students' fourth year (2026). The current article, however, focuses on students' vocabulary development in their first two years at university.

Population Sample

Overall, 643 students participated in the study between 2023 and 2024. Voluntary participation in both pre- and post-tests explains the large drop-off between pre- and post-test sample sizes (see Table 1). Full cohorts of students in the BED programme typically range between 700 and 800 students per year, indicating a relatively high initial participation rate in this study at PRE level (of between 35 and 45%). This is likely due to the novelty effect of university studies — first-year students tend to be keen to participate in all activities as they enter university but become less eager as the novelty wears off. In total, 180 students completed both PRE and POST1, while 49 students completed both of these in addition to POST2.

Table 1

Number of Participants for Each Test, Overall and Per Cohort

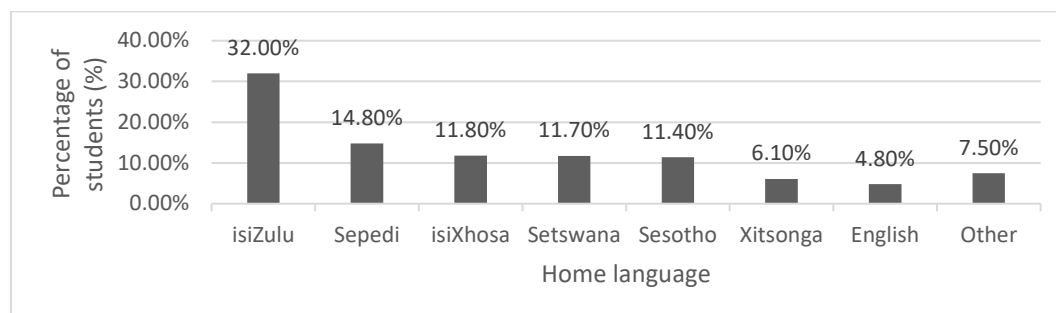
	Overall	2023 cohort	2024 cohort
PRE	643	283	360
POST1	180	94	86
POST2	49	49	N/A

To determine whether the 2023 and 2024 cohorts were sufficiently similar to treat as one sample, they were compared using age, gender, ethnicity, home language, school quintile, Grade 12 average, and students' scores at the L3000, L5000, and UWL levels of the VLT. Differences between the two cohorts across variables were either not statistically significant or showed small effect sizes. Consequently, the cohorts were considered together for subsequent analysis ($N = 643$).

Of this overall group, 63.9% were between 18 and 19 years old, followed by students in the 16–17 years age group (19.9%), students aged 20–21 (12%), and students older than 21 years (4.2%). The group overwhelmingly comprised female students (75.4%), indicating that teaching still seems to be considered primarily a female occupation in South Africa. The group consisted mainly of black students (95.3%; not including students of colour from other races). The group was also very multilingual, with 32% of students speaking isiZulu as their home language and only 4.8% identifying English as their home language (see Figure 1 for a breakdown of home languages).

Figure 1

Home Language of Participants Overall and Per Cohort



Instrument

Vocabulary knowledge was assessed using the productive VLT (Laufer & Nation, 1995). This widely used vocabulary test, based on the BNC/COCA corpora, provides an estimate of English language learners' general and academic English vocabulary breadth. As indicated earlier, the VLT (Laufer & Nation, 1995; Nation, 2022) assesses students' vocabulary knowledge at L2000, L3000, L5000, UWL, and L10000 levels. Vocabulary scores were calculated for each word level and categorised using the mastery threshold of 85%, based on Laufer and Nation's (1999) recommendation.

Data Analysis

Students' vocabulary levels were analysed by considering their total scores per frequency level. One mark was awarded for partially correct words, and two marks for words that contained no spelling or grammar errors. Extreme divergences in spelling that impeded meaning were marked incorrectly. Each word level consisted of 18 test questions.

Data were first analysed through descriptive statistics that report the results of all participants: PRE ($n = 643$), POST1 ($n = 180$), and POST2 ($n = 49$). In addition, a smaller paired sample subset of participants who wrote a) both PRE and POST1 ($n = 180$) and b) PRE, POST1, and POST2 ($n = 49$) were analysed inferentially using SAS software. The change in L3000, L5000, and UWL level scores from PRE to POST1 and from POST1 to POST2 was assessed by the paired samples t -test for continuous scores, reported as p -values. For continuous scores, Cohen's d was used as the effect size measure; for categorised (mastery) scores analysed using the McNemar test, Cohen's g was reported instead as it is better suited to paired binary outcomes. The following scale of interpretation was used for effect sizes: ≥ 0.80 = large effect; 0.50 – 0.79 = moderate effect; 0.20 – 0.49 = small effect; below 0.20 = near zero effect.

Results

The results are presented in relation to the two research questions: (1) To what extent do university students' vocabulary levels increase a) from the beginning to the end of their first year, and b) from the end of their first year to the end of their second year at university? and; (2) At which vocabulary levels do these students achieve mastery at pre-test, post-test 1, and post-test 2 stages?

Results are first presented in tabular form and thereafter discussed per word level. Data in Table 2 partially address Research Question 1. This table represents descriptive statistics using data for all students who participated at each data collection point. For each data collection point, the median percentage score is given, followed in brackets by the interquartile range (IQR). The IQR is expressed as the median percentage scores at the 25th percentile (i.e., the 25% of students who scored the lowest in the assessment) and the 75th percentile (i.e., the 25% of students who scored the highest in the assessment). For L2000, for example, the overall median at PRE stage is 86%, with the strongest quartile of students achieving a median of 92% and the weakest quartile achieving a median of 78%. Medians rather than means are used because the data contained significant outliers.

Table 2

Descriptive Statistics: Median (IQR) at PRE, POST1, and POST2 (Total Scores, All Participants)

Word level	PRE ($n = 643$) (%)	POST1 ($n = 180$) (%)	POST2 ($n = 49$) (%)
L2000	86 (78–92)	89 (81–94)	92 (89–97)
L3000	67 (56–78)	72 (63–81)	80 (67–83)
L5000	47 (36–58)	56 (44–67)	57 (50–69)
UWL	58 (50–67)	64 (54–72)	69 (58–75)
L10000	6 (0–14)	8 (6–17)	14 (6–28)
Overall average	53 (46–61)	57 (51–64)	63 (53–69)

Table 3 indicates the improvement from data point 1 (PRE) to data point 2 (POST1) and from data point 2 (POST1) to data point 3 (POST2), though this time using paired samples only drawing on data from students who wrote both sets of tests in each case. The change in mean between data points is also represented, with its significance and effect size indicated by p - and d -values, respectively.

Table 3

Improvement Between a) PRE and POST1 (n = 180) and b) POST1 and POST2 (n = 49)

Word level	Improvement between PRE, POST1 (n = 180)				Improvement between POST1, POST2 (n = 49)			
	PRE (%) Median (IQR)	POST1 (%) Median (IQR)	Change (%) mean	Significance (p) Effect size (d)	POST1 (%) Median (IQR)	POST2 (%) Median (IQR)	Change (%) mean	p-value (Cohen's d)
L2000	86 (75–92)	89 (81–94)	4.4	< .001* (d = .39)	89 (83–94)	92 (89–97)	4.2	< .001* (d = .55)
L3000	67 (57–78)	72 (63–81)	4.7	< .001* (d = .35)	75 (61–81)	80 (67–83)	4.6	.002* (d = .48)
L5000	47 (36–58)	56 (44–67)	6.9	< .001* (d = .48)	56 (50–69)	57 (50–69)	3.4	.049* (d = .30)
UWL	58 (49–67)	64 (54–72)	4.4	< .001* (d = .31)	64 (56–69)	69 (58–75)	4.5	.020* (d = .35)
L10000	6 (0–14)	8 (6–17)	3.2	< .001* (d = .25)	11 (6–22)	14 (6–28)	3.7	.070 (d = .27)
Overall average	53 (46–60)	57 (51–64)	4.8	< .001* (d = .47)	59 (52–67)	63 (53–69)	4.1	< .001* (d = .56)

Note. * Significant at 5% level.

Table 4 indicates mastery per word level using the paired samples, while also showing whether any change was statistically significant, signified by *p*-values. If a change is significant, this is followed by an effect size (*g*). Results of all participants closely resemble these figures.

Table 4

Percentage of Students Who Achieved the Mastery Benchmark of 85% (Paired Samples)

Word level	Improvement between PRE, POST1 (n = 180)			Improvement between POST1, POST2 (n = 49)		
	PRE (%) (n = 180)	POST1 (%) (n = 180)	Significance (%) (p)	POST1 (%) (n = 49)	POST2 (%) (n = 49)	Significance (%) (p)
L2000	50.6	67.2	< .001* (g = .28)	69.4	83.7	.002* (g = .30)
L3000	10.0	13.9	.180	12.1	22.4	.096
L5000	0.0	2.8	.025* (g = .50)	2.0	8.2	.083
UWL	1.7	3.3	.260	2.0	8.2	.083
L10000	0.0	0.6	.317	0.0	0.0	N/A

Note. * Significant at 5% level.

L2000

The results for the overall group (Tables 2 and 3) indicate mastery at L2000 PRE stage, with a median of 86% for both total participants and those in the paired sample. The spread between the 25th and 75th percentile scores for this metric of 17% within the paired sample (Table 3) shows some variability between lower- and higher-performing students. Only 50.6% of students reached the 85% mastery benchmark at L2000 at the beginning of their university careers (Table 4).

At POST1, the median for L2000 increases by 3 percentage points to 89%, a statistically significant improvement with a small effect size ($p < .001$, $d = .39$). However, there is a notable difference in students who reached the 85% benchmark for mastery

at POST1 — 67.2% of students reached this benchmark in the paired sample, showing an increase of 16.6 percentage points. At POST2, a further improvement in the overall median (to 92%) is evident, with the greatest gain seen in the weakest quartile of students. This is also reflected in the increased proportion of students who reached the mastery benchmark of 85% at POST2 stage in the paired sample. At this stage, 83.7% of students in the paired sample reached this benchmark — a statistically significant, yet relatively small, effect size ($p = .002$, $g = .30$; Table 4).

L3000

At the L3000 PRE stage (Tables 2 and 3), the cohort as a whole achieved a median of 67% for both total participants and those in the paired sample; not even the strongest quartile of students reached mastery at this level. The difference between the 25th and 75th percentile scores increases at the L3000 PRE stage, with a gap of 21 percentage points in the paired sample. Table 4 indicates that in the paired sample of 180 students, only 10% reached mastery at L3000.

At POST1 stage, the median for the overall group (both total participants and those in the paired sample) rises by 5 percentage points to 72% (Tables 2 and 3), with greater gains for the 25th percentile than for the 75th percentile. Among the 180 students who completed both PRE and POST1, scores rise by an average (mean) of 4.7 percentage points ($p < .001$, $d = .35$; Table 3). This small increase is not enough to significantly impact mastery levels; in the paired sample, there is only an increase of 3.9% of students who achieved mastery between PRE and POST1 ($p = .18$; Table 4).

At POST2 stage, a further overall increase in median to 80% can be seen (both total participants and those in the paired sample; Tables 2 and 3). This improvement between POST1 and POST2 for the paired sample shows a statistically significant moderate effect size ($p = .002$, $d = .48$; Table 3). For the 49 students who completed all three tests, this translates to an improvement of 10.3 percentage points in the percentage of students who achieved mastery at L3000, though unlike the improvement at L2000 for mastery, this improvement was not found to be statistically significant.

L5000

At L5000, the cohort overall (both total participants and those in the paired sample) achieved a median of 47%, and even the 75th percentile score fell far short of the mastery benchmark of 85% (Tables 2 and 3). The difference between the 25th and 75th percentile scores remains large at L5000, with a spread of 22 percentage points, similar to that at L3000. In the paired sample of 180 students, no students reached mastery for L5000 (Table 4).

At POST1 stage, the median increases by 9 percentage points to 56%, with slightly greater gains for the 25th percentile than for the 75th percentile (Tables 2 and 3). Among the 180 students who completed both PRE and POST1, statistically significant moderate improvements in mean scores of 6.9 percentage points ($p < .001$, $d = .48$) are observed (Table 3). Only 2.8% of students in the paired sample ($n = 180$) reached mastery at L5000 at POST1, from no students at PRE ($p = .025$, $g = .50$) (Table 4).

There is only a 1% increase in median from POST1 to POST2, with no movement in either the 25th or the 75th percentile in the paired sample. The change in average of 3.4%, however, shows a small, marginally significant effect size ($p = .049$, $d = .30$) (Table 3). While there is a jump in students who achieved mastery, from 2.0% at POST1 to 8.2% at POST2 for the paired sample of 49 students, this improvement is not statistically significant ($p = .083$; Table 4).

UWL

At the PRE stage, the group overall achieved a median of 58% — midway between their performance at L3000 and L5000 (Tables 2 and 3). Median scores do not indicate mastery overall, nor at the 75th percentile. The difference between the 25th and 75th percentile scores is slightly smaller than at L3000 or L5000, at 17 percentage points. Only 1.7% of students reached mastery at PRE (Table 4).

At POST1, an increase in overall median of 6 percentage points to 64% is seen (Tables 2 and 3), with similar increases evident at both the 25th and 75th percentiles (Table 3) for the 180 students who wrote both PRE and POST1. Within this group, statistically significant increases in mean (with a small effect size) of 4.4 percentage points are seen ($p < .001$, $d = .31$; Table 3). A small increase in mastery level of 1.6% is seen, though this is not statistically significant (Table 4).

An increase of 5 percentage points can be seen between POST1 and POST2 (Tables 2 and 3), with a greater improvement in the 75th percentile than in the 25th percentile in the paired sample of 49 students (Table 3). Within that same sample, a statistically significant increase of 4.5%, though with a small effect size, is seen ($p = .020$, $d = .35$; Table 3). Identical to L5000, there is a statistically non-significant improvement of 6.2% in students who reached mastery at this level.

L10000

This level is not discussed in detail, as students performed very poorly throughout. At PRE, a median of 6% was achieved. This increases by only 2% at POST1 and a further 6% at POST2 — the improvement in each case is statistically significant, though with very small effect sizes. In the paired sample of 180 students, there is a 0.6% increase in mastery from PRE (at 0% mastery) to POST1. However, among the 49 students who wrote all three tests, no students achieved mastery at this level.

Discussion

The lack of documented vocabulary interventions at higher education level indicates that universities likely assume exposure to academic reading will incidentally build vocabulary. While this might be true in print-rich contexts where students habitually read extensively and have a strong foundation in high-frequency vocabulary levels, this does not seem to be the case within the multilingual LMIC context of South Africa. Moreover, incidental learning is more likely in cases where students have already met the lexical threshold of between 95% and 98% (Hu & Nation, 2000; Laufer & Ravenhorst-Kalovski, 2010). This does not seem to be the case for the students in this study (certainly regarding reading academic texts), most of whom had not yet mastered the high-frequency L3000 (and in many cases, not even L2000). Within the context of massification and widening participation in universities worldwide (also for students from LMICs), these data would indicate that we cannot merely rely on historical assumptions about incidental English vocabulary acquisition for students in higher education.

Within the South African context where English is used as a lingua franca and from Grade 4, university students might appear competent because they have (largely) mastered the high-frequency L2000 words typical in spoken language. However, they hit a barrier at L3000, which will prevent them from accessing authentic disciplinary texts. This drop-off at L3000 and at the UWL level (which draws on both higher-frequency spoken words likely to be encountered in academic contexts, as well as words encountered in academic English) has implications for academic survival. The gains shown after both one and two years of studies, while statistically significant, remain small to moderate, and are not enough to help students achieve mastery at the L5000 level, which is necessary for mediated engagement with authentic academic texts (Nation, 2006).

The persistently low vocabulary levels documented in this study have implications beyond vocabulary development itself. As established in the literature, vocabulary knowledge is a strong predictor of reading comprehension (Biseko, 2023; Jeon & Yamashita, 2014), which in turn predicts academic performance (Scheepers, 2016). The stagnation observed in the current study, particularly at L3000 and L5000, suggests that students in this cohort face compounding risks. Insufficient vocabulary knowledge that remains below mastery thresholds will likely constrain reading comprehension, which in turn is likely to limit meaningful engagement with academic texts across the curriculum. Moreover, this vocabulary plateau is likely to be an intergenerational threat, considering that this cohort of students are pre-service teachers. If teachers have not mastered certain levels of vocabulary, they also cannot mediate vocabulary acquisition at these levels for their future learners. Ultimately, there must be a congruence between teachers' discourse and their learners' needs (Creighton, 2022; Pretorius & Murray, 2025). This shows the need for vocabulary interventions across subjects which specifically target the levels at which students are not yet displaying mastery.

Limitations

The results of this study pertain to BEd students at one university in South Africa who commenced their studies in 2023 and 2024. Results can therefore not be generalised to other contexts. A high attrition rate of students who completed post-tests due to voluntary participation could further have skewed the results, with the possibility of more conscientious students volunteering to complete post-tests. However, as indicated in this article, the smaller post-groups were compared to the initial pre-group to ensure that groups were still comparable.

Conclusion

Given the predictive value of vocabulary knowledge on reading ability, this group of BEd students' limited English vocabulary knowledge suggests that the reading crisis faced at foundation and intermediate phase level in South African schooling (Department of Basic Education, 2023) persists during at least the first two years of higher education. These multilingual learners complete the majority of their schooling through the medium of English. Despite this, incidental exposure to English vocabulary in these contexts does not lead to sufficient organic improvement in vocabulary levels to equip students with the necessary vocabulary knowledge to enable them to comprehend authentic texts. Even for low-frequency words (specifically L3000), only the strongest quartile of students approach mastery halfway through their four-year degrees, and gains in vocabulary knowledge as students progress through their university careers are incremental rather than dramatic. Furthermore, gains are not enough to help students gain mastery at L5000, which would be needed for them to access academic texts with support. Pre-service teaching programmes within similar contexts therefore need to move away from relying on passive exposure and instead consider targeted interventions to support students' vocabulary acquisition. This is important for all students, but even more so for students in teacher education. If future teachers cannot command the most fundamental level of English vocabulary with high accuracy, their ability to provide rich linguistic input to their own learners is severely compromised. Future research should measure the impact of such interventions and continue with longitudinal studies into the reading levels of diverse students in higher education institutions up to postgraduate level. In addition, student pass rates remain higher than the results of this study would suggest likely, indicating that students might find ways around reading texts to still engage with information — this also warrants further investigation. However, the limits of osmosis are clear: mere immersion in English educational environments within multilingual LMIC contexts such as that of South Africa is insufficient to build the vocabulary breadth required to access authentic, and particularly academic, texts.

Acknowledgements

This research was made possible through the Oppenheimer Memorial Trust Award under the title “Literacy education: Reading in Higher Education” (OMT Ref. 2025-7973) as well as the Wits School of Education Thematic Research Grant under the title “Multilingual Reading Interventions in Higher Education: Design, Pilot and Implementation”.

We further want to thank Dr Petra Gaylard for her invaluable assistance with the statistical analysis for this paper, as well as Prof. Elizabeth Pretorius for acting as a critical reader. Their contributions greatly strengthened this paper.

References

- Andrianatos, K. (2018). *First year university students' reading strategies and comprehension: Implications for academic reading support* [Doctoral thesis, North-West University]. NWU Institutional Repository. <https://repository.nwu.ac.za/items/6f34f0b8-b4cf-4f5f-8126-4f72651faa6d>
- Biseko, J. M. (2023). Vocabulary size and comprehension of academic texts by Tanzanian university students: An exploratory study. *Southern African Linguistics and Applied Language Studies*, 41(4), 496–513. <https://doi.org/10.2989/16073614.2022.2151480>
- Carter, J., Podpadec, T., Pillay, P., Babayiğit, S., & Gazu, K. A. (2024). A systematic review of the effectiveness of reading comprehension interventions in the South African multilingual context. *Educational Research and Evaluation*, 29(1-2), 69–103. <https://doi.org/10.1080/13803611.2024.2314522>
- Cheng, J., & Matthews, J. (2018). The relationship between three measures of L2 vocabulary knowledge and L2 listening and reading. *Language Testing*, 35(1), 3–25. <https://doi.org/10.1177/0265532216676851>
- Cliff, A. (2014). Entry-level students' reading abilities and what these abilities might mean for academic readiness. *Language Matters*, 45(3), 313–324. <https://doi.org/10.1080/10228195.2014.958519>
- Creighton, G. (2022). Congruence between teachers' spoken discourse and students' vocabulary levels: Is the gap too wide? *Language Matters*, 53(1), 63–80. <https://doi.org/10.1080/10228195.2022.2066157>
- Cronje, M. M. (2022). The role of higher education institutions in addressing South Africa's reading crisis in view of sustainable development. *Perspectives in Education*, 40(3), 181–196. <https://doi.org/10.18820/2519593X/pie.v40.i3.12>
- Department of Basic Education. (2023). *Progress in International Reading Literacy Study 2021: South African preliminary highlights report*. Department of Basic Education. https://www.education.gov.za/Portals/0/Documents/Reports/2023/PIRLS%202021_Highlights%20Reports.pdf?ver=2023-05-25-100949-603
- Dong, Y., Tang, Y., Chow, B. W.-Y., Wang, W., & Dong, W.-Y. (2020). Contribution of vocabulary knowledge to reading comprehension among Chinese students: A meta-analysis. *Frontiers in Psychology*, 11, Article 525369. <https://doi.org/10.3389/fpsyg.2020.525369>
- Dubravac, V., Brdarević-Čeljo, A., & Bećirović, S. (2024). English vocabulary development in the global south context. *Language Teaching Research*. Advance online publication. <https://doi.org/10.1177/13621688241273844>
- Hu, M., & Nation, P. (2000). Unknown vocabulary density and reading comprehension. *Reading in a Foreign Language*, 13(1), 403–430. <https://doi.org/10.64152/10125/66973>
- Jeon, E. H., & Yamashita, J. (2014). L2 reading comprehension and its correlates: A meta-analysis. *Language Learning*, 64(1), 160–212. <https://doi.org/10.1111/lang.12034>
- Krashen, S. (1989). We acquire vocabulary and spelling by reading: Additional evidence for the input hypothesis. *The Modern Language Journal*, 73(4), 440–464. <https://doi.org/10.1111/j.1540-4781.1989.tb05325.x>
- Laufer, B., & Nation, P. (1995). Vocabulary size and use: Lexical richness in L2 written production. *Applied Linguistics*, 16(3), 307–322. <https://doi.org/10.1093/applin/16.3.307>
- Laufer, B., & Nation, P. (1999). A vocabulary-size test of controlled productive ability. *Language Testing*, 16(1), 33–51. <https://doi.org/10.1177/026553229901600103>
- Laufer, B., & Ravenhorst-Kalovski, G. C. (2010). Lexical threshold revisited: Lexical text coverage, learners' vocabulary size and reading comprehension. *Reading in a Foreign Language*, 22(1), 15–30. <https://doi.org/10.64152/10125/66648>
- Nation, I. (2006). How large a vocabulary is needed for reading and listening? *The Canadian Modern Language Review*, 63(1), 59–82. <https://doi.org/10.3138/cmlr.63.1.59>
- Nation, I. (2016). *Making and using word lists for language learning and testing*. John Benjamins Publishing Company. <https://doi.org/10.1075/z.208>
- Nation, I. (2022). *Learning vocabulary in another language* (3rd ed.). Cambridge University Press. <https://doi.org/10.1017/9781009093873>
- Nizonkiza, D., & Van Dyk, T. (2015). Academic literacy of South African higher education level students: Does vocabulary size matter? *Stellenbosch Papers in Linguistics*, 44, 147–174. <https://doi.org/10.5774/44-0-159>
- 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. <http://hdl.handle.net/11660/3616>
- Pretorius, E. J., & Murray, S. (2025). *A vocabulary journey with Foundation Phase teachers*. Zenex Foundation. <https://www.zenexfoundation.org.za/working-article-a-vocabulary-journey-with-foundation-phase-teachers/resources/>
- Pretorius, E. J., & Stoffelsma, L. (2021). The potential for word growth in Grade 3 in disadvantaged urban schools: Factors that facilitate or impede the transition to Grade 4. In C. van der Walt & V. Pfeiffer (Eds.), *Multilingual classroom contexts: Transitions and transactions* (pp. 17–44). SUN PRESS.

- Scheepers, R. (2016). The importance of vocabulary at tertiary level. *Journal for Language Teaching*, 50(1), 53–77. <https://doi.org/10.4314/jlt.v50i1.3>
- Schmitt, N., Schmitt, D., & Clapham, C. (2001). Developing and exploring the behaviour of two new versions of the Vocabulary Levels Test. *Language Testing*, 18(1), 55–88. <https://doi.org/10.1177/026553220101800103>
- Spaull, N., & Pretorius, E. (2019). Still falling at the first hurdle: Examining early grade reading in South Africa. In N. Spaull & J. D. Jansen (Eds.), *South African schooling: The enigma of inequality: A study of the present situation and future possibilities* (pp. 147–168). Springer. https://doi.org/10.1007/978-3-030-18811-5_8
- Statistics South Africa. (2023). *Census 2022 statistical release* (Report P0301.4). https://census.statssa.gov.za/assets/documents/2022/P03014_Census_2022_Statistical_Release.pdf
- Stoffelsma, L., Wilsenach, C., & Mutema, F. (2026). Perspectives on oral reading fluency, vocabulary and reading comprehension in bilingual learners from Ghana, South Africa and Zimbabwe. *Reading & Writing-Journal of the Literacy Association of South Africa*, 17(1), Article 583. <https://doi.org/10.4102/RW.v17i1.583>
- Teng, M. F., & Reynolds, B. L. (Eds.). (2024). *Researching incidental vocabulary learning in a second language*. Routledge. <https://doi.org/10.4324/9781003270782>
- Van Hees, J., & Nation, P. (2017). *What every primary school teacher should know about vocabulary*. NZCER Press.
- Wang, Z., Sabatini, J., O'reilly, T., & Weeks, J. (2019). Decoding and reading comprehension: A test of the decoding threshold hypothesis. *Journal of Educational Psychology*, 111(3), 387–401. <https://doi.org/10.1037/edu0000302>
- Webb, S. (2019). Incidental vocabulary learning. In S. Webb (Ed.), *The Routledge handbook of vocabulary studies* (pp. 225–239). Routledge. <https://doi.org/10.4324/9780429291586-15>
- Weder, N. (2022). *The effect of different interventions on the inferencing abilities of academically vulnerable first-year students* [Doctoral thesis, University of Pretoria]. UPSpace Repository. <https://repository.up.ac.za/handle/2263/89324>
- Xing, P., & Fulcher, G. (2007). Reliability assessment for two versions of Vocabulary Levels Tests. *System*, 35(2), 182–191. <https://doi.org/10.1016/j.system.2006.12.009>
- Zhang, S., & Zhang, X. (2022). The relationship between vocabulary knowledge and L2 reading/listening comprehension: A meta-analysis. *Language Teaching Research*, 26(4), 696–725. <https://doi.org/10.1177/1362168820913998>

Please cite this article as:

Fouché, I., & Prozesky, M. (2026). The osmosis plateau: The limits of incidental vocabulary acquisition in South African education students. *Student Success*, Vol 17(2), 92-103. <https://doi.org/10.63608/ssj.4026>

This article has been accepted for publication in *Student Success*. Please see the Editorial Policies under the ‘About’ section of the Journal website for further information.

Student Success: A journal exploring the experiences of students in tertiary education.



Except where otherwise noted, content in this journal is licensed under a [Creative Commons Attribution 4.0 International Licence](https://creativecommons.org/licenses/by/4.0/). As an open access journal, articles are free to use with proper attribution. ISSN: 2205-0795