

2013

Recognition vocabulary skill as a predictor of academic English performance and academic achievement in English

Thomas Roche
Southern Cross University

Michael Harrington
University of Queensland

Publication details

Postprint of: Roche, T & Harrington, M 2013, 'Recognition vocabulary skill as a predictor of academic English performance and academic achievement in English', *Language Testing in Asia*, vol. 3, no. 12, pp. 133-144.

Published version available from:

<http://dx.doi.org/10.1186/2229-0443-3-12>

© 2013 Roche and Harrington This is an open access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/2.0>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

This Provisional PDF corresponds to the article as it appeared upon acceptance. Fully formatted PDF and full text (HTML) versions will be made available soon.

Recognition vocabulary knowledge as a predictor of academic performance in an English as a foreign language setting

Language Testing in Asia 2013, **3**:12 doi:10.1186/2229-0443-3-12

Thomas Roche (thomas.roche@scu.edu.au)
Michael Harrington (m.harrington@uq.edu.au)

ISSN 2229-0443

Article type Research

Submission date 21 March 2013

Acceptance date 21 March 2013

Publication date 30 April 2013

Article URL <http://www.languagetestingasia.com/content/3/1/12>

This peer-reviewed article can be downloaded, printed and distributed freely for any purposes (see copyright notice below).

For information about publishing your research in *Language Testing in Asia* go to

<http://www.languagetestingasia.com/authors/instructions/>

For information about other SpringerOpen publications go to

<http://www.springeropen.com>

Recognition vocabulary knowledge as a predictor of academic performance in an English as a foreign language setting

Thomas Roche^{1*}

* Corresponding author

Email: thomas.roche@scu.edu.au

Michael Harrington²

Email: m.harrington@uq.edu.au

¹ Southern Cross University, Australia & Sohar University, Queensland, Australia

² University of Queensland, Queensland, Australia

Abstract

This paper presents findings of a study of recognition vocabulary knowledge as a predictor of written Academic English Proficiency (AEP) and overall Academic Achievement in an English medium higher education program in an English-as-a-Foreign-Language (EFL) context. Vocabulary knowledge was measured using a Timed YES/NO (TYN) test. AEP was assessed using an academic writing test based on IELTS. Performance on these measures was correlated with Grade Point Average (GPA) as a measure of academic achievement for Arabic L1 users (N=70) at an English-medium College of Applied Sciences in the Sultanate of Oman. Vocabulary size and speed correlated with both academic writing and GPA measures. The combined vocabulary and writing measures were also examined as predictors of academic achievement. The TYN test is discussed as reliable, cost and time effective general measure of AEP and for showing if students have the necessary vocabulary knowledge to undertake study in a tertiary level English medium program.

Keywords

Placement test, Language proficiency, Academic achievement, Recognition vocabulary, English medium education

Background

English as a medium of instruction

English has become the dominant medium of instruction used in higher education internationally. In 2006 English was used as a medium of instruction in 103 countries, whereas the second most commonly used university classroom language was French in 42 countries (Ammon, 2006)^a. This growth is particularly noticeable in countries defined by Kachru (1985) as being in the outer- and expanding-circles of English use where English is traditionally considered a foreign language (Bashir, 2007).

In Kachru's notion of language circles, the national context of English use can be classified as situated within inner-, outer- or expanding-circles, according to a nation's historico-political relationship with Anglo-American hegemonic power. Inner-circle nations are those where English is the primary language of the state (e.g. New Zealand). Outer-circle nations consist of former British dominions where English is still widely used as a second language to communicate between different groups (e.g. India). The expanding-circle refers to those nations without direct experience of British rule, where English plays an important role, be that in commerce, education or elsewhere (e.g. Oman).

Research exploring the challenges facing L2 students in English medium programs primarily focus on L2 English users in Anglophone countries, such as Australia and USA (see, Murray, 2012; Read, 2008), rather than English L2 users at universities and colleges in countries where English is not the primary language outside the classroom^b. Those studies concerned with L2 English users studying at English medium universities in expanding- and outer-circle contexts indicate that such students experience a number of language proficiency related difficulties. In one self-assessment questionnaire study of students' perceptions of their study experience in English medium faculties of Social Science, Humanities, Engineering and Business at Norwegian and German universities, 42% of the Norwegian sample and 72% of the German students reported substantial comprehension difficulties (Hellekjaer, 2010). In another self-report study of students in English medium higher education programs in China, students identified comprehending lectures, understanding specialist vocabulary and writing in an appropriate academic style, as the major difficulties associated with studying through English as an L2 (Evans & Morrison, 2011)^c. Research by Kırkgöz (2005) in Turkey, similarly found among first and fourth year English L2 undergraduate students that the incomprehensibility of both lectures and reading material were amongst students' greatest concerns. Though little empirical research exists into the relationship between L2 comprehension and performance in these expanding- and outer-circle educational contexts, one study of 2000 bilingual Arab students in an English medium tertiary level science program, found that proficiency in pre-faculty study preparatory year English correlated significantly with academic performance in a Calculus course (Yushau & Omar, 2007). The current research aims to contribute to the literature by investigating the relationship between academic English proficiency and overall academic performance of students studying in English medium tertiary level programs in the Sultanate of Oman, a country on the Arabian Peninsula with no direct experience of British rule but where English plays a key role in education and commerce, and as such will be considered as being in the expanding-circle of English use for the purpose of this study.

Vocabulary and success at english medium universities

Research in the 1980s and 90s lead to the recognition that vocabulary knowledge is a precondition for most other language abilities, (see, Alderson & Banerjee, 2001, for a review of that research); and, the emergence of lexical approaches to language learning (Willis, 1990; Lewis, 2002; McCarthy, 2003). Corpus linguistics studies reported that the most frequent 2,000 headwords used in English account for between 80-85% of the words of any spoken or written English text, depending on the text type (Nation, 1990; Nation & Waring, 1997; Nation & Newton, 1997). The educational corollary of this research is that L2 learners should first master these 2000 most frequently occurring words before attempting to study content through English (Meara, Lightbown & Halter, 1997).

Unassisted comprehension of English texts requires vocabulary knowledge greater than the most 2,000 frequently occurring headwords. In this research the headword *build*, forms the core of a word-family including related words such as *builder*, *building*, and *built*. Between 95-98% of the word-families in a text must be comprehensible for unassisted comprehension of that text (Waring & Nation, 2004; Hsueh-Chao & Nation, 2000). Knowledge of only the 2000 most frequent word-families would typically mean that approximately one in every five words encountered in an academic English text will still be unknown and thereby render the text largely unintelligible (Nation & Waring, 1997). More recent studies suggest that learners need knowledge of 8,000 to 9,000 word-families for unassisted comprehension of written texts, and 6,000 to 7,000 for spoken texts (Nation, 2006). However, this figure of 8,000-9,000 word-families is not a gatekeeping figure for undertaking higher education studies in English medium environments. Although some researchers insist that a vocabulary of 10,000 word-families is necessary for successful study in linguistically demanding higher degree courses (Hazenberg & Hulstijn 1996 in Schmitt, Schmitt & Clapham, 2001), there is general agreement in the literature (Nation, 1990, Schmitt, 2000, Meara et al., 1997) that an L2 learner's passive vocabulary should reach a minimum threshold of 5,000 word-families before L2 users undertake studies in English medium universities and colleges.

The greater the learner's vocabulary knowledge, the less cognitive demands are placed on a learner. More developed levels of vocabulary knowledge enable learners to read with less effort (Segalowitz & Segalowitz, 1993; Segalowitz, Segalowitz & Wood, 1998) and result in better performance on comprehension tests (Chen, 2011; Miller & Peleg, 2010; Nassaji, 2003; Qian, 1999). The preponderance of the evidence indicates that in order for an L2 to become a vehicle for learning, vocabulary knowledge must first be sufficiently developed. This in turn suggests that measures of vocabulary knowledge could predict written AEP proficiency and academic performance.

Learners' vocabulary size has been assessed using both traditional multiple-choice checklist formats involving yes/no judgments (Meara & Buxton, 1987; Mochida & Harrington, 2006). A widely used paper-based test of vocabulary size is the *Vocabulary Levels Test* (Nation, 1990). The test uses a multiple matching format in which items are banded by frequency levels, allowing size to be inferred from performance across decreasing frequency levels. An alternative measure of vocabulary size is the Yes/No test first introduced by (Meara & Buxton, 1987; Meara & Jones, 1988). The test presents a mix of frequency-banded words as well as phonologically-possible nonwords. The learner simply checks which words are known and which are not. The inclusion of nonwords is designed to control for guessing, with nonword errors subtracted from correct word performance for an overall score (Mochida & Harrington, 2006). Although the validity and usefulness of the correcting for guessing in educational testing has been questioned (Ebel, 1979), nonword performance provides an important control variable for the format, especially when guessing varies across participants. The use of non words does raise a number of issues for the validity and reliability of the format, and these have received significant attention from researchers (Beeckmans, et al., 2001; Cameron, 2002; Eyckmans, 2004; Huibregtse, Admiraal, & Meara 2002; Mochida & Harrington, 2006). Although the YN test does not directly elicit vocabulary knowledge, it has been shown to correlate highly with standard measures of vocabulary knowledge (Mochida & Harrington, 2006). The *Timed Yes/No* test used here differs from earlier versions in that it also collects participants' response time, thus providing an additional, implicit measure of L2 vocabulary knowledge.

The TYN test measures vocabulary knowledge as a discrete, context-free (trait) entity (Read & Chapelle, 2001). The use of both vocabulary size and speed of access provides, in principle, a measure of vocabulary ability that will be more valid than either dimension alone (Chapelle, 2006). The instrument has been piloted with tertiary EFL learners in Singapore, and has also been tested extensively in Australia on students who were newly arrived in that country, and thus had EFL equivalent proficiency at the time of testing. Test findings, including validity and reliability data have been reported (Harrington & Carey, 2009; Harrington, 2006; Harrington, 2007; Mochida & Harrington, 2006). The effectiveness of this timed Yes/No response test has never been established in an outer- or expanding circle EFL setting, and this is one of the reasons why this study is important.

A secondary aim of this study is to establish whether decontextualized word recognition tests are valid tools for assessing Arabic L1 users' English language proficiency. Schmitt, Schmitt & Clapham (2001, 64) point out that the linguistic, cognitive and cultural background of participants influences their response (time and/or accuracy) to vocabulary tests. Research by Fender (2003; 2008) and Ryan & Meara (1991) indicates that Arab L1 users experience greater difficulty than proficiency equivalent ESL learners generally in reading tasks and in isolated word recognition tasks. Fender (2008) suggests that nature of Arabic literacy and a difference in decoding skills used in English and Arabic may account for these difficulties. Across the Arab world online and print materials of the mass media are typically written in Modern Standard Arabic, an alphabet that does not encode short vowels thus requiring that readers not only utilize the explicit phonological information in the orthography, but also extra-lexical information such as morpho-syntactic knowledge and discourse context (Fender, 2008, 26; Abu Rabia & Seigel, 1995). In terms of culture, universal school education began in the Sultanate of Oman in the late sixties, prior to which there were only three schools and no universities in Oman (Roche, 2009), as a result of which the widespread illiteracy of the 1960's dropped to below 19% in the early twenty-first century (CIA, 2011). Al-Amrani (2009) found that current Omani EFL university students are reluctant readers who use less effective bottom-up reading strategies than proficiency equivalent non-Omani EFL student peers. Given this background the present study also attempts to establish whether word-recognition tasks can be used with Arabic L1 users to determine their Academic English Proficiency.

The study

This study explores the relationship between vocabulary knowledge, written Academic English proficiency and academic performance of EFL students in Oman. The first objective of the study, is to establish whether:

- 1) Academic performance by Arabic L1 users at an English medium institution in Oman can be predicted by two key measures of written academic English proficiency.

Academic Performance will be measured by Grade Point Average (GPA) taken from students' academic transcripts, and written Academic English Proficiency (AEP) will be measured through scores on a mock IELTS academic writing test. The second hypothesis is a refinement of the first:

- 2) The TYN Test is an effective predictor of written academic English proficiency for the participants.

The motivation for this study is to examine the relationship between vocabulary recognition knowledge, written academic English proficiency, and student academic performance in an English medium institution in an expanding-circle context for Arabic

L1 users. Of particular interest is the potential use of measures of vocabulary recognition knowledge as a tool for identifying learners who may lack the needed written English proficiency to undertake tertiary studies in English medium institutions in this context. To the extent the effectiveness of the tool can be established, it will provide these institutions an efficient and cost-effective method for identifying students who may require further academic English language support before attempting tertiary study in English.

Methods

Participants

The participants from this study (n= 70) were students in an English language medium higher education institution in the Sultanate of Oman, Rustaq College of Applied Science. These students were from the first and fourth year of Faculty study and all reported their first language to be Arabic. The data collection period took place within one week with students in their 13th and 14th week of the second 15-week semester for that academic year. Participation was voluntary and the study, was carried out in accordance with the College's ethical guidelines.

Materials

Academic English written proficiency was assessed using a task adapted from practice IELTS materials (Cambridge University Press, 2005). Each paper was rated on a ten-point scale reflecting proficiency in grammar, vocabulary, coherence and response to the task. Each student was directed to write a 250 word short essay on the topic "Oman in the past, Oman in the future". Students were given 40 minutes to complete the task.

A discrete item computerized timed YES/NO response test was used to determine vocabulary knowledge. Two versions of 100 items each were used. Each test consisted of 72 words (18 each from four frequency levels) and 28 nonwords. The words are drawn from bands of the 1000 (1K), 2000 (2K), 3000 (3K), 5000 (5k) and 10,000 (10K) most frequently occurring words in the British National Corpus (Harrington & Carey, 2009). In the first set, Word Test A, less frequent words were used from the 2K, 3K, 5K, 10K bands; in Word Test B, more common words from the 1K-5K frequency bands were tested. Nonwords used are phonologically permissible English words (e.g. *blurge*) as opposed to non-permissible items (e.g. *rbgeul*). The TYN test measures participants' reaction time and accuracy, responding "YES" to words and "NO" to nonwords. "YES" responses to nonwords (false-alarms) resulted in a reduction in the participants' score. See Table 1 for the matrix of possible responses used to test vocabulary knowledge. For each set of words, each item appears for 5000 milliseconds (5 seconds), if a participant fails to respond within that time the word disappears to be replaced by a new word. Response time score is taken from the mean response of each participant to the presented items. The test was administered using *LMAP*, a web-based testing tool developed at the University of Queensland, Australia, www.languagemap.com.

Table 1 Matrix of possible responses, where UPPER CASE = correct responses

	Word	Non-word
<i>Yes</i>	HIT	False Alarm
<i>No</i>	Miss	CORRECT REJECTION

The academic performance of the students was measured by the participants' Grade Point Average (GPA) in their studies for the current semester. The GPAs were provided with the students' permission, by the Dean's Office.

Procedure

Data collection took place during the 13th and 14th weeks of the 15-week second semester. The tests were administered by the first author and collaborating staff. All testing was done in a computer lab. Students first completed the writing tests, followed by the computer-based vocabulary tests. Instructions in Modern Standard Arabic were given in written form and were also read aloud by an Omani research assistant. The testing format was explained and students did a set of practice items for each test. In both tests they were encouraged to work as quickly and as accurately as possible, as both accuracy and response time measures were being collected.

For the Vocabulary test students were warned that clicking "YES" for nonwords would result in lower scores. They were also told that each item would appear on the screen for only 5000 milliseconds (5 seconds) and then disappear. If there was no response it was counted as incorrect. There were very few no responses, comprising less than 0.05% of the total response set. In addition to being an additional window on underlying proficiency, the inclusion of the response time condition discouraged strategic and reflective processing on the part of the students, thus providing a more direct measure of vocabulary knowledge.

Results

The academic writing papers were marked by trained and practicing IELTS examiners, with a random sample of essays (20%) marked by both raters. Interrater reliability was assessed by using the intra-class coefficient (ICC) statistic, which measures the consistency between the raters' judgment (Field, 2009). $ICC = .88, p < .001$; 95% confidence interval, 67.00 - 96.10. This indicates a high level of consistency across the two raters.

Score reliability for the vocabulary measures was assessed using Cronbach's Alpha. Table 2 reports the reliability coefficients for accuracy and response time measures for the Advanced and Basic versions.

Table 2 Cronbach's Alpha reliability coefficients for advanced and basic word tests

	Words (72 items)	Nonwords (28 items)	Combined (100 items)
Advanced test			
Accuracy	.89	.89	.85
Response time	.96	.92	.97
Basic test			
Accuracy	.89	.88	.84
Response time	.95	.92	.94

Separate coefficients for the word and nonword items were calculated as the recognition of words and the rejection of nonwords arguably reflect different dimensions of underlying lexical knowledge (Mochida & Harrington, 2006). Combined reliability values were also calculated.

The descriptive statistics for the language measures and grade-point averages are presented in Table 3. Mean performance on the Basic test was better than on the Advanced tests, for both accuracy and response time, presumably reflecting the difficulty of the items in the respective tests.

Table 3 Descriptive statistics (Means, Standard Deviations, Range) for advanced and basic word tests, grade point average and IELTS writing scores, N = 70

	M^a	SD	Range	
Advanced Word				
Accuracy	35.96	12.05	8.93	58.93
Response time ^b	1590	371	946	3200
False alarm rate	23.70	13.46	3.57	60.71
Basic Word				
Accuracy	53.23	15.83	8.33	80.95
Response time	1571	372	909	3120
False alarm rate	22.57	11.03	3.57	50.00
IELTS writing	5.18	.86	3.50	7.00
Grade point average	2.79	.55	1.00	3.70

^a Corrected for guessing score, proportion of yes responses to words (hits) minus proportion of yes responses to nonwords (false alarms).

^b Milliseconds (ms).

A pair wise t-test on the mean accuracy differences for Test A and B was significant, $t(69) = 12.43$, $p < .001$, Cohen's $d = 1.22$, the latter indicating a strong effect size. The mean response times were log-transformed for all statistical analyses reported here. The response time differences were not significant. False alarm rates for both word tests compare with pre-faculty English language students in Australia whose mean false alarm scores were 25% for beginners and 10% for advanced learners (Harrington & Carey, 2009). There were a small number of students with extremely high false alarm rates that could have arguably been removed as outliers. However, given that a motivation for the study is to assess the effectiveness of the TYN Test in an authentic testing context, these potentially distorting data points were not removed, as they might be in a typical psychology or laboratory setting study.

The strength of association among the word, writing and GPA measures are evaluated by first reporting on the bivariate correlations between the measures. See Table 4. Also reported is a composite word measure. This measure combines the individual participants z-scores for accuracy and response times into a single measure in an attempt to provide a more stable measure of underlying word skill (Ackerman & Cianciolo, 2000). The effectiveness of the composite score as a predictor of academic achievement will be assessed in both the bivariate correlations and a regression analysis that evaluates the relative contributions the word and writing measures make in predicting academic performance.

Table 4 Bivariate correlations for language measures and grade-point-average

	2. Adv RT	3. Adv Comp	4. Basic Acc	5. Basic RT	6. Basic Comp	7. Acad Writ	8. GPA
1. Advanced Word Accuracy	.02	.70**	.68**	-.06	.47**	.35*	.14
2. Advanced Word Response Time	1	-.70**	-.16	-.17	.57**	-.20	-.27*
3. Advanced Word Composite		1	.60**	-.49**	.69**	.38*	.29*
4. Basic Word Accuracy			1	-.25*	.79**	.51**	.34**
5. Basic Word Response Time				1	-.79**	-.11	-.38**
6. Basic Word Composite					1	.36*	.43**
7. Academic writing						1	.40**
8. Grade-point-average							1

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

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

As is evident in Table 4 all the language measures, with the exception of the results for Advanced Word Accuracy, had a moderately strong correlation with GPA. The lack of correlation between accuracy and response time for the respective tests indicated there was no systematic speed-accuracy trade-off by the participants.

In general, faster and more accurate word recognition skill had a significant correlation with GPA, as did performance on the writing task. The composite scores had a stronger correlation with GPA than the accuracy and response time results for the respective tests. It was also evident that vocabulary skill was a moderately strong predictor of writing outcomes.

Regression analyses were also carried out to assess how well the vocabulary and writing measures predicted GPA when the two variables were considered in combination. Given the inter-correlations between vocabulary, writing and GPA, it is not clear if better vocabulary performance is due to being a better writer, or better vocabulary skills determine writing performance, the similar correlations ($r = .3$ and $-.4$) with GPA indicating that the two measures tap the same underlying knowledge. Alternatively, the vocabulary and writing scores may make relatively independent contributions to academic achievement. If the former, then the two measures would be interchangeable as tests of academic English skill, though with important practical differences between the two in terms of administration and scoring. Alternatively, if the vocabulary (accuracy and response time) and writing measures each account for a substantial amount of additional variance in GPA, this would indicate that two complement each other in indexing learner proficiency levels. The regression models assess how much overall variance the measures together account for, and the relative contribution of each measure to this amount.

Table 5 reports four analyses. The Advanced and Basic Word results are evaluated separately. For each word test two models were developed. The first assess the contribution

of the word measures to predicting the criterion GPA after the writing measures are entered. The second model will enter the vocabulary measures first and then the writing measures. The composite scores, which were generated from the raw measures, will not be analysed.

Table 5 Hierarchical regression analyses of the advanced word and writing measures with GPA as criterion variable and writing and word measure as predictors

Advanced Word Models	R^2	Adjusted R^2	R^2 change	B	SEB	β
Writing	.160	.147	.160***	.230	.077	.399
Advanced Accuracy	.160	.135	.000	.001	.005	.021
Advanced Response time	.198	.162	.028	-1.171	.659	-.200
Advanced Accuracy	.02	.005	.02	.001	.005	.021
Advanced Response time	.091	.063	.071*	-1.172	.659	-.2
Writing	.198	.162	.108**	.230	.077	.357
Basic Word Models						
Writing	.16	.147	.160***	.215	.079	.334
Basic Accuracy	.184	.160	.024	.003	.004	.093
Basic Response time	.267	.234	.083**	-1.708	.624	-.299
Basic Accuracy	0.215	0.079	0.334	0.003	0.004	0.093
Basic Response time	0.003	0.004	0.093	-1.708	0.624	-0.299
Writing	-1.708	0.624	-0.299	0.215	0.079	0.334

B, Unstandardized coefficient; SEB, Standard Error B; β , Standardized coefficient.

F significant at * $p < .05$, ** $p < .01$, *** $p < .001$.

$N = 70$.

To summarize the regression analyses, as expected from the bivariate correlations, the word measures and writing measures predicted significant variance in GPA differences. The model based on the Basic word measures accounted for nearly 25% of the GPA variance (total adjusted $R^2 = .234$) while the Advanced word model only accounted for 16% (.162). This is due to both the accuracy and response time results serving to discriminate between GPA differences. Although the word and writing measures have similar correlations with GPA in the bivariate comparisons, writing is a better predictor when the two variables are considered together.

Discussion

The first objective of the study was to assess the ability of Arabic L1 users' written Academic English Proficiency (AEP) to predict academic performance at tertiary education institutions in an expanding-circle context. Results indicate that AEP, both academic writing skills and vocabulary knowledge are good predictors of overall academic performance in this context. The TYN Test is a less sensitive measure than the academic writing tests but nonetheless predicts academic performance. These results give support to research outlining the importance of English proficiency (Yushau & Omar, 2007), and in particular, vocabulary knowledge as a prerequisite for academic success in English medium programs (Meara, et al. 1997; Waring & Nation, 2004; Hsueh-Chao & Nation, 2000).

It is of note that the results gathered on the TYN Test present a high number of false alarms. As noted above, lower false alarm rates were evident in studies in Anglophone countries. The results of the regression analysis indicate that the TYN Test can be used as a screening tool

with fairly comparable effectiveness to a writing task. However, the results here differed noticeably from those obtained from previous studies with English L2 university students in English speaking countries. As noted above, the mean false alarm rates were much higher and the group means (as a measure of vocabulary size) were also lower, with students in this study performing at the level of pre-tertiary students in Australia. It is possible that the results are influenced by the participants' L1 Arabic. Arab students have been shown to encounter greater difficulty with English spelling and word processing than proficiency equivalent ESL learners from other L1 backgrounds. This effect has been attributed to the influence of Arabic orthography and literacy practices (Fender, 2008; Ryan & Meara, 1991; Milton, 2009). Performance on the written TYN Test requires knowledge of spelling and word meaning, and it is not clear the extent to which the poorer performance on both the words and nonwords (the latter conforming to English spelling rules) can be reduced to spelling difficulties. This is a question for future research.

This overall higher false alarm rate might be because the instructions were not clear and participants were not aware that they were being penalized for incorrect guesses; or, that these were acquiescent responses (Dodorico-McDonald, 2008) - random clicks that brought the test to an un-taxing end rather than reflecting their knowledge or lack of knowledge of the items presented. For improved results the test results should be meaningful to the students, for example we would expect fewer false alarms if the test acted as gateway to further study or was integrated into courses and contributed towards students marks. Further research might also look at how test performance interacts with instructional variables such as time of semester. The comparatively slower response times may indicate that not all students were aware that speed of response was also being measured, another recommendation of the study is that the instructions should be delivered online in a standardized format with a video demonstration on the computer to improve the clarity of the instructions.

The results also show that vocabulary recognition test scores serve as a predictor of written academic English language proficiency. There was a strong correlation between *Language MAP* test scores and academic writing scores, particularly with Word Test B, which represents words from the 1-5K most commonly occurring words, which authors such as Nation (1990), Schmitt (2000) and Meara et al., (1997) claim are essential for reading English texts and are the minimum knowledge required to begin study in English medium programs. Once again, the findings here are not as strong as the correlations resulting from studies using *Language MAP* and placement tests for students studying in inner-circle settings, such as pre-Faculty courses in Australia (Harrington & Carey, 2009) but the results are nonetheless significant. The weaker the student's vocabulary knowledge the poorer they are likely to perform both in measures of their academic English proficiency, which is fundamental to success in their studies, and in overall academic performance. This highlights the TYN Test's potential for use as an English language proficiency placement tool, or for tracking and monitoring changes in written academic English proficiency among matriculated students for Post-Enrolment Language Assessment (PLA) purposes.

Conclusion

This study contributes to a growing body of research stressing the fundamental importance of L2 English proficiency for achieving success in English medium tertiary education programs (Murray, 2012; Read, 2008; Harrington & Roche, submitted). The results presented here recommend the TYN Test as a useful predictor of both overall academic achievement and written academic English language proficiency for Arabic L1 users, despite concerns raised

to the contrary in the literature concerning the written language processing strategies used by Arabic L1 users of English (Ryan & Meara, 1991; Fender, 2008; Abu Rabia & Seigel, 1995), and in particular their difficulties in reading English electronic texts (Al-Amrani, 2009). The results show that visual word recognition is a good measure of L2 proficiency. The TYN Test is especially attractive given the limited resources needed to administer the test and generate results. As for future directions, research is underway that assesses the predictive power of the test as a diagnostic for readiness to take IELTS by learners in expanding-circle settings like the one examined here.

Endnotes

^aOf approximately 6,000 living languages, 82 are used in tertiary education, 39 languages are used in 2 or more countries' universities and only 13 are used in three or more (Ammon, 2006, 557).

^bFor example, research shows that non-English speaking background (NESB) students' English language proficiency correlates positively with academic success in Australian Universities (Feast, 2002)

^cThe students' perceptions fit well with Lin & Morrison's (2010) vocabulary levels research which reported that the majority (76.1%) of Hong Kong students at English medium universities do not have sufficient vocabulary to comprehend lectures in English or undertake tertiary study.

Competing interest

The authors declare that they have no competing interest.

Authors information

Thomas Roche is the Director of Studies of the English Language Centre at Southern Cross University, Australia, an Associate Professor at Sohar University, Sultanate of Oman, and an Associate Research Fellow in the School of Languages & Comparative Cultural Studies at the University of Queensland, Australia. His research interests include individual difference in foreign language learning and language testing.

Michael Harrington is a Senior Lecturer in Second Language Acquisition in the School of Languages & Comparative Cultural Studies at the University of Queensland, Australia. He has published in areas including lexical processing and the measurement of L2 vocabulary skills.

Authors' contribution

Both authors read and approved the final manuscript.

Acknowledgements

This research was supported by a grant from the Omani Research Council [Grant number ORG SU HER 10 003]. The authors would like to thank the anonymous reviewer for useful comments and suggestions.

References

Abu Rabia, S., & Seigel, L. S. (1995). Different orthographies, different context effects: The effects of Arabic sentence context in skilled and poor readers. *Reading Psychology, 16*, 1–19.

Ackerman, P. L., & Cianciolo, A. T. (2000). Cognitive, perceptual-speed, and psychomotor determinants of individual skill acquisition. *Journal of Experimental Psychology: Applied, 6*, 259–290.

Alderson, J. C., & Banerjee, J. (2001). Language testing and assessment (Part 1) “State of the art review. *Language Testing, 18*, 213–236.

Al-Amrani, S. (2009). Strategies for reading on-line and printed texts by Omani EAP students. In T. Roche (Ed.), *Orientations in language learning and translation* (pp. 41–60). Muscat, Oman: Al Falaj Press.

Ammon, U. (2006). The language of tertiary education. In E. K. Brown (Ed.), *Encyclopaedia of languages & linguistics* (pp. 556–559). Amsterdam: Elsevier.

Bashir, S. (2007). *Trends in international trade in higher education: implications and options for developing countries education working papers series*. Washington: World Bank.

Beeckmans, R., Eyckmans, J., Janssens, V., Dufranne, M., & Van de Velde, H. (2001). Examining the yes-no vocabulary test: Some methodological issues in theory and practice. *Language Testing, 18*, 235–274.

Cambridge University Press. (2005). *Cambridge IELTS 4. Examination papers from University of Cambridge ESOL Examinations*. Cambridge, UK: Cambridge University Press.

Cameron, L. (2002). Measuring vocabulary size in English as an additional language. *Language Teaching Research, 6*, 145–173.

Chapelle, C. (2006). L2 vocabulary acquisition theory: The role of inference, dependability and generalizability in assessment. In M. Chalhoub-Deville, C. A. Chapelle, & P. Duff (Eds.), *Inference and generalizability in applied linguistics: multiple perspectives* (pp. 47–64). Amsterdam: John Benjamins.

Chen, K. Y. (2011). The impact of EFL students’ vocabulary breadth of knowledge on literal reading comprehension. *Asian EFL Journal, 51*. Retrieved from <http://asian-efl-journal.com/teaching-articles/2011/04/21/the-impact-of-efl-students-vocabulary-breadth-of-knowledge-on-literal-reading-comprehension>.

CIA. (2011). *World factbook: Middle East: Oman*. Retrieved from <https://www.cia.gov/library/publications/the-world-factbook/geos/mu.html>.

Dodorico-McDonald, J. (2008). Measuring personality constructs: The advantages and disadvantages of self-reports, informant reports and behavioural assessments. *Enquire, 1*(1). Retrieved from www.nottingham.ac.uk/shared/McDonald.pdf.

Ebel, R. L. (Ed.). (1979). *Essentials of educational measurement*. Englewood Cliffs, NJ: Prentice-Hall.

Evans, S., & Morrison, B. (2011). Meeting the challenges of English-medium higher education: The first-year experience in Hong Kong. *English for Specific Purposes, 30*(3), 198–208.

Eyckmans, J. (2004). *Measuring receptive vocabulary size*. Utrecht: LOT.

Feast, V. (2002). The impact of IELTS scores on performance at University. *International Education Journal, 3*(4), 70–85.

Fender, M. (2003). English word recognition and word integration skills of native Arabic- and Japanese-speaking learners of English as a second language. *Applied Psycholinguistics, 24*, 289–315.

Fender, M. (2008). Spelling knowledge and reading development: Insights from Arab ESL learners. *Reading in a Foreign Language, 20*(1), 19–42.

Field, A. (2009). *Discovering statistics using SPSS*. London, UK: SAGE Publications.

Harrington, M. (2006). The lexical decision task as a measure of L2 lexical proficiency. *EUROSLA Yearbook, 6*, 147–168.

Harrington, M. (2007). *The coefficient of variation as an index of L2 lexical processing skill. University of Queensland working papers in linguistics*. Brisbane, Australia: School of English, Media Studies and Art History, The University of Queensland.

Harrington, M., & Carey, M. (2009). The on-line Yes/No test as a placement tool. *System: An International Journal of Educational Technology and Applied Linguistics, 37*, 614–626.

Harrington, M., & Roche, T. (2009). *Identifying academically at-risk students at an English-medium university in Oman: Post-enrolment language assessment in an English-as-a-foreign language setting*. submitted.

Hellekjaer, G. O. (2010). Lecture comprehension in English-medium higher education. *Hermes, 45*, 11–34.

Hsueh-Chao, M. H., & Nation, P. (2000). Unknown vocabulary density and reading comprehension. *Reading in a Foreign Language, 13*, 403–30.

Huibregtse, I., Admiraal, W., & Meara, P. (2002). Scores on a yes-no vocabulary test: correction for guessing and response style. *Language Testing, 19*, 227–245.

- Kachru, B. B. (1985). Standards, codification and sociolinguistic realism: The English language in the outer circle. In R. Q. A. H. Widdowson (Ed.), *English in the world: Teaching and learning the language and literatures* (pp. 11–36). Cambridge, UK: Cambridge University Press.
- Kırkgöz, Y. (2005). Motivation and student perception of studying in an English-medium university. *Journal of Language and Linguistic Studies*, 1(1), 101–123.
- Lewis, M. (2002). *Implementing the lexical approach*. Boston: Thomson and Heinle.
- Lin, L. H. F., & Morrison, B. (2010). The impact of the medium of instruction in Hong Kong secondary schools on tertiary students' vocabulary. *Journal of English for Academic Purposes*, 9(4), 255–266.
- McCarthy, M. (2003). *Vocabulary*. Oxford: Oxford University Press.
- Meara, P., & Buxton, B. (1987). An alternative multiple-choice vocabulary test. *Language Testing*, 4, 142–145.
- Meara, P., & Jones, G. (1988). Vocabulary size as a placement indicator. In P. Grunwell (Ed.), *Applied Linguistics in Society*. London, UK: CILT.
- Miller, P., & Peleg, O. (2010). Doomed to read in a second language: Implications for learning. *Journal of Psycholinguistic Research*, 39(1), 51–65.
- Milton, J. (2009). *Measuring second language vocabulary acquisition*. Bristol, England: Multilingual Matters.
- Mochida, A., & Harrington, M. (2006). The Yes-No test as a measure of receptive vocabulary knowledge. *Language Testing*, 23, 73–98.
- Meara, P., Lightbown, P., & Halter, R. (1997). Classrooms as lexical environments. *Language Teaching Research*, 1, 28–47.
- Murray, N. (2012). Ten “Good Practice Principles” ten key questions: considerations in addressing the English language needs of higher education students. *Higher Education Research & Development*, 31(2), 233–246.
- Nassaji, H. (2003). Higher-level and lower-level text processing skills in advanced ESL reading comprehension. *The Modern Language Journal*, 87, 261–276.
- Nation, P. (1990). *Teaching and learning vocabulary*. Boston: Heinle & Heinle.
- Nation, P. ((2006). *The Canadian Modern Language Review / La revue canadienne des langues vivantes*, 63(1), 59–81. September/septembre 2006.
- Nation, P., & Newton, J. (1997). Teaching vocabulary. In J. Coady & T. Huckin (Eds.), *Second language vocabulary acquisition*. Cambridge, UK: Cambridge University Press.

- Nation, P., & Waring, R. (1997). Vocabulary size text coverage. In N. A. Schmitt (Ed.), *Vocabulary: description, acquisition and pedagogy*. Cambridge, UK: Cambridge University Press.
- Qian, D. D. (1999). Assessing the roles of depth and breadth of vocabulary knowledge in reading comprehension. *Canadian Modern Language Review/ La Revue canadienne des languesvivantes*, 56(2), 282–308.
- Read, J. (2008). Identifying academic language needs through diagnostic assessment. *Journal of English for Academic Purposes*, 7(3), 180–190.
- Read, J., & Chapelle, C. A. (2001). A framework for second language vocabulary assessment. *Language Testing*, 18(1), 1–32.
- Roche, T. (2009). Introduction. In T. Roche (Ed.), *Orientations in language learning and translation* (pp. 7–8). Muscat, Oman: Al Falaj Press.
- Ryan, A., & Meara, P. (1991). The case of the invisible vowels: Arabic speakers reading English words. *Reading in a Foreign Language*, 7, 531–540.
- Schmitt, N. (2000). *Vocabulary in language teaching*. Cambridge, UK: Cambridge University Press.
- Schmitt, N., Schmitt, D., & Clapham, C. (2001). Developing and exploring the behaviour of two new versions of the vocabulary levels test. *Language Testing*, 18, 55–88.
- Segalowitz, N., & Segalowitz, S. (1993). Skilled performance, practice, and the differentiation of speed-up from automatization effects: Evidence from second language word. *Applied Psycholinguistics*, 14, 369–385.
- Segalowitz, S. J., Segalowitz, N. S., & Wood, A. G. (1998). Assessing the development of automaticity in second language word recognition. *Applied Psycholinguistics*, 19, 53–67.
- Waring, R., & Nation, I. S. P. (2004). Second language reading and incidental vocabulary learning. *Angles on the English Speaking World*, 4, 97–110.
- Willis, D. (1990). *The lexical syllabus: A new approach to language teaching*. London, UK: Collins COBUILD.
- Yushau, B., & Omar, M. H. (2007). Preparatory year program courses as predictors of first calculus course grade. *Mathematics and Computer Education*, 41(2), 92–108.

	Word	Non-word
<i>Yes</i>	HIT	False Alarm
<i>No</i>	Miss	CORRECT REJECTION

Figure 1