

12

How Much Lexis is Necessary for Reading Comprehension?

BATIA LAUFER

1. BACKGROUND

In *Reading in a foreign language: a reading problem or a language problem?* Alderson (1984) ascribes poor reading in L2 to four possible causes: (a) poor reading ability in the first language; (b) inadequate knowledge of the foreign language; (c) incorrect strategies for reading in the foreign language; (d) reading strategies in the first language not being employed in the foreign language, due to inadequate knowledge of the foreign language.

After surveying the empirical evidence for each one of the above hypotheses, Alderson's conclusion is that reading in L2 is both a reading problem and a language problem, but firmer evidence supports the latter hypothesis, namely that 'some sort of threshold or language competence ceiling has to be attained before existing abilities in the first language can begin to transfer' (Alderson, 1984: 20).

Another important question is what the nature of this threshold is. According to Ulijn (1984), Ostyn and Godin (1985), and Ostyn, Vandecasteele, Deville and Kelly (1985), reading in L2 depends mainly on the meaning of words and the knowledge of the subject matter. Laufer and Sim (1985a, 1985b) found that, in interpretation of texts, learners relied on word meaning first, then on their knowledge of the subject and least of all on syntax. Thus, it seems that the nature of the language threshold for reading purposes is largely lexical.¹ This conclusion is insufficient for practical purposes such as syllabus design, teaching and testing, since it does not specify the extent of the lexical knowledge necessary for reading comprehension. Nation and Coady (1987) suggest that, in considering the optimal ratio of unknown to known words in a text which is to be read, 'it is still wise to follow the guideline suggested by West (1941: 24) of a ratio of no more than one unknown word to fifty known words (2 per cent).' Studies by Deville (1985) and

Laufer (1989) claim that reading comprehension at an academic level requires 95 per cent lexical coverage, i.e. the knowledge of 95 per cent of word tokens in a given text. Yet a more important finding would be the number of words the reader must possess in his lexicon to be able to read in L2, namely the number of words constituting the threshold vocabulary which will ensure the transfer of reading skills from the L1 to the L2.² This is the issue addressed in the present study.

2. THE STUDY

2.1. Research questions

The purpose of the study was to explore the relationship between passive vocabulary size and the comprehension of academic texts. Specifically, the following questions were addressed:

- How do vocabulary size scores correlate with reading comprehension scores?
- What is the threshold level, i.e. the lowest vocabulary size level at which there are more 'readers' than 'non-readers'?'³
- How does an increase in vocabulary size (above the threshold level) affect the increase in reading scores?

2.2. Subjects

Ninety-two subjects took part in the study. They were first-year university students in the Faculty of Arts and Social Sciences, who were taking a course in English for Academic Purposes. The aim of the course was to improve the students' reading comprehension in English. The mother tongue of the majority of learners was Hebrew; some were native speakers of Arabic. Their English level was a near-equivalent of the Cambridge First Certificate of English.⁴

2.3. Procedure

As the study was aimed at exploring the relationship between vocabulary size and reading, two scores were calculated for each student: a reading score and a vocabulary size score.

Reading comprehension was measured by two standardised reading tests (one with a reliability of 0.81, the other 0.92). Sixty students received one test – the reading component of *Examen Hoger Algemeen Vortgezet Onderwijs* (produced by CITO in the Netherlands); this consisted of two texts and 20 multiple-choice comprehension questions. Thirty-two students received the English subtest of the Israeli university psychometric entrance test, which included texts with 40 multiple-choice comprehension items. Vocabulary level was measured by two vocabulary size tests. Sixty-two students took the Vocabulary Levels Test (Nation, 1983b) and 32 the Eurocentres Vocabulary Test (Meara and Jones, 1989).

The reading scores were arranged in five groups, on the basis of the learners' vocabulary scores. One group consisted of the reading scores of students whose vocabulary level was below 2,000; the second one was of those whose level was 2,000; the third group of reading scores comprised students whose vocabulary level was 3,000; the fourth, 4,000; the fifth, 5,000. This arrangement was necessary in order to find (a) the vocabulary level (of the five groups) at which there were more 'readers' than 'non-readers' (for which a chi square test was used), and (b) the vocabulary levels between which the differences between the reading scores were significantly different (for which ANOVA was used). Additional statistical procedures were: Pearson product moment correlation between the reading scores and the vocabulary size scores, and linear regression in order to predict the reading scores from the vocabulary size scores.

2.4. Results

Highly significant correlations were found between reading and vocabulary scores: 0.5 ($p < 0.0001$) in the case of the Vocabulary Levels Test and 0.75 ($p < 0.0001$) in the case of the Eurocentres Vocabulary Test.

The minimal vocabulary level at which there were significantly more 'readers', i.e. learners who scored 56 per cent and above on the reading test, was 3,000 for both the Vocabulary Levels Test (chi square = 4.08, $p < 0.05$) and for the Eurocentres Vocabulary Test (chi square = 4.20, $p < 0.05$).

The difference between the reading means of the different vocabulary levels (less than 2,000, 3,000, 4,000, 5,000) was found significant ($F = 5.97$, $p < 0.0005$ for the Vocabulary Levels Test; $F =$

14.42, $p < 0.0001$ for the Eurocentres Vocabulary Test). As for the comparison between individual vocabulary level groups, the most significant difference was between the reading scores of the 2,000 and 3,000 vocabulary level groups. There were differences between the reading scores of the 3,000 and 4,000 levels; 4,000 and 5,000; 3,000 and 5,000. None of these differences, however, was significant.

Regression analysis showed that a 3,000 vocabulary level would predict a reading score of 56 per cent (intercept = 35, slope = 0.007, standard error = 4.73; thus $35 + (0.007 \times 3,000) = 56$). A 4,000 vocabulary level would result in a reading score which is 7 points higher, i.e. 63 per cent; 5,000 words would yield a reading score of 70 per cent.

3. DISCUSSION

Since correlations between the two factors do not imply a causal relationship between them, the results in this study (correlations between reading and vocabulary scores) do not automatically mean that a poor or good reading ability is a direct result of low or high vocabulary size. There are obviously other factors which contribute to reading comprehension. But by measuring one variable, vocabulary size, we can predict the reading score.

This was indeed the conclusion reached by some studies on reading in the native language. Davis (1968, 1972), Thorndike (1973) and Spearritt (1972) found that of all the factors contributing to the ability to read, vocabulary was the most important and had the strongest effect. In L2 research, Meara and Buxton (1987) found that the scores on the Eurocentres Vocabulary Test produced very high correlations with an extended test of overall ability in EFL. The results of the present study support the claim that vocabulary is a good predictor of the reading comprehension level in a foreign language.

The results of chi square and ANOVA tests suggest that the minimal number of words constituting the lexical threshold is 3,000. In this study, Nation's definition of a 'word' was adopted: a word was taken to be a word family. This means that the knowledge of *observe*, for example, subsumes the knowledge of *observation*, *observable*, *observant*, *observance* and all their inflections. Yet different meanings of the same form, e.g. *pupil*, would be different

'words'. If we represented the same number (3,000 word families) in terms of lexical items, or dictionary entries, the result would be $3,000 \times 1.6 = 4,800$ (for the conversion formula see Nation, 1983b). Thus, the level at which the number of 'readers' was significantly higher than the number of 'non-readers' was approximately 5,000 lexical items. Below this level, there was a higher number of 'non-readers'. Unlike the chi square test, which compared the number of below-56-per-cents and above-56-per-cents on the reading test, ANOVA provided a comparison of the actual reading scores across the different vocabulary levels. These scores became significantly higher at the transition from the 2,000 to the 3,000 vocabulary level. The reading scores of the 4,000-level learners were higher than those of the 3,000-level learners, but not significantly so. The same was true for the 5,000 level *vis-à-vis* 4,000 and for the 5,000 *vis-à-vis* 3,000. The differences in the reading scores were not significant. This suggests that the turning point of vocabulary size for reading comprehension is 3,000 word families.⁵ In other words, good L1 readers can be expected to transfer their reading strategies to L2 when they have reached this level. Until then, reading in L2 will be hampered by an insufficient knowledge of vocabulary.

A possible counter-argument to this conclusion would be that the study examined the knowledge of general vocabulary only, since the vocabulary size tests are based on frequency lists; in the reading of academic texts, however, the learner relies heavily on the technical vocabulary of the particular discipline the text deals with, and such knowledge may compensate for a general vocabulary below 3,000 word families. However, while it is true that the knowledge of technical vocabulary is helpful, its value should not be overestimated: empirical evidence (Cohen, Glasman, Rosenbaum-Cohen, Ferrara and Fine, 1979) showed that it was the general not the technical vocabulary that was most problematic for learners.

The results pertaining to the third research question (how an increase in vocabulary size affects the reading scores) should be interpreted with caution. The regression analysis showed that an increase in 1,000 words would result in an increase of 7 per cent on a comprehension test. Thus the knowledge of 3,000 word families would result in a reading score of 56 per cent 4,000 in 63 per cent 5,000 in 70 per cent 6,000 in 76 per cent and so on. These figures are correct if the relationship between reading and vocabulary size is always linear. It is possible, however, that when the learner

reaches a certain vocabulary level (most probably an advanced one of about 7,000), the improvement of the reading score will decrease and finally level off.

Even if the results are not conclusive for all vocabulary levels, they do provide a general idea of how reading progresses above the threshold level of 3,000 word families. This might have a practical implication for teachers and syllabus designers. If the optimal reading level is considered to be, for example, 70 per cent then the vocabulary size to aim at will be 5,000 word families; if 63 per cent is taken as a passing grade, then 4,000 will suffice.

Notes

1. Judging by the above-mentioned studies, this seems to be the case with at least adult learners who have to read argumentative prose.
2. Deville (1985) and Laufer (1989) suggest that the knowledge of 5,000 lexical items will result in 95 per cent of lexical coverage. But their studies did not investigate the relationship between the number of words known to the learner and his reading competence, and it is precisely this kind of investigation that is crucial to the understanding of an L2 lexical threshold.
3. In this study, a subject was defined as 'reader' if his score was 56 per cent and above (56 is the minimal passing grade in our EFL courses). It was found that students with this passing grade were able to read satisfactorily the bibliography in their discipline.
4. This is the level of the Israeli Matriculation Examination in English which is taken at the end of high school.
5. The Eurocentres test does not claim to be based on word families, yet it yields results similar to those of the Vocabulary Levels Test. Several explanations are possible: (a) the Eurocentres test regards several meanings of one word form, e.g. *fair*, as belonging to the same word; so while on one test the increase beyond the forms tested is in terms of derivatives, in the other test it is in terms of additional word meanings; (b) in comprehension, it may not matter very much whether one word of a family, or the entire family is 'known'. It can be assumed that, in most cases, if the basic word is familiar, then the regular derivatives will be understood too, provided the morphological rules are known to the learner.

References

- Alderson, J.C. (1984). 'Reading in a foreign language: a reading problem or a language problem?', in Alderson, J.C. and Urquhart, A.H. (eds), *Reading in a Foreign Language*. London: Longman, pp. 1-27.

- Cohen, A.D., Glasman, B., Rosenbaum-Cohen, P.R., Ferrara, J. and Fine, J. (1979). 'Reading English for specialized purposes: discourse analysis and the use of student informants', *TESOL Quarterly*, Vol. 13, pp. 551-64.
- Davis, F.B. (1968). 'Research in comprehension in reading', *Reading Research Quarterly*, Vol. 4, pp. 499-545 (quoted in Nation and Coady, 1987).
- Davis, F.B. (1972). 'Psychometric research on comprehension in reading', *Reading Research Quarterly*, Vol. 7, pp. 628-78 (quoted in Nation and Coady, 1987).
- Deville, G. (1985). *Measuring a FL learner's lexical needs*. Paper presented at the 5th LSP Symposium, Leuven, Belgium.
- Laufer, B. (1989). 'What percentage of lexis is essential for comprehension', in Lauren, C. and Nordman, M. (eds), *From Humans Thinking to Thinking Machines*. Clevedon: Multilingual Matters, pp. 316-23.
- Laufer, B. and Sim, D.D. (1985a). 'Taking the easy way out: non-use and misuse of contextual clues in EFL reading comprehension', *English Teaching Forum*, Vol. 23, No. 2, pp. 7-10.
- Laufer, B. and Sim, D.D. (1985b). 'Measuring and explaining the threshold needed for English for Academic Purposes texts', *Foreign Language Annals*, Vol. 18, pp. 405-13.
- Meara, P. and Buxton, B. (1987). 'An alternative to multiple choice vocabulary tests', *Language Testing*, Vol. 4, pp. 142-54.
- Meara, P. and Jones, G. (1989). *Eurocentres Vocabulary Test 10 KA*. Zurich: Eurocentres.
- Nation, P. (1983a). 'Testing and teaching vocabulary', *Guidelines 5 (RELC supplement)*, pp. 12-24.
- Nation, P. (1983b). *Learning and Teaching Vocabulary*. Wellington, New Zealand: Victoria University.
- Nation, P. and Coady, J. (1987). 'Vocabulary and reading', in Carter, R. and McCarthy, M. (eds), *Vocabulary and Language Teaching*. London: Longman, pp. 97-110.
- Ostyn, P. and Godin, P. (1985). 'RALEX: an alternative approach to language teaching', *Modern Language Journal*, Vol. 69, pp. 346-53.
- Ostyn, P., Vandecasteele, M., Deville, G. and Kelly, P. (1985). 'Towards an optimal programme of FL vocabulary acquisition', in Cornu, A.M., Vanparijs, J., Delaheyne, M. and Baten, L. (eds), *Beads or Bracelet? How do we approach LSP?* Oxford University Press, pp. 292-305.
- Spearritt, D. (1972). 'Identification of subskills of reading comprehension by maximum likelihood factor analysis', *Reading Research Quarterly*, Vol. 8, pp. 92-11 (quoted in Nation and Coady, 1987).
- Thorndike, E.L. (1973). 'Reading as reasoning', *Reading Research Quarterly*, Vol. 9, pp. 135-47 (quoted in Nation and Coady, 1987).
- Ulijn, J.M. (1984). 'Reading for professional purposes: psycholinguistic evidence in a cross-linguistic perspective', in Pugh, A.K. and Ulijn, J.M. (eds), *Reading for Professional Purposes*. London: Heinemann, pp. 66-81.
- West, M.P. (1941). *Learning to Read in a Foreign Language*. London: Longman (quoted in Nation and Coady, 1987).

Objective Lexical and Grammatical Characteristics of L2 Written Compositions and the Validity of Separate-Component Tests

PIERRE J.L. ARNAUD

1. INTRODUCTION

Discrete-item tests of separate components of language went out of fashion in the 1970s with the trend towards communicative assessment. They were, however, summarily condemned and, in Farhady's (1983) words, '... the discrete-point approach became the object of serious and sometimes unreasonable attacks from scholars in the field.' This does not mean that they are entirely devoid of practical or theoretical interest, and the present study is an attempt at validating discrete-item tests of vocabulary and grammar against objective characteristics of English compositions written by French learners.

A methodological remark is necessary at the outset. In a validation study, the criteria against which the tests are being validated should be valid themselves, otherwise one is drawn into a vicious circle. Circularity of argument cannot be avoided altogether here, however, since research on the qualities of compositions has not solved all problems or produced unquestionable measures to date. Nevertheless, reasonably high correlations between the tests and the objective composition measures can be considered as evidence that the tests provide at least a moderate, but useful, prediction of the productive behaviour of learners in a communicative situation. The words *reasonably high* were chosen to underline the fact that whereas a test represents an *ad hoc*, concentrated measurement situation, a composition cannot provide very reliable measures and