University of
Reading

Word frequency and trends in the development of French vocabulary in lower intermediate students during Year 12 in English schools

Article

Accepted Version

It is advisable to refer to the publisher's version if you intend to cite from the work.

www.reading.ac.uk/centaur

CentAUR

Central Archive at the University of Reading

Reading's research outputs online

Word frequency and trends in the development of French

vocabulary in lower intermediate students

during Year 12 in English schools

BRIAN RICHARDS, DAVID MALVERN

and

SUZANNE GRAHAM

Institute of Education

University of Reading

Address for Correspondence.

Brian Richards

University of Reading

Institute of Education

Earley, Reading, RG6 1HY, UK

e-mail: b.j.richards@reading.ac.uk

ABSTRACT

This article reports research into the development of vocabulary in lower intermediate

level learners of French as a foreign language in Year 12 in 20 schools in the south of

England. The focus of attention is the role of less common, or non-basic, vocabulary

at a stage in students' learning when they have just moved to a much more advanced

syllabus that makes particular demands on their vocabulary knowledge. Results from

a test of receptive vocabulary and analyses of learners' writing confirms that the

acquisition of less common vocabulary makes a major contribution to students'

progress during their Year 12 course and that it is an important factor in individual

differences in overall achievement. Evidence from longitudinal data shows that

students make significant progress on both receptive and productive vocabulary over a

period of two terms, but that, in spite of this, students' perception of their own

vocabulary knowledge declines between GSCE and AS Level, a phenomenon

attributable to the discontinuity between these two examination courses.

INTRODUCTION

A recent survey of attitudes to reading among languages undergraduates in the UK (Gallagher-Brett, 2006) unequivocally identified vocabulary as the major cause of difficulty in reading in a foreign language. The survey uncovered "a sense of frustration about vocabulary" (p.17), and vocabulary was mentioned more than three times as often as its closest rival (grammar) as causing problems. These tendencies were particularly marked in those students who had most recently left school. Vocabulary development is widely acknowledged to be a key factor in successful language learning (e.g. Harley, 1995); according to a recent overview by Daller, Milton and Treffers-Daller (2007, p.8) "Vocabulary is now considered to be integral to just about every aspect of language knowledge". Moreover, it is also a key factor in unsuccessful learning. In one study of secondary school learners in New Zealand (McLauchlan, 2006, cited in McLauchlan, 2007), lack of progress in vocabulary was the most frequently cited reason for discontinuing language study.

Evidence relating to the UK suggests that pupils in secondary schools lag behind their counterparts on the European mainland in vocabulary size and rate of vocabulary development, even when inequalities in lesson time are taken into account. Using a test of receptive vocabulary, Milton (2006a) has proposed that, on average, pupils learn 170 words per year during the first five years of secondary schooling, that is, from Years 7 to 11, at which point they take the General Certificate of Education Examination (GCSE) at the age of 16. In the two years following the GCSE, where students of French are working at lower-intermediate level [1], the number of words learnt increases to 535 per annum. Research into vocabulary development at the lower intermediate level (Years 12 and 13) is sparse, but in a large-scale study of post-compulsory language learning in England, Graham (1997) found that pupils starting

4

courses in Year 12 tended to be overwhelmed by the extent of new vocabulary they were faced with. This is consistent with Milton's (2006a) results. It was also significant that these students did not have access to a wide range of strategies for dealing with vocabulary, a point that is echoed in Gallagher-Brett's (2006) survey.

As Milton (2006a) has pointed out, our knowledge of progress in vocabulary learning by pupils at UK schools is very limited compared with what we know about learners of English as a foreign language, and this is particularly true of post-GCSE students. In this article we report findings from three data sets collected from Year 12 students: a small pilot study in a selective (grammar) school; a one-year study of four non-selective (comprehensive) schools; and a quasi-experimental study conducted in fifteen comprehensive schools. In a series of analyses (Graham, Richards & Malvern, forthcoming; Malvern & Richards, 2006a, 2006b, forthcoming; Richards & Malvern, 2005, 2007), we have explored the rate and nature of vocabulary development during Year 12 and compared it with other areas such as listening and writing. Here, we present some new analyses that focus on the relationship between word frequency and vocabulary learning and on the relationship between students' actual progress and developments in their perceived vocabulary progress.

WORD FREQUENCY

The ability to display a lexical command that goes beyond a basic core of everyday native speaker vocabulary is frequently regarded as an indicator of proficiency in a speaker or writer. The use of lower frequency vocabulary, or less common words, can be regarded as one dimension of lexical richness that is often referred to as "lexical sophistication" (e.g. Read, 2000). Some notion of lexical sophistication is often incorporated either implicitly or explicitly into analytic or holist rating scales for essays and oral interviews under more general descriptors such as "lexical command",

"breadth of language", "adequacy of vocabulary" (e.g. Chambers & Richards, 1992) or "sophisticated range" versus "adequate" or "limited range" (Jacobs et al., 1981, quoted in Nation, 2007, p.43).

For first language development, the importance of growing up in a linguistic environment that is rich in low frequency vocabulary has been highlighted by an ongoing series of investigations conducted in the United States as part of the "Home-school study of language and literacy development" (Dickinson & Tabors, 2001), a long-term longitudinal project into children from low-income families. Results suggest that early exposure to "rare vocabulary" in the home and pre-school predict later language and literacy levels and help to develop useful tools for educational achievement such as ability with formal definitions. For children acquiring their first language, in both English-speaking and certain other European contexts, command of less frequent lexis gives access to words of classical (rather than Germanic) origin that are particularly important in educational contexts. For both first and second language learners, however, words from lower down the frequency continuum are respected inter alia as having greater precision (Read, 2000), giving access to technical terminology, being more literary, abstract and with more complex meanings (Weizman & Snow, 2001), as well as being longer (Zipf, 1949), harder to spell, having a more complex derivational morphology, learned later and perceived as being more difficult by learners.

In addition to the more subjective assessment of rare vocabulary in learners' language production referred to above, there have been many attempts to conduct assessment objectively by subjecting digital files of students' writing or transcripts of speaking to automated word counts that assign each lexical item to the category "basic" or "non-basic" as defined by frequency of usage in large language corpora,

usually from native speakers, and then computing some kind of simple proportion, for example, the proportion of different words that are rare (e.g. Snow, Tabors & Dickinson, 2001). A more sophisticated approach categorizes each word according to the frequency band it belongs to, as for example in Laufer and Nation's (1995) Lexical Frequency Profile.

Just as rare words are often assessed as part of *productive* vocabulary, word frequency also underpins tests of *receptive* vocabulary, including tests of vocabulary size, by ensuring that vocabulary items are included from a range of frequency bands such as the thousand most common words in the language, the second thousand, the next thousand, and so on (see Daller, Milton & Treffers-Daller, 2007, for examples). There is good evidence that this approach is valid, in that learners do indeed tend to learn the higher frequency words earlier as their lexical knowledge moves from the easy to the more difficult (Milton, 2007). Not only are the total scores a useful reflection of vocabulary size, but the individual scores across the frequency band can provide a useful diagnostic profile. Nevertheless, as Milton points out, "frequency information drawn from a wide variety of native speaker sources may not be relevant to foreign language learners who….only have textbooks to draw on" (Milton, 2007, p.49), especially as textbooks tend to be structured thematically. This mode of testing may therefore be inappropriate during the early stages of learning when pupils' limited range of vocabulary is difficult to sample. In addition, it goes without saying that word frequency is on a contiunuum, and that the studies of "rare vocabulary" or non-basic, i.e. "advanced" or "sophisticated" words quoted above have drawn such distinctions at different, and fairly arbitrary, points on the continuum. It may well be the case that tests of vocabulary knowledge based on such cut-off points are unreliable with beginners or lower intermediate learners simply because their vocabulary has not

yet extended sufficiently beyond a basic, high frequency core. It remains to be seen whether an adjustment of the cut-off between "basic" and "advanced" towards the more frequent end of the spectrum would provide a reliable basis for testing.

RESEARCH QUESTIONS AND PILOT STUDY

The above issues were explored in a series of related studies that focused on Year 12 learners of French in English secondary schools. The beginning of Year 12 is a crucial transition point. It marks the beginning of non-compulsory education, and of specialization in a smaller number of school subjects. Foreign language study moves from a basic survival syllabus to much more demanding and complex topics that require rapid learning of a large, more difficult and more abstract vocabulary. These factors and the issues raised in the foregoing literature review prompted the following questions:

1. to what extent is the receptive and productive vocabulary knowledge of Year 12 students related to word frequency as defined by native speaker corpora?

2. assuming that students make measurable progress in vocabulary proficiency during Year 12, to what extent does this progress depend on acquiring lower frequency or non-basic lexis?

3. given reports of students feeling overwhelmed by the quantity of vocabulary to be learned in Year 12 (Graham, 1997), how does the more objective measurement of students' vocabulary progress compare with their own perceptions?

For the assessment of receptive vocabulary, the intention was to use the computer version of X_Lex (v. 2.02), a test developed from native speaker frequency counts by Meara and Milton (2003). X_Lex is a "yes/no test" (Meara & Buxton, 1987) that samples vocabulary knowledge from five frequency bands (1K to 5K). Test-takers

have to click one of two buttons to state whether or not they understand the meaning of 120 words that appear on the screen (see Fig. 1). In order to guard against respondents over-stating their knowledge, 20 items do not actually exist in French, but look plausible (e.g. "clabrer", "muce"). Thus, the software presents 20 words from each frequency band plus 20 nonsense words. Respondents who claim to know the meaning of the latter are heavily penalized, to the extent that for each accepted pseudo-word the total score is adjusted downwards by 250 points, while each accepted real word scores 50 points.

Insert Fig. 1 about here

Because we were not entirely confident that the difficulty level was appropriate for Year 12 and did not want to risk students being demoralized, we piloted X_Lex on a small group of Year 12 students attending a grammar school who volunteered to try out the software during their lunch break. The test worked well, in as far as the twelve volunteers were highly motivated by the computerized format and were eager to have a second attempt, exploring the consequences of taking risks and trying different parallel versions of the vocabulary items. The test was efficient to administer with no misunderstandings of what was required and could be completed in between four and seven minutes. One concern had been that students' scores would be distorted by guessing, that is to say their raw scores would be drastically adjusted downwards as a result of claiming to know the nonsense words. As Milton (2007) points out in a discussion of Greek learners' results on the English version of the test, too much guesswork can seriously endanger the validity of the results. In fact, one student in the pilot study did guess extensively, claiming to know eight of the nonsense words. Nevertheless, the other 11 students did not rely on guessing, making, on average, less than one "error" out of a possible 20. Final adjusted scores (i.e. taking

guessing into account) were surprisingly high for students who would be regarded as being at only "lower intermediate level". The adjusted score across all 12 learners ranged from 1,900 to 4,150 with a mean of 3,129. An analysis of scores out of 20 across the five frequency bands showed a clear and statistically significant trend downwards as the words became less frequent (see Richards & Malvern, 2007), indicating that the results were sensitive to word frequency.

STUDY I

To discover whether the results of the pilot study were valid in a more representative sample of students, and to test whether differences in students' productive vocabulary also showed effects of word frequency, we collected data from twenty-three Year 12 students attending four comprehensive schools encompassing Year 11 to Year 18 in the Oxford and Reading areas. Like most students learning French in Year 12, they were following the AS (Advanced Supplementary) Level course, and, because they were tested towards the end of the school year (4 months later than the grammar school pupils), had been studying French for nearly 6 years.

Method

Receptive vocabulary was assessed by X_Lex, and productive vocabulary was assessed through a narrative, to be written in 30 minutes, that described the sequence of events in six pictures [2] (see the Appendix for an example of writing produced in response). Vocabulary in the essays were scored in two ways. First, vocabulary was rated on a 15 point scale for appropriate usage and effective range and register (see Graham et al., forthcoming). This was carried out independently by two experienced experts. Inter-rater agreement was very high ($r = .86$; $p < .001$) and the two sets of scores were averaged to provide the final mark. Second, all the writing samples were transcribed into computer analyzable format and processed through the CLAN

(Computerized Language Analysis) software (MacWhinney, 2000) to obtain objective vocabulary measures based on the total number of words, number of *different* words, and the number of rare words. File contents were tagged to ensure that only valid words were entered into the analyses. This meant that all of the following were excluded: illegible or unintelligible words, words used inappropriately, proper names that are the same in English, words from languages other than French, numbers written as digits. Files were also transcribed so that the identification of *different* words was based on the dictionary headword rather than their various inflected forms. Thus "jouer", "joue", "jouons", etc. were treated at one word type.

"Rare" words (i.e. words that extended beyond a basic vocabulary) were identified by reference to Le Français fondamental 1er degré (FF1), the list of core French vocabulary developed by Gougenheim, Rivenc, Michéa, and Sauvageot (1964). Although FF1 was derived from oral interviews and conversation rather than writing, it was considered to be appropriate here because of its high correspondence with GCSE vocabulary. Nevertheless, we did have to update the list to include some words commonly used by the students such as "super". A computer file of the final list of 1,370 basic words was created so that the process of identifying lower frequency or non-basic words could be fully automated. Finally, we used the outputs from the CLAN software to compute a measure of the diversity of non-basic words. Space does not allow a full account of this measure here, but it builds on previous work by Malvern and Richards on the measurement of vocabulary diversity using mathematical modelling techniques (Malvern, Richards, Chipere & Durán, 2004) by subtracting a value for the diversity of FF1 words from the diversity of all vocabulary. Thus, the measure depends on a simple two-way comparison between basic and non-basic words (see Malvern & Richards, 2006a, 2006b, forthcoming).

Results: X_Lex

Twenty-three students were tested for receptive vocabulary. As in the pilot study with grammar school pupils, the comprehensive school pupils responded well to X_Lex and their guessing levels were acceptable. The mean adjusted score was 2,437 (range 1250-3500) which is substantially lower than for the grammar school (3,129) even though the comprehensive school pupils were tested four months further into their course. The difference between the means is statistically highly significant ($t = 2.93$, d.f. = 16.9, $p = .009$).

There is clear evidence that vocabulary knowledge at this stage *is* related to word frequency. Average raw scores out of 20 across the frequency bands show a gradual decline from 18.0 (1K), through 13.6 (2K), 9.9 (3K), 9.3 (4K) to 8.5 (5K). This is a highly significant downward group trend. that can be fitted by a linear ($r^2 = .87$) or logarithmic ($r^2 = .97$) function. Moreover, individual curve-fitting for each student showed that the trend was statistically significant in 15 out of 23 cases (r^2 range = .55-.96). The reliability of the total score is very high (Cronbach's alpha = .92) although the 1K band is the least reliable, presumably because scores on this, more basic, vocabulary are approaching ceiling.

A question of particular interest is the relative importance of each frequency band in relation to individual differences in learners' overall language performance. We explored this question by computing rank order correlations between frequency band scores and learners' French GCSE grades (from a year earlier) and their predicted grades for AS Level. (see Table 1). What is interesting here is the contrast between the correlations with the total adjusted X_Lex score in the first row of the table, and the pattern for the separate frequency bands in the following five rows. Whereas total scores are equally related to GCSE and predicted AS grades, it is only

12

the 1K band that is related to GCSE performance. In other words, knowledge of high frequency words is strongly related to GCSE grade. By contrast, predicted AS results are significantly associated with all bands but relatively weakly for the high frequency lexis. The strongest relationships are for the *low* frequency words in the 4K and 5K bands, particularly at 4K. This changing pattern neatly reflects the shift in importance from high frequency to low frequency vocabulary in explaining individual differences in achievement during Year 12.

Insert Table 1 about here

Results: written narratives

On average, the pupils wrote 162 words (range 83-350) in 30 minutes. Of these, 144 (88%) were valid French words used correctly and appropriately according to the criteria above. The number of different words per student was 68 (range 37-121). The average number of non-basic words was 11 (range 3-33), and the number of *different* non-basic words was 9 (range 3-26). There is clearly huge variation between students on all these variables, including how much they wrote. Although all students used some low frequency lexis, this ranged from minimal to evidence of more extensive advanced vocabulary.

Given the degree of individual differences between students, it is useful to examine how measures of productive vocabulary in writing, and particularly overall lexical diversity and diversity of non-basic words, relate to the variation in GCSE and predicted AS grades investigated above. The results are shown in Table 2, and to put these in perspective, we also include the results relating to the proportion of words in each essay that were valid (a proxy index of accuracy in vocabulary usage) as well as raters' subjective assessment of vocabulary. All relationships are statistically significant; both lexical diversity and diversity of non-basic words are related to both

13

GCSE and predicted AS grades, and the relationship between rare words and the latter is the stronger, which is consistent with the results for X_Lex. On the other hand these correlations are relatively weak when compared with raters' overall impression of vocabulary and with accuracy and appropriateness of usage.

Insert Table 2 about here

STUDY II

This far, we have shown how the level of achievement of Year 12 students attending a number of different schools relates to their knowledge of vocabulary from different frequency bands. Data from two different samples of students have provided snapshots of performance at two points in the school year. What has been lacking so far, however, is a longitudinal dimension that allows analyses to be related to the progress of individual pupils during the course of Year 12. In this section, therefore, we introduce data from a much larger sample of 150 Year 12 pupils attending fifteen comprehensive schools in England whose performance was assessed at the beginning of the academic year and two school terms later (henceforth pre-test and post-test). Students were tested on X_Lex and completed a picture narrative at both time points, with a different but comparable set of pictures used at each point [2]. These were rated as in Study I. Raters assessed the essays after each wave of data collection and, as the sets of pictures were different, they would have been aware which had been completed at pre-test and post-test. However, as these procedures were conducted eight months apart, any influence of the first assessment on the second is likely to be minimal. In order to gauge participants' own sense of progress they were also asked to complete a self-assessment of their vocabulary knowledge on a four-point scale ("poor", "a bit weak", "good", "very good") at pre-test and post-test.

Results

As we have reported elsewhere (Graham et al., forthcoming), students made substantial and statistically significant progress in both receptive and productive vocabulary during the course of two terms: mean X_Lex adjusted scores rose from 2,282 to 2,700, and for the vocabulary ratings of the written narratives from 7.8 to 9.6 out of a possible 15. For both these variables there were highly significant correlations between their scores at pre-test and post-test.

With regard to the distribution of average scores across frequency bands, the expected fall as words become rarer can be seen clearly in Figure 2. This is particularly striking for the pre-test results, but curve-fitting shows that the downward trends are linear and statistically significant at both time points (pre-test: $r^2 = .91$, $p = .012$, d.f. = 3; post-test: $r^2 = .82$, $p = .034$, d.f. = 3).

Insert Fig. 2 about here

What is most interesting, however, is the *comparison* of the profiles at pre-test and post-test in Figure 2, which demonstrate a clear shift towards the rarer words. This pattern is illuminated by examining the mean values and differences in average score that are presented in Table 3. It can be seen that students make significant progress on all bands except 2K where there is, in fact, a slight, though non-significant fall. The largest gains, however, are to be found in the lower frequency bands as shown in column 4 of the table, with the largest of all being in the 4K band. It is possible that further learning in Year 13 will now shift the emphasis to the 5K band, although further research will be needed to verify this.

Insert Table 3 about here

In the light of the encouraging results reported above relating to the amount of progress made by the learners on productive and receptive vocabulary, and particularly with regard to the amount of more advanced vocabulary they have

successfully assimilated, we might expect the students to have a positive opinion of their vocabulary knowledge. However, the results for students' self-rating are initially somewhat puzzling because not only does this fail to increase, but it actually falls. At the beginning of the study both the median and mode rating for vocabulary knowledge were 3 out of a possible 4, and 75% of students rated themselves as "good" or "very good". By their third school term of more advanced French, however, both the mean and mode had fallen to 2 and less than 50% of students rated themselves as "good" or "very good". These differences are significant on a Wilcoxon signed ranks test ($z = 3.43$, $N = 74$, $p = .001$). Furthermore, in stark contrast to the results from X_Lex and the writing data, there is no relationship between the self-ratings at the two time points (Kendall's tau-b = .172, $N = 74$, $p = .106$) and they are not significantly related to the other variables in the study such as GCSE scores. This and the striking discrepancy between the researchers' assessment of vocabulary progress and pupils' own perceptions will be discussed below.

DISCUSSION

Previous investigations indicate that foreign language vocabulary is a problematic area for learners in the post-compulsory phase of British schools and for those continuing with their language study in higher education. Research by Milton (2006a) suggests that pupils in British schools lag behind their European counterparts even when differences in the time devoted to language study are taken into account. Graham's (1997) study presents a picture of students being overwhelmed by the vocabulary demands of Year 12, something that is given statistical support by Milton's (2006a) finding that more than three times the amount of vocabulary learning takes place per annum in Years 12 and 13 than in each of the previous years.

As noted in our introductory discussion of word frequency and its relationship with proficiency, a key issue in any language, and for both first and additional languages, is access to a wide range of vocabulary, and in particular a range of sophisticated, advanced, low frequency vocabulary that can be understood and used in appropriate contexts. One would predict that this would become increasingly important from Year 12 onwards as learners are exposed to more difficult texts on complex topics and are expected to achieve more advanced functions in language production. On the other hand, the move beyond a basic, core vocabulary rooted in a limited range of GCSE topics and geared to coping with situations encountered during a visit abroad may have been a very gradual one. Achieving lexis that could be defined as complex, advanced or sophisticated by reference to *native speaker* norms may lie a long way in the future. It was therefore not obvious that, at this stage in their learning, methods of assessing either receptive or productive vocabulary using either a range of frequency bands or the distinction between basic and non-basic lexis would be sensitive to advances in knowledge that depended on lower frequency words, nor that word frequency effects would be important in explaining individual differences in achievement. We set out to answer three questions concerning: the extent to which vocabulary knowledge was related to word frequency in Year 12; the extent to which *progress* in vocabulary knowledge depended on acquiring lower frequency or non-basic words; and how students' own perceptions compared with our own assessment of their progress.

Results from students in a total of 20 schools showed a consistent pattern. Even during their first term of more advanced French, the contribution of word frequency phenomena to individual differences in achievement was apparent, and receptive vocabulary data from four different points during the school year showed

significant trends across frequency bands. Two findings are illustrative of the shift in importance during Year 12 from basic, high frequency to lower frequency vocabulary. The first is that results from X_Lex show that it is only the highest frequency band that is related to individual differences in GCSE results, whereas the lower frequency bands are the ones that are most strongly correlated with predicted AS Level grades. Second, in the longitudinal data there is a clear shift in profile over the frequency bands towards increasing scores on low frequency vocabulary rather than the high frequency bands. Both these analyses suggest strongest effects in the 4K band, and one would predict that further study would later shift the emphasis into the 5K band.

There was wide variation on all the vocabulary variables obtained from the writing task. All students used non-basic words but for some this was minimal. By contrast, there were other students who demonstrated a command of a relatively wide range of more advanced vocabulary. These individual differences in the diversity of non-basic words were significantly, though relatively weakly, correlated with GCSE grades and slightly more strongly with predicted AS performance.

Finally, in spite of independent evidence that students did indeed make good progress on both receptive and productive vocabulary over a period of two school terms and that a major part of this progress was attributable to words of lower frequency, their self-assessment of their vocabulary knowledge actually fell. Following their GCSE, students were very positive about their standard of knowledge, but appear to become demoralized later. This seems to be a reflection of the lack of continuity between GCSE and AS Level—an abrupt transition from more basic transactional, survival language with a concrete lexis to more complex linguistic functions. Not only does this entail a dramatic increase in the quantity of vocabulary required, but also in its quality. As noted above, low frequency vocabulary tends to be

more difficult: words are among other things longer, harder to spell and to read, more complex in structure, more abstract, and convey more specific rather than general meanings.

The disjuncture that many pupils experience on continuing language study after GCSE has been a concern for two decades, but there is no sign that it has diminished. Graham (1997) points out that students have a lack of strategies for dealing with vocabulary, and the recent survey of undergraduates by Gallagher-Brett (2006) has as one of its four main recommendations "Raising awareness of strategies for dealing with vocabulary difficulties" (p.5). Whether this means strategies for learning vocabulary or coping strategies to compensate for a lack of vocabulary is not entirely clear, but obviously there is a necessity for both. Unfortunately, in spite of a revival of interest in vocabulary in the L2 field in general, this does not seem to have influenced modern foreign language teaching in UK schools. The selection of books published by CILT, the National Centre for Languages "the Government's recognised centre of expertise on languages" (www.cilt.org.uk/about.htm, 7/01/08) only appears to have one publication ("Doing it for themselves: focus on learning strategies and vocabulary building" by Harris & Snow, 2004) with vocabulary in the title or subtitle and the "vocabulary building" section of this book has been severely criticized by Milton, (2006b) for perpetuating myths about vocabulary learning. That is not to say that other CILT publications do not deal with vocabulary issues—the excellent resource file by Biriotti (2001) on nouns, gender and adjectives offers an attractive blend of enjoyment, structure, rigour and systematic learning, but this is geared towards younger learners and is not underpinned by theory. If the quotation at the beginning of this article that "Vocabulary is …. integral to just about every aspect of language knowledge" (Daller et al., 2007, p.8) is true, then this is an area that

deserves much more attention in both research and teaching in the future. Clearly, there is a lack of coherence between theory and practice. From our experience, it appears that insufficient emphasis is placed on vocabulary learning in initial teacher education and that not only teachers and teacher educators, but also the learners themselves need to be better informed about the successful acquisition of vocabulary.

Finally, it seems to us that the above has implications, not just for language teachers, but also for writers of textbooks and materials, and those who make policy related to curriculum and examinations. The fact that the students in our research could be shown objectively to make significant progress in vocabulary acquisition, and yet subjectively rate themselves more negatively over time draws attention to the consequences of the lack of continuity between GCSE and advanced level courses and is likely to be one of the causes of the high drop-out rate in Modern Languages following the AS examination. A recent comparison of learners of French in three European countries (UK, Greece and Spain) by Milton and Alexiou (forthcoming) suggests that the increase in vocabulary knowledge between GCSE and A Level is much greater for students in the UK than the equivalent transition on the Common European Framework for Languages for students elsewhere in Europe. This is at least partly due to much lower vocabulary knowledge in UK students at the age of 16 and suggests that vocabulary enrichment pre-GCSE would benefit pupils who progress to more advanced courses.

This problem has arguably been exacerbated by the introduction of the AS examination. Instead of a more gradual two-year development of knowledge and skill towards A Level, students are, at least in theory, expected to achieve a comparable standard in less than a year, given that the purported value of AS Level is half that of A Level. This notion may be ill-conceived for a hierarchical subject like language

learning. However, as there can be no guarantee that the examination system will change, at least there are steps that teachers and materials writers can take to motivate students and give them a greater sense of progress through a more graded approach to learning that is less influenced by the washback effect of the final examination. Our impression from observing in Year 12 classrooms is that a lot of time is spent practising advanced tests in areas like speaking and listening rather than using graded texts as vehicles for learning. Even with their own course books, learners are overwhelmed by the difficulty of the texts they are faced with, becoming frustrated by the number of words they have to look up in a dictionary. The density of unknown words is far greater than the recommended level of 95% to 98% of known words necessary for comprehension and learning (see Schmitt, 2008). The provision of carefully graded texts, therefore, seems essential.

In a recent review of research into instructed second language vocabulary learning, Schmitt (2008) sums up the most important principle as being "to increase the amount of engagement learners have with lexical items" (p. 329). This entails not simply repeated exposure to each item, but also intentional learning, explicit attention and manipulation. For teachers and materials writers this implies the systematic revisiting and recycling of lexis in order to maximize exposure, attention and use. Zhang (2008) has demonstrated the rate at which Chinese learners of English in senior high school forget vocabulary if teachers and textbooks fail to do this, although there are considerable individual differences between students. Learners, teachers and authors of materials need to be aware that learning vocabulary is more than knowing a one-to-one mapping between L1 and L2. Much attention has been paid to depth (as opposed to breadth) of knowledge. As Schmitt (2008) points out, there are many facets to this qualitative aspect of vocabulary knowledge that need to be taken account

of: "a learner must also know a great deal about each [lexical] item in order to use it well" (p. 333). During a period where vocabulary development depends to a large extent on lower frequency, often more difficult and more complex vocabulary the planned repeated exposure and engagement with items becomes even more important, especially where matters of culture and idiom make mappings between languages more complicated (see Macaro, Guo, Chen & Tain, forthcoming).

1. In Years 12 and 13, pupils in England generally prepare for the Advanced Supplementary and Advanced Level examinations respectively. In international terms, however, their level is best described as lower-intermediate. At the time of the studies described in this article, foreign language learning was compulsory for pupils up to GCSE level (now compulsory only up until age 14).

2. In Study I, the six pictures showed a family visiting a stately home and being given a guided tour. The son is very bored and has disappeared by the end of the tour. The parents find him lying on a four-poster bed reading a comic. In Study II the same pictures were used at post-test. At pre-test the pictures showed a grandfather taking his grandson out for the day. The grandfather drinks too much and ends up seeing four grandchildren and buying four tickets for the train journey home.

APPENDIX

Sample of writing at pre-test from a high-scoring student:

Un jour un grandpère est sorti avec son petit fils qui s'appellait Pière. Le grandpère habitait normalement dans une maison pour les vieux donc il était vraiment heureux d'être libre

      "Pière, je suis vraiment heureux à te vois" dit il

      "Les gens dans la maison sont tellement ennuyeux et.... et vieux! Moi, je ne sui pas du tout vieux. Regard, je n'ai pas de besoin de ma canne"

      Et il donnait sa canne au pière et commençait à courir ey ajouter

      "Eeeaa! Je suis libre, je suis libre!!! "

Pière regardait pour un moment la canne dans ses mains et après pense que peut être son grandpère serait mieux dans la maison. Il a décidé que c'était un bon chose que son grandpère était tellement heureux et il a lui suivi.

      Ensuite ils ont trouvé dans un champ une table avec un panneau qui dit "BAR" et un grand maître d'hotel sérieux qui se tiendit la derrière

      "Bonjour monsieur, qu'est-ce que vou voulez? "

dit il...

      Quelques heures suivant le grandpère étit endormi sous un arbre et Pière s'ennuyait.

      "Grandpère, " chucotait il "reveilles tu! "

Le grandpère se reveillait subitement

      "Hein?! Je suis où? Ah, Pière... je suis desolée. Maintenant on va au cinéma. "

      D'une manière ou d'une autre ils sont arrivés au cinéma. Le grandpère etait tellement ivre qu il a demande quatre billets pour Pière.

Sample of writing at post-test from the same high-scoring student:

Le samedi dernier le jeune garçon Pière a dû visiter un vieux château célèbre avec ses parents. Il ne voulait pas y aller, il voulait rester chez lui pour lire sa nouvelle BD mais ses parents lui ont dit qu'il devait visiter le chateau parce qu'il était mal en histoire et il devait améliorer ses notes au collège. Pière a accepté de y aller parce que ses parents ont ajouté que si il refusait, its cacheraient toutes ses BD et lui donneraient des choux trois fois par jour pendant un an. Alors, quand ils sont arrivés au château ils devaient attendre le guide pendant longtemps. Un autre garçon jouait avec le panneau ce qu'il Pière trovait marrant, mais quand le guide est arrivé elle s'est fâché et elle a demandé que le garçon et sa famille devraient partir.

Le guide a parlé depuis longtemps des vieux tableaux dans toutes les vieilles salles e Pière s'ennuyait. Il regardait le jardin par la fenêtre et il a vu un fantôme ce qu'il a trouvé très interessant pendant presque cinq minutes – et ensuite il s'ennuyait encore. Il a décidé de trouver une chambre vide où il pourrait lire sa BD en paix. Il se détendait sur un lit énorme jusque a ses parents se sont rendu compte que son fils avait disparu et l'ont cherché.

REFERENCES

Biriotti, L. (2001). *Getting the basics right: nouns, gender and adjectives*. London: CILT.

Chambers, F., & Richards, B. J. (1992). Criteria for oral assessment. *Language Learning Journal*, 6, 5-9.

Daller, H., Milton, J., & Treffers-Daller, J. (2007). Editors' introduction. In H. Daller, J. Milton, & J. Treffers-Daller (Eds.), *Modelling and assessing vocabulary knowledge* (pp. 79-92). Cambridge: Cambridge University Press.

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

Gallagher-Brett, A. (2006). *Hard going but worth it: a snapshot of attitudes to reading among languages undergraduates*. Southampton: Subject Centre for Languages, Linguistics and Area Studies, University of Southampton.

Gougenheim, G., Rivenc, P., Michéa, R., & Sauvageot, A. (1964). *L'élaboration du français fondamental (1er degré): étude sur l'établissement d'un vocabulaire et d'une grammaire de base* (2nd ed.). Paris: Didier.

Graham, S. (1997). *Effective language learning: positive strategies for advanced level language learning*. Clevedon: Multilingual Matters.

Graham, S., Richards, B. J., & Malvern, D. D. (forthcoming). Progress in learning French vocabulary in a one-year advanced course at school. *Journal of French Language Studies*.

Harley, B. (1995). *Lexical issues in language learning*. Amsterdam: Benjamins.

Harris, V., & Snow, S. (2004). *Doing it for themselves: focus on learning strategies and vocabulary building*. London: CILT, the National Centre for Languages.

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

Macaro, E., Guo, T., Chen, H., & Tian, L. (forthcoming). Can differntial processing of L2 vocabulary inform the debate on teacher codeswitching behaviour: the case of Chinese learners of English. In B. J. Richards, H. Daller, D. D. Malvern, P. Meara, J. Milton, & J. Treffers-Daller (Eds.), Vocabulary studies in first and second language acquisition: the interface between theory and application. Basingstoke: Palgrave MacMillan.

MacWhinney, B. (2000). The CHILDES project: Tools for analyzing talk (3rd ed., Vol. 1: Transcription format and programs). Mahwah, NJ: Erlbaum.

Malvern, D. D., & Richards, B. J. (2006a). Measuring the deployment of rare words in the writing of lower intermediate students of French as a foreign language. Paper presented at the ESRC Seminar on "Testing and teaching vocabulary in a second language setting", University of Swansea, June 2006.

Malvern, D. D., & Richards, B. J. (2006b). Extending the vocd progam of CHILDES to measuring the deployment of rare or infrequently used words in language assessment. Paper presented at the symposium on "Using CHILDES to support the French SLA research agenda" at the Joint AAAL/CAAL conference, Montreal, June 2006.

Malvern, D. D., & Richards, B. J. (forthcoming). Assessing the usage of low frequency lexical items in foreign language writing. In B. J. Richards, H. Daller, D. D. Malvern, P. Meara, J. Milton, & J. Treffers-Daller (Eds.), *Vocabulary studies in first and second language acquisition: the interface between theory and application*. Basingstoke: Palgrave Macmillan.

Malvern, D. D., Richards, B. J., Chipere, N., & Durán, P. (2004). *Lexical diversity and language development: quantification and assessment*. Basingstoke: Palgrave Macmillan.

McLauchlan, A. (2007). Bilingual dictionaries in written language examinations: why New Zealand needs to catch up with Australia. Available at http://www.encyclopedia.com/doc/1G1-165691970.html (accessed 21 December 2007).

Meara, P., & Buxton, B. (1987). An alternative to multiple choice vocabulary tests. *Language Testing*, 4, 142-154.

Meara, P. M., & Milton, J. L. (2003). XLex: Swansea Vocabulary Levels Test (Version 2.02) [Computer software]. Swansea: Lognostics.

Milton, J. (2006a). Language lite? Learning French vocabulary in school. *Journal of French Language Studies*, 16, 187-205.

Milton, J. (2006b). Review of K. S. Folse, Vocabulary myths: Applying second language research to classroom teaching. Ann Arbour: The University of Michigan Press, 2004; V. Harris & D. Snow. Doing it for themselves: focus on learning strategies and vocabulary building. London: CILT, 2004. *Language Learning Journal*, 33, 75-77.

Milton, J. (2007). Lexical profiles, learning styles and the construct validity of lexical size tests. In H. Daller, J. Milton, & J. Treffers-Daller (Eds.), *Modelling and assessing vocabulary knowledge* (pp. 79-92). Cambridge: Cambridge University Press.

Milton, J., & Alexiou, T. (forthcoming). Vocabulary size and the Common European Framework for Languages. In B. J. Richards, H. Daller, D. D. Malvern, P. Meara, J. Milton, & J. Treffers-Daller (Eds.), *Vocabulary studies in first and second*

language acquisition: the interface between theory and application. Basingstoke: Palgrave MacMillan.

Nation, I. S. P. (2007). Fundamental issues in modelling and assessing vocabulary knowledge. In H. Daller, J. Milton, & J. Treffers-Daller (Eds.), *Modelling and assessing vocabulary knowledge* (pp. 35-43). Cambridge: Cambridge University Press.

Read, J. (2000). *Assessing vocabulary*. Cambridge: Cambridge University Press.

Richards, B. J., & Malvern, D. D. (2005). French vocabulary assessment at lower intermediate level. Paper presented at The 14th World Congress of Applied Linguistics, Madison, Wisconsin, July 2005.

Richards, B. J. and Malvern, D. D. (2007). Validity and threats to the validity of vocabulary assessment. In: H. Daller, J. Milton and J. Treffers-Daller (eds.), *Modelling and assessing vocabulary knowledge*. Cambridge: Cambridge University Press, pp. 79-92.

Schmitt, N. (2008). Instructed second language vocabulary learning. *Language Teaching Research*, **12**, 329-363.

Snow, C. E., Tabors, P. O., & Dickinson, D. K. (2001). Language development in the preschool years. In D. K. Dickinson & P. O. Tabors (Eds.), *Beginning literacy with language: Young children learning at home and at school* (pp. 1-25). Baltimore, MD: Paul H. Brookes Publishing.

Weizman, Z. O., & Snow, C. E. (2001). Lexical input as related to children's vocabulary acquisition: Effects of sophisticated exposure and support for meaning. *Developmental Psychology*, 37, 265-279.

Zhang, H. (2008). *An investigation into the forgetting of vocabulary by Chinese EFL learners in senior high school*. Unpublished MA Thesis, University of Reading, Reading, UK.

Zipf, G. K. (1949). *Human behavior and the principle of least effort*. Reading, MA: Addison Wesley.

Table 1: Spearman rank order coefficients and statistical significance of correlations

between X_Lex scores and GCSE and predicted AS grades (N = 23)

	GCSE Grade		Pred. AS Grade	
Adjusted total score	.43	(.020)	.43	(.020)
Raw score on 1st 1000	.55	(.004)	.35	(.049)
Raw score on 2nd 1000	.29	(ns)	.43	(.021)
Raw score on 3rd1000	.32	(ns)	.35	(.050)
Raw score on 4th 1000	.32	(ns)	.58	(.002)
Raw score on 5th 1000	.23	(ns)	.49	(.009)

Table 2: Spearman rank order coefficients and statistical significance of correlations between writing vocabulary measures and GCSE and predicted AS grades

	GCSE Grade (N = 31)		Predicted AS (N = 29)	
Lexical diversity (D)	.40	(.014)	.40	(.016)
Diversity of non-basic words	.41	(.011)	.44	(.008)
Proportion of valid words	.61	(<.001)	.62	(<.001)
Vocabulary ratings	.66	(<.001)	.68	(<.001)

Table 3: Mean raw scores for X_Lex on each frequency band at pre-test and post-test,

and statistical significance of the change (t-tests) (N = 107)

Frequency	Pre-test	Post-test	Increase	t	Significance
1K	17.64	18.24	0.06	2.97	.004
2K	13.08	12.94	-0.14	0.41	ns
3K	11.04	12.80	1.76	5.11	<.001
4K	9.61	11.47	1.86	6.06	<.001
5K	8.55	10.17	1.62	4.78	<.001

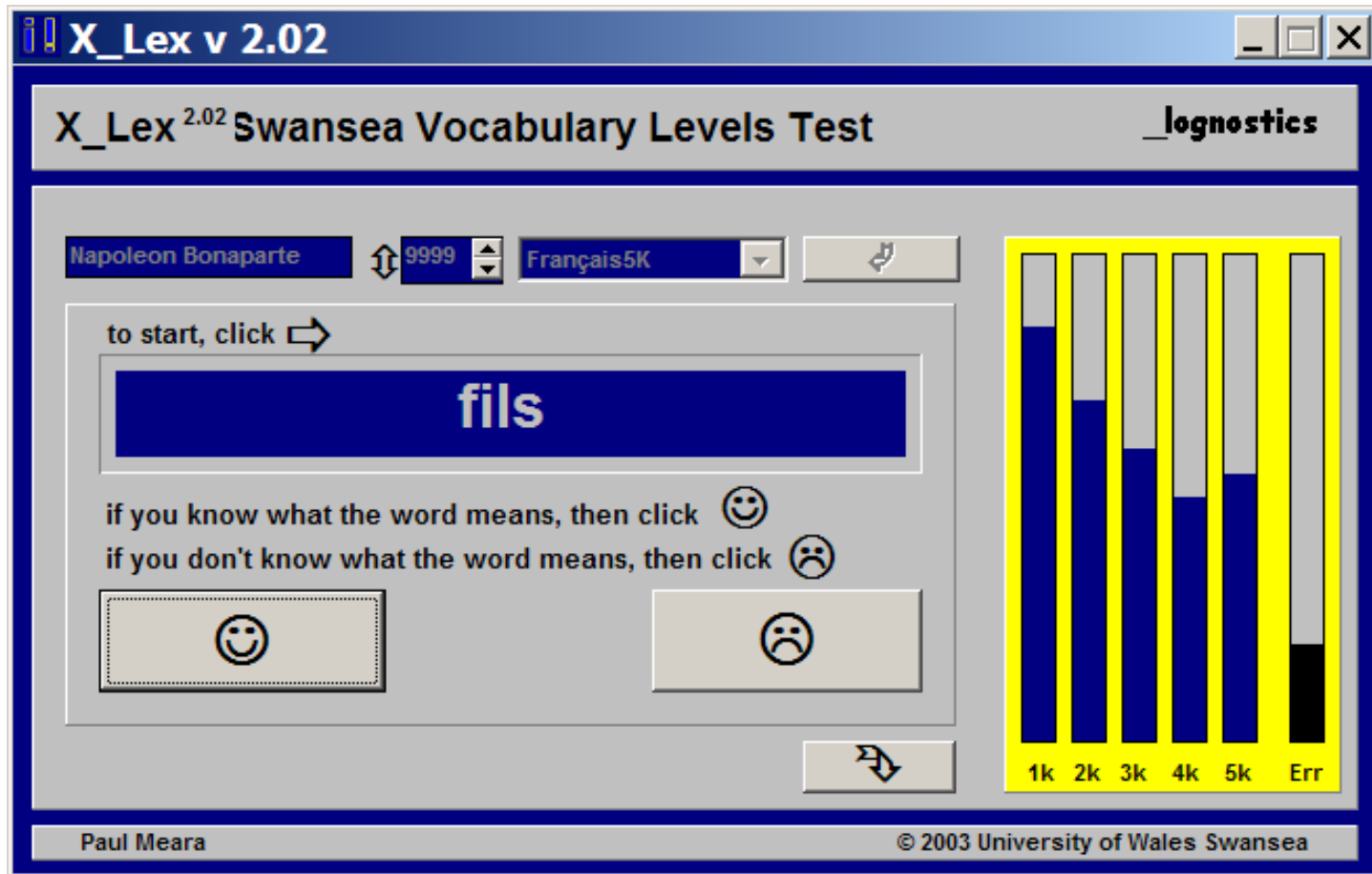Figure 1: Opening screen of the X_Lex test of receptive vocabulary size

Figure 2: Distribution of average X_Lex scores of 107 students at the beginning of Year 12 and after two school terms