University of Reading

Progress in learning French vocabulary in a one-year advanced course at school

Article

Accepted Version

Graham, S. J. ORCID: https://orcid.org/0000-0002-7743-3977, Richards, B. J. and Malvern, D. D. (2008) Progress in learning French vocabulary in a one-year advanced course at school. Journal of French Language Studies, 18 (3). pp. 349-364. ISSN 0959-2695 doi: 10.1017/S0959269508003499 Available at https://centaur.reading.ac.uk/8010/

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

To link to this article DOI: http://dx.doi.org/10.1017/S0959269508003499

Publisher: Cambridge University Press (CUP)

All outputs in CentAUR are protected by Intellectual Property Rights law, including copyright law. Copyright and IPR is retained by the creators or other copyright holders. Terms and conditions for use of this material are defined in the End User Agreement.

www.reading.ac.uk/centaur

CentAUR

University of Reading

Central Archive at the University of Reading

Reading's research outputs online

Progress in learning French vocabulary in a one-year advanced course at school

SUZANNE GRAHAM, BRIAN RICHARDS

AND

DAVID MALVERN

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

ACKNOWLEDGEMENTS

This research forms part of the project 'Strategy training in Year 12 French' funded by the Economic and Social Research Council (Grant no. RES-000-23-0324) and conducted by Ernesto Macaro, Bob Vanderplank (Oxford University) and Suzanne Graham, Brian Richards and Kristyan Spelman-Miller (Reading University). Special thanks go to Lynn Erler, Shirley Lawes and Denise Santos for their help with data collection and analysis. We would also like to thank Jim Milton for his suggestions following an earlier oral presentation of these data and three reviewers for their helpful comments.

ABSTRACT

This article reports on vocabulary learning in French during Year 12 by 150 pupils in 15 English comprehensive schools. Vocabulary progress over two school terms was investigated in the context of learner strategy training that targeted listening and writing. This intervention was carried out in eleven schools and results compared with a comparison group of four schools. Recognition vocabulary and writing vocabulary measures taken during the students' course are compared with listening and writing scores, and examination results. Results suggest that interventions which target listening and writing skills do not lead to significant gains in vocabulary through incidental learning and the deployment of vocabulary in writing tasks. However, it also appears that extra time and effort devoted to developing listening and writing is not at the expense of vocabulary learning.

1 INTRODUCTION

The decline in the numbers of pupils studying French beyond the age of 16 in England has been well documented (Fisher, 2001; Graham, 1997, 2002, 2004; CILT, 2005). Reasons for this are complex, but important factors are the difficulty learners experience, both at the General Certificate of Secondary Education (GCSE) and in the first year of post-16 study, and their perceived lack of progress when faced by the more stringent linguistic demands of the Advanced Level syllabus. The transition from Year 11 (GCSE) entails a steep increase in difficulty, arising partly from exposure to large amounts of new and more complex vocabulary as students move from a basic transactional and survival-based syllabus to more extended and abstract texts and topics. There are also indications that learners have insufficient awareness of how to make progress (Graham, 2004, 2006). In one of the few large-scale studies conducted into students' perceptions of language-learning post 16 in England, Graham (1997) found that many students reported difficulties learning the large amounts of new lexis they encountered from the start of their A-level course. In addition, she found that students used a narrow range of strategies for learning vocabulary, which mainly consisted of fairly shallow strategies such as list-making and 'hide-cover-write-check'.

Graham's (1997) study does not, however, attempt to chart the development of learners' vocabulary knowledge during advanced level

study. Indeed, in comparison with what we know about how learners of English develop vocabulary, little research has been conducted into learners of French in schools in England in general and at A-level in particular. Milton (2006a) represents probably the only study to address this question. Using data drawn from X_Lex (Meara and Milton, 2003), a vocabulary recognition test, he estimates that pupils learn on average 170 words per year in the first five years of secondary schooling (Years 7 to 11). More progress seems to occur in Year 12, however, where the mean number of words known rises from 852 in Year 11 to 1,555. Milton (2006a) also reports a strong relationship between vocabulary levels and the A-Level grade learners achieve in French, which underlines the importance of learners' vocabulary levels for overall progress.

Milton's (2006a) study concludes that the rate of vocabulary growth among learners of French in England is slow when compared with learners of other languages in different countries, even when the smaller number of lessons typically received by pupils in English schools is taken into account. Reasons for this are unclear, although Milton suggests that an analysis of the textbook contents used by English learners might provide insights.

Milton's research looks at vocabulary development under 'normal' classroom conditions, which presumably involve a combination of explicit and incidental vocabulary learning. Schmitt (2000: 116) defines the former as 'the focused study of words', and the latter as learning 'through exposure

when one's attention is focused on the use of language, rather than the learning itself' (e.g. through reading or listening). That both explicit and implicit modes play some role in vocabulary acquisition is generally acknowledged (e.g. Min, 2008). Nevertheless, the relative importance of these two modes of learning has been the subject of controversy (Laufer, 2006), with a recent gulf emerging between research evidence and semi-official advice (see Milton's, 2006b, review of the Harris and Snow, 2004, book on vocabulary building published by The National Centre for Languages, London).

A number of studies have been conducted into both processes. In terms of explicit learning, one area of research has been how learning can be enhanced through word-focused activities such as sentence construction, sentence completion and using target vocabulary in essays (Laufer, 2006). In addition, studies of instruction in vocabulary learning strategies have provided generally positive results (e.g. Burgos-Kohler, 1991, for learners of Spanish in a US university, and Lawson and Hogben, 1998, for secondary school learners of Italian in Australia. See also Nyikos and Fan, 2007, for a review of vocabulary intervention studies).

Regarding incidental learning, there is a sizeable body of research (reviewed in Pigada and Schmitt, 2006, and in Swanborn and de Glopper, 1999) which suggests that vocabulary growth can occur through extensive reading, although gains tend to be small. As both Pigada and Schmitt (2006)

and Swanborn and de Glopper (1999) found, there is a lack of clarity surrounding such research, in terms of differences in the measurement instruments used and exactly what factors have led to the vocabulary gain. Is it frequency of occurrence of words in the text, for example? If so, to what extent is this dependent on textual support mechanisms such as marginal glosses or dictionary use? (see, for example, Hulstijn, Hollander and Greidanus, 1996). Few studies, however, have looked at incidental L2 vocabulary acquisition through listening. Vidal (2003) located only three such studies in non-laboratory settings, ranging from studies of young beginners (Schouten-Van Parreren, 1989) to university-level Japanese learners of English (Toya, 1992), all of whom made gains in vocabulary through listening activities. Vidal herself conducted a study of vocabulary gains among ESP students in a Spanish university, to see whether vocabulary was gained through listening to academic lectures. Before watching and listening to video-taped lectures, learners were tested on their knowledge of 36 technical, academic and low-frequency words contained in the lectures. This test was then repeated after listening, and again as a delayed post-test. Results showed significant gains in knowledge of the selected items at post-test, but this declined at delayed post-test.

Clearly, incidental learning of L2 vocabulary *can* occur through listening and reading. Nevertheless, it is also true that second language learners fail to learn words they encounter (Hulstijn et al., 1996) and that

incidental learning will be influenced by factors that impact on the learner's ability to infer the meaning of unknown words such as the ratio of known to unknown words (Laufer, 1989, 1992). Of particular relevance here seems to be the 'quality of information processing' (Laufer and Hulstijn, 2001: 12) involved in the task through which incidental vocabulary learning may occur. Interventions that seek to improve learners' performance in listening or reading skills, and which typically involve learners in greater depth of processing as learners consciously apply strategies to tasks, would therefore be likely to have an additional benefit for their vocabulary acquisition. We located only one published study (Fraser, 1999), however, that looked at this by-product of strategy instruction, and that was in reading. Fraser investigated the impact of instruction in 'lexical processing strategy' (LPS) use on vocabulary learning among eight university-level Francophone learners of English. Participants received training over a period of two months in three forms of LPS: ignore the word and continue reading; consult (a dictionary or another person); and infer (via linguistic or contextual clues). Data were collected on eight occasions, with four measurement points to determine the extent of LPS use: a baseline period; after the metacognitive strategy training; after the focused language instruction; and one month after the treatment. Learners' use of LPS was ascertained through a think-aloud procedure, whereby they identified unknown words in a text they were asked to read and then explained how

they had dealt with these unknown words. A week after this procedure, each learner was presented with 10 of the words identified as unknown and asked to indicate whether they knew them or not. The mean 'retention rate' (Fraser, 1999: 238) was 28%, although there was much individual variation and a standard deviation of 12%. Fraser argues that this retention rate is higher than reported in previous studies of incidental vocabulary learning through reading. She also reports that learners' rate of ignoring unknown words decreased and their rate of success when making inferences increased. Thus the instruction seemed to have improved the underlying conditions for vocabulary learning. As there was no comparison group, however, and the number of participants small, the results need to be interpreted cautiously.

The limited amount of research in this area thus suggests that the potential impact on vocabulary acquisition of strategy interventions in other skill areas requires further exploration. One possible argument might be that the whole process of strategy awareness raising and a focus on alleviating students' problems in listening and writing would have beneficial spin-offs. In relation to listening tasks, these could result from increased incidental vocabulary learning through mechanisms such as including improved word segmentation, identification and comprehension. In addition, students might benefit from multiple spaced repetition in listening texts (Nyikos and Fan, 2007) and the recycling of previously encountered words known only

superficially. The use of writing strategies might benefit vocabulary through, among other things, deeper processing, more efficient use of feedback, redrafting, and vocabulary-focused planning and monitoring. On the other hand, a counter-argument might be that strategy training would not benefit lexical learning because a tight focus on other skill areas such as listening and writing would be at the expense of curricular time devoted to vocabulary learning and promotion of vocabulary-specific strategies.

The study described here forms part of a project to evaluate the impact of strategy instruction on Year 12 learners' performance in these two skills. Results show a clear-cut positive effect of the intervention for listening (Graham and Macaro, forthcoming, 2008). The benefits for writing were less clear, however, with much a smaller effect (Macaro, Graham, Richards, Spelman-Miller and Vanderplank, 2006). Against this background, and in view of the paucity of studies looking at vocabulary development as a by-product of strategy instruction in other skill areas, we identified the following questions specific to vocabulary:

1. In view of the increased vocabulary demands post-GCSE and Milton's (2006a) findings of a vocabulary 'spurt' in Year 12 (see above) together with the difficulty of finding reliable indices of vocabulary growth over short periods (Tonkyn, 2006), can this progress be measured over two school terms in students' writing as well as their recognition vocabulary?

2. What student variables at pre-test predict post-test scores and progress in vocabulary over this period?
3. If vocabulary progress is measurable (Q1), how much progress is made compared with listening and writing?
4. What is the effect of strategy intervention that targets listening and writing? There are three possibilities:
    a. there are incidental benefits to vocabulary;
    b. the extra attention paid to listening and writing is at the expense of other areas of development such as vocabulary learning;
    c. there is no effect on vocabulary.

2 METHOD

2.1 Participants

The Year 12 students were aged 16 to 17 and were in the first year of post-compulsory education, having elected to continue with French following their GCSE. These participants were preparing for the Advanced Subsidiary (AS) examination at the end of Year 12, with the option of continuing their studies into Year 13 and sitting the Advanced Level (A2) examination. Typically they had already studied French for five years, receiving between 400 and 600 hours of instruction. At the outset, a total of 150 pupils (120 females, 30 males) took part. By the time of the post-test, however, the number of participants had reduced, and inevitably there was some

absenteeism on days when data were collected, leaving 107 pupils who completed both pre- and post-tests for listening, writing and vocabulary. Such a high attrition rate is not unusual at this level and reflects a number of factors including pupils moving to different schools or colleges, changes of subject choice and, above all, the high drop-out rates for languages.

Background data on grades indicate that the 150 participants were high achievers in their GCSE French examinations at the end of the compulsory phase of their education: 42.2% obtained an A* grade (the highest), 39.1% an A, 12.5% B, 5.5% C and 0.8% (one student) was awarded a D. This reflects the typical profile of students in AS level classes.

2.2 The schools

Students attended 15 mixed-sex comprehensive schools. The comparison group consisted of four schools, with the remaining eleven schools allocated to one of two treatment groups (see below). We sought to obtain a stratified sample with matched pairs, i.e. allocation of schools was conducted in such a way as to obtain three groups that were as well-matched as possible for type, location, and make-up of school (e.g. general level of pupil achievement). Randomised allocation was neither ethically nor logistically possible within the framework of a state school setting. The comparison schools were not located in the same counties as the treatment schools.

2.3 Strategy instruction

Schools receiving strategy instruction were placed in one of two groups: high scaffolding (HSG, 5 schools) or low scaffolding (LSG, 6 schools). Both groups received two initial one-hour sessions from researchers, in which modelling of selected strategies for listening and writing took place. Over the following six months, class teachers then led further modelling and practice activities (five main activities were provided for teachers in each skill) in normal class time. Additionally, students were encouraged to use the strategies that had been introduced whenever they were engaged in writing and listening tasks through the use of strategy prompt sheets and record sheets. Detailed instruction notes were provided for teachers as well as briefing meetings to guide them through the implementation of the strategy instruction. For the HSG only, scaffolding of strategy use was provided in the form of additional awareness-raising and reflection about strategy use in the initial researcher-led sessions, a diary in which to record reflections on strategy use, and feedback from researchers both on these diary entries and on the strategy record sheets that learners submitted along with the language tasks that they accompanied (see Graham and Macaro, forthcoming, 2008). Students in the comparison group received no strategy instruction but simply followed their normal French classes.

Selection of strategies was based on the problems in strategy use exhibited by a different, but comparable, sample of students. For listening, these included: poor use of prediction and inferencing; lack of monitoring;

and difficulties with identifying familiar words and word boundaries, the latter being a particular problem for English-speaking learners of French. Materials for developing effective use of prediction, inferencing and monitoring, and clusters of strategies were therefore created, along with those aimed at improving students' perception skills. For writing, students had exhibited problems at the 'formulation' stage of composing, i.e, the point at which they wanted to turn their ideas into French. Instruction materials therefore focused on formulation strategies (for example, re-combining or restructuring known phrases) but also included planning (e.g. 'mind-mapping' or 'brainstorming' of known French that fitted the task requirements), monitoring and using feedback. For details and examples of materials in both skills, see Graham and Macaro (2007).

2.4 Assessing students

At the beginning of their AS course (pre-test) and after two school terms (post-test), students' listening was assessed through a written recall protocol after they had listened to four short passages on the theme of holidays. A different set of passages was used each time, with the difficulty level held constant in terms of length, percentage of unknown words, speech rate, and judgements of a group of students who had listened to passages during piloting. Students were instructed to write down in English everything they had understood in each passage. Responses were written during the two hearings of each passage.

Recall protocols were scored by two raters, who used a banded rating score (four bands) to assess how many idea units had been recalled (in the form of words or phrases) across all four passages. There was a high level of agreement for the total scores of the two raters: .95 at Time 1, at .96 at Time 2 (Pearson correlations). Differences in scores were resolved by discussion.

Students' performance in writing was assessed through a narrative writing task. They were given a six-picture narrative and asked to write a past-tense account of approximately 200 words in 30 minutes. A different but comparable set of pictures was used at each time point. Consulting a bilingual dictionary was allowed because bilingual dictionary use was included in the strategy training. Scoring was conducted using a six-dimensional analytical marking scheme adapted from Jacobs, Zingraf, Wormuth, Hartfiel, and Hughey (1981) and Weir (1993): Content (max. = 20), Organisation (20), Local coherence (20), Vocabulary (15), Grammar (15), Mechanics (10). Organisation includes the control of genre conventions, Local Coherence covers cohesion and the development and integration of ideas, and Mechanics focuses on spelling and punctuation. Each dimension was divided into bands of marks with descriptors for each band. For vocabulary, these were: 'inadequate vocabulary, basically translation (0-2)'; 'frequent lexical inappropriacies, circumlocution, and/or repetition (3-7)'; 'some lexical inappropriacies and/or circumlocution (8-

12)'; 'almost no inadequacies in vocabulary for the task, effective range of vocabulary and appropriate register (13-15)'. Scoring was carried out independently by two expert raters and interrater reliability across the six categories ranged from .69 to .77 (Pearson correlations). Discrepancies were resolved by negotiation. While the reliability of writing assessment is often improved by including more than one task, the pressures of working in busy school contexts meant that this was out of the question. Fortunately, this constraint had no adverse effect on the total writing score whose reliability (Cronbach's apha) was .957 at pre-test and .955 at post-test.

In addition to vocabulary production in narratives, students' receptive vocabulary was assessed using X_Lex (Version 2.02), the Swansea Vocabulary Levels Test (Meara and Milton, 2003). This is a computerized 'yes/no test' that asks respondents to indicate whether they know the meaning of a series of 120 words that appear on screen. One hundred of these words are real words taken from five frequency bands (1K to 5K) based on Baudot's (1992) frequency count. In order to control for false positives resulting from guessing, overconfidence or cheating, twenty of the items are non-existent but plausible words that follow the phonotactic rules of French, for example 'clabrer' or 'muce'. An adjusted score is calculated that takes account of guessing (see Milton, 2006a). A pilot study (Richards and Malvern, 2007) had shown that X_Lex was appropriate for Year 12 pupils and that their results across the five bands were sensitive to

word frequency. In the intervention study reported here, students did two parallel forms of the test both at pre- and post-test. The intention was to use the second set of results because we feared that students would overstate their knowledge at their first attempt before realising how heavily this would be penalized in the adjusted score.

We have referred above to 'vocabulary production' and 'receptive vocabulary'. As Read (2000, pp. 154-157) has pointed out, however, there is much confusion about the distinctions between receptive and productive vocabulary. It is important, therefore, to note that our two measures actually address subsets of receptive and productive knowledge. Using Read's terminology, these entail *recognition* of pre-selected, decontextualized L2 words and ratings of contextualized *use* of L2 vocabulary in writing constrained only by the content of the picture stimuli.

3. RESULTS

3.1 Progress on vocabulary measures (the three groups combined)

This section addresses the first research question by examining whether, for the whole sample of students, their progress on both vocabulary indices is measurable. In doing so, we will need to consider the relative validity of the students' two attempts at X_Lex on each occasion.

It will be recalled that students had two attempts at X_Lex in order to benefit from a predicted practice effect. In practice, however, some students lost motivation on their second attempt, with several achieving

surprisingly low scores and others who failed to finish (see Ns and minimum scores in Table 1). Table 1 shows the results for the four attempts including Cronbach's alpha coefficients assessed from the raw totals from each of the five frequency bands as an indicator of internal consistency. These all indicate high reliability. Tests of normality indicate that the first three sets of results were normally distributed with zero skew but the second post-test attempt was strongly negatively skewed (skew = -.558, s.e. skew = .228, $z = -2.45$, $p < .05$) and there were two outliers with low scores. Parallel forms reliability was therefore estimated using Spearman rank order correlations. These were also satisfactory, though less impressive: .636 ($N = 130$) for the two pre-test attempts and .687 ($N = 112$) for the two post-test attempts.

A comparison of the average scores for the two attempts in Table 1 using the Wilcoxon signed ranks tests shows that, contrary to expectations, there was no advantage for the second attempt at either time point (pre-test: $z(129) = .469$; post-test: $z(111) = .169$; $ps > .05$). However, on both attempts, the students performed substantially better at the second time point (first attempt: $z(107) = 5.34$, $p < .001$; second attempt: $z(99) = 4.65$, $p < .001$), thus demonstrating measurable progress on receptive vocabulary over the period of the study.

Because of concerns over loss of motivation and students failing to finish their second attempt, the analyses that follow include only the data

from their first attempt at pre-test and post-test. This has the additional advantage of allowing the use of parametric statistics.

Insert Table 1 about here

Raters' scores for vocabulary in students' writing at pre-test and post-test are shown in Table 2. Both variables are normally distributed and a paired samples t-test shows gains over the course of the study ($t(112) = 5.75$, $p < .001$, $Eta^2 = .228$).

Substantial progress can thus be reliably demonstrated for both productive and receptive vocabulary. For the latter, however, analysis of the two successive attempts suggests that students tend to demonstrate a fatigue effect rather than a practice effect.

Insert Table 2 about here

3.2 Inter-relationships between variables at each time point

In this section we explore the relationships among the vocabulary, listening and writing variables and participants' success at GCSE as a precursor to identifying factors that predict success in vocabulary learning (Research Question 2). For the continuous variables we examine Pearson correlations at both time points. All are highly significant (Table 3), with inter-relationships being slightly weaker at the second point. Productive and receptive vocabulary are related, but of particular interest are the strong correlations between writing vocabulary and the total writing score (.942 and .933), as well as the moderately strong correlations between writing

vocabulary and listening (.603 and .561) and between listening and writing scores (.601 and .544). It must be remembered, of course, that the writing scores were not independent of vocabulary production in the sense that 15% of the total marks allocated were assigned to rating of vocabulary. A new variable was therefore created at both time points which excluded the vocabulary ratings from the total writing scores. While this does not ensure total independence because of possible halo effects between different dimensions that contribute to the total, it does provide mathematical independence. Correlations between the vocabulary rating and these new, more independent, writing variables were .919 at pre-test and .905 at post-test, consistent with a highly important contribution of vocabulary to success in writing at both time points.

When using examination results in educational research it is common to convert the grades into points and to treat the resulting scale as ordinal or even interval level of measurement (see, for example, Croll, 1995). The skewed distribution of GCSE grades in our data, however, makes this inappropriate and the restricted range makes even an analysis based on rank orders questionable. Participants were therefore divided into a high grade and low grade group. In practice this meant that the high group consisted entirely of those whose grade was A* while all other grades were allocated to the low group. Univariate ANOVAs were then carried out to determine relationships with the continuous variables that entered into the

correlational analysis above. At pre-test all differences were substantial and highly reliable (ps < .001): for X_Lex means were 2083.8 for the low group and 2576.5 for the high group ($F(1,120) = 17.64$, $Eta^2 = .128$); for writing vocabulary the means were 6.32 and 10.06 respectively ($F(1,121) = 76.20$, $Eta^2 = .386$); for the total writing score the means were 40.55 versus 65.24 ($F(1,121) = 82.59$, $Eta^2 = .406$); and for listening they were 14.54 versus 22.98 ($F(1,121) = 48.69$, $Eta^2 = .287$). As can be seen from the Eta^2 values the largest effect sizes are for the two writing variables and the smallest is for the receptive vocabulary test.

3.3 Predictors of success in vocabulary learning

Having examined inter-relationships at each time point, the next step was to discover which variables predicted success in receptive and productive vocabulary after two terms, and the amount of progress made (Research Question 2). Predicting success and predicting progress are, of course, entirely different procedures. Correlations between pre-test and post-test variables will predict *success*, but may only tell us that students who did well at Time 1 also did well at Time 2. Of more interest is *progress* and this needs to be assessed by the *gain* made from time one to time two. The measurement of gains is problematic, however. Simple or 'crude' gain scores (Time 2-Time 1) or percentage gain scores ((T2-T1)/T1 x 100) are not independent of Time 1 scores, and tend to be negatively correlated with them (Barnes, Gutfreund, Satterly and Wells, 1983). A solution is to use

residual gain scores (O'Connor, 1972), that is to say the difference between actual and predicted scores obtained from the regression of the post-test scores on the pre-test scores. By definition, these are independent of Time 1 scores. Unfortunately, residual gains are affected by similar problems to other gain scores, as their reliability is not only a function of the reliability of both the Time 1 and Time 2 scores but also of the correlation between them. The *stronger* the correlation between Time 1 and Time 2 the *lower* the reliability (see Ross, 1998). Ideally, therefore, our vocabulary scores need to be significantly correlated between the two occasions in order to justify the calculation of residuals, but not strongly enough to impair reliability. As can be seen from Table 4, this optimum level of correlation was the case for both X_Lex and writing vocabulary, and so residual gains were computed for these two variables.

Pearson correlations were computed between pre-test and post-test continuous variables (Table 4). For GCSE grade, ANOVAs were conducted with GCSE group as the independent variable. As can be seen, all the pre-test variables are highly significant predictors of post-test receptive and productive vocabulary (all ps < .001), although predictors for writing vocabulary tend to be rather stronger than those for X_Lex. Correlations with residual gains indicate how strongly the pre-test variables predict later success *over and above* what would be predicted by the students' original pre-test status on that measure. These associations can therefore be expected

to be weaker than those discussed above, and this is indeed the case. Nevertheless, both receptive and productive vocabulary at the first time point are significantly associated with later vocabulary gains, as is pre-test listening. However, there is no relationship between pre-test total writing scores and gains for vocabulary. This can be accounted for by the high correlation at pre-test between writing and writing vocabulary: in other words, writing does not account for a significant amount of additional variance beyond that already explained by pre-test writing vocabulary.

Results of univariate ANOVAs with GCSE group (upper versus lower) as the independent variable follow a similar pattern: GCSE predicts post-test X_Lex scores for receptive vocabulary ($F(1,100) = 6.23$, $p = .014$, $Eta^2 = .059$) and, much more strongly, productive vocabulary in writing ($F(1,106) = 47.47$, $p < .001$, $Eta^2 = .309$). GCSE also weakly predicts gain scores for writing vocabulary ($F(1,102) = 5.66$, $p = .019$, $Eta^2 = .053$) but not for X_Lex ($p > 05$).

3.4 Comparing progress within and between groups

To compare how much progress the students made on vocabulary, listening and writing (Research Question 3), a series of repeated measures ANOVAs was conducted comparing pre-test and post-test scores. In order to remove any confounding effects of the intervention, these were performed on the comparison group only. Significant progress could be detected on all variables except receptive vocabulary: X_Lex ($F(1,41) = 3.51$, $p = .068$),

writing vocabulary ($F(1,45) = 9.24$, $p < .001$, $Eta^2 = .170$), listening ($F(1,37) = 42.85$, $p < .001$, $Eta^2 = .537$), writing ($F(1,44) = 5.66$, $p = .022$, $Eta^2 = .114$). Although we are not able to show that effect sizes differ significantly from each other, it is clear that by far the largest increase is for listening. Progress on writing and written vocabulary is similar with a slight advantage for writing. No effect size is reported for receptive vocabulary as there was no significant difference between pre-test and post-test for this group.

In order to test whether the intervention programme had focused on listening and writing to the detriment of vocabulary or, conversely, whether it had provided positive spin-offs for vocabulary (Research Question 4), an ANCOVA was conducted with group as the independent variable, pre-test receptive vocabulary as a covariate, and post-test receptive vocabulary as the dependent variable. This was not statistically significant ($F(2,104) = 1.31$, $p = .275$). A corresponding analysis for writing was also carried out, again with group as the independent variable, but entering pre-test writing vocabulary as the covariate and post-test writing vocabulary as the dependent variable. There was no significant effect of group ($F(2,110)$, $= .556$, $p = .569$), indicating that the intervention had neither positive nor negative effects on students' vocabulary.

4. DISCUSSION

The research reported above investigated four research questions about the measurement, characteristics and extent of development of vocabulary proficiency over a period of two terms in English comprehensive schools. The first question considered whether students' progress could be measured reliably over such a short period. Even though Year 12 is a period of increased vocabulary growth compared with Years 7 to 11 (Milton, 2006a), it was not clear that measures of general vocabulary proficiency, particularly the X_Lex test of receptive vocabulary, would be sensitive to developments across fifteen schools using different textbooks and different examination syllabuses whose varied range of topics and linguistic content lead to different expectations concerning vocabulary. Nevertheless, for our measures of both receptive and productive vocabulary, reliable scores were obtained from which significant progress could be measured even over a relatively short period.

One potential measurement problem had been the issue with yes/no tests of how to prevent respondents from wrongly claiming to know the meaning of a word. X_Lex attempts to control for this by interspersing the stimuli with highly plausible nonsense words and heavily penalizing those who claim to know them. For each such 'error', the raw score out of 5000 is adjusted downwards by 250 points. Thus a student who really did know all the genuine words but also claimed to know all 20 nonsense words would receive a raw score of 5000 and an adjusted score of zero. We were

concerned that students prone to guessing or risk-taking would fail to realize the impact of lack of caution and obtain scores that greatly underestimated their knowledge. We therefore allowed two attempts on each occasion in order that students could benefit from practice. In the event, this provided no advantage and second attempt scores were no better than for the first attempt. If anything, some students became bored and demotivated. Nevertheless, concerns remain about the effect of test-taking style or test-taker personality on the validity of such tests, and this is an area that needs further investigation.

The second research question addressed predictors of success and progress in vocabulary. Firstly, however, concurrent intercorrelations at pre-test and post-test between measures of vocabulary, listening and writing were examined. These were all substantial and highly reliable, as are their relationships with GCSE grades at pre-test. What was particularly striking, however, were the very high correlations (.94 and .93) between productive vocabulary and writing at both time points, providing evidence of the central role of vocabulary in the writing process at this stage.

This research question made an important distinction between success *after* two terms and progress, i.e. gains, *over* two terms and asked which student pre-test variables predicted these for vocabulary. Both pre-test vocabulary measures and pre-test writing and listening scores were highly significant predictors of post-test vocabulary, as were GCSE grades

(particularly of productive vocabulary). Our measure of progress used residual gain scores to obtain an index for each vocabulary measure that reflected progress over and above what would be predicted from the pre-test score on that measure. As would be expected, these were predicted less strongly, but were still significantly related to pre-test productive vocabulary, X_Lex and, especially, listening. GCSE grade was also a significant predictor of productive vocabulary.

With regard to our third research question concerning the relative amount of progress made in vocabulary, listening and writing, this analysis was carried out on the comparison group only in order to control for the effects of the intervention programme. It was clear that students made progress in listening, writing and writing vocabulary. The result for X_Lex, however, was not significant. It seems to be the case that it is more difficult to make measurable progress on receptive vocabulary, particularly when compared with listening which had the largest effect size. It may well be that the skill element of listening undergoes rapid development in Year 12, even for students who do not receive strategy training, and this has a large impact on test scores. Writing, on the other hand, which offers more opportunity for reflection, planning, and self-direction may progress more gradually. With regard to receptive vocabulary versus vocabulary in writing, it seems likely that students would be able to demonstrate progress more easily on the latter when they are able to choose the lexical items they will

use rather than relying on any chance correspondence between the vocabulary they have learnt and the items contained in X_Lex's dictionary.

Finally, we considered the three possible effects on vocabulary learning of an intervention that addressed listening and writing skills. The first possibility was that students' vocabulary in the intervention groups would benefit indirectly from the attention to listening and writing strategies, possibly through incidental learning from more effective listening, and practice and depth of processing in writing. The second possibility was that the intervention could harm lexical development by diverting attention from it, and the third possible result was that the intervention made no difference. Although we had expected vocabulary to benefit from the experimental programme, analyses found no difference between the groups. It may be, as Walters (2006: 238) argues, that the effects of strategy instruction on vocabulary development take time to emerge fully, perhaps because of the 'incremental nature of incidental vocabulary development'. It is also possible that strategy instruction that had focussed more sharply on word identification or inferencing might have led to greater vocabulary gains in the intervention groups. These are questions that need to be explored more fully in future research. Nevertheless, the fact that the strategy instruction had a positive impact on students' listening skills, while not jeopardising their vocabulary development suggests that strategy instruction is a fruitful avenue to explore

in terms of pedagogy and improving learners' attainment in French at lower-intermediate level.

REFERENCES

Barnes, S., Gutfreund, M., Satterly, D., & Wells, C. G. (1983). Characteristics of adult speech which predict children's language development. *Journal of Child Language*, 10.1: 65-84.

Baudot, J. (1992). *Fréquence d'utilisation des mots en français écrit contemporain*. Montréal: Presses de Montréal.

Burgos-Kohler, N. (1991). The effect of a selected group of language learning strategies upon language development. Unpublished PhD dissertation, University of Texas at Austin.

CILT. (2005). *Languages yearbook 2005*. London: The National Centre for Languages.

Croll, P. (1995). Early linguistic attainment, family background and performance in 16+ examinations. *Educational Studies*, 21: 13-28.

Fisher, L. (2001). Modern foreign languages recruitment post-16: The pupils' perspective. *Language Learning Journal*, 23: 33-40.

Fraser, C. (1999). Lexical processing strategy use and vocabulary learning through reading. *Studies in Second Language Acquisition*, 21.2: 225-241.

Graham, S. (1997). *Effective Language Learning. Positive Strategies for Advanced Level Language Learning*. Clevedon: Multilingual Matters.

Graham, S. (2002). Experiences of learning French: a snapshot at Years 11, 12 and 13, *Language Learning Journal*, 25: 15-20.

Graham, S. (2004). Giving up on modern foreign languages? Students' perceptions of learning French. *The Modern Language Journal*, 88.2: 171-191.

Graham, S. (2006). Listening comprehension: The learners' perspective. *System*, 34.2: 165-182.

Graham, S. and Macaro, E. (2007). Designing Year 12 strategy training in listening and writing: from theory to practice. *Language Learning Journal*, 35.2: 153-173.

Graham, S. and Macaro, E. (forthcoming, 2008). Strategy instruction in listening for lower-intermediate learners of French. *Language Learning*.

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

Hulstijn, J., Hollander, M., and Greidanus, T. (1996). Incidental vocabulary learning by advanced foreign language students: The influence of marginal glosses, dictionary use, and reoccurrence of unknown words. *The Modern Language Journal*, 80: 327-339.

Jacobs, H., Zingraf, S., Wormuth, D. Hartfiel, V., and Hughey, J. (1981). *Testing ESL composition: a practical approach*. Rowley, MA: Newbury House.

Laufer, B. (1989). What percentage of text lexis is essential for comprehension? In: C. Lauren and M. Nordman (eds.), *Special*

language: From humans to thinking machines. Clevedon: Multilingual Matters, pp. 316-323.

Laufer, B. (1992). How much lexis is necessary for reading comprehension? In: P. J. L. Arnaud and H. Bejoint (eds.), *Vocabulary and applied linguistics*. London: Macmillan, pp. 126-132.

Laufer, B. (2006). Vocabulary acquisition in a second language: Do learners really acquire most vocabulary by reading? Some empirical evidence. *Canadian Modern Language Review*, 59: 567-587.

Laufer, B. and Hulstijn, J. (2001). Incidental vocabulary acquisition in a second language: the construct of task-induced involvement. *Applied Linguistics*, 22.1: 1-26.

Lawson, M. and Hogben, D. (1998). Learning and recall of foreign-language vocabulary: effects of a keyword strategy for immediate and delayed recall. *Learning and Instruction*, 8.2: 179-194.

Macaro, E., Graham, S., Richards, B. J., Spelman-Miller, K. and Vanderplank, R. (2006). Strategy training in year 12 French: listening and writing. Final report to the Economic and Social Research Council (Project RES-000-23-324).

Meara, P. M., Milton, J. L. (2003). X_Lex: 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; and V. Harris and D. Snow. Doing it for themselves: focus on learning strategies and vocabulary building. London: CILT, 2004. Language Learning Journal, 33: 75-77.

Min, H-T. (2008). EFL vocabulary acquisition and retention: reading plus vocabulary enhancement activities and narrow reading. *Language Learning*, 58.1: 73–115

Nyikos, N. and Fan, M. (2007). A review of vocabulary learning strategies: focus on language proficiency and learner voice. In: A.D. Cohen and E. Macaro (eds.), *Language Learner Strategies: Thirty Years of Research and Practice*. Oxford: Oxford University Press, (pp.251-273).

O'Connor, E. F. (1972). Extending classical test theory to the measurement of change. *Review of Educational Research*, 42, 73-97.

Pigada, M. and Schmitt, N. (2006). Vocabulary acquisition from extensive reading: A case study. *Reading in a Foreign Language*, 18.1: 1-28. Retrieved December 7, from http://nflrc.hawaii.edu/rfl

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

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.

Ross, S. J. (1998). *Measuring gain in language programs: theory and research*. Sidney: National Centre for English Language Teaching and Research, Macquarie University.

Schmitt, N. (2000). *Vocabulary in Language Teaching*. Cambridge: Cambridge University Press.

Schouten-van Parreren, C. (1989). Vocabulary learning through reading: which conditions should be met when presenting words in texts? *AILA Review*, 6: 75-85.

Swanborn, M. S. L. and de Glopper, K. (1999). Incidental word learning while reading: a meta-analysis. *Review of Educational Research*, 69.3: 261-285.

Tonkyn, A. (2006). *Towards progress-sensitive measures of L2 speaking proficiency: a case study of instructed learners*. Unpublished Ph.D. Dissertation, University of Reading.

Toya, M. (1992). *Form of explanation in modification of listening input in L2 vocabulary learning*. Unpublished MA in ESL thesis, Department of ESL, University of Hawai'i, Honolulu.

Vidal, K. (2003). Academic listening: a source of vocabulary acquisition? *Applied Linguistics*, 24.1: 56-89.

Walters, J. M. (2006). *Beyond exhortation: training students of English as a second language to infer meaning from context.* Unpublished PhD thesis, University of Nottingham, UK.

Weir, C. J. (1993). *Understanding and Developing Language Tests*. Hemel Hempstead: Prentice Hall.

Table 1. Descriptive statistics for the students' four attempts at X_Lex (maximum possible score = 5,000) and internal consistency of scores across frequency bands (Cronbach's alpha).

	N	Mean	SD	Min.	Max.	Alpha
Pre-test 1st attempt	138	2281.9	674.3	550	4300	.903
Pre-test 2nd attempt	130	2297.3	839.7	100	4550	.923
Post-test 1st attempt	115	2700.0	697.5	1050	4600	.907
Post-test 2nd attempt	112	2723.7	719.7	100	4350	.896

Table 2. Descriptive statistics for ratings of vocabulary in writing at pre-test and post-test (maximum possible score = 15).

	N	Mean	SD	Min.	Max.
Pre-test	139	7.79	3.03	0	14
Post-test	121	9.58	2.53	3	15

Table 3. Intercorrelations between vocabulary, listening and writing variables at both time points.

	1	2	3	4
		Pre-test		
1. X_Lex	--	.455	.410	.441
2. Vocabulary in writing		--	.942	.603
3. Total writing score			--	.601
4. Total listening score				--
		Post-test		
1. X_Lex	--	.342	.327	.340
2. Vocabulary in writing		--	.933	.561
3. Total writing score			--	.544
4. Total listening score				--

Note. All ps <.001

Table 4. Pearson correlations between pre-test variables and success and progress in receptive and productive vocabulary.

	Post-test measures			
Pre-test measures	X_Lex score	X_Lex residual gains	Writing vocabulary score	Writing vocabulary residual gains
X_Lex	.489***	--[a]	.424***	.225*
Vocabulary in writing	.371***	.217*	.640***	--[a]
Total writing score	.317***	.180	.641***	.062
Total listening score	.426***	.304**	.574***	.339***

Note. [a]Correlations between residual gain scores and the independent variable from which they were calculated are always zero.

*p <.05. **p <.01. ***p <.001.