

Language Teaching

<http://journals.cambridge.org/LTA>

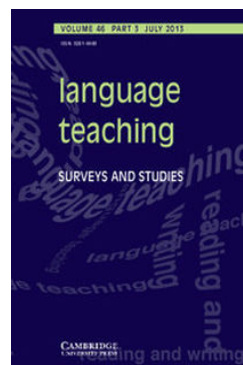
Additional services for *Language Teaching*:

Email alerts: [Click here](#)

Subscriptions: [Click here](#)

Commercial reprints: [Click here](#)

Terms of use : [Click here](#)



Vocabulary knowledge in relation to memory and analysis: An approximate replication of Milton's (2007) study on lexical profiles and learning style

Paul Booth

Language Teaching / Volume 46 / Issue 03 / July 2013, pp 335 - 354

DOI: 10.1017/S0261444813000049, Published online: 05 June 2013

Link to this article: http://journals.cambridge.org/abstract_S0261444813000049

How to cite this article:

Paul Booth (2013). Vocabulary knowledge in relation to memory and analysis: An approximate replication of Milton's (2007) study on lexical profiles and learning style. *Language Teaching*, 46, pp 335-354 doi:10.1017/S0261444813000049

Request Permissions : [Click here](#)

Replication Study

Vocabulary knowledge in relation to memory and analysis: An approximate replication of Milton's (2007) study on lexical profiles and learning style

Paul Booth Kingston University London
p.booth@kingston.ac.uk

This paper presents an approximate replication of Milton's (2007) study on lexical profiles and learning style. Milton investigated the assumption that more frequent words are acquired before less frequent ones. Using a vocabulary recognition test (X-Lex) to measure vocabulary size, Milton found that L2 English group profiles show a linear relationship between greater knowledge of high frequency words and lesser knowledge of low frequency items. The profiles also showed variability in individual profiles. Milton hypothesised that the individual differences in profiles are partly attributable to different approaches to learning, as elicited via language aptitude tests of memory and analysis. Learner profiles that showed a linear relationship with vocabulary frequency scored higher on analysis; learners who had irregular profiles scored higher on memory. The aim of this replication is to confirm whether learning style helps to determine what L2 lexis is learnt. It duplicates the vocabulary size test and the memory and analysis tests. However, the replication uses regression analyses, rather than a single ANOVA, to determine whether memory or analysis contributes to vocabulary knowledge. Moreover, the participants' L1 backgrounds are mixed, and they are older. The results from this replication do not support Milton's findings, but a post-study supports the notion that at low proficiency there is a relationship between memory and vocabulary size. It is concluded that neither memory nor analysis is related to patterns in lexical profiles, but that memory contributes to vocabulary size.

1. Introduction

1.1 Vocabulary size and competence

The more frequent a word, the more likely it is that a language learner will acquire it. As long ago as the 1960s, Mackey (1965) argued that words have a 'natural selection' (p. 176). That is to say, words which are useful occur frequently in use. Even before that, West (Palmer & West 1937) argued that vocabulary selection of the most frequent words, and so by implication the most useful, means that few words can be taught well, rather than a large number of words taught ineffectively (p. 137). The premise is that learners need very frequent lexis in order to

master a language. Indeed, L2 learners tend to acquire more of the very frequent lexis than the less frequent (Laufer & Nation 1995). More recently, Lindqvist, Bardel & Gudmundson (2011) found that in other languages (French and Italian), advanced learners' oral production matches that of native speakers in terms of lexical rarity.

Lexical rarity is one measure of lexical sophistication and, by implication, is an indication of the size of the lexicon. Meara (1996) argues that, in order to be proficient in an L2, vocabulary size is one of the most important dimensions in language skills. However, determining how many words a learner knows is by no means as straightforward as it might seem. Indeed, one of the fundamental issues in vocabulary testing is how to test vocabulary size. It is prohibitively impractical to test all the possible words which a learner could know. Nevertheless, testing learners on a sample of words from different frequency bands is a way around this problem.

1.2 Vocabulary size and development

Meara & Buxton (1987) argue that vocabulary recognition from various frequency bands is an efficient way of quantifying the number of words L2 learners actually know. The Yes/No test is based on whether the testee recognises a lexical item or not. The test then extrapolates from a random sample of words in a given frequency band to give an overall estimate of the amount of words a person can recognise. This test made some reasonable predictions of candidates' exam success in the Cambridge First Certificate language test (Meara & Buxton 1987: 150). The computer version of the Yes/No test (Meara & Milton 2003a) also showed strong correlations with placement measures of listening, grammar and writing, indicating that it was sensitive to differences in learner proficiency (Harrington & Carey 2009). The empirical data confirm the notion that vocabulary size is related to language proficiency. However, many factors influence lexical learning.

How learner vocabulary grows as the learner gains proficiency shows an interesting relationship with vocabulary size. Milton & Meara's (1995) study found that exchange students who had the largest vocabulary knowledge made less gain than students with smaller vocabularies. Students with the lower starting vocabularies made the greatest progress in vocabulary gain. There seems to be an inverse relationship between vocabulary size and lexical development. Another reason may be that the classroom environments in which learners are placed may be highly idiosyncratic with respect to which words are acquired.

Vocabulary growth is dependent, to some extent, on the incidental learning of vocabulary and, in particular, what type of words learners are exposed to. Milton & Vassiliu (2000) found that, in elementary foreign language textbooks, function words (words that show grammatical relationships) seem to be relatively infrequent. They found that nearly 50% of lemmas (the headword and some of its inflected and reduced forms) found in elementary textbooks fell outside the first and second thousand-word frequency bands. This is what led to Milton's notion that some learners' vocabulary may follow the content of the textbook with relatively little knowledge of high frequency words, whilst others may make up for any lexical shortcomings. Moreover, Milton argues, if a vocabulary size test is based on a methodology which is dependent on learners knowing the highly frequent words before being tested on rarer words, the validity of size tests may be called into question.

Differences in lexical profiles, Milton (2007) hypothesised, could be attributable in part to learning style. One of the ways in which learners are thought to differ is that they may deviate from the regular pattern of knowing more words at successively higher frequency bands, showing an irregular pattern instead. He puts forward the notion that varying lexical profiles may be the product of varying aptitude strengths. When group profiles are examined there is a linear pattern between frequency and lexical knowledge. However, on an individual level, learners can deviate from this. For example, Milton highlights the fact that some learners can show an unusual dip in the one thousand frequency band in which function words occur. Milton hypothesised that learners who find function words particularly difficult to acquire may be memory-orientated learners, whereas learners who do not have difficulty acquiring function words could be analysis-orientated learners. This dual mode of learning underpins Skehan's memory–analysis framework (see below).

1.3 Learner types

Skehan (1998) proposed a memory–analysis framework for learning style. Informed by a dual-mode model of L2 performance and development, memory-orientated learners are thought to store a wide range of lexicalised exemplars readily available for use, whereas analysis-orientated learners are thought to prize rule-based language and engage in regular restructuring. Harley & Hart (1997: 395) found that vocabulary size, as measured by Meara's Yes/No test, was also related to learning characteristics. They argue that for early immersion students, predictors of language success are more likely to be associated with a memory measure. L2 success may be more reliant on analytical language ability for late starters than for early starters. It supports the notion that learners can take different paths to vocabulary development and that these paths can, to a certain extent, be attributed to a particular learning style or aptitude strengths.

Reiterating this idea in their work with the Yes/No test format, Meara & Milton (2003b: 9) found that some learners produced profiles with unexpected dips in the 1,000 (1k) frequency band where we find functional vocabulary, but performed relatively better in identifying rarer lexis. It was this insight that provided the motivation for Milton's (2007) study. From a vocabulary acquisition perspective, we can see how Milton used Skehan's learning style framework to help explain why learners may have unexpected dips in their lexical profiles. Meara & Milton (2003b) suspect that some learners' relatively low scores on functional vocabulary may be due to their learning strategy. A learner who has a memory-based approach may find words in the rarer bands more concrete and imageable – and therefore more learnable. Analytic learners should display a normal profile because, as Milton (2007: 52) argues, their more rule-based, bottom-up approach may make them more aware of any gaps in lexical knowledge.

1.4 Why a replication is needed

Milton used an ANOVA to detect whether there was a relationship between lexical profiles and the memory–analysis framework. The results pointed to a relationship between learners

classified into normal and level two (the second thousand-word frequency band) deficit profiles and test results of memory and analysis. However, on closer inspection, the differences in scores on the analysis test to detect grammatical patterns were minimal. Learners with normal profiles appeared to score only marginally better on the analytic (LAT C) test than learners with level two deficit lexical profiles. Therefore a replication is needed to verify whether lexical dips in profiles are indicative of differences in learning style.

Milton also used L1 Greek learners of English as participants. The fact that some English words have Greek origins could be advantageous for Greek learners over learners whose L1 is not cognate with English. Greek learners may have identified more of the low frequency words. This advantage could have skewed the profiles to be more irregular, which would be based on their L1 knowledge rather than their learning style. Therefore, an approximate replication is needed to determine whether Milton's findings are valid for a more heterogeneous group of participants. This replication duplicates the vocabulary and learning style tests of Milton's study but differs in terms of participants. Moreover, the approximation study analyses the data collected from the tests by regression analysis rather than a single ANOVA.

2. The original study: Milton (2007)

Milton designed the first part of his research study to test the commonly held view that, in learning a second language (L2), many of the most frequent words are acquired first. He describes the profile of a typical language learner as showing the greatest knowledge of words in the highest frequency band (first 1,000 words, 1k), with progressively less knowledge in the less frequent bands. He also highlights learners who appear to buck the trend with 'irregular' profiles that show dips in the high frequency bands. Milton's study introduces learning style and argues that this factor may be associated with the occurrence of 'irregular' profiles.

First, Milton's study set out to investigate learners' lexical profiles and determine whether the acquisition of words matches the regular profiles described above, or whether the profiles display variability: is the frequency of occurrence of words a reliable predictor of acquisition? Second, Milton examined the variability of acquisition: a sub-group of learners was examined individually to determine the variability of profiles within the group. Milton then considered whether a regular vocabulary profile, which he speculates may be indicative of an analytic learning style, can be an advantage in vocabulary acquisition and, ultimately, language proficiency.

Milton used a computerised version of the Yes/No test, *X-Lex* (Meara & Milton 2003a), which presents learners with a sample of words from each of the first five 1,000 frequency bands separately, so learners with any lexical dips in the first or second 1,000 most frequent words are not penalised too much for lexical size. The results of all the groups demonstrated that the learners' profiles do in fact follow a predictable pattern: the frequency of words is related to their order of acquisition. However, the mean scores of these learners hid huge variability between individuals, so Milton looked at a sub-group of learners. One type of profile, in which learners have a higher mean score of vocabulary knowledge at each successively higher frequency level, could be broadly described as 'regular'. Milton suggests that these learners may have been more adept at using strategies which help them to overcome

any deficiencies in their language-learning environment. He notes that they seemed to have a high proportion of words in the highest frequency band, i.e. the band in which function words occur (Nation 2001: 206). Milton suggests that learners with an analytic learning style may acquire more of the highest frequency words because structural or function words appear in the highest frequency band. Another type of profile described as ‘irregular’ displayed relatively large lexical deficiencies at the 2,000 (2k) frequency band. Milton speculates that this type of deficit is influenced by the frequency of lexis in the beginner-level learning texts that students encounter in class (Milton & Vassiliu 2000).

The notion that learning styles (memory and analysis) influence what and how vocabulary is acquired became the basis of the second part of Milton’s study. A sub-sample of learners was tested using two language aptitude tests (LAT B and C) (Meara, Milton & Lorenzo-Duz 2001). Although designed mainly to measure aptitude, these tests were used to determine the sub-sample’s language learning style in terms of a memory–analysis framework. LAT B is a test of memory and LAT C is a test of language analysis. Milton examined the results to see whether the ‘normal’ and ‘irregular’ profiles were linked to certain learning styles. He found that learners with ‘normal’ profiles scored higher on the analytical dimension, and were broadly confirmed as being more predisposed towards an analytical learning style. These learners are presumed to have acquired more function words at the top end of the frequency scale and so do not show any deficit. Learners with level 2 deficit profiles scored higher on the memory dimension, so were broadly confirmed as more predisposed towards a learning style that relies heavily on memory. It is presumed that these learners have not acquired as many of the words which fall into the first two 1,000-word frequency levels and that a learner’s style may influence, to some extent, the words that are learnt.

This replication study explores Meara & Milton’s (2003b) notion, broadly supported by Milton (2007), that learning style may influence the lexical profiles of learners. The Yes/No method of calculating vocabulary knowledge has been examined by various researchers to test its validity. Mochida & Harrington (2006) compared the Yes/No test to the Vocabulary Levels Test (Nation 1990), a multiple-matching test of receptive vocabulary, and found that the Yes/No test is a reliable measure. Since Milton’s study, researchers have used various formats of the Yes/No test as a placement tool for different groups of L2 learners of various L2 languages (David 2008; Harrington & Carey 2009; Fairclough 2011). In parallel with Milton’s study, low-level learners of L2 French displayed irregular profiles, which David also argues could indicate a lack of exposure to the high frequency lexis to which native speakers are exposed. This variability points to the need to further examine individual learner differences in relation to lexical knowledge and establish whether irregular profiles are, in fact, indicative of a particular learning style.

More empirical work is needed to illuminate Skehan’s (1986, 1998) learning style framework: this is based on aptitude strengths that he maps onto processing stages in SLA development (Skehan 1998: 203). Since its early days, for example in Robinson (1997, 2001), there has been a resurgence of interest in research that examines individual differences in aptitude in relation to learning conditions of an L2. More recently, Ben Maad’s (2010) study examined a dual-mode model of language performance, rule-based and lexically based, from oral production in story-telling, which builds upon Skehan’s theoretical framework of processing modes. Nevertheless, more work is needed in other areas of L2 knowledge. Tight

(2010) looked at the relationship between learning style and lexis and argued that matching the teaching of lexis to the learners' preferred mode of perception did not necessarily produce the best results in terms of lexical retention over time. The theoretical framework of learning style in Tight's study was based on the mode of perception, e.g. auditory rather than visual, which has been subject to criticism (Coffield et al. 2004). This replication of Milton builds upon Skehan's dual-mode learning style framework, which is more concerned with how learners process language and how it relates to lexical knowledge.

On a wider level, this replication has sought further evidence as to whether individual differences in learning can result in differences in lexical profiles. Schmitt & Meara (1997) found individual differences in vocabulary development: although group means may display a linear pattern in vocabulary development, high frequency words are not necessarily learnt before less frequent lexis in a linear manner (p. 25). From a Dynamic Systems perspective, De Bot, Lowie & Verspoor (2007: 14) characterise a language learner as a dynamic subsystem with his/her own cognitive ecosystem. Memory and analysis could be regarded as part of this ecosystem alongside the learner's L1 and L2 lexis. The interaction of these systems is what gives rise to development, and variability is part of this non-linear developmental process. De Bot et al. criticise 'grand sweep' studies which do not embrace variability in the L2 development process. Aptitude strengths which characterise learner types can be one way of understanding how learners differ, and thus provide a greater understanding of SLA development in vocabulary. Skehan's learning style framework may provide a way for researchers to understand the heterogeneous nature of L2 vocabulary profiles.

The present study tests Milton's conclusions as to whether regular or irregular lexical profiles can be attributable in part to learning strengths. A difference in the replication study is that it uses a more heterogeneous group of learners. In Milton's study, the learners were young Greek L2 learners of English, while the participants in the replication were older and came from a wide variety of L1 backgrounds; some had a third language, and they were also more heterogeneous in age of learning L2 English. These participants were used to determine whether Milton's findings can be generalised to a more heterogeneous group of L2 learners of English.

3. The replication study

This replication study aims to test Milton's conclusions that normal profilers are stronger in analysis (LAT C) than level two deficit profilers who, it is suggested, are stronger in memory (LAT B). On a broader level, the aim is to examine the relationship between the words L2 learners acquire and the manner in which they acquire them. This study sets out with the same research questions as Milton:

1. What is the incidence in variability of lexical profiles of learners and in what proportions do they occur?
2. Are there any relationships between particular types of lexical frequency profiles and learners' respective performances on the LAT B memory and LAT C analysis tests?
3. Is there a relationship between memory and irregular profiles (level two deficit), and analysis and normal profiles?

The first question deals with the variability that is inherent in researching lexical profiles. The expected overall pattern is that learners know more words in the high frequency bands and that at each successively rarer frequency level, show progressively less knowledge. The second deals with the relationship between types of profiles and learning style. Based on Milton's findings, we would expect that learners who score higher on the LAT C (analysis) test would display normal profiles, whereas learners who score higher on the LAT B (memory) test would display irregular profiles. The third and final question asks whether a level two deficit profile is indicative of a memory-orientated approach to learning. The learners in my replication are less homogenous in terms of proficiency and L1 background than those in Milton's study, and these two combined factors had implications for the results obtained.

3.1 Method

3.1.1 Participants (first cohort)

The participants were taken from two different colleges in London. They comprised 20 elementary, 6 pre-intermediate, 14 intermediate and 1 upper intermediate. Classifications were based on a computer level test or their English teacher's judgment. This group of learners differs from the participants in Milton's study. Firstly, the nationalities and L1 of the learners in this study are mixed. Secondly, in this group the minimum age is 18, whereas in Milton's study the learners were under 18. Thirdly, and perhaps more importantly, the low English proficiency level of some of these learners had some consequences for one of the learning style tests. Whilst these differences affect the basis for comparison to some extent, this new cohort of participants offers a more diverse range of L2 learners.

3.1.2 The tests

A background questionnaire was administered before the tests were conducted. In all cases the tests were demonstrated to the learners via a projection screen. The learners also had copies of written instructions to refer to if necessary. The tests were held in various computer labs and the participants were seated as far as possible from each other to avoid communication. All 41 learners were tested on their vocabulary knowledge using the *X-Lex* (v.2.00) Swansea Vocabulary Levels Test (Meara & Milton 2003a) to determine their lexical profiles. This vocabulary test, the same as in Milton's work, presents learners with a set of 120 words. As each word is presented to the learners they are asked if they know the meaning of the word. The words are taken from five frequency bands (1k–5k). The scores are extrapolated to give an overall score of known words to a maximum of 5,000. The *X-Lex* (v.2.00) also gives a profile of words known at each frequency level, using a set of word files based on Nation's (1984) word lists and Hindmarsh (1980).

The test presents real words and pseudowords to the testee. If the learner claims to know any of the pseudowords, the overall score is reduced in proportion to the amount of guesswork. Out of the 41 learners who took the vocabulary test, 11 learners scored over 5% on the error

dimension. These learners (approximately 27%) were discounted from the study because, as Meara & Milton (2003b) claim, their profile scores cannot be taken at face value because they might not reliably represent the learners' vocabulary knowledge. The learners who overestimated their vocabulary knowledge were mostly at elementary level, so they may be particularly uncertain of their vocabulary knowledge and more prone to take wild guesses than higher proficiency level students. In Milton's sub-sample, approximately eight out of 29 participants were also discounted because their profiles were unstable.

After learners completed this vocabulary test they were tested on their learning style: memory (visual memory of paired associates) and analysis (ability to detect grammatical rules) (Meara et al. 2001). LAT B (memory) is a timed test in which learners are asked to recall the meaning of ten words in a fictional language. LAT C (analysis) was used to measure the ability to infer grammatical rules. After initial learning of sentences in a new language, learners were given sets of two similar sentences and asked to choose which one is grammatically correct. The participants used in my research were at various levels of proficiency, so it is important that the language presented is at the same level of difficulty for all participants, i.e. a fictional language which nobody had been given the chance to study before. These two tests are built upon the Modern Language Aptitude Tests pioneered by Carroll, who identified four components of language aptitude (Carroll 1981). Two of these components – memory and analysis – form the learning style framework outlined by Skehan (1998).

3.2 Results

3.2.1 Research question 1: Distribution of profiles

The mean profile scores for the group mirrored Milton's findings in that, overall, the learners displayed a greater knowledge of the more frequent words and less knowledge of the less frequent words, as can be seen in Figure 1. In fact, 13 students, or approximately 43% of learners, had normal profiles. In Milton's study of Greek learners, approximately 60% of learners displayed normal profiles.

Within the group, though, there were other profiles which did not fall into the normal pattern. In my study, eight learners (approximately 26%) displayed a level 2 (2k) deficit profile (see Figure 2); in Milton's study 25% had a level 2 deficit profile. However, none of the learners in my study displayed a level 1 (1k) deficit (i.e. the type of learners who have lower scores at level 1 than at level 2), whilst in Milton's main study, 10% of learners fell into a level 1 deficit category.

The remaining nine learners (approximately 30%) displayed profiles that had erratic dips at various frequency bands. In Milton's first study, only a very small proportion also displayed profiles that could not be classified. In my study, learners with unclassifiable profiles have been in the UK approximately the same amount of time (on average, around three years) as the rest of the cohort, which suggests that time spent in the country where the L2 is spoken does not appear to make profiles more 'normal'. These results are also reflected by Milton (2007), who observes that individuals are highly varied not only in their profiles but also in their mean vocabulary scores.

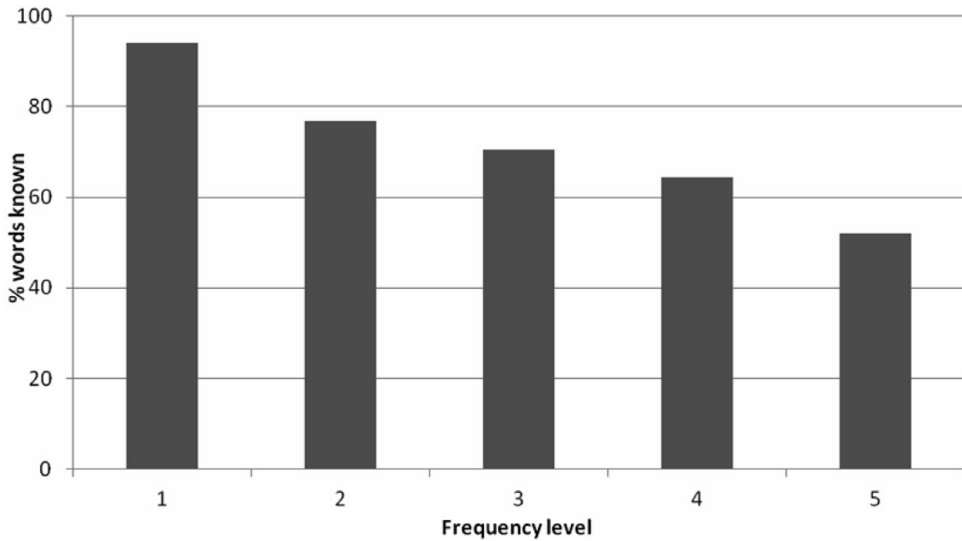


Figure 1 Whole group mean scores for frequency bands

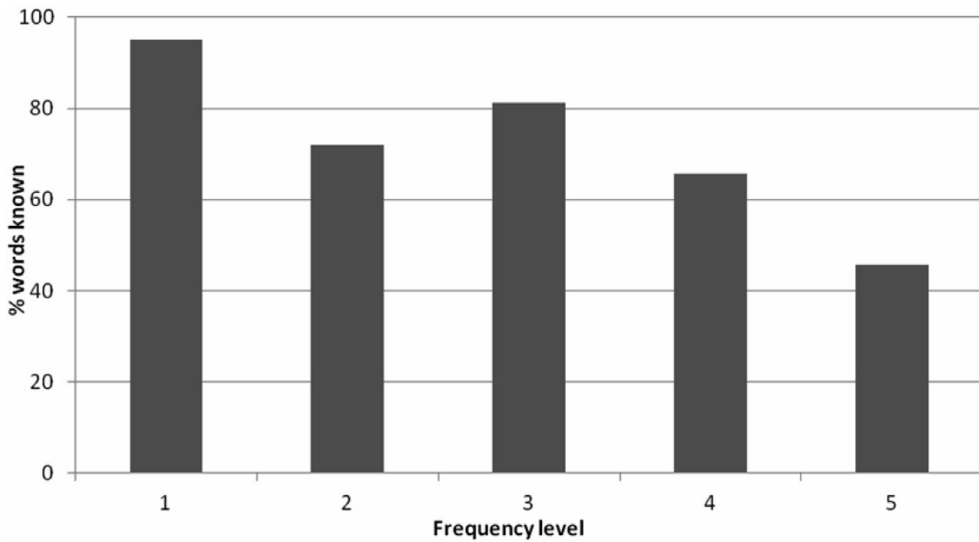


Figure 2 Mean score for level 2 (2k) deficit profile

3.2.2 Research questions 2 and 3: Profiles (normal and level 2 deficit) and learning style

If my results are to mirror those of Milton (2007) (see Table 1), we would expect learners with normal profiles to score higher on the analysis dimension (LAT C) and lower on the memory (LAT B) than irregular profilers.

Table 1 Summary of Milton (2007)

| Lexical profile | Learning style | |
|-----------------|----------------|------------------|
| | Memory (LAT B) | Analysis (LAT C) |
| Normal | low | high |
| Level 2 deficit | high | low |

Table 2 Lexical profile and mean LAT B and C scores

| Lexical profile | Learning style | | | |
|---------------------------|----------------|-------|------------------|------|
| | Memory (LAT B) | | Analysis (LAT C) | |
| | Mean | SD | Mean | SD |
| Normal (N = 13) | 45.23 | 30.26 | 48.85 | 9.60 |
| Level two deficit (N = 8) | 37.25 | 25.94 | 53.13 | 9.23 |

In Milton's study, learners with normal profiles scored considerably lower on the memory dimension than learners with level two deficit profiles. Learners with normal profiles scored just under 50% on the analysis dimension, whilst learners with level two deficit profiles scored marginally less. The scores in Table 2, however, do not mirror Milton's findings. Learners with normal profiles do not score significantly differently on the memory or analysis dimensions from level two deficit profilers.

The learners with unclassifiable profiles have not been included because four out of the nine did not report their memory and/or analysis scores. This may have been because they found the tests too difficult and abandoned them. It is interesting to note that the standard deviation (SD) for the memory scores is high. There were some very high memory scores (e.g. 86%) and some very low scores (e.g. 0%), while the analysis test did not produce such a wide variability.

In this section, a binary logistic regression is used to analyse the data to see which independent variable (memory and/or analysis) predicts the category of a normal or level two deficit profile. The Wald statistic is used to test the significance of individual coefficients (B) in the model. Table 3 shows that the contributions of memory and analysis on the binary dependent variable (i.e. 0 = normal or 1 = level two deficit) fail to reach significance. The constant has no practical interpretation but is retained in the model.

Interestingly, the coefficient of memory is negative, which is the reverse of what would be expected from Milton's argument. More importantly, neither the memory nor the analysis coefficient is significantly different from zero, indicating that memory and analysis are not causal factors in detecting normal or level two deficit profiles. In light of these results, I decided to conduct a follow-up study with more participants, to give more weight to any statistical analysis, also categorising them more strictly by their English language level.

Table 3 Binary regression analysis assessing the contributions of memory and analysis to lexical profile

| | | B | S.E. | Wald | df | Sig. |
|---------------------|----------|--------|-------|-------|----|------|
| Step 1 ^a | Constant | -3.391 | 2.799 | 1.467 | 1 | .226 |
| | Memory | -.019 | .018 | 1.035 | 1 | .309 |
| | Analysis | .072 | .057 | 1.601 | 1 | .206 |

^a Variables entered on step 1: Memory, Analysis

3.3 Follow-up study (second cohort)

In my first study, some of the learners were at intermediate level and so had a good knowledge of lexis from the higher frequency bands. This meant that they knew all of the words in the high frequency bands, which resulted in a ceiling effect. The test is therefore not sensitive enough to any subtle differences in their profiles. However, there is scope for *X-Lex* to be used as a test of vocabulary size in relation to memory and analysis strengths. Because vocabulary size is associated with language proficiency, a new cohort of learners, of various L1 backgrounds, was recruited, and knowledge of their language proficiency was obtained.

The 60 participants in this study were from the pre-session English language course at Kingston University. The participants were not a random sample of L2 learners but they were fairly representative of the international students who study at the university. The students in this sample were mostly of Asian origin and the predominant nationalities were Thai, Korean, Chinese and Japanese. The mean age of the participants was 24.1 years (the age range was between 18 and 37) and females outnumbered the males by nearly two to one (39 females and 21 males).

The background data on this new cohort of participants was used to separate learners into high and low proficiency. IELTS 5.5 and below were classified as low proficiency, and IELTS 6.0 and above as high. This was done because the LAT C test requires a good level of English to complete without wild guesswork. All learners who scored over 5% error on the vocabulary test were discounted from the study, which left 22 low proficiency and 30 high proficiency learners. Instead of focusing on normal and level two deficit profiles, a multiple regression analysis of the data was performed in order to understand the contribution of memory and analysis to lexical knowledge.

3.4 Results

The effect size is the R Square, which is the proportion of variance in the vocabulary scores (*X-Lex*) which can be accounted for by the variance in the memory and analysis scores. In relation to low proficiency learners, Table 4 shows that the R Square is .246, which is a large effect size (Kinneer & Gray 2004: 324).

Overall, the regression does a reasonable job of modelling the vocabulary size. About 15% of the variation in *X-Lex* is explained by the model.

Table 4 Multiple linear regression (low proficiency)

| R | R Square | Adjusted R Square | Standard Error of the Estimate |
|-------------------|----------|-------------------|--------------------------------|
| .496 ^a | .246 | .167 | 417.395 |

^a Predictors: (Constant), Analysis, Memory

Table 5 Multiple linear regression assessing the independent contributions of memory and analysis to vocabulary knowledge (low proficiency)

| Coefficients of regression model ^a | | | | | |
|---|-----------------------------|------------|---------------------------|-------|------|
| | Unstandardised Coefficients | Std. Error | Standardised Coefficients | t | Sig. |
| | B | | Beta | | |
| (Constant) | 3029.885 | 399.636 | | 7.582 | .000 |
| Memory | 9.270 | 3.742 | .505 | 2.477 | .023 |
| Analysis | 1.531 | 5.380 | .058 | .285 | .779 |

^a Dependent variable: *X-Lex*

Table 6 ANOVA of the linear regression, modelling *X-Lex* as a linear function of memory and analysis (low proficiency)

| ANOVA ^a | | | | | |
|--------------------|----------------|----|-------------|-------|-------------------|
| | Sum of Squares | df | Mean Square | F | Sig. |
| Regression | 1080760.176 | 2 | 540380.088 | 3.102 | .068 ^b |
| Residual | 3310148.915 | 19 | 174218.364 | | |
| Total | 4390909.091 | 21 | | | |

^a Dependent Variable: *X-Lex*

^b Predictors: (Constant), Analysis, Memory

In Table 5, the beta coefficients (B) are rather more illuminating. We can see that memory makes the bigger contribution to the vocabulary scores (*X-Lex*) and that it reaches significance, whereas the analysis coefficient does not. The model indicates that there is an increase of approximately nine words for every one percent increase on LAT B (memory).

The regression ANOVA assesses the significance of the linear relationship between the variables. In Table 6, the value of F is not significant. This tells us that when the variables are analysed together, *X-Lex*, memory and analysis are not significantly related in a linear manner.

Because of the results shown in Table 6, I decided to run the model again but excluding the analysis value as a predictor since it was not significant in the previous model. In Table 7,

Table 7 One variable regression (low proficiency)

| R | R Square | Adjusted R Square | Standard Error of the Estimate |
|-------------------|----------|-------------------|--------------------------------|
| .428 ^a | .183 | .144 | 417.721 |

^a Predictors: (Constant), Analysis, Memory

Table 8 Coefficient for the regression of the vocabulary score on Memory (low proficiency)

| Coefficients of regression model ^a | | | | | |
|---|-----------------------------|------------|---------------------------|--------|------|
| | Unstandardised Coefficients | | Standardised Coefficients | t | Sig |
| | B | Std. Error | Beta | | |
| (Constant) | 3177.637 | 213.342 | | 14.895 | .000 |
| Memory | 7.613 | 3.510 | .428 | 2.169 | .042 |

^a Dependent Variable: *X-Lex*

Table 9 Multiple linear regression (high proficiency)

| R | R Square | Adjusted R Square | Std. Error of the Estimate |
|-------------------|----------|-------------------|----------------------------|
| .290 ^a | .084 | .016 | 373.525 |

^a Predictors: (Constant), Analysis, Memory

we can see that the R Square is .183, which is a large effect. Overall, it is approximately similar to the previous model when analysis was included.

The beta coefficient (B) in Table 8 is significant at the 5% level and the value of the standardised coefficient shows that a change of one standard deviation on memory produces a change of 0.428 standard deviation on vocabulary score (*X-Lex*).

When the focus is switched to high proficiency learners the results are not statistically significant. Table 9 shows that, for this group, the R Square is .084, which is a medium effect size.

However, neither the memory nor the analysis coefficients are significant at the 5% level.

In Table 11, the regression ANOVA shows that the value of F is not significant. The data for the high proficiency group shows that there is no significant linear dependence of the *X-Lex* score on memory and analysis, so I decided not to run any one-variable regression analyses.

Table 10 Multiple linear regression assessing the independent contributions of memory and analysis to vocabulary knowledge (high proficiency)

| Coefficients of regression model ^a | | | | | |
|---|-----------------------------|------------|---------------------------|--------|------|
| | Unstandardised Coefficients | | Standardised Coefficients | t | Sig. |
| | B | Std. Error | Beta | | |
| (Constant) | 3744.909 | 353.336 | | 10.599 | .000 |
| Memory | 4.346 | 2.991 | .269 | 1.453 | .158 |
| Analysis | 2.261 | 4.873 | .086 | .464 | .646 |

^a Dependent Variable: *X-Lex*

Table 11 ANOVA of the linear regression, modelling *X-Lex* as a linear function of memory and analysis (high proficiency)

| ANOVA ^a | | | | | |
|--------------------|----------------|----|-------------|-------|-------------------|
| | Sum of Squares | df | Mean Square | F | Sig. |
| Regression | 344682.951 | 2 | 172341.476 | 1.235 | .307 ^b |
| Residual | 3767067.049 | 27 | 139521.002 | | |
| Total | 4111750.000 | 29 | | | |

^a Dependent Variable: *X-Lex*

^b Predictors: (Constant), Analysis, Memory

3.5 Discussion

3.5.1 Lexical profiles

The findings from these results are unexpected in light of Milton's (2007) work. In my study, learners with normal or level two deficit profiles do not score any differently on either memory or analysis. The hypothesis that normal profilers may have an analytical learning style and that level 2 deficit profilers may have a memory-orientated style has not been supported. In Milton's study, the difference between memory and analysis scores for level 2 deficit profilers was small; memory was only slightly higher.

One reason why learners with different profiles did not score significantly differently on the tests of memory and analysis in my replication may be the heterogeneity of the learners in the study. In Milton's study, the learners were following the same course and had the same L1 (Greek). In contrast, learners in the replication came from diverse language learning backgrounds and had different L1s. Their L2 learning contexts would be highly diverse. This may have created more 'noise' in the replication data; in other words, the results may be more variable, so any subtle patterns may have been cancelled out in the data.

It is possible that the learners in Milton's study followed the same course and course book, which may have accounted for any lexical dips in their profiles. They may not have been exposed to the wide variety of lexis to which learners are exposed when they are in the country where the target language is spoken (Foster 2009). The implication from Milton is that an analytical approach could somehow make up for any lexical dips (Milton 2007: 52). If learners, as in my study, are exposed to a greater amount of L2 vocabulary, any differences in learning style may be difficult to detect from their profiles. Another reason why no relationship was found between types of profiles and learning style is that some of the learners from my replication may have been at a higher proficiency level than Milton's cohort, which could cancel out any early developmental patterns in lexical knowledge.

We have to reconsider the hypothesis of Milton (2007: 52–53) and Meara & Milton (2003b: 9): that learners with an analytical approach may compensate for a lack of function words, and that learners with a memory-based approach may be influenced by the lexical frequency and content of some course books. This notion may still be valid, but the tests used in this study are not sensitive enough to pick up this relationship between lexis and style with a more heterogeneous group of learners. In my replication study, however, the issues surrounding function words remain unresolved, though in productive profiles we find that analysis and stability of function words are related (Booth 2011).

3.5.2 Methodological issues: Testing for learning style

Informal interviews with a sub-sample of students revealed that they found the vocabulary profile test (*X-Lex*) user-friendly and enjoyable to do. However, the test for an analytic learning style (LAT C) proved to be more problematical for some of the students: they found the test to detect grammatical rules more taxing, especially the elementary learners, who struggled to understand the various grammatical constructions in English. From the qualitative feedback on the test provided by informants, many expressed the opinion that the analysis (LAT C) test was harder than the memory (LAT B) test. Four of the analysis test scores are missing, which could indicate that the test was too hard for these learners to complete. What is more, these tests are all in English (apart from the pseudo language on which the tests are based). Learners need a good level of English in order to complete them. The memory test simply presents new pseudowords for the learners to remember, whereas the analysis test presents sentence structures in both of the languages for learners to detect grammatical patterns, which could be too taxing for the learners. Testing learners' analytical style via LAT C requires them to have a certain amount of proficiency in English grammar in order to detect the grammatical rules in the new language.

The results showed that there was a much higher standard deviation for the memory test than for the analysis. This may have had something to do with the scoring systems for both tests. With the memory test it is possible to get every word wrong, which may explain why there was a higher SD for the memory test than the analysis test. The analysis test, however, offers a choice of two answers – one is right, the other wrong – so a learner could score 50% simply by guessing the answer. Meara et al.'s (2001) normative data for this test puts a score of 0–49 into the bottom 10% of all scores for this test. This calls into question Milton's

interpretation of his results; for example, the mean score for learners with normal profiles (i.e. those who do well on tests of analytic ability) is under 50%. The mean score is in the bottom 10% and suggests that, taken as a group, they do not score particularly well.

3.5.3 Methodological issues: *X-Lex*

The vocabulary test also needs to be examined to ensure that there is a sufficient proportion of function words in the 1k frequency band for the learners to be tested on. A quick calculation from the list of English words used in the test shows that there are 18 function words¹ out of 100 lexical words in the 1k frequency band and one function word in the 2k frequency band. Because the test randomises the words used, it is possible that the learners are not tested on any function words. On average, though, learners would be tested on about four function words each time they took the test, so from the distribution of function words in the test, a learner profile that displays a level two deficit is probably not indicative of a deficiency in functional vocabulary.

In this replication study, learners were classified with a level 2 deficit even if there was only a one-word deficit between the 2k frequency band and the 3k frequency band. Learners are tested on 20 words for each frequency band, so knowing only a single word more in the 3k band than the 2k band would be enough to tip the balance. Milton (2007) also recognised the problem with the test insofar as a single shift in the mark can change the profile. Milton tried to eradicate this problem by testing the learners' vocabulary knowledge twice, so as to obtain stable profile scores. The majority of learners, 21 out of 29, retained stable profiles in his study. With my cohort of learners there was insufficient time to test learners twice so, because Milton's study showed that learners with unstable profiles tended to have higher error scores than those with stable profiles, learners with error scores over 5% were eliminated.

3.5.4 Memory and its association with lexical size

From the post-replication study, a multiple linear regression showed that it is memory, rather than analysis, that can contribute to vocabulary size. However, the effects were only significant with low proficiency learners. One possible explanation for this is that L2 learners who are less proficient in L2 are more dependent on memory-based learning for increasing their vocabulary size. Analysis has a less direct relationship to vocabulary size for either set of learners. A reasonable assumption is that the ability to memorise words has a clearer relationship to lexical knowledge than the ability to infer grammatical patterns, and that this relationship becomes non-linear with greater proficiency.

There is no doubt that learners differ in the vocabulary they acquire. However, it could be factors such as L1 to L2 mapping of the meaning and form, and the salience of the word (Ijaz 1986; Sjöholm 1998) that influence which words are acquired, and when. This may explain why profiles obtained from learners of a similar L1 background may differ from profiles taken

¹ Based on Nation (2001: 430–431).

from other learners of mixed L1 backgrounds, even if they are at a similar proficiency level. Estimates of lexicon size based on frequency measures are only as good as the frequency lists on which they are based. For example, a word which is frequent in one context could be rare in another, which is why more recent studies are starting to use word frequency lists based on corpora from a wide variety of contexts (Lindqvist, Bardel & Gudmundson 2011). Measures of vocabulary size based on the word lists used by Meara & Milton's *X-Lex* (2003a) may still give a valid score out of the most frequent 5,000 words, but it is unwise to relate any dips in the one or two thousand frequency bands to learning style differences.

In fact, a limitation of this study is that no account has been taken of the qualitative aspect of vocabulary knowledge. Learners are tested on vocabulary size but not on how they use lexis. In other words, vocabulary size is just one dimension of vocabulary knowledge. How learners use lexis in context could be quite different from one learner to another even if they have lexicons of similar size. Knowledge of rare lexis does not necessarily mean that a learner has a sophisticated understanding of words. In English, one of the features of sophisticated language use is the complex usage of function words to grammaticise language.

3.6 Conclusions

This replication study sought to establish a relationship between learners' lexical profiles and learning styles. Although Milton's (2007) study broadly supported a relationship between normal profiles and analysis, and level two deficit profiles and memory scores, the replication did not. Instead, my study found that the normal or irregular profilers did not differ significantly in mean memory or analysis scores. A follow-up study to the replication highlighted how memory, rather than analysis, contributes to vocabulary knowledge. In this study, I argued that an estimate of the number of words a learner recognises and memorises shows that memory and vocabulary size are more likely to be related because the figure is based on the total score. The scaling-up of known words at each frequency level may exaggerate any irregular dips in profiles.

Good associative memory could be helpful for L2 vocabulary knowledge but may not tell us whether words are used accurately or appropriately. Although high memory learners can have a deficit in their lexical profiles, they may simply be able to acquire rare lexis more easily, which could tip the balance in favour of a higher 3k band than a 2k band when tested by *X-Lex*. Milton's explanation was that they lacked function words; however, it is more likely that they just have more random profiles, regardless of their ability to learn function words. The *X-Lex* test of vocabulary size is not adequate for detecting deficits in functional vocabulary; instead, it performs better when estimating vocabulary size. In light of the results, we can say that vocabulary size and memory for paired associates are related and that this relationship is sensitive to the learners' responses to the test items.

Further work is needed in order to establish whether the relationship between memory and vocabulary size exists in other languages. Skehan (1998: 218) argues that exceptional L2 learners are those with very good memories. This replication has given some evidence to support the notion that memory, however difficult it is to measure, is an important component of learning style in relation to vocabulary size. There may well be a certain threshold in

memory ability below which some learners will find it difficult to acquire large vocabularies. More evidence could help us to know whether learners with exceptional memories can become native-like in their store of lexical items and use of language.

Acknowledgements

I would like to thank Clarissa Wilks, Vince Daly and Gordon Hunter for their helpful suggestions. In addition, I wish to thank Graeme Porte for his support and the three *Language Teaching* reviewers for their insightful comments.

References

- Ben Maad, M. R. (2010). Holistic and analytic processing modes in non-native learners' performance of narrative tasks. *System* 38, 591–602.
- Booth, P. (2011). The interplay between lexis and learning: A study of second language vocabulary profiles and learning style. Ph.D. dissertation, Kingston University.
- Carroll, J. B. (1981). Twenty five years of research on foreign language aptitude. In K. Diller (ed.), *Individual differences and universals in language learning aptitude*. Rowley, MA: Newbury House, 83–118.
- Coffield, F., D. Moseley, E. Hall & K. Ecclestone (2004). Learning styles and pedagogy in post-16 learning: A systematic and critical review. *Learning and Skills Research Centre*. www.leerbeleving.nl/wp-content/uploads/2011/09/learning-styles.pdf
- David, A. (2008). Vocabulary breadth in French L2 learners. *The Language Learning Journal* 36.2, 167–180.
- De Bot, K., W. Lowie & M. Verspoor (2007). A Dynamic Systems Theory approach to second language acquisition. *Bilingualism: Language and Cognition* 10.1, 7–21.
- Fairclough, M. (2011). Testing the lexical recognition task with Spanish/English bilinguals in the United States. *Language Testing* 28.2, 273–297.
- Foster, P. (2009). Lexical diversity and native-like selection: The bonus of studying abroad. In B. Richards, H. M. Daller, 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, UK: Palgrave Macmillan, 91–106.
- Harley, B. & D. Hart (1997). Language aptitude and second language proficiency in classroom learners of different starting ages. *Studies in Second Language Acquisition* 19, 370–400.
- Harrington, M. & M. Carey (2009). The on-line Yes/No test as a placement tool. *System* 37, 614–626.
- Hindmarsh, R. (1980). *Cambridge English lexicon*. Cambridge: Cambridge University Press.
- Ijaz, I. H. (1986). Linguistic and cognitive determinants of lexical acquisition in a second language. *Language Learning* 36, 401–451.
- Kinnear, P. R. & C. D. Gray (2004). *SPSS 12 made simple*. Hove, UK: Psychology Press.
- Laufer, B. & I. S. P. Nation (1995). Vocabulary size and use: Lexical richness in L2 written production. *Applied Linguistics* 16.3, 307–322.
- Lindqvist, C., C. Bardel & A. Gudmundson (2011). Lexical richness in the advanced learner's oral production of French and Italian L2. *International Review of Applied Linguistics in Language Teaching* 49.3, 221–240.
- Mackey, W. (1965). *Language teaching analysis*. London: Longman.
- Meara, P. M. (1996). The dimensions of lexical competence. In G. Brown, K. Malmkjaer & J. Williams (eds.), *Performance and competence in second language acquisition*. Cambridge: Cambridge University Press, 33–53.
- Meara, P. M. & B. Buxton (1987). An alternative to multiple choice vocabulary tests. *Language Testing* 4, 142–154.
- Meara, P. M. & J. L. Milton (2003a). *X_Lex: The Swansea Vocabulary Levels Test. v2.00*. Swansea, UK: Lognostics.

- Meara, P. M. & J. L. Milton (2003b). *The Swansea Vocabulary Levels Test: The Manual*. Newbury, UK: Express Publishing.
- Meara, P. M., J. L. Milton & N. Lorenzo-Duz (2001). *Language aptitude tests*. Newbury, UK: Express Publishing.
- Milton, J. L. (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*. Cambridge: Cambridge University Press, 47–58.
- Milton, J. L. & P. M. Meara (1995). How periods abroad affect vocabulary growth in a foreign language. *ITL Review of Applied Linguistics* 107/108, 17–33.
- Milton, J. L. & P. Vassiliu (2000). *Frequency and the lexis of low level EFL texts*. Proceedings of the 13th International Symposium on Theoretical and Applied Linguistics, Aristotle University, Greece, 444–455.
- Mochida, A. & M. Harrington (2006). The Yes/No test as a measure of receptive vocabulary knowledge. *Language Testing* 23.1, 73–98.
- Nation, I. S. P. (1984). *Vocabulary lists: Words, affixes and stems*. English Language Institute, Victoria University of Wellington, New Zealand.
- Nation, I. S. P. (1990). *Teaching and learning vocabulary*. New York: Heinle & Heinle.
- Nation, I. S. P. (2001). *Learning vocabulary in another language*. Cambridge: Cambridge University Press.
- Palmer, H. E. & M. West (1937). Discussion: Word frequency. *Modern Languages* 81.3, 136–138.
- Robinson, P. (1997). Individual differences and the fundamental similarity of implicit and explicit adult second language learning. *Language Learning* 47.1, 45–99.
- Robinson, P. (2001). Individual differences, cognitive abilities, aptitude complexes and learning conditions in second language acquisition. *Second Language Research* 17.4, 368–392.
- Schmitt, N. & P. M. Meara (1997). Researching vocabulary through a word knowledge framework. *Studies in Second Language Acquisition* 19.1, 17–36.
- Sjöholm, K. (1998). A reappraisal of the role of cross-linguistic and environmental factors in lexical L2 acquisition. In K. Haastrup & Å. Viberg (eds.), *Perspectives on lexical acquisition in a second language*. Lund, Sweden: Lund University Press, 209–237.
- Skehan, P. (1986). Cluster analysis and the identification of learner types. In V. Cook (ed.), *Experimental approaches to second language acquisition*. Oxford: Pergamon Press, 81–94.
- Skehan, P. (1998). *A cognitive approach to language learning*. Oxford: Oxford University Press.
- Tight, D. G. (2010). Perceptual learning style matching and L2 vocabulary acquisition. *Language Learning* 60.4, 792–833.

PAUL BOOTH is a Senior Lecturer at Kingston University in the UK, where he teaches courses in TESOL. His research includes second language vocabulary and individual differences in language learning. His doctoral work focussed on second language vocabulary and its relationship to learning style. His work has been published in the *Malaysian Journal of ELT Research* (2010) and as a chapter in a book entitled *Issues in second language proficiency* (Continuum, 2009).

Comment from the author of the original study

Professor James Milton University of Swansea, UK

J.L.Milton@swansea.ac.uk

Much of our understanding about the learning of foreign language vocabulary is drawn from the analysis of the performance of groups of learners. The initial purpose of the original paper, on which this approximate replication is based, was to demonstrate the widely-made assertion that the learning of the foreign language lexicon is heavily influenced by word frequency and that frequent words tend to be learned before less frequent ones. The frequency profile, demonstrating the linear relationship between frequency and learning, is easy to

demonstrate – in groups of learners. But researchers in vocabulary learning are all too aware that general assertions like this disguise very considerable individual variation, and we have much less research to explain why individuals vary in the way that they do. The second focus of the original paper, therefore, was much more speculative in attempting to investigate whether learning style – a memory facility as opposed to an analytic facility – might influence the words that were learned and the frequency profile that individual learners produced. The results suggested that learning style could influence which words were learned. But it was a very small study, based on lexical uptake among young learners in one school in one country. It is always difficult to generalise convincingly from this kind of data even if the conclusions appear to make sense in so far as we understand the learning process. Repeat experiments and replications are therefore a very valuable activity.

The approximate replication reported here is very interesting, not least because it fails to reproduce the results and conclusions of the original experiment. The relationship whereby learning style can govern in part the words learners acquire need not be as simple as the first experiment suggested, and it challenges us again to find a better explanation for the variation in learning which is such a feature of the development of the lexicon.

There is an important way, however, in which the results reported here are not in complete conflict with the results and conclusions of the original study. The original paper noted that the uneven lexical profiles being investigated were very much a feature of very low-level learners. It might be speculated that exposure to the words of the foreign language at this level might be sufficiently small and idiosyncratic that the normal frequency patterns we associate with larger-scale corpora are not so marked, allowing learners far greater latitude in their learning than might otherwise be the case. As learning progresses, and exposure to a wider variety of texts increases, normal frequency profiles assert themselves. The learners in the replication are rather older, are living in the foreign language environment, which gives them huge access to normal English, and appear to be at a significantly higher level of foreign language ability than the learners in the original study. It is a group in which, perhaps, one would not expect to see the effects of learning style as described in the original experiment. That is probably a useful caveat to the original study and this replication usefully points up the importance of the relationship between learning and frequency of occurrence which underlies both studies.