

Retention of foreign vocabulary learned using the keyword method: a ten-year follow-up

Alan Beaton *University of Wales, Swansea,*
Michael Gruneberg *University of Wales, Swansea and*
Nick Ellis *University of Wales, Bangor*

This article assesses one individual's level of recall for foreign vocabulary learned ten years previously using the keyword method. Without any revision at all, he remembered 35% of the test words with spelling fully correct and over 50% with only very minor errors of spelling. After 10 minutes spent looking at a vocabulary list, recall increased to 65% and 76% respectively. After a period of revision lasting a further 1½ hours, recall was virtually 100%. This level of recall was maintained for at least one month. The results indicate 1) that the keyword method (as incorporated in Linkword courses) may be used to learn a large list of vocabulary; and 2) this method of learning is not inimical to retention in the long term. Some theoretical aspects of the findings are discussed.

I Introduction

A major problem facing the second-language learner is to acquire a sufficiently large vocabulary to be able to communicate effectively. This applies however the language is taught, whether by total immersion, rote learning or any other method or combination of methods, and whatever the purpose of learning the language. Any method that can promote the learning and retention of vocabulary beyond what can be naturally 'picked up' is likely to be of help to those learners who can use it.

One method of vocabulary learning which has received a good deal of experimental study is known as the keyword method. In essence, this involves using interactive imagery to link the sound of a word in one's native language (the keyword) to the sound of some foreign word which has to be learned. For example, the Italian for frog is *rana*. A learner might be required to imagine in his or her mind's eye that he/she *ran* a mile after seeing a horrible *frog*. The idea is that in attempting to retrieve the Italian word for frog a

Address for correspondence: Dr A. Beaton, Department of Psychology, University of Wales, Swansea, Singleton Park, Swansea SA2 8PP, UK.

© Edward Arnold 1995

native English speaker would remember the relevant image and this would lead to recall of the Italian word *rana*. A large number of published studies show that this technique is superior to rote learning in enhancing the retention of foreign words (Pressley *et al.*, 1982; Cohen, 1987; McDaniel *et al.*, 1987; Gruneberg and Jacobs, 1991: see Gruneberg, 1992, for a review). Two points should be noted at this juncture. First, it is not essential that the keyword has exactly the same sound as the foreign word to be learned. A near approximation may be enough to stimulate recall of the foreign word. Secondly, the experience of the many learners known to the authors of this article is that the keyword method assists in the early stages of learning when one attempts to remember previously unknown words. Once these words have been learned sufficiently well, learners no longer need rely upon the image specified by the keyword.

It is not the purpose here to argue for the merits of the keyword system or any other method, and we certainly would not wish to argue that the keyword method should not be used in conjunction with, for example, 'communicative' methods of language teaching. Our purpose in writing this article is to show that where large amounts of vocabulary have been learned in a short period of time using the keyword method, this vocabulary can be recalled several years later either 'on demand' or after comparatively little time spent in revision. All studies reported to date on the keyword method concern short retention intervals, usually one month or less. The number of vocabulary items is typically less than 40. In contrast, our study involves retention of a 350-word vocabulary over 10 years.

II Case study

The article reports a single case study of a 47-year-old male university lecturer (hereafter referred to by the initials N.P.) who had reason to learn some Italian ten years previously. The method he adopted was to work through the Linkword Italian course (subsequently published by Gruneberg, 1987) which took him about 10 hours. This course uses the keyword method of vocabulary learning integrated with basic grammar. Learners using the Linkword book are presented with successive groups of Italian vocabulary items together with sentences relating each item to its English translation by means of a suitable keyword; at the end of each group of words instructions are given for recalling each of the items in turn. Simple sentences which incorporate the grammar are also presented for translation at intervals. The learner works through the material in

sections and then uses the accompanying audiocassette tape to listen to the correct pronunciation of each Italian word presented in the preceding section. N.P. did not use the audiocassette tape and thus his performance can be regarded as being based on an incomplete exposure to the course material. Nonetheless, he is confident that at the time of initial learning in 1983 his retention of the entire vocabulary of approximately 350 words was very close to 100%. In the event, N.P. had no opportunity to use Italian and had no further reason to rehearse what he had learnt from the Linkword course prior to this study.

In 1993 N.P. made a brief trip to Italy. He was interested to know how much Italian he could remember and therefore set himself the task of working through the list of vocabulary given at the end of the Linkword course he had read ten years previously. The vocabulary provides a list of 350 (English) words. N.P. looked at each English word, while keeping the Italian translations covered with a sheet of blank paper, and attempted to write down the Italian for each word in turn (Recall 1). After working through the entire list, N.P. spent 10 minutes studying the list of vocabulary, this time with the Italian words uncovered, before again covering the Italian words and making a second attempt to write them all down on seeing the English equivalents (Recall 2). This was followed by N.P. spending 1½ hours silently working through the entire course, including grammar sections and sentences for translation. Finally, N.P. made a third attempt to write down all the Italian words in the vocabulary list (Recall 3). The following day, and by now having arrived in Italy, N.P. made a fourth recall attempt before leaving his hotel. As this (and a subsequent recall attempt made one month later) yielded data that were almost identical to Recall 3, only data from the first three recall attempts are provided in Tables 1, 2 and 3.

N.P. obtained a good pass in A-level French in 1966 and has used French intermittently since. He has also used the Linkword method at various times since 1985 to learn some German, Spanish and Greek prior to going on holiday. He has no specialist knowledge of languages and does not regard himself as being especially gifted in terms of language learning ability.

III Results

The written recall attempts made by N.P. were scored in two ways. The first used the strict criterion that each word had to be fully correct. The second criterion was slightly more lenient in allowing minor errors of spelling (e.g., giving double letters – such as two *lls* – where one was correct, or vice versa, or substitution of ‘similar

sounding' letters – e.g., writing *v* instead of *b* in *cavallo*, *c* for *g* in *segretaria* or *a* for *o* in *castello*, or omission of one letter – e.g., of the *i* in *piede* – provided the correct word would probably have been guessed correctly by a listener).

Because of their similarity to French (and other languages) and the distinct possibility of being correctly guessed, the days of the week, months of the year and numbers were excluded from the list.

The data for nouns, verbs, adjectives and other parts of speech combined are shown separately in Tables 1, 2 and 3. The number in brackets refers to the value as a percentage of the maximum possible. Table 1 gives data for Recall 1 (after 10 years), Table 2 for Recall 2 (after looking at the vocabulary list for 10 minutes following Recall 1) and Table 3 shows data for Recall 3 (after reading through the entire course following Recall 2).

Table 1 Number of words correctly recalled – Recall 1

| | Completely correct: criterion 1 | Almost correct: criterion 2 |
|------------|---------------------------------|-----------------------------|
| Nouns | 74 (35.1%) | 113 (53.6%) |
| Verbs | 6 (22.2%) | 12 (44.4%) |
| Adjectives | 21 (48.8%) | 23 (53.5%) |
| Other | 11 (35.5%) | 17 (54.8%) |
| Total | 112 (35.2%) | 165 (51.9%) |

Table 2 Number of words correctly recalled – Recall 2

| | Completely correct: criterion 1 | Almost correct: criterion 2 |
|------------|---------------------------------|-----------------------------|
| Nouns | 146 (69.2%) | 170 (80.6%) |
| Verbs | 14 (51.9%) | 18 (66.7%) |
| Adjectives | 33 (76.7%) | 33 (76.7%) |
| Other | 12 (38.7%) | 17 (54.8%) |
| Total | 205 (65.7%) | 238 (76.3%) |

Table 3 Number of words correctly recalled – Recall 3

| | Completely correct: criterion 1 | Almost correct: criterion 2 |
|------------|---------------------------------|-----------------------------|
| Nouns | 199 (94.3%) | 211 (100.0%) |
| Verbs | 24 (88.9%) | 25 (92.6%) |
| Adjectives | 42 (97.7%) | 43 (100.0%) |
| Other | 30 (96.8%) | 31 (100.0%) |
| Total | 295 (94.6%) | 310 (99.4%) |

IV Discussion

The data in Table 1 show that after a period of ten years, without any exposure whatever to Italian, N.P. recalled 112 words (of a maximum possible of 312) entirely correctly. If allowance is made for very minor spelling errors, the number of words recalled is 165 with every grammatical class represented. It is interesting to note that before commencing this exercise N.P. estimated that he would recall 'about two dozen' Italian words. This reflects the fact that individuals tend to underestimate the size of their vocabulary because they do not appreciate *which* words they are capable of retrieving until they are put to the test.

The data in Table 2 show that after spending only 10 minutes looking at a vocabulary list, N.P.'s level of recall increased to a total of 205 words (or 238 on the more lenient criterion). In other words, total correct recall increased by 83%. Since it is unlikely that N.P. relearned 93 words from scratch in just 10 minutes, these data suggest that the amount of information at one's disposal may be greater than is revealed by conventional recall attempts. A distinction may be made between storage and retrieval of knowledge. With appropriate retrieval cues, such as are provided by reading a list, access can be made to knowledge that has been *retained* but not necessarily *recalled*.

The data provided by this case study strongly imply that one of the criticisms often made of rapid learning methods such as Linkword, namely that material that is quickly learned is quickly forgotten, is misguided. What is probably more to the point is not how *quickly* a vocabulary has been learned but how *well*. To judge from the outcome of this study, material that has been learned well initially can be relearned in a fraction of the time required for the original learning. Of course, it is not claimed that all language learners would do as well in absolute terms as N.P., nor that N.P. would not have done equally as well using some other method or combination of methods. What *is* claimed is that use of the keyword method of vocabulary learning, far from being detrimental to subsequent relearning, appears to promote rapid relearning as and when the material is required.

Our results can be compared with those of Bahrick (1984), who studied memory for Spanish learned as a second language at school many years prior to testing. Bahrick showed different degrees of forgetting during the first 3–6 years after learning but then virtually no loss of material for up to 25 years thereafter. The level of recall varied with the degree of initial competence in Spanish, but amount of retention was not related to opportunities for rehearsal. Bahrick

(1984: 23) argued that though recall of vocabulary declines relatively quickly at first, and then stabilizes 'very significant portions of semantic memory remain perfectly accessible for decades without being used at all'. The findings of our own study provide some support for Bahrick's conclusion.

We wish now to discuss briefly an aspect of the data that is not usually considered in studies of vocabulary learning. Does anything distinguish those words that N.P. remembered at the first attempt from those that he did not recall? For this purpose we shall consider only nouns and will use the data from the more lenient criterion.

A feature of the Linkword courses is that, where the foreign word to be learned is very similar to the native language equivalent (e.g., *banca* for bank), a particular convention is adopted. In the case of Italian, the learner is instructed to imagine a scene in which spaghetti is associated in some way with the item to be learnt (e.g., a bank stuffed full of spaghetti!). Of the 113 Italian nouns recalled by N.P. before looking at the vocabulary list, 27 sounded similar to their English translations, and imagery instructions therefore involved spaghetti. Of the 109 nouns not remembered, seven involved spaghetti. However, the difference almost certainly has nothing to do with the imageability of Italian pasta! Other authors have pointed out the relative ease with which cognate words may be learned (Banta, 1981; Nation, 1982). While this implies that similarity to English can explain N.P.'s level of 'spontaneous' recall for a good proportion of Italian words, clearly it can not account for more than 25% of them.

Given that N.P. initially used the keyword method (incorporated into the Linkword course) to learn an Italian vocabulary, it occurred to us that some difference in the effectiveness of the keywords might explain why some words were recalled and others were not.

Ellis and Beaton (1993; 1994) have shown that nouns tend to be more effective as keywords than verbs. Once 'spaghetti' images are excluded (plus five words such as *vino* which N.P. might have been expected to know even without learning Italian), 44 of the remaining 81 items recalled by N.P. had noun keywords compared with 49 of 102 nouns not recalled. The percentage of noun keywords was thus very similar for each of the two groups of words. Some other factor, therefore, such as the idiosyncratic memorability of the keyword images, presumably explains why some words were recalled and others were not.

It is likely that spontaneous associations between the Italian word and similar sounding English words having a semantic relation to the Italian word facilitate recall. Examples are *mano* [hand]

– manacles; *cavallo* [horse] – cavalry; *avvocato* [lawyer] – advocate; *tappeto* [carpet] – tapestry. When asked if he could think of such ‘obvious’ semantic associations, N.P. found that he was sometimes unable to produce appropriate associations rapidly. Examples where this occurred included 30 of the words he had not remembered initially (e.g., *topo* [rat] and *cucchiaio* [spoon]). On the other hand, he could think readily of semantic associations for all but 12 words (such as *ucello* [bird]; *crizzo* [jack]; *specchio* [mirror] he had remembered. N.P. stated that he remembered the keywords for these latter words as well as for some with more ‘obvious’ associations. This is unsurprising since the Linkword system is deliberately designed to capitalize, where appropriate, on pre-existing associations.

It is also probable that in some cases N.P.’s knowledge of French helped him to recall Italian words that were similar to their equivalents in French but not in English (e.g., *mano* [hand] – French *main*; *freno* [brake] – French *frein*; *armadio* [cupboard] – French *armoire*). However, there was an equally high proportion of Italian words with similar French translations among those words that N.P. failed to remember as among those that he did recall. In addition, foreign equivalents can mislead as well as assist. For instance, N.P. recalled the Italian for cow as ‘vacca’ (from *vache*?) rather than *mucca* and glass as ‘verro’ (from *verre*?) rather than *bicchiere*. There were also some intrusions from a Spanish course that N.P. had learned (in 1985). Indeed, one of the examples just given (*vacca*) of mistakes made by N.P. might derive from Spanish rather than, or in addition to, French.

The overall point we wish to make is that it would not be reasonable to attribute what, in our view, is N.P.’s impressive level of recall after an interval of ten years simply to his existing framework of knowledge. On the other hand, we would not want to deny that this contributed to his performance. What we do maintain is that, as a result of his learning an Italian vocabulary of over 350 words using the keyword method, N.P. was able to remember approximately one half of these words after an interval of ten years. Furthermore, he reached a level of nearly 100% correct recall after little more than 1½ hours revision, and this level was sustained for at least a further month. This seems to us evidence against the view that methods for learning large amounts of vocabulary quickly are of little or no benefit in the long term.

We shall end on a speculative note. Ellis and Beaton (1994) reported that there was a significant negative correlation between the time taken by one group of British adults to pronounce a list of German words and recall of these same words by a second group who were asked to learn them. The more quickly words were pro-

nounced, the more easily they were learnt. This was not simply a reflection of word length. In a longitudinal study, Gathercole and Baddeley (1989) demonstrated that 5-year-old children's native vocabulary scores were predicted by their ability to repeat non-words one year earlier. In a later study the same authors showed that children who were poor at nonword repetition were slower at learning a novel vocabulary than those children who were good at nonword repetition (Gathercole and Baddeley, 1990). Taken together, what the three studies suggest is that (all else being equal) those *learners* who can most efficiently articulate novel material are those who will learn it most quickly and, conversely, the more readily unfamiliar *material* can be pronounced the more quickly it will be learned. The question therefore arises as to whether any relationship can be demonstrated between the speed with which items of vocabulary can be articulated and the level of retention for material that can be considered to be well established in long-term memory.

N.P. was asked on four occasions to read aloud as fast as possible 1) the list of nouns which had remembered at the initial recall attempt (Recall 1) and 2) those nouns which he failed to recall. The lists were read in balanced order. Average reading time for words recalled was consistently shorter than for the other words. Averaged over the four occasions, the mean reading time for the list of words recalled was 633 milliseconds per word and the mean reading time for the unremembered words was 646 milliseconds per word. When Italian words that were similar in sound to their English translations were omitted from each list, the difference in reading time still favoured the list of words recalled by N.P. Although the mean number of syllables per word was 2.25 for words recalled and 2.71 for words not recalled, the median for both lists was 3 and the distributions of words classified by syllable length did not differ significantly (chi square <1) between the two lists.

Clearly, we would not wish to place too much emphasis on these findings from one individual. It would be as appropriate to argue that the words of the first list were read more quickly *because* they were remembered well as to argue that they were better retained in memory because they could be easily or quickly articulated. Nonetheless, in the context of the above discussion, the findings might be seen as lending weight to the idea that phonotactic factors are not only important in the short-term acquisition of material but also have some influence on the retention of material over the long term as well.

V References

- Bahrlick, H.** 1984: Semantic memory content in permastore: fifty years of memory for Spanish learned in school. *Journal of Experimental Psychology: General* 113, 1–29.
- Banta, F.G.** 1981: Teaching German vocabulary: the use of English cognates and common loan words. *Modern Language Journal* 65, 129–36.
- Cohen, A.D.** 1987: The use of verbal and imagery mnemonics in second-language vocabulary learning. *Studies in Second Language Acquisition* 9, 43–62.
- Ellis, N. and Beaton, A.** 1993: Factors affecting the learning of foreign language vocabulary: imagery keyword mediators and phonological short-term memory. *Quarterly Journal of Experimental Psychology* 46A, 533–56.
- 1994: Psycholinguistic determinants of foreign language vocabulary learning. *Language Learning* 43, 559–617.
- Gathercole, S. and Baddeley, A.** 1989: Evaluation of the role of phonological STM in the development of vocabulary in children: a longitudinal study. *Journal of Memory and Language* 28, 200–213.
- 1990: The role of phonological memory in vocabulary acquisition: a study of young children learning new names. *British Journal of Psychology* 81, 439–54.
- Gruneberg, M.** 1987: *Linkword Italian*. London: Corgi.
- 1992: The practical application of memory aids: knowing how, knowing when and knowing when not. In Gruneberg, M. and Morris, P., editors, *Aspects of memory. Vol. 1. The practical aspects*. London: Routledge, 168–95.
- Gruneberg, M. and Jacobs, G.** 1991: In defence of Linkword. *Language Learning Journal* 3, 25–29.
- McDaniel, M.A., Pressley, M. and Dunay, P.K.** 1987: Long-term retention of vocabulary after keyword and context learning. *Journal of Educational Psychology* 79, 87–89.
- Nation, I.S.P.** 1982: Beginning to learn foreign vocabulary: a review of the research. *RELC Journal* 13, 15–36.
- Pressley, M., Levin, J.R. and Delaney, H.D.** 1982: The keyword method and foreign word acquisition. *Review of Educational Research* 52, 61–91.