



# Should the Keyword Method be introduced in Tertiary Foreign Language Classrooms?<sup>1</sup>

**Kazuyo Taguchi**

([kazuyo.taguchi@unisa.edu.au](mailto:kazuyo.taguchi@unisa.edu.au))

University of South Australia & University of Adelaide, Australia

---

## Abstract

Studies on the effect of the keyword method, a mnemonic technique that uses two links (acoustic and imagery) to facilitate memory for acquiring foreign words, are numerous. The majority of past studies found it effective for various age groups of learners for learning words of different languages. The findings however, have been inconclusive in: (a) its application to classroom settings; (b) its usefulness for older or experienced learners; and (c) its effectiveness for a longer period of time. These three inconclusive results indicate a missing link between everyday teaching practice and this theory-driven method.

The main aim of the present study is to fill this gap by investigating whether the keyword method could be successfully applied to an extant university Japanese language classroom for a period of one semester (half-year). The research questions are: (1) Does the use of the keyword method improve vocabulary test scores?; (2) Do tertiary students take up the keyword method for their vocabulary learning?

The data collected were from (a) two questionnaires and (b) the average scores of five vocabulary tests while the procedures used for the statistical analysis are: (a) the Rasch concurrent equating process; and (b) Hierarchical Linear Model.

Application of the technique to the classroom context of university students was found to be a predictor for the productive mode of Japanese words with a 0.1 significance level. This level was judged to be beneficial for the method to be introduced in language classrooms. Half of the treatment group and a quarter of the delayed-treatment group students used the method. The reasons these learners gave for using or not using the technique was 'the rule of economy'. That is, if the efforts necessary to use this method were less than that of self-devised techniques, then they were prepared to use it.

---

## 1 Introduction

The keyword method is a mnemonic technique that uses two links for the identification of a keyword: (a) an acoustic link; and (b) an imagery link. The keyword sounds like a part of the word to be learnt, and the imagery link is a visual image of the keyword that interacts with the meaning of the word (Bruning et al., 1999).

The keyword method has been empirically found to enhance the vocabulary learning of learners of various age groups for a variety of foreign languages, such as Spanish (Rough & Atkinson, 1975), German (Ott, Butler, Blake & Ball, 1973), Italian (Lawson & Hogben, 1996, 1998), French (Wang, Thomas & Ouellette, 1992), Russian (Atkins & Rough, 1975), Old English (McDaniel & Pressley, 1989), Tagalog (Wang, Inzana, Primicerio & Thomas, 1993), and Greek (Toumoutzoglou, 1998). Some doubt has been raised, however, about the effectiveness of the keyword

method for memory retrieval in the following two contexts: (a) the classroom application, and (b) with experienced or older language learners. Fuentes (1976) and Levin, Pressley, McCormick, Miller and Shriberg (1979) found that experienced language learners used strategies that were as elaborate and effective as the keyword method and, therefore, the explicit classroom instruction on the keyword method had little effect, although Lawson and Hogben (1998) have refuted this finding. In addition to these two contexts, long-term effect has also been questioned. Long-term forgetting was greater for learners who used the keyword method than for learners engaged in rote rehearsal in the study by Wang and Thomas (1992). The reason for inconclusive results of the method's application to intact second year high school Spanish classes was given by Fuentes as: "What occurs in the classroom is markedly different from the isolation of the psychological laboratory. Apparently, there are too many unknown variables [...]" (1976, p. 88) After replicating the same study, Levin et al. stated: "[...] we are left with a conclusion similar to that of Fuentes: the keyword method, as conceptualised to date cannot be translated directly into practice in actual high school foreign language classes." (1979, p. 590) In the same study (Levin et al., 1979), however, it was found that the method was effective for younger (Grade 5) students as well as for small groups of high school students. The majority of these studies measured the receptive mode (target language → native language) of learning and involved only nouns. In all these aforementioned experiments, learners were required by the experimenter to use the keyword method for between a few days and four weeks.

The present study was conceived when the researcher, working as a teacher, realised that her tertiary students were experiencing great difficulty in learning a vast number of new words in Japanese, and therefore, they needed to learn how to learn effectively. Laufer's (1986, p. 70) statement that "learners themselves claim that lexis is their greatest difficulty in the second language" was echoed by the students in the current study when they indicated that they found learning new words was the most difficult aspect of their Japanese language study. Since lexical knowledge is vital, as suggested by Widdowson (1978), and a message containing correct lexis with wrong grammar is understood better by native speakers than a message with correct grammar and wrong words, meta-strategic knowledge for vocabulary acquisition should be considered part of language learning instruction.

The current study was designed to investigate whether the keyword method should be introduced in tertiary Japanese language classrooms as a response to the suggestion posed by Levin et al. (1979) for the classroom application of a longer duration (a semester or a year) of this mnemonic technique in order to find its pedagogical effects. The following two questions are set as specific research questions for this study.

- (1) Does the use of the keyword method improve vocabulary test scores?
- (2) Do tertiary students take up the keyword method for their vocabulary learning?

The significance of the study lies in the following four areas. First, this study was the first to apply the keyword method in intact tertiary Japanese language classrooms as part of the curriculum. All the other class activities were left unchanged as if the investigation were not taking place, including the timing and content of vocabulary tests consisting of both receptive (Japanese→English) and productive (English→Japanese) modes of learning, as well as of various parts of speech (as with most of the vocabulary tests in language courses) rather than restricting to nouns only (Appendix B). This is to ensure that the usual language curriculum will stay intact unlike laboratory-like experiments. The only addition to the routine was the introduction of the keyword mnemonic and its trial sessions of a short duration (5 to 10 minutes) each week.

Using extant classes limited the reliability of the study, since randomisation of its subjects was not possible, nor was there control over the frequency and length of time students spent on vocabulary learning. These factors may have been the 'many unknown variables operating in classrooms' that Fuentes (1976, p. 88) and Levin et al. (1979, p. 64) had alluded to. The rationale for including the variables (a) time, (b) gender, (c) previous achievement, (d) perceived difficulty, and (e) perceived importance of vocabulary in the current study was to examine some of the plausible 'unknown' factors. Perceived importance and difficulty of vocabulary were chosen since it is possible to hypothesise that students who perceive vocabulary to be important and difficult might devote

more time to vocabulary learning. The gender and students' previous semester's achievements were also examined, since non-randomised groups meant that these factors were not controlled in the study.

The second significant aspect of the study was that this was the first to investigate whether, when left to their own choosing, tertiary students would take up this method for their Japanese vocabulary learning. The keyword method was introduced in class as one of a number of vocabulary-learning strategies, but its use (or non-use) outside the classroom was left to the students themselves.

The third significant aspect of the study was the application of the keyword method to vocabulary acquisition (involving both phonetic and orthographic knowledge) of a character-based language. The effect of this mnemonic had been examined with Chinese (Lu, Mei-Yan, Webb, Krus & Fox, 1999; Wang, Thomas & Ouellette, 1992, 1995), and Japanese (Kuwahara 2000). The focus of these past studies was, however, learning the characters called *kanji*, many of which had pictographic origins, while the pronunciation of the characters, that is, the phonetic knowledge of the words, was not required of the learners. These aforementioned studies on character-based languages did not use the keyword method as defined by Brunning et al. (1999). In the responses to the questionnaire (Appendix A), the participants in the present study indicated that the greatest difficulty in learning Japanese was the acquisition of new words because of the strangeness of their sounds, rather than difficulty with their orthography, including *kanji*. Therefore, an effective strategy was needed most for learning the sounds and pronunciation of new words to which the keyword method is designed to contribute by the use of a phonetic link.

The last significant aspect of the study was its duration, namely, one semester (half a year) of a tertiary academic year. This longitudinal nature was also, at the same time, one of the limitations of this study. The long duration facilitated the possibility of leakage of the keyword method technique from the treatment to the delayed-treatment group members. The other limitations of this present study are: (a) the homogeneous nature and small size of the sample; (b) the fact that neither the delayed effect nor the interactive effects between variables were measured; and (c) missing data and lost cases, inevitable because of the longitudinal nature of the study and the use of intact classes.

## 2 Method

### 2.1 *The participants and their groupings*

The participants in the study comprised some 50 students enrolled in the second semester of a first-year Japanese language course at a university in South Australia. They had done one semester of Japanese as beginners before they enrolled in the course. There were some overseas students: eight from China, and one each from Singapore, Hong Kong, Korea, and Thailand. Students who were literate in the Chinese language found their knowledge of ideographic characters called *kanji* beneficial, since most of them were identical to Japanese in meaning. However, their pronunciation was quite different in Japanese. Students were given a choice of which group out of Class 1 (delayed-treatment) and Class 2 (treatment) to join; therefore, in a way, the groups were formed in a quasi-random fashion. Class 2 received explicit instruction on the use of the keyword method at the beginning of the investigation, followed by weekly trials on one or two words. The ethics committee regulations required the researcher to introduce the method to the control group subjects as well. Thus, the control group was not left controlled and, therefore, Class 1 in this study is labelled as 'delayed-treatment group'.

### 2.2 *Data*

Two sets of information were collected: (a) the scores of vocabulary tests, and (b) responses to two questionnaires (Appendix A) administered to all participants before the beginning of the investigation (Questionnaire 1) and also at the completion of their fourth test (Questionnaire 2).

These questionnaires were designed to examine the following three aspects of the use of vocabulary learning strategies: (a) students' perceptions of the importance and the difficulty of vocabulary knowledge and learning; (b) vocabulary learning strategies that students had used before as well as since the beginning of the investigation; and (c) students' use of the keyword method, namely the frequency of use and reasons for using or not using this method.

The procedure and timing of the investigation are listed below.

(1) Lectures begin	July 29
(2) Questionnaire 1	July 31
(3) Introduction of the keyword method (treatment group)	August 7
(4) Test 1	August 12
(5) Test 2	August 26
(6) Test 3 (= mid-semester written examination)	September 16
(7) Introduction of the keyword method (delayed-treatment group)	October 10
(8) Test 4	October 14
(9) Questionnaire 2	October 21
(10) Lectures finish	November 1
(11) Test 5 (= final written examination)	November 11

All the vocabulary tests (Appendix B) were in written form. Test 5 (item 11), the final written examination, which included all the vocabulary items learnt in the semester, could have served as the delayed test. However, this idea was abandoned because students would most likely be (re-)learning words fresh for this examination and, therefore, it would not be measuring the delayed effect. Three out of these five tests contained 8 to 10 identical vocabulary items that served two purposes: (a) they were used as anchor items in calibration; and (b) they demonstrated to students the importance of cumulative learning.

The main concern for the timing of the introduction of the keyword method to each of the groups was to 'control' the delayed-treatment group as long as possible. Thus the treatment group was introduced to the method in the second vocabulary class of the term, while the introduction to the delayed-treatment group took place in the last vocabulary lecture of the term. This enabled the treatment group to have eight more classes (and two weeks of mid-semester break) to use this method compared to the delayed-treatment group.

The following qualitative and quantitative data were collected in order to obtain between-student variables involved in the investigation (Appendix A). Codes used in the analyses are given in capital letters.

- (a) results of 5 written vocabulary tests of English → Japanese items (scored either 0, 1, or 2 with 48 items), and Japanese → English items (scored either 0 or 1 with 47 items) with some common items,
- (b) TREATMENT: member of the treatment group,
- (c) GENDER: gender,
- (d) PRIOR: prior achievement (the average of vocabulary tests for the previous semester),
- (e) IMPT: the relative importance placed on vocabulary knowledge (most important = 1, all other responses = 0)
- (f) KEYWORD: use of the keyword method for vocabulary learning (Category 0 = not used, Category 1 = used but rarely, Category 2 = used often),
- (g) DIFFT: perception of the difficulty of learning new Japanese words (most difficult = 1, other responses = 0),
- (h) strategies that were used to learn new Japanese words and the frequency of their use,
- (i) reasons for using or not using the keyword method, and
- (j) students' attitudes to the keyword method and vocabulary learning in general.

Of the ten items of data listed above, the data (a)-(g) were used in statistical analyses to answer Research Question 1 and qualitative data (h)-(j) were examined to answer Research Question 2. The average of five vocabulary test scores, that is (a), is the outcome variable, and the remaining, (b)-(g), are hypothesised predictors. In addition to these six variables, Time (student growth over time) was included as a hypothesised within-student (Level 1) predictor for the following three reasons. First is the longitudinal nature of the current study, second is the capacity of Hierarchical Linear Model (HLM) to analyse variables at different levels simultaneously (Bryk & Raudenbush, 1992; Raudenbush, Bryk, Cheong & Congdon, 2000), and the third comes from a language acquisition theory, which suggests that the higher the learners' linguistic proficiency, the easier their language acquisition will be (Nation, 1982, p. 80).

The following is an example of the introduction of the keyword method in class.

### **2.3 Introduction and practice of the keyword method in class**

A set of some 20 new words were introduced by reading them aloud (checking their corresponding meaning on the page at the same time) one by one. When some words proved difficult, concentrated practice of imitating the teacher was repeated phoneme by phoneme. This was followed by a brief discussion on different vocabulary learning techniques. The additional session given only to the treatment group was the introduction of the keyword method. The class was asked which words on the list seemed difficult to learn. Students chose the word れいぞうこ /re:zo:ko/ [refrigerator] and gave the 'strange sound' as their reason. When asked if any part of this word sounded like any English word, a student said that the first two syllables れいぞう /re:zo:/ sounded like a 'razor'. Having identified the keyword *razor* by an acoustic link, students were then to find an imagery link. The teacher asked the class 'Imagine that here are a refrigerator and a razor. Can we somehow put these two images together so that they are remembered together?' One student suggested that it was easily done by placing a sharp-edged razor in the refrigerator. The whole process took less than 7 minutes, after which the class moved on to the next phase, grammar practice. This short session of choosing a word or two and practising the keyword method with them took place every Wednesday. The words that students chose for the class practice (Appendix C) contained a variety of parts of speech because the criterion for selecting the words was their perceived difficulty.

### **2.4 Data analyses**

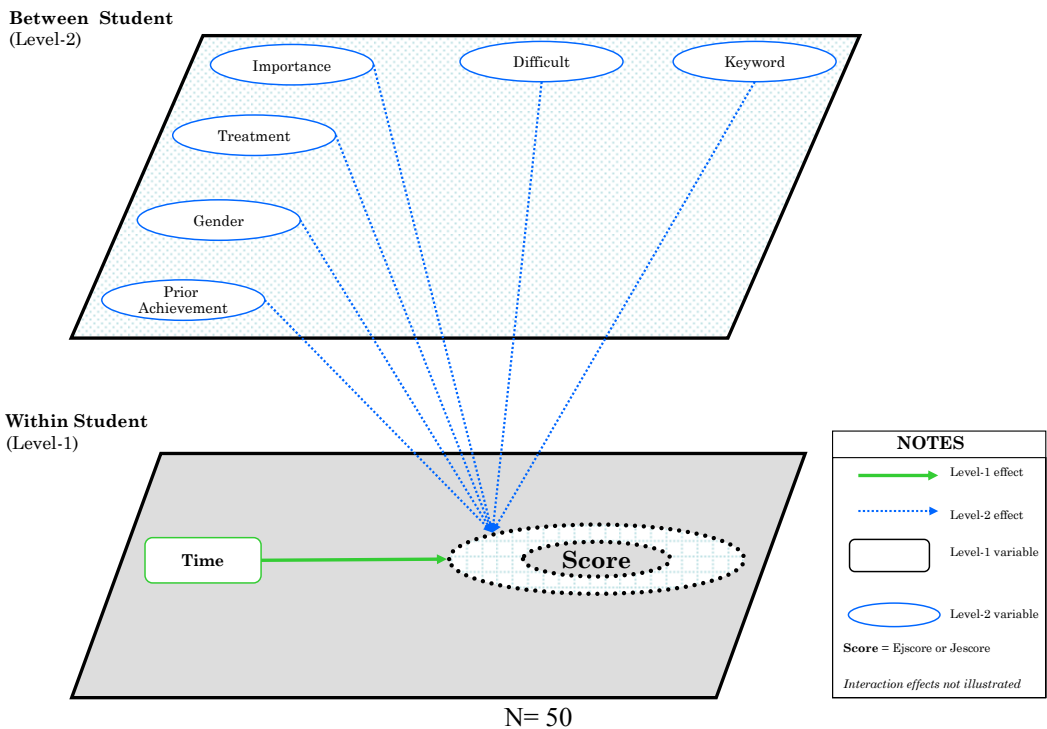
Since the five vocabulary tests whose scores are to be used as the outcome variable in the investigation, were not created to be equal in difficulty levels, the raw scores were concurrently equated using *QUEST* (Adams & Khoo, 1993). Items that did not statistically fit the model, that is, items whose Infit Mean Square Value was not between 0.77 and 1.30, were deleted systematically in progressive runs until all items were within this acceptable level (McNamara, 1996). Omitted items and non-reached items were ignored in this calibration. The English→Japanese items, which were given partial marks of 0, 1, or 2, had 37 items left. Of the original 47 Japanese→English items, most of which were scored dichotomously, the estimates of 38 items were identified as acceptable and used for the second stage of the statistical analyses, HLM.

In a two-level HLM model, Level 1 investigated the within-student change in performance associated with time, while Level 2 examined the between-student difference associated with predictors that were hypothesised to influence the vocabulary test scores as in Table 1.

Level	Variables of Interests	Variables Tested in HLM
1 (within-student)	Time	TIME
2 (between-student)	Prior achievement	PRIOR
	Gender	GENDER
	Treatment group	TREATMENT
	Importance of vocabulary 1	IMPT 1
	Importance of vocabulary 2	IMPT 2
	Difficulty of vocabulary 1	DIFFT 1
	Difficulty of vocabulary 2	DIFFT 2
	Keyword use	KEYWORD

**Table 1: The hypothesised predictors and corresponding codes used at each level**

TREATMENT means that a student was in the treatment group and IMPT1 and IMPT2 indicated how important a student perceived vocabulary learning for successful learning of the Japanese language. Since there was an identical question in two questionnaires, IMPT 1 and IMPT 2 stood for the responses to Questionnaires 1 and 2 respectively. The same held for DIFFT 1 and DIFFT 2 (how difficult students perceived learning new words), and KEYWORD stood for how frequently a student used the keyword method in private vocabulary learning. The hypothesised model is shown in Figure 1.



**Fig. 1: Hypothesised two-level hierarchical model for both English→Japanese and Japanese→English vocabulary test scores**

Figure 1 indicates that it is hypothesised that the seven identical predictors influence the outcome variable for both directions of learning, English→Japanese and Japanese→English. Using the capability of HLM to compare relationships between the variables of Levels 1 and 2 as in a nesting structure, the models were constructed as shown below and then tested starting from Level 1.

The null model for both English-Japanese and Japanese-English outcome is:

$$\text{Level 1 (within-student)} \quad Y = B0 + R$$

Where  $Y$  is the vocabulary test score of each student at each test (or occasion), while  $B0$  is an intercept 1, that is, a student's average performance across five occasions, and  $R$  is an error element (or departure from average for each student).

$$\text{Level 2 (between-student)} \quad B0 = G00 + U0$$

Where  $G00$  is the intercept 2, which is the overall mean of vocabulary test scores of all five occasions) and  $U0$  is an error element.

The hypothesised models given below are used for both Japanese→English and English→Japanese scores to examine the effect of each element.

$$\begin{aligned} \text{Level 1 (within-student)} \quad & Y = B0 + B1 (\text{TIME}) + R \\ \text{Level 2 (between-student)} \quad & B0 = G00 + G01 (\text{PRIOR}) + G02 (\text{GENDER}) + G03 (\text{TREATMENT}) \\ & + G04 (\text{IMPT1}) + G05 (\text{IMPT2}) + G06 (\text{DIFFT1}) + G07 (\text{DIFFT2}) \\ & + G08 (\text{KEYWORD}) \end{aligned}$$

### 3 Results

#### 3.1 Does the use of the keyword method improve vocabulary test scores?

The first research question of this study is whether the use of the keyword method improved students' vocabulary test scores. Tables 2 and 3 below show raw scores of five vocabulary tests.

Tests	No. of Items	Full mark	Average marks (Standard deviation)	
			Delayed-Treatment Group	Treatment Group
1	14	28	22.4 (5.3)	22.7 (6.1)
2	15	30	20.3 (9.5)	21.9 (7.6)
3	8	16	8.6 (4.3)	7.4 (4.8)
4	15	30	16.7 (9.9)	15.9 (8.3)
5	10	20	11.6 (7.3)	12.3 (7.1)
Total	62	124	79.6 (7.3)	80.2 (6.8)

**Table 2: Average raw scores of five vocabulary tests (English→Japanese items)**

Tests	No. of Items	Full mark	Average marks (Standard deviation)	
			Delayed-Treatment Group	Treatment Group
1	14	14	11.8 (2.9)	12.0 (2.7)
2	15	15	10.4 (5.2)	11.1 (3.6)
3	7	7	3.7 (2.4)	4.2 (2.9)
4	15	15	9.5 (5.2)	9.4 (4.7)
5	10	10	6.8 (3.7)	6.9 (3.7)
Total	62	62	42.2 (3.9)	43.6 (3.5)

**Table 3: Average raw scores of five vocabulary tests (Japanese→English items)**

Although the treatment group performed slightly better in the first two tests, in the third test this group's performance was worse in the English→Japanese test than the delayed-treatment group. When considering the fact that the delayed-treatment group students were introduced to the keyword method between tests 3 and 4, it is interesting to note that the performance of this group

in test 4 was slightly better than that of the treatment group in both English→Japanese and Japanese→English tests. In order to discover whether these apparent differences are in fact statistically significant, and by how much, and if the hypothesised factors had any influence on the outcome, these scores were equated, then the null model was constructed to examine the variances available at levels 1 and 2 and tested as recommended by Darmawan & Keeves: “Estimating the null model is an important preliminary step in a hierarchical analysis. It produces a point estimate and confidence interval for the grand mean,  $Y_{00}$ .” (Darmawan & Keeves, 2002, p. 51)

The subsequent HLM runs examined all the hypothesised variables (cf. Table 1) and eliminated non-significant variables one at a time. As a result, significant predictors were left in the final models as shown in Tables 4 and 5, and Figures 2 and 3. Since the study yielded different results depending on the direction of the learning, they are reported separately with English→Japanese first, followed by Japanese→English learning mode.

Fixed effect	English→Japanese				
	Coefficient	Standard Error	T-ratio	d.f.	P-value
For INTRCPT 1, B0					
INTRCT 2, G00	0.58	0.15	5.58	48	0.00
PRIOR, G01	0.06	0.02	3.88	48	0.00
KEYWORD, G02	0.34	0.19	1.82	48	0.07

**Table 4: Final estimation of fixed effects for English→Japanese**

The final models were constructed based on these results.

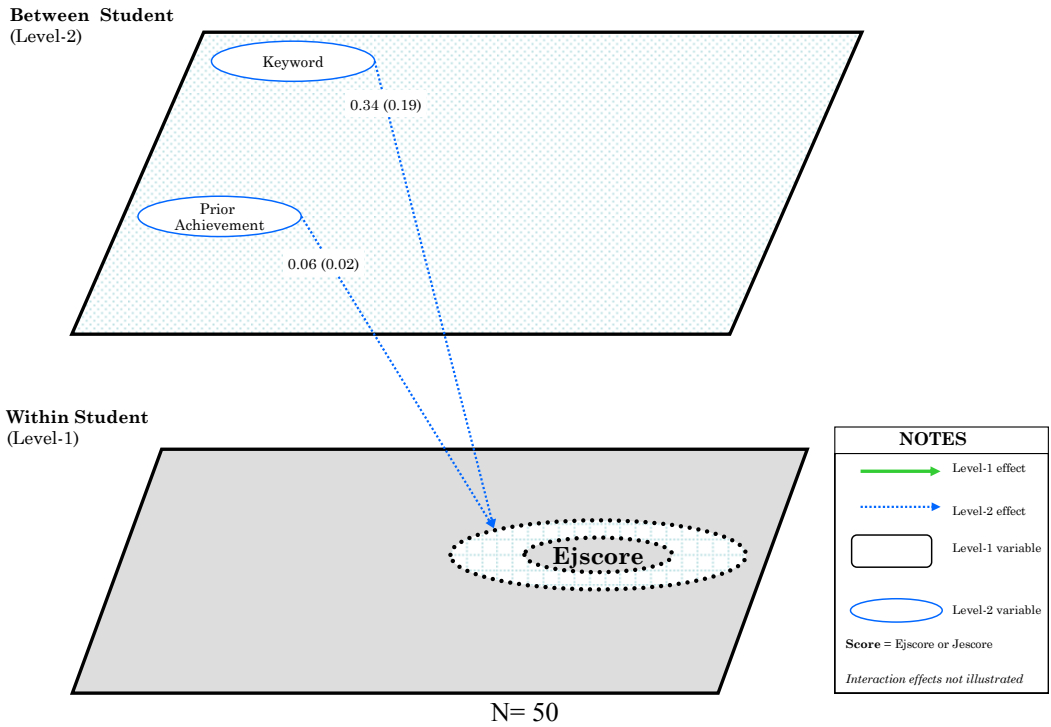
#### English→Japanese models

$$\text{Level 1} \quad Y = B0 + R$$

$$\text{Level 2} \quad B0 = G00 + G01 (\text{PRIOR}) + G02 (\text{KEYWORD}) + U0$$

Figure 2 is the visual display of these models.





**Fig. 2: The significant predictors for English→Japanese scores**

These results indicate that for English →Japanese vocabulary scores, only two factors influenced the outcome: (a) the use of the keyword method, and (b) the vocabulary test scores of the previous semester (0.05 level of significance). That is, the most powerful predictor was student achievement in the previous semester. The use of the keyword method was the only other factor, although the significance level was 0.1. The within-student variable, growth over time, did not prove to be a predictor. This indicates that the difficulty level of vocabulary learning after a semester remained just as high as at the beginning of the semester for this cohort of students. In other words, the proficiency level of these subjects did not improve enough to lessen their effort for learning new words (Nation, 1982). The remainder of the variables examined, namely gender, how important the students perceived vocabulary learning, how difficult students perceived learning new words, and being a member of the treatment group, did not influence the outcome.

Japanese →English					
Fixed effect	Coefficient	Standard Error	T-ratio	d.f.	P-value
For INTRCPT 1, B0					
INTRCT 2, G00	1.63	0.33	4.89	48	0.00
PRIOR, G01	0.07	0.02	3.47	48	0.00
KEYWORD, G02	0.61	0.43	1.77	48	0.08
For TIME slope, B1					
IMTRCPT 2, G10	0.08	0.26	3.32	50	0.00

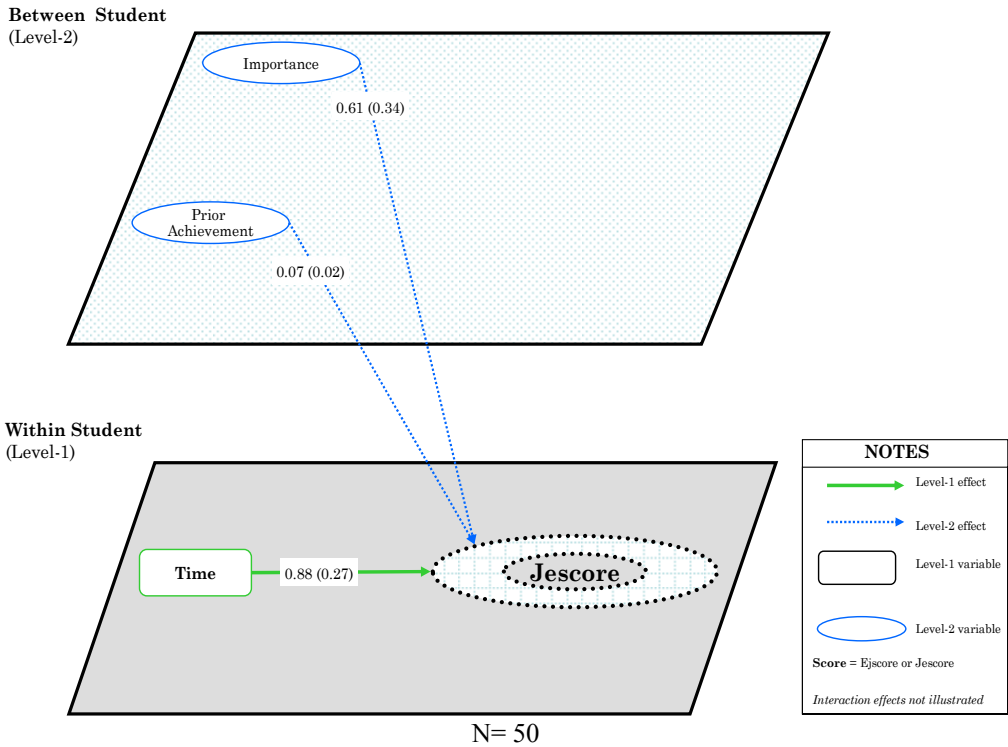
**Table 5: Final estimation of fixed effects for Japanese→English**

**Japanese→English models**

Level 1  $Y = B0 + B1 (TIME) + R$   
 Level 2  $B0 = G00 + G01 (PRIOR) + G02 (IMPT) + U0$

$$BI = GI0 + UI$$

Figure 3 below is the visual display of these results.



**Fig. 3: The significant predictors for Japanese→English scores**

For the receptive mode, Japanese→English, which is believed to be easier (Nation, 1982) than the productive mode, the within-student factor TIME (i.e. a student’s growth over time) was found to be a factor with a 0.1 level of significance. Unlike the productive mode, as the students continued with the language study, their vocabulary learning (for the direction of Japanese to English only) became less difficult, indicating that their proficiency level was high enough to facilitate it. At the same time, the perceived importance of learning new words, which did not influence the productive mode of learning, affected the outcome of the receptive mode of learning (although at 0.1 level). The only other factor found to be statistically significant was the previous semester’s achievement (0.05 level), as was the case with the productive mode of learning.

With regard to the fitness of the models used in this study (Darmawan & Keeves, 2002), the running of null models showed that the amount of variance that could be explained in the analyses are for Japanese→English scores 62.7% for Level 1 (within-student), 37.3% for Level 2 (between-student), while for English→Japanese 53.9% and 64.1% respectively.

Synthesising all these results, the answer to Research Question 1 is not straightforward. The use of the keyword method has been found to improve only the English→Japanese learning process marginally (at 0.1 level only) but not the Japanese→English learning process. The study also found that two variables, namely student’s growth over time and their perceived importance of vocabulary knowledge, marginally influenced the outcome for the receptive learning, Japanese to English, but not the opposite direction. The only variable found to be statistically significant enough to be confidently called a predictor was the previous semester’s vocabulary scores, while

other hypothesised factors such as gender, perceived difficulty, and being in the treatment group, were found to make no difference.

### 3.2 Do tertiary students take up the keyword method?

The second research question was whether tertiary students would take up the keyword method for learning new Japanese words.

Question 3a in Questionnaire 2 (Appendix A) asked students' opinions about the use of the keyword method. Table 6 summarises the findings.

Category	Delayed-treatment group	Treatment group
1 (used the method)	9	6
2 (used the method but rarely)	9	4
3 (used the method often)	6	10
Total	24	20

**Table 6: Use of the keyword method for vocabulary learning (N=44)**

Exactly half of the treatment group subjects took up this method readily, while only a quarter of the delayed-treatment group students did so. Some responses to this question are quoted below.

#### *Delayed-treatment group responses*

Category 3 [used readily]

S1: Very helpful. Often, whenever possible

S2: I think it's a good idea but I tried to make up my own, as they're easier to remember if I have come up with them. Yes, when a word is difficult to remember, usually it works.

Category 2 [used only rarely]

S4: Was all right at the time, but not as good as repetition with a Japanese word with straight-out English translation.

S5: Not too bad, but perhaps some of the associations are a little far fetched. I think that it should be stressed that this method can only be used when realistic. [Used] only a few times. Sort of [worked].

Category 1 [did not use]

S12: It confused me, but I think it worked well for other people. Not too often because after a while it became confusing.

S13: Appeared to be effective for most students. No, my head doesn't work that way.

#### *Treatment group responses*

Category 3 [used readily]

S14: It is a helpful and useful method to grasp the Japanese word because it is triggered by an association with something familiar in English. [Used it] for a few more difficult words when I recap the lesson in the textbook.

S16: They work on most words. But for some, it's harder to remember the association than the word in Japanese. Yes, [used] for half the words.

S17: Good technique, but it won't work in everyone's case. It works better if people do this individually, not everyone's memory works the same. Yes, [used it] for four or five words per lesson. It seems to [work].

S19: It often works well. I used this method before. I find it easier, usually, to remember the links that I myself came up with. It is a good method. It works unless my stories were very obscure, then I forget.

S20: I find that I can't use this technique on more than about five words every week.

Category 2 [used only rarely]

S23: It's OK, and helpful in some situations, but not my preferred method — it takes too long for so few words. [Used it] occasionally only for a few words. I found the keyword method to be a very exhausting exercise. While it was useful in some circumstances, I found it not as effective as other methods.

Category 1 [did not use]

S26: It's a good idea but at the same time it can be off-putting or make you remember the word incorrectly. [Used it] not often as I don't want to spend so much time thinking of associating words.

S27: This does not work for myself.

S28: Personally, it is more difficult for me. I remember better by writing and reading.

Even the students who did not find the keyword method useful endorsed the idea of introducing this method to Japanese language learners in the future. To Question 3b, *Shall I introduce this keyword method to the first year Japanese students next year?* every respondent gave a positive answer. One delayed-treatment group subject, who did not use the method, answered "Yes, but a little bit earlier".

Students' decision to adopt this method seemed to depend on two factors: (a) economy of time, and (b) the strength of association. The first was the students' judgment about the worthiness of the necessary effort to create two links, acoustic and imagery. Some students must have felt that an equal amount of energy could be better spent by learning words using their proven and preferred methods as expressed in their comments. It is worth noting that even those who found the technique useful remarked that they did not use it for every word. This was because either some words (a) did not need any association since they were easy to learn, or (b) were found to be too difficult to conjure the necessary associations no matter how far their imagination was stretched.

This last point focuses on the second main reason for students using or not using the keyword method, namely, the strength of the links. In cases where association was so unrealistic that it taxed their memories, which also meant that there was a danger of incorrect recall, students did not believe it was worthwhile to use the keyword method. As an answer to Research Question 2 then, tertiary students would use the keyword method if the following three conditions were met: first, the word was difficult to learn by simple rehearsal and therefore required an additional technique to remember it; second, time and energy required of a learner to generate acoustic and imagery links were less than committing the word to memory by rehearsal; and third, they could generate links close enough that no further load was placed on their memory in order to retrieve the newly-learned word correctly. It should be noted that students who were given opportunities to experiment with this method in class used it twice as much as those who were not. The inconclusive results on the effect of the keyword method of the past studies might be an artefact of their design, namely, not all these three conditions were present. For example, many past investigations measured the receptive mode of learning only, that is, the words may not have been difficult enough to require this method.

## 4 Discussion and conclusion

### 4.1 Implications for classroom application

The study, an application of the keyword method to foreign language curriculum, added two pieces of evidence to inconclusive results of past studies. The first is an indication that this method may facilitate foreign language vocabulary learning of experienced and older learners. Since the use of this mnemonic improved only the productive mode of learning with a 0.1 level of significance, statistically its merit can be claimed only as an indication or possibility. However, from the pedagogical perspective, the author believes that the level is deemed strong enough to be used as one form of association that Carter and McCarthy recommended: "Learners should be encouraged to make their own lexical associations when they are actively learning new vocabulary. However, at present we do not know which kind of associations are the most useful in aiding retention." (1988, p. 94)

The second piece of evidence uncovered by the study is that tertiary language learners are prepared to take up this method for learning new words after minimal practice sessions in class. This is an additional reason for this mnemonic to be introduced in language classrooms.

### ***4.2 Implications for theories***

This method was found to benefit only the productive mode of learning and not the receptive mode. The author speculates about the reasons for this result. First, as learners reach a certain linguistic proficiency level and learn how to learn effectively (with time and the accumulation of learning experience), vocabulary acquisition becomes less demanding (Nation, 1982; Thorndike, 1908; Webb, 1962). Second, productive learning of a word is more demanding than receptive learning (Nation, 1982). Third, the students in the current study may be at some mid-point in their proficiency between the two levels, that is, their proficiency is high enough to facilitate receptive learning without using any special or new strategy, but is not high enough to do the same for productive learning. The author hypothesises that as these students' proficiency level improves, learning Japanese words productively becomes easier and as a result the use of the keyword method loses its effectiveness. Since no study has examined this aspect, only future studies can reveal whether this is the case, and if it is, at what proficiency level this occurs.

### ***4.3 Implications for future studies***

As has been described, many more aspects need to be investigated to discover which classroom variables (and in what combinations) affect the effectiveness of the keyword method. Some variables that this study has brought to the surface that could shed more light on the merits of this mnemonic are: (a) the difficulty level of the words to be learned, including various parts of speech, their abstractness, and the phonetic distance from the native language; (b) the proficiency level of learners; and (c) the age and foreign language learning experience of learners. These variables, separately or in combination, may also hold the key to the usefulness of the technique.

In addition, this study has demonstrated that future studies can confidently rely on the HLM for analysing data that contain lost cases and missing data, which are inevitable in longitudinal educational studies such as this one.

Many more classroom-based research studies are required to verify what the current study has indicated, and also to consider in which types of classroom environments the keyword method can be used most effectively. In the meantime, because the learning and teaching of foreign vocabulary must continue, even though how it happens in our memory is not clear, it is hoped that at least learners and instructors are made aware of up-to-date information about effective strategies that are available to learn new words in the shortest possible time.

---

#### **Notes**

<sup>1</sup> An earlier version of this paper was presented at the 14<sup>th</sup> biennial conference of the Japanese Studies Association of Australia, 3–6 July 2005, Adelaide, South Australia.

#### **Acknowledgements**

Of the many who supported and helped me with this project, I would like to express my special gratitude to Professor Keeves, Professor Lawson, Dr. Allen, Mrs Worthley, my students, and my family.

#### **Appendices**

##### ***Appendix A: Questionnaire 2***

[administered towards the end of the study]

**Japanese vocabulary learning**

Questionnaire 2

Name: \_\_\_\_\_

Dear Student,

As I have explained, for many years as a language teacher, I have been searching for methods that would help students learn new words. That is why I am interested in your ideas and feeling concerning vocabulary learning of Japanese. I appreciate your help in this endeavour. Please answer the following questions.

**Question 1.** How would you rank-order the importance of the following aspects of the language in order to master Japanese? Indicate your ranking by placing 1 in the brackets next to the most important, 2 to the second important and so on.

Rank	Aspects
( )	Grammar (word order, particles, etc.)
( )	Kanji
( )	Vocabulary
( )	Pronunciation
( )	Fluency
( )	Other (please specify) _____
( )	Other (please specify) _____

**Question 2.** Of those aspects listed above in Question 1,  
(a) which do you find the most difficult? \_\_\_\_\_  
(b) which one do you find the easiest? \_\_\_\_\_

**Question 3.** What do you think of the keyword method that we have tried out in class (associating the sound-like English word with the picture showing the meaning of the Japanese word)?

\_\_\_\_\_  
\_\_\_\_\_  
\_\_\_\_\_

**3a.** Have you tried this keyword method? Yes/No

If yes, how often? Did it work?

If no, why not?

\_\_\_\_\_

**3b.** Shall I introduce this keyword method to the first year Japanese students next year?

Yes/No

Why/Why not?

\_\_\_\_\_

**Question 4.** Please write if you have any comment to make regarding Japanese vocabulary learning.

---



---



---

Thank you very much,

**Appendix B: List of words examined in five vocabulary tests**

**1. English→Japanese**

- |                                       |                                    |
|---------------------------------------|------------------------------------|
| 1. a sandwich                         | 2. to be, to exist                 |
| 3. a fixed place to catch taxis, etc. | 4. counter for frequency (~ times) |
| 5. to take (time)                     | 6. Australia                       |
| 7. an envelope                        | 8. left (hand side)                |
| 9. a tree                             | 10. a woman                        |
| 11. only                              | 12. my own older brother           |
| 13. refrigerator                      | 14. three weeks                    |
| 15. a map                             | 16. to take off (clothes)          |
| 17. to use                            | 18. straight                       |
| 19. to make, produce                  | 20. to put, place                  |
| 21. how to read, way of reading       | 22. to sit down                    |
| 23. hair                              | 24. a foreign student              |
| 25. to lose                           | 26. illness                        |
| 27. to worry                          | 28. important                      |
| 29. in what way, how                  | 30. leg, foot                      |
| 31. animal                            | 32. dentist                        |
| 33. to wash                           | 34. slow, late                     |
| 35. salt                              | 36. to get tired                   |
| 37. especially                        | 38. maple tree, autumn leaves      |
| 39. a few (people), a little          | 40. heavy                          |
| 41. fine arts                         | 42. to send, post (letter)         |
| 43. by far                            | 44. wide, spacious                 |
| 45. to get married                    | 46. thirsty                        |
| 47. examination                       |                                    |

**2. Japanese→English**

- |                              |                                    |
|------------------------------|------------------------------------|
| 1. むかえます (to welcome)        | 2. せんもん (a major subject)          |
| 3. もうすこし (a little more)     | 4. せいひん (goods)                    |
| 5. いそぎます (to hurry)          | 6. つけます (to turn on)               |
| 7. とこや (a barber)            | 8. もっていきます (to take along)         |
| 9. そくたつ (special delivery)   | 10. こたえ (an answer)                |
| 11. くすり (medicine)           | 12. おとうとさん (younger brother)       |
| 13. はらいます (to pay)           | 14. わかい (young)                    |
| 15. せがたかい (tall)             | 16. でかけます (to go out)              |
| 17. まだまだです (long to go)      | 18. まど (a window)                  |
| 19. がいこく (a foreign country) | 20. しゃちょう (president of a company) |
| 21. けいざい (economy)           | 22. たちます (to stand up)             |
| 23. のど (a throat)            | 24. おしえます (to teach)               |
| 25. てつだいます (to help)         | 26. はる (spring)                    |
| 27. せかい (the world, globe)   | 28. たくさん (a lot, much)             |

- |                              |                               |
|------------------------------|-------------------------------|
| 29. からい (spicy)              | 30. はじめて (for the first time) |
| 31. おそい (late)               | 32. つかれます (to get tired)      |
| 33. くやくしよ (municipal office) | 34. せまい (narrow)              |
| 35. りょうほう (both)             | 36. とうろく (registration)       |
| 37. すぐ (straight away)       | 38. ぜひ (by all means)         |

**Appendix C: Examples of keywords practised in class (keyword in capitals)**

1. そくたつ/soktΛts/ [special, speedy delivery] SOCK (that Santa delivers)
2. まっすぐ/mAssωg/ [straight] MUSCLES of a strong man that are straightening a bent bar
3. ずっと/zωtto/ [by far] ‘ZOOOH’ the noise made by a skidding bike on the road to go ‘far’
4. でかけます/dekΛkemΛs/ [to go out] CAR to take a family for outing (you are all in a car)
5. なくします/nΛkω□εmΛs/ [to lose something] A CUSHION (you are sitting on one) behind which you lose a small object
6. かるい/kArωI/ [light] (KANGA) ROO that is very light and can jump up high
7. もみじ/momI□□I/ [autumn leaves or maple tree] A MIDJIT who is as tall as a maple leaf is standing next to a maple leaf.
8. まがります/mΛgArImΛs/ [to turn] MARGARITA (a drink) that you put in a drink mixer and turn and shake to make a glassful
9. もうすこし/mo:sωko□/ [a little bit more] SQUASH (to squash something to make it smaller)
10. おもい/omoI/ [heavy] OMO --washing powder—(‘heavy’ duty washing powder in a huge, heavy, 10 kilo box)
11. あまい/ΛmΔI/ [sweet] MINE –my sweetheart—(Mine is very sweet—two lovers)
12. どのぐらい/donogωrΔI/ [about how far] DO’KNOW (don’t know how far it is from here to that mountain top)

**References**

- Adams, R., & Khoo, S-T. (1993). *QUEST: The Interactive Test Analysis System*. Melbourne: ACER.
- Atkinson, R., & Raugh, M. (1975). An application of the mnemonic keyword method to the acquisition of a Russian vocabulary. *Journal of Educational Psychology: Human Learning and Memory*, 104(2), 126–133.
- Bruning, R., Schraw, G., & Ronning, R. (1999). *Cognitive psychology and instruction* (3<sup>rd</sup> ed.). New Jersey: Prentice-Hall.
- Bryk, A., & Raudenbush, S. (1992). *Hierarchical Linear Models: applications and data analysis methods*. London: Sage Publications.
- Carter, R., & McCarthy, M. (1988). *Vocabulary and language teaching*. London: Longman.
- Darmawan I.G.N., & Keeves, J. (2002). Two-level model of information technology adoption in local government of Bali. *International Education Journal*, 3(1), 47–60.



- Fuentes, E. (1976). An investigation into the use of imagery and generativity in learning a foreign language vocabulary (Doctoral dissertation, Stanford University, 1976). *Dissertation Abstracts International*, 37(5A), 2694.
- Kuwahara, Y. (2000). Imagery mediation strategy and learning Japanese characters: paired-associate learning of characters and English translations. *Japanese Journal of Educational Psychology*, 48, 389–399.
- Laufer, B. (1986). Possible changes towards vocabulary acquisition research. *International Review of Applied Linguistics*, 24, 69–75.
- Lawson, M., & Hogben, K. (1996). The vocabulary-learning strategies of foreign-language students. *Language Learning*, 46(1), 101–135.
- Lawson, M., & Hogben, K. (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.
- Levin, J., Pressley, M., McCormick, C., Miller, G., & Shriberg, L. (1979). Assessing the classroom potential of the keyword method. *Journal of Educational Psychology*, 71, 583–594.
- Lu, M.-Y., Webb, J., Krus, D., & Fox, L. (1999). Using order analytic instructional hierarchies of mnemonics to facilitate learning Chinese and Japanese Kanji characters. *Journal of Experimental Education*, 67(4), 293–303.
- McDaniel, M., & Pressley, M. (1989). Keyword and context instruction of new vocabulary meanings: effects on text comprehension and memory. *Journal of Educational Psychology*, 81(2), 204–213.
- McNamara, T.F. (1996). *Measuring second language performance*. New York: Longman.
- Nation, P. (1982). Beginning to learn foreign vocabulary: a review of the research. *RELC Journal*, 13(1), 14–37.
- Ott, C.E., Butler, D.C., Blake, R.S., & Ball, J.P. (1973). The effect of interactive-image elaboration on the acquisition of foreign language vocabulary. *Language Learning*, 23, 197–206.
- Raudenbush, S., Bryk, T., Cheong, Y.F., & Congdon, R. (2000). *HLM 5: Hierarchical Linear and Nonlinear Modeling* (Statistical software manual). Lincolnwood, IL: Scientific Software International.
- Raugh, M., & Atkinson, R. (1975). A mnemonic method for learning a second-language vocabulary. *Journal of Educational Psychology*, 67, 1–16.
- Thorndike, E. (1908). Memory for paired associates. *Psychological Review*, 15, 122–138.
- Touloumtzoglou, J. (1998). *Effectiveness of the keyword mnemonic in modern Greek (second language) vocabulary acquisition*. Unpublished MA thesis. Flinders University of South Australia, Adelaide, South Australia, Australia.
- Wang, A., Thomas, M., & Ouellette, J. (1992). Keyword mnemonic and retention of second-language vocabulary words. *Journal of Educational Psychology*, 84(4), 520–528.
- Wang, A., Thomas, M., & Ouellette, J. (1995). Effect of keywords on long-term retention: help or hindrance? *Journal of Educational Psychology*, 87, 468–475.
- Wang, A., Thomas, M., Inzana, C., & Primicerio, L. (1993). Long-term retention under conditions of intentional learning and the keyword mnemonic. *Bulletin of the Psychonomic Society*, 31(6), 545–547.
- Wang, A.Y., & Thomas, M.H. (1992). The effect of imagery-based mnemonics on the long-term retention of Chinese characters. *Language Learning*, 42(3), 359–376.
- Webb, W.B. (1962). The effects of prolonged learning on learning. *Journal of Verbal Learning and Verbal Behaviour*, 1, 173–182.
- Widdowson, H.G. (1978). *Teaching language as communication*. Oxford: Oxford University Press.