An Experimental Investigation of HAM, a Novel Mnemonic Technique for Learning L2 Homonyms and Homophones

Charles M. Mueller
Fuji Women's University, Sapporo, Japan
https://doi.org/10.7820/vli.v07.1.mueller

Abstract

Over the past 40 years, extensive research has examined the effectiveness of mnemonics for vocabulary learning. Much of this research has investigated the keyword method (Atkinson & Raugh, 1975), which involves linking an image related to a to-be-learned L2 word with an image related to a similar-sounding L1 word. Whereas most research has shown the keyword method to be effective (Webb & Nation, 2017) with impressive long-term retention rates (Beaton, Gruneberg, & Ellis, 1995), some have questioned its usefulness, particularly due to the quality of the resulting lexical representations and extended latencies associated with recall (Barcroft, Sommers, & Sunderman, 2011; Van Hell & Candia Mahn, 1997). Other drawbacks of the keyword technique are the equating of dissimilar L1 and L2 phonemes and the difficulty in creating associations for languages with markedly different phoneme inventories. The current study presents a novel approach called the Homonym/Homophone Association Method (HAM). It overcomes some of the drawbacks of the keyword method by associating meanings of L2 homonyms or homophones, one known by the learner and one unknown. Because the pronunciations of the L2 target words are identical (or nearly identical), learners only need to associate two distinct meanings. A quasi-experiment (*N* = 71) employing a within-subjects design compared the effectiveness of (1) HAM using researcher-generated associations and images, (2) HAM using self-generated associations, and (3) production practice that involved writing target words in sentences. Results on an unannounced posttest given 3 weeks after instruction suggest an advantage for HAM using researcher-generated associations.

Keywords: HAM; keyword method; homonyms; homophones; vocabulary learning; mnemonics; puns; humor

1. An Experimental Investigation of HAM, a Novel Mnemonic Technique for Learning L2 Homonyms and Homophones

In SLA research, deliberate vocabulary learning has not always received the attention it deserves (Boyd Zimmerman, 1997). This is unfortunate as a "well-chosen basic vocabulary" is essential to successfully complete everyday tasks in an L2 (second language) environment (De Groot & Van Hell, 2005, p. 9) and is highly predictive of L2 acquisition (Laufer, 1997).

At advanced levels of proficiency, vocabulary learning presents a significant challenge. For example, a learner of English may get by with knowledge of 4000 to 5000 word families, but knowledge of around 9000 word families is thought to be optimal for learners who need to independently read authentic English texts (Nation, 2015). To get a sense of the learning difficulty this "optimal" lexical knowledge poses, consider learning through extensive reading. It is estimated that learners must encounter a word around a dozen times to learn it through incidental exposure (Nation, 2014). Based on this, Nation (2014) predicts that a corpus of 2 427 807 tokens (roughly the number of words in 20 novels) would be needed for a learner to encounter the first 8000 word families of English an average of 12 times.

To assist learners with the formidable task of L2 lexical acquisition, incidental learning through massive amounts of input is ideally supplemented by deliberate learning (Webb & Nation, 2017). Language instructors are thus in need of pedagogical techniques to establish initial form-meaning links, especially as these relate to words in the fourth to ninth most frequent 1000 word families of English (i.e., "mid-frequency" words in Schmitt & Schmitt, 2014).

Over the last 40 years, researchers interested in investigating approaches to deliberate vocabulary learning have conducted extensive research on mnemonics. These memory techniques have been employed widely in education (for an overview and discussion, see Putnam, 2015). For example, Japanese school children, when learning a historical fact such as the date when Christopher Columbus arrived in the Americas, often associate the phrase *iyokuni moeru* ("[He] was burning with passion") with Columbus since the first four syllables sound like the Japanese words for 1, 4, 9, and 2 (cp. Fontana, Scruggs, & Mastropieri, 2007).

In SLA, a large portion of mnemonics-related research has investigated the keyword method (Atkinson & Raugh, 1975; Pressley, Levin, & Delaney, 1982), which involves linking a target L2 word to a similar-sounding L1 word via an image incorporating both words. For example, an English L1 speaker who wants to memorize the Japanese word for cat (i.e., *neko*) might create an image of a man with a cat clinging to his neck. Research has generally shown the key word method to be effective (Campos, Rodríguez-Pinal, & Pérez-Fabello, 2014) for L1 vocabulary learning, particularly for children with learning disabilities (Urberti, Scruggs, & Mastropieri, 2003). Applied to second language acquisition, this method has been shown to result in impressive long-term retention rates (Beaton et al., 1995), outperforming conventional deliberate vocabulary learning activities by around 20% (Webb & Nation, 2017, p. 115), although some studies (e.g., Hall, Wilson, & Patterson, 1981; Wei, 2015) have failed to show such advantages. Moreover, it has been found to be effective for learning a wide range of L2s, including English (Avila & Sadoski, 1996; Elhelou, 1994; Rodriguez & Sadowski, 2000), Galician (Campos et al., 2014), German (Desrochers, Wieland, & Coté, 1991; Fritz, Morris, Acton, Voelkel, & Etkind, 2007), Korean (Griffith, 1981), Russian (Atkinson & Raugh, 1975), Spanish (Gruneberg & Pascoe, 1996), and Tagalog (Wang, Thomas, & Ouellette, 1992). While most of the keyword studies have tested the method's effectiveness for college-aged adults, it has also been found to be useful for early (Elhelou, 1994; Pressley, 1977) and older L2 learners (Gruneberg & Pascoe, 1996).

Even so, some research has questioned the suitability of the keyword method on various grounds. A key concern stems from findings that suggest that the quality of the resulting lexical representations is poor. A related concern is that the encoding of extra information as a cue (i.e., the imagery linking the two words) appears to result in extended latencies associated with recall (Barcroft et al., 2011; Van Hell & Candia Mahn, 1997). An additional drawback of the keyword technique is that it involves equating dissimilar L2 and L1 phonemes. A Japanese learner, for example, may choose to learn the English word *tar* by imagining a *taru* (the Japanese word for barrel) full of tar. Yet, this may inadvertently reinforce the Japanese learner's tendency to equate the English /r/ sound with the Japanese intervocalic [r] sound.

At a practical level, the keyword method is difficult to apply in the case of languages in which the L1 phoneme inventory and phonotactic constraints differ markedly from those of the L2 (cp. comments about difficulty of applying the keyword method with monosyllabic L1s in Webb & Nation, 2017, p. 116). For example, Japanese learners using the keyword method to learn English may find it difficult to learn English words with phonemes that are not present in Japanese. English learners may also find it difficult to apply the method to languages such as Japanese which, unlike English, contains many open vowels and unfamiliar phonemic contrasts such as vowel length or pitch accent.

The current study presents a novel mnemonic approach to vocabulary learning called the Homonym/Homophone Association Method (HAM). Languages typically contain words with distinct meanings that share the same written and/or spoken form. Words with unrelated meanings that share both spoken and written forms are commonly referred to as *homonyms*. An example would be *bill* (something you pay) and *bill* (a duck's beak). In many cases, homonyms actually have an etymological relationship that has become lost to language users (and is probably even less accessible for L2 learners). Since the current focus is on pedagogical approaches in L2 acquisition, words will be regarded as homonyms if the meaning relationship is not likely to be accessible to the L2 learner. In reality, the boundary separating homonyms (words with unrelated meanings) from polysemes (words with related meanings) is often fuzzy and scalar (for a practical measure of degrees of semantic relatedness, see Nagy & Anderson, 1984). Unlike homonyms, homophones (words like *hair* and *hare*) share the same sound but have different spellings. In some languages such as Japanese, the distinction between homonyms and homographs is also fuzzy. For example, many words are homonyms when written in *hiragana* (a phonetic script) but can be considered homographs when written in *kanji* (Sino-Japanese characters).

HAM is designed as an approach to L2 learning of homonyms and homophones, although it could, with some adaptations, perhaps be applied to homographs as well. While more limited in potential applicability, it overcomes some of the drawbacks of the keyword method by associating two L2 homonyms or homophones, one of which is known by the learner and one of which is unknown. Because the pronunciations of the L2 target words are identical, learners only need to associate two distinct meanings.

As an example of the technique, consider a learner of Japanese who wants to remember the low-frequency Japanese word "salmon roe" (*ikura*). The learner, even at an elementary stage of acquisition, is likely to know the homonym (*ikura*)

that means "how much". Although the words are often distinct in Japanese writing, they are pronounced the same. Consequently, to learn the low-frequency meaning of *ikura*, the L2 learner could form a picture of someone in a market holding up a carton of salmon eggs while asking how much they cost (*Ikurawa ikura desuka*?)

Instructors creating pedagogical materials for instruction using HAM can readily find extensive materials from Internet sites (particularly, social networking sites) featuring humorous pictures. A common genre of humor on the Internet involves images in which a low-frequency member of a homonym or homophone pair appears within a context in which the high-frequency member would be expected. For example, in one such comic, an image shows an elderly man and woman with the woman saying, "You look quite distinguished with that grey hare on your head, Jim." In the picture, in place of the man's gray hair (expected based on context), there is a gray hare. This use of irony in humor can be explained in terms of Giora's (1997) graded salience hypothesis, which maintains that more salient meanings of words (e.g., meanings that are more prototypical, familiar, frequent, conventional, and so on) will be retrieved from the lexicon more rapidly so that they temporally precede the processing of ironic interpretations. The hypothesis has received support in experiments that have compared reading times or response times of literal or ironic meaning (Giora & Fein, 1999; Giora, Fein, & Schwartz, 1998).

The idea of associating homonyms and homophones to aid memory is not new. It has been used in general education, for example, to recall definitions of key terms in history classes (Fontana et al., 2007). However, it has not received systematic treatment in SLA. Webb and Nation (2017), in their discussion of the keyword method, briefly allude to the possibility of associating the target word with L3 words (i.e., words from other foreign languages) or with other L2 words, yet they do not explicitly discuss homonyms or homographs.

In SLA, much of the research on the pedagogical potential of homonyms and homographs has focused on puns. For example, Lucas (2005) examined learner engagement with comic strips that featured puns. She found that the task focused learners' attention on the phonological, morphological, and syntactic aspects of language and thus led to an increase in comprehension of the target meanings. Tocalli-Beller and Swain (2007), in a longer study, similarly demonstrated that pair-work in which L2 learners shared and explained riddles that involved homonyms led to sizable and durable gains in knowledge of the target lexical items. Bell (2012), in a qualitative study of spontaneous humorous interaction in an L2 classroom, concluded that humor facilitated the recall of language items, particularly in the case of word meanings. While these studies suggest that humor promotes the establishment of enduring form-meaning links, they did not explore whether the focus on homonyms (in the form of puns) promoted learning more effectively than alternative tasks.

In sum, previous research on homonyms and homographs has mostly focused on language play within classroom contexts, whereas research on mnemonics has focused on associations between L1 and L2 forms. The mnemonic potential of L2 homonyms and homographs has remained largely unexplored. This is unfortunate as the HAM technique would appear to have certain advantages. As with the keyword method, it should be particularly useful as a means

of forming initial form-meaning links, especially for words that are unlikely to be encountered frequently enough to be initially learned or retained. The use of a mnemonic should also help learners more readily recall a target word so that it can be used in production. By explicitly focusing on the phonological or written similarity of the to-be-learned word and a known word, HAM should also lead to more detailed encoding of the formal aspects of words. Finally, the method, by limiting its scope to equivalences between L2 word forms, completely avoids the potential problem of equating L2 and L1 phonemes (a possible objection to the keyword method). Finally, it adheres to recommendations that new words be learned through connections with pre-existing knowledge (Sökmen, 1997).

To determine whether HAM is an effective elaborative technique for teaching low-frequency vocabulary, the current study reports the results of an experiment comparing this approach with a more conventional vocabulary learning technique. The technique chosen for the comparison condition involved elaboration of an L2 item by using the target word in an English sentence. Boers (2015) points out that this word memorization technique has the advantage of stimulating learner engagement with both the form and meaning of the target word. Research has generally shown that writing a sentence with an L2 word is an effective means of learning the word as long as the word form is clearly linked to the associated meaning during production (cp. Barcroft, 2006). Coomber, Ramstad, and Sheets (1986), for example, found that college-aged learners of 10 artificial words who learned the words through sentence composition showed superior retention relative to participants who learned through rehearsal of definitions or exposure to examples (see also Zou, 2017).

In the current experiment, HAM instruction was further divided into two types to determine whether learners benefitted more from self-generated associations or instructor-generated associations. It is known that cognitive effort during encoding generally has a positive effect on recall (Tyler, Hertel, McCallum, & Ellis, 1979). While this would suggest that learners who generate their own keyword associations should show better retention, several studies that have examined the keyword method have failed to find significant differences between the two conditions (e.g., Colón & Rodriguez, 2004–2005; Hall, 1988; Shapiro & Waters, 2005). One possible explanation for the lack of a facilitative effect is that the benefits of effortful encoding may be offset by the time spent in generating (versus memorizing) the keyword mnemonic. The experiments in the current study explore whether this is also the case when the mnemonic technique targets L2 homonyms and homographs.

2. Method

A quasi-experiment employing a within-subjects design compared the effectiveness of (1) HAM using researcher-generated associations and images, (2) HAM using self-generated associations, and (3) production practice involving use of the target words in sentences.

Participants. The participants (N = 71) were first-year Japanese-L1 EFL learners from three intact classes at two universities in Japan. One class (n = 16) consisted of female students in a department focusing on linguistics and

literature at a private women's university. Those in the other two classes (*n* = 27, *n* = 28) were male and female students from a diverse range of majors who were first-year students at a large public university. Based on subjective impressions of participants' performance in class, the they were, in terms of the Common European Framework of Reference (Council of Europe, 2001), mostly at Level B1 (threshold or intermediate), although some may have been at Level A2 (waystage or elementary).

Materials. The experiment targeted 30 words from various parts of speech. To avoid ceiling effects on the pretest, the senses of the targeted homonyms and homophones were low-frequency lexical items that the participants were unlikely to know.

The materials for the researcher-generated HAM (hereafter, "HAM Given") condition consisted of an 11-page hand-out. All instructions were provided in Japanese. The first page introduced the Homonym Association Method (HAM). The initial example showed a toad seated like a human being (i.e., with legs hanging down in front) on a bench at a bus stop with the caption: "Just waiting for the bus cause my car got toad". The text explained that the humor of the text involved the replacement of *towed* (the word that would naturally come to mind in this situation) with the homophonous word *toad*, and the ludicrous notion that the toad would drive or wait for a bus. The following 10 pages had similar humorous pictures associating two L2 words. In the materials for this condition and the other two conditions (discussed below), four of the ten target words involved homonyms, three involved homophones, and three involved "near" homonym/homophone associations in which both the spelling and spoken form were slightly different (e.g., *dessert* and *deserted*). Below the picture on each page, the associated L2 words were given with their Japanese translations. Below these words, there were seven lines. Participants were asked to explain in Japanese why the picture was humorous, modeling their responses on the explanation on the first page of the hand-out (the page with the toad at the bus stop).

The materials for the HAM Self-Generated condition introduced the method in Japanese using the example discussed in the introduction (i.e., the association of the two meanings of *ikura* using an image of a shopper asking how much salmon roe costs at a fish market). In addition to this example, targeting the association of homonyms, an example was provided of a homophone association in which the words *die* and *dye* were linked by means of a humorous picture showing an Easter egg dipped in dye saying, "I don't want to die in this dye!" It was pointed out that HAM could also be used when there was a slight mismatch between the two words being associated (as is the case, for example, with *dessert* and *deserted*). The following pages of the hand-out showed two to-be-associated words with their Japanese translations. Below the two words were blank lines. Participants were asked to think of an image associating the two words and then write a sentence linking the two words, modeling their responses on the examples provided on the first page (i.e., the *ikura* example and the die/dye example).

In the Sentence Composition condition, participants were given an English word with a Japanese translation. Below the words, there were blank lines where participants were to write a sentence using the targeted sense of the English word.

To ensure that the relative difficulty of the words would not bias the results, the three sets of 10 words were counterbalanced so that each set appeared in a different condition for each class. The pretest and posttest measures were identical and consisted of a translation task in which the L1 word had to be translated into the L2 target. The test consisted of 30 items ($\alpha = .77$). To prevent participants from translating the L1 word with an untargeted L2 word, the number of letters of the target word (shown with bolded underlines) and some of the letters were provided.

Procedure. The experiment was conducted in participants' regular class during regular class times with their regular instructor (the author). Participants received the pretest forms and were given 20 minutes to take the test. After completing the test, they learned the 30 target words, 10 words at a time, as they completed learning tasks in each of the three conditions. To ensure that sequencing of the conditions did not influence results, the order was counterbalanced so that the three classes did the three conditions in different sequences. Thus, the (A) HAM Given condition, (B) the HAM Self-Generated condition, and (C) the Sentence Composition condition were performed in differing sequences (i.e., ABC, BCA, and CAB) by each class.

Participants received hand-outs for one of the experimental conditions, and the instructor briefly went over the directions with the class. The participants then had 20 minutes to complete the task for that condition. The same procedure was followed for the other two conditions. Three weeks after the pretest and intervention, participants were given an unannounced posttest during which they were given 20 minutes to answer the questions. It should be noted that there was no immediate posttest and instead, the "posttest" was given later, at the point when a delayed posttest is typically given in SLA experiments. This format was adopted to avoid the confounding practice effects that can occur when an immediate posttest precedes a delayed posttest (for another methodological approach to this issue, see Avila & Sadoski, 1996; Wang et al., 1992).

Scoring. To ensure that the test measure captured partial learning, participants were given credit for the percentage of each answer that was correct. For example, the test item *hoarse* appeared on the test with five blanks (the "s" was already provided). Since there were five blank spaces for that item, a correct letter occurring at each correct position was counted as worth 20% of the total points for that item.

3. Results

The participants appeared to have ample time to complete the pretest and posttest, thus time pressure was unlikely to have affected the results. The pretest and posttest scores for the three conditions are shown in Table 1. For all tests, the maximum possible score was 30 points.

As can be seen from the pretest results, the participants ($N = 71$) knew only around 10% of the target words prior to the intervention. On the posttest, the scores for the HAM Self-Generated and Sentence Composition conditions were virtually identical at just over 20%, whereas the HAM Given condition led to scores just

Table 1. Pretest and Posttest Scores for the Three Within-Subjects Conditions

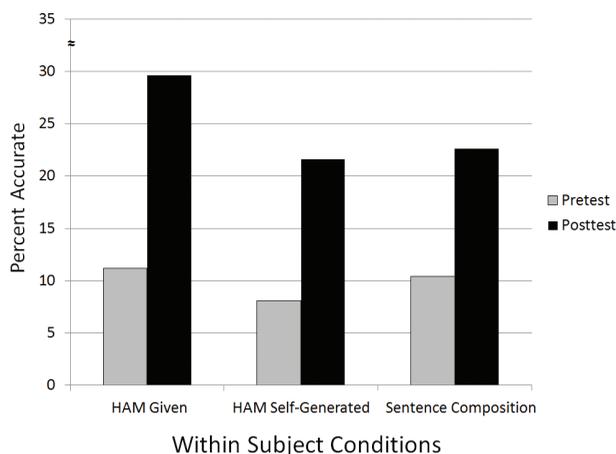| HAM given		HAM self-generated		Sentence composition	
Pretest	Posttest	Pretest	Posttest	Pretest	Posttest
M (*SD*)	*M* (*SD*)	*M* (*SD*)	*M* (*SD*)	*M* (*SD*)	*M* (*SD*)
Range	Range	Range	Range	Range	Range
3.35 (2.6)	8.89 (6.0)	2.44 (2.3)	6.48 (4.2)	3.13 (2.9)	6.78 (4.3)
0.0–9.7	0.0–22.8	0.0–7.7	0.0–21.2	0.0–12.0	0.0–18.0

Figure 1. Percentage of Items Answered Correctly on the Pretest and Posttest for each of the Three Within-Subjects Conditions.

under 30%. The higher SDs on the HAM Given posttest suggests greater variability in participants' ability to take advantage of this type of instruction. Figure 1 shows the same within-subject scores in terms of percentage accurate.

To determine whether the differences between scores in the three conditions were statistically significant, an ANCOVA was conducted with Type of Instruction as the independent variable with three levels (HAM Given, Ham Self-Generated, and Sentence Composition), with posttest scores as the dependent variable, and with pretest scores as the covariate. An alpha of .05 was used for all tests. The results of the ANCOVA indicated that there were significant differences among the three adjusted means, $F(2,209) = 4.00$, $p = .020$, $\eta_p^2 = .037$. Post hoc tests were conducted to determine which conditions differed. The confidence intervals were all adjusted for multiple comparisons using a Sidak correction. Scores on target words learned in the HAM Given condition were significantly higher than those for words learned in the Sentence Composition condition ($p = .025$, 95% CI [0.6, 12.1]), but the difference between the HAM Given and HAM Self-Generated scores fell short of statistical significance ($p = .090$, 95% CI [−0.6, 11.0]). The negligible difference between the HAM Self-Generated and Sentence Composition scores was not significant ($p = .956$).

Partial scoring had only a minor effect on the results. Most responses were 100% accurate or were completely inaccurate, in the latter case, usually due to

participants' failure to provide a response. Fewer than 7% of participants' responses involved partial scores greater than zero and less than 75%. (It should be noted that partial scores over 75% usually reflected minor spelling errors).

4. Discussion

The current experiment sought to determine whether a mnemonic vocabulary learning technique using either researcher-generated or self-generated images to associate the different senses of homonyms and homophones (i.e., HAM) would be more effective than Sentence Composition, a conventional vocabulary technique used to promote elaborative encoding of form-meaning associations. The results suggest that HAM was more effective than Sentence Composition when the associations and images were provided to participants. When participants had to generate their own images and associations, HAM and Sentence Composition appeared to be equally effective.

In the current study, the locus of the effect is not entirely clear. A straightforward explanation would be that HAM, like the keyword method, is effective due to the facilitative effects of imagery on lexical acquisition (Campos, Amor, & González, 2004; Farley, Ramonda, & Liu, 2012; Paivio, 1986). In the current experiment, participants achieved more learning in the condition in which they were provided with images. It could be that these images were a crucial factor promoting greater learning in the HAM Given condition (cp. Thomas & Wang, 1996).

Another possibility is that the associations learned in the HAM Given condition were better retained due to humor. While the effects of humor on learning have been investigated within the general field of education (e.g., Martin, Preiss, Gayle, & Mike, 2006), they have only begun to be explored in the area of vocabulary acquisition (for an overview, see Bell, 2017). Bell (2012) found superior retention for lexical items when the incidental focus on the word involved humor. Along similar lines, some research on the keyword method (e.g., Campos, Amor, & González, 2002) has found that bizarre images facilitate learning more than conventional images. Future research on HAM and related mnemonic techniques may therefore gain greater insights if the presence of humor is treated as an independent variable manipulated within the research design.

In terms of practical implications, the results of the current study are quite positive, suggesting that HAM provides a highly effective alternative to conventional vocabulary instruction techniques. It should be noted that pedagogical applications of the technique will require some adjustments. In the current experiment, the key criterion for selecting target words was that they would be unknown to participants prior to the experiment; hence, most of these words would not be useful to typical students learning English. When applied to actual pedagogical situations, the selection would need to be motivated by students' needs. HAM is likely to be particularly effective when students have good knowledge of one sense of a homonym and can benefit from learning a different sense of the same homonym, or when one member of a homophone pair (e.g., *hair*) is known and the other (e.g., *hare*) unknown. While vocabulary acquisition in the current experiment was fairly modest (an increase from just over 11% to nearly 30% for the HAM Given condition), it should be kept in mind that long-term retention of vocabulary

from a single brief exposure is rare. Typically, deliberate learning of words requires around seven repetitions to result in enduring form-meaning associations (Webb & Nation, 2017). Viewed in this light, the durable learning for words studied in the HAM Given condition in the current experiment was quite remarkable. These retention rates could probably be even further enhanced by combining HAM with some additional techniques such as flashcard review of target words using spaced repetition (Griffin & Harley, 1996; Roediger & Karpicke, 2010).

It must be acknowledged that the homonym association method, being applicable primarily to homonyms and homophones (and to a lesser extent, to near-homonyms, near-homophones, and homographs) is of much more limited scope than the keyword method, which can be used on a large range of words. That said, homonyms and homophones are quite frequent in many languages. According to Parent (2012), 3.3% of English words (75 words) on the roughly 2000-word General Service List (West, 1953) are homonyms, whereas 6.4% of the words (147 words) are homophones. Parent furthermore estimates that around 10% of the 2284 most frequent words in English are homonyms, homophones, or homographs. Wang Ming-tzu and Nation (2004) likewise claim that around 10% of the Academic Word List (Coxhead, 2000) consists of homonyms and homographs.

Future research will need to verify the effects for HAM observed in the current study. Although time on task was controlled, it is not clear whether HAM would be superior to other alternatives such as a word association condition (e.g., rote memorization of L2–L1 pairs), especially if the word pairs were presented multiple times at a fast rate (cp. Hall, Owens, & Wilson, 1987). Although multiple exposures to words within a short session lead to diminishing returns, the gains are usually superior to single exposures (Nakata, 2017). It could be that a paired association learning condition involving multiple retrievals would be more effective than HAM if time on task were controlled or if learning were operationalized in terms of items acquired per fixed time interval (e.g., words successful learned per minute of study). It could also be that the HAM would be even more effective if combined with other conventional techniques. Future research should therefore examine HAM in comparison with (and in combination with) other deliberate learning approaches.

The current study found that self-generated mnemonics were markedly inferior to mnemonics given to the participants. Some general educational research suggests that mnemonics are more effective when they are generated by learners if the learners are adequately trained in the method and have sufficient time during encoding (Putnam, 2015). The current study strictly controlled time on task, so it is possible that participants were unable to generate effective mnemonic associations within the allotted time. The current results are actually in line with previous SLA research on the keyword method, which has found that conditions in which learners generate their own associations lead to similar or less learning than conditions in which the associations are provided (Campos et al., 2002).

Future research should also confirm whether HAM results in representations that are qualitatively inferior (cp. Barcroft et al., 2011) to those developed using other methods. The issue of lexical representation is complex, as there are

currently a number of competing models (Heredia & Brown, 2013). Some models assume that a bilingual's languages are stored separately but are linked to a common conceptual system. Based on this assumption, the Revised Hierarchical Model (Kroll & Stewart, 1994) further claims that lexical links from L2 to L1 words are established prior to the formation of strong semantic links between L2 words and their corresponding concepts. L1 words, on the contrary, have been shown to be strongly linked to their corresponding concepts. This asymmetry, typical of less proficient learners, is said to be especially strong in the case of production (Kroll, Van Hell, Tokowicz, & Green, 2010). L2 representations could also be examined based on distributional models (e.g., Van Hell & de Groot, 1998).

While a detailed discussion of models of bilingual lexical representation is beyond the scope of the current paper, one fruitful avenue of future research would be to use theory-based psycholinguistic behavioral measures to explore the type of representations created through HAM- versus keyword-based instruction. Ideally, such research would take into account learning over long time intervals. Like the keyword method, HAM would presumably lead to representations that are initially mediated by extraneous information (i.e., the image associating the two senses of the homonym or homograph), and this mediation would be expected to slow recall. In the long term, however, HAM and the conventional keyword method may lead to more enduring initial form-meaning links that allow for the development of unmediated form-meaning links as the target word is subsequently encountered incidentally in input and the form-meaning link is processed.

Finally, HAM should, in future research, be compared directly with the conventional keyword method. The keyword method has been described by some researchers as being of "low utility" in practice due to drawbacks related to acceleration of forgetting and difficulty of application (Dunlosky, Rawson, Marsh, Nathan, & Willingham, 2013, p. 24). It would be interesting to know if HAM provides some advantages in this regard.

References

Atkinson, R.C., & Raugh, M.R. (1975). An application of the mnemonic keyword method to the acquisition of a Russian vocabulary. *Journal of Experimental Psychology: Human Learning and Memory, 1*(2), 126–133. https://doi.org/10.1037/0278-7393.1.2.126

Avila, E., & Sadoski, M. (1996). Exploring new applications of the keyword method to acquire English vocabulary. *Language Learning, 46*(3), 379–395. https://doi.org/10.1111/j.1467-1770.1996.tb01241.x

Barcroft, J. (2006). Can writing a new word detract from learning it? More negative effects of forced output during vocabulary learning. *Second Language Research, 22*(4), 487–497. https://doi.org/10.1191/0267658306sr276oa

Barcroft, J., Sommers, M.S., & Sunderman, G. (2011). Some costs of fooling Mother Nature: A priming study on the Keyword Method and the quality of developing L2 lexical representations. In P. Trofimovich & K. McDonough (Eds.), *Applying priming methods to L2 learning, teaching and research:*

Insights from psycholinguistics (pp. 49–72). Amsterdam, The Netherlands: John Benjamins.

Beaton, A., Gruneberg, M.M., & Ellis, N.C. (1995). Retention of foreign vocabulary learned using the keyword method: A ten-year follow-up. *Second Language Research, 11*(2), 112–120. https://doi.org/10.1177/026765839501100203

Bell, N.D. (2012). Comparing playful and nonplayful incidental attention to form. *Language Learning, 62*(1), 236–265. https://doi.org/10.1111/j.1467-9922.2011.00630.x

Bell, N.D. (2017). Humor and second language development. In S. Attardo (Ed.), *The Routledge handbook of language and humor* (pp. 444–455). New York: Routledge.

Boers, F. (2015). Words in second language learning and teaching. In J.R. Taylor (Ed.), *The Oxford handbook of the word* (pp. 582–596). Oxford, UK: Oxford University.

Boyd Zimmerman, C. (1997). Historical trends in second language vocabulary instruction. In J. Coady & T. Huckin (Eds.), *Second language vocabulary acquisition* (pp. 5–19). Cambridge, UK: Cambridge University Press.

Campos, A., Amor, Á., & González, M.Á. (2002). Presentation of keywords by means of interactive drawings. *The Spanish Journal of Psychology, 5*(2), 102–109. https://doi.org/10.1027/1618-3169.51.2.125

Campos, A., Amor, Á., & González, M.Á. (2004). The importance of the keyword-generation method in keyword mnemonics. *Experimental Psychology, 51*(2), 125–131. https://doi.org/10.1027/1618-3169.51.2.125

Campos, A., Rodríguez-Pinal, M.D., & Pérez-Fabello, M.J. (2014). Receptive and productive recall with the keyword mnemonics in bilingual students. *Current Psychology, 33*(1), 64–72. https://doi.org/10.1007/s12144-013-9197-y

Colón, M., & Rodriguez, M. (2004–2005). The effects of self-generated versus supplied interactive images on the long-term retention of EFL vocabulary by college students. *Lenguas Modernas, 30*, 191–200.

Council of Europe. (2001). Common European framework of reference for languages: Learning, teaching, assessment. Cambridge, UK: Cambridge University.

Coxhead, A. (2000). A new academic word list. *TESOL Quarterly, 34*(2), 213–238. https://doi.org/10.2307/3587951.

De Groot, A.M.B., & Van Hell, J.G. (2005). The learning of foreign language vocabulary. In J.F. Kroll & A.M.B. De Groot (Eds.), *Handbook of bilingualism: Psycholinguistic approaches* (pp. 9–29). New York: Oxford University.

Desrochers, A., Wieland, L.D., & Coté, M. (1991). Instructional effects in the use of the mnemonic keyword method for learning German nouns and their grammatical gender. *Applied Cognitive Psychology, 5*(1), 19–36. https://doi.org/10.1002/acp.2350050103.

Dunlosky, J., Rawson, K.A., Marsh, E.J., Nathan, M.J., & Willingham, D.T. (2013). Improving students' learning with effective learning techniques: Promising

directions from cognitive and educational psychology. *Psychological Science in the Public Interest, 14*(1), 4–58. https://doi.org/10.1177/1529100612453266

Elhelou, M.-W.A. (1994). Arab children's use of the keyword method to learn English vocabulary words. *Educational Research, 36*(3), 295–302. https://doi.org/10.1080/0013188940360308

Farley, A.P., Ramonda, K., & Liu, X. (2012). The concreteness effect and the bilingual lexicon: The impact of visual stimuli attachment on meaning recall of abstract L2 words. *Language Teaching Research, 16*(4), 449–466. https://doi.org/10.1177/1362168812436910

Fontana, J.L., Scruggs, T.E., & Mastropieri, M.A. (2007). Mnemonic strategy instruction in inclusive secondary social studies classes. *Remedial and Special Education, 28*(6), 345–355. https://doi.org/10.1177/07419325070280060401

Fritz, C.O., Morris, P.E., Acton, M., Voelkel, A.R., & Etkind, R. (2007). Comparing and combining retrieval practice and the keyword mnemonic for foreign vocabulary learning. *Applied Cognitive Psychology, 21*(4), 499–526. https://doi.org/10.1002/acp.1287

Giora, R. (1997). Understanding figurative and literal language. The graded salience hypothesis. *Cognitive Linguistics, 7*(1), 183–206. https://doi.org/10.1515/cogl.1997.8.3.183

Giora, R., & Fein, O. (1999). Irony comprehension: The graded salience hypothesis. *Humor, 12*(4), 425–436. https://doi.org/10.1515/humr.1999.12.4.425

Giora, R., Fein, O., & Schwartz, T. (1998). Irony: Graded salience and indirect negation. *Metaphor and Symbol, 13*(2), 83–101. https://doi.org/10.1207/s15327868ms1302_1

Griffith, D. (1981). An evaluation of the key-word technique for the acquisition of Korean vocabulary by military personnel. *Bulletin of the Psychonomic Society, 17*(1), 12–14. https://doi.org/10.3758/BF03333652

Griffin, G., & Harley, T.A. (1996). List learning of second language vocabulary. *Applied Psycholinguistics, 17*(4), 443–460. https://doi.org/10.1017/S0142716400008195

Gruneberg, M.M., & Pascoe, K. (1996). The effectiveness of the keyword method for receptive and productive foreign vocabulary learning in the elderly. *Contemporary Educational Psychology, 21*(1), 102–109. https://doi.org/10.1006/ceps.1996.0009

Hall, J.W. (1988). On the utility of the keyword mnemonic for vocabulary learning. *Journal of Educational Psychology, 80*(4), 554–562. https://doi.org/10.1037/0022-0663.80.4.554

Hall, J.W., Owens, W.L., & Wilson, K.P. (1987). Presentation rates and keywords in vocabulary learning. *Bulletin of the Psychonomic Society, 25*(3), 179–181. https://doi.org/10.3758/BF03330321

Hall, J.W., Wilson, K.P., & Patterson, R.J. (1981). Mnemotechnics: Some limitations of the mnemonic keyword method for the study of foreign language vocabulary. *Journal of Educational Psychology, 73*(3), 345–357. https://doi.org/10.1037/0022-0663.73.3.345

Heredia, R.R., & Brown, J.M. (2013). Bilingual memory. In T.K. Bhatia & W.C. Ritchie (Eds.), *The handbook of bilingualism and multilingualism* (2nd ed., pp. 269–291). Malden, MA: Blackwell.

Kroll, J.F., & Stewart, E. (1994). Category interference in translation and picture naming: Evidence for asymetric connections between bilingual memory representations. *Journal of Memory and Language, 33*(2), 149–174. https://doi.org/10.1006/jmla.1994.1008

Kroll, J.F., Van Hell, J.G., Tokowicz, N., & Green, D.W. (2010). The revised hierarchical model: A critical review and assessment. *Bilingualism: Language and Cognition, 13*(3), 373–381.

Laufer, B. (1997). The lexical plight in second language reading: Words you don't know, words you think you know, and words you can't guess. In J. Coady & T. Huckin (Eds.), *Second language vocabulary acquisition* (pp. 20–34). Cambridge: Cambridge University Press.

Lucas, T. (2005). Language awareness and comprehension through puns among ESL learners. *Language Awareness, 14*(4), 221–238. https://doi.org/10.1080/09658410508668838

Martin, D.M., Preiss, R.W., Gayle, B.M., & Mike, A. (2006). A meta-analytic assessment of the effect of humorous lectures on learning. In B.M. Gayle, R.W. Preiss, N. Burrell, & M. Allen (Eds.), *Classroom communication and instruction processes* (pp. 295–313). Mahwah, NJ: Erlbaum.

Nagy, W.E., & Anderson, R.C. (1984). How many words are there in printed school English? *Reading Research Quarterly, 19*(3), 304–330. https://doi.org/10.2307/747823

Nakata, T. (2017). Does repeated practice make perfect? The effects of within-session repeated retrieval on second language vocabulary learning. *Studies in Second Language Acquisition, 39*(4), 653–679. https://doi.org/10.1017/S0272263116000280

Nation, P. (2014). How much input do you need to learn the most frequent 9,000 words? *Reading in a Foreign Language, 26*(2), 1–16.

Nation, P. (2015). Which words do you need? In J.R. Taylor (Ed.), *The Oxford handbook of the word* (pp. 568–581). Oxford, UK: Oxford University.

Paivio, A. (1986). *Mental representation: A dual coding approach*. New York: Oxford University.

Parent, K. (2012). The most frequent English homonyms. *RELC Journal, 43*(1), 69–81. https://doi.org/10.1177/0033688212439356

Pressley, M. (1977). Children's use of the keyword method to learn simple Spanish vocabulary words. *Journal of Educational Psychology, 69*(5), 465–472.

Pressley, M., Levin, J.R., & Delaney, H.D. (1982). The mnemonic keyword method. *Review of Educational Research, 52*(1), 61–91. https://doi.org/10.3102/00346543052001061

Putnam, A.L. (2015). Mnemonics in education: Current research and applications. *Translational Issues in Psychological Science, 1*(2), 130–139. https://doi.org/10.1037/tps0000023

Rodriguez, M., & Sadowski, M. (2000). Effects of rote, context, keyword, and context/keyword methods on retention of vocabulary in EFL classrooms. *Language Learning, 50*(2), 385–412. https://doi.org/10.1111/0023-8333.00121

Roediger, H.L., III, & Karpicke, J.D. (2010). Intricacies of spaced retrieval: A resolution. In A.S. Benjamin (Ed.), *Successful remembering and successful forgetting: Essays in honor of Robert A. Bjork* (pp. 23–47). New York: Psychology Press.

Schmitt, N., & Schmitt, D. (2014). A reassessment of frequency and vocabulary size in L2 vocabulary teaching. *Language Teaching, 47*(4), 484–503. https://doi.org/10.1017/S0261444812000018

Shapiro, A.M., & Waters, D.L. (2005). An investigation of the cognitive processes underlying the keyword method of vocabulary learning. *Language Teaching Research, 9*(2), 129–146. https://doi.org/10.1191/1362168805lr151oa

Sökmen, A.J. (1997). Current trends in teaching second language vocabulary. In N. Schmitt & M. McCarthy (Eds.), *Vocabulary: Description, acquisition and pedagogy* (pp. 237–257). Cambridge, UK: Cambridge University.

Thomas, M.H., & Wang, A.Y. (1996). Learning by the keyword mnemonic: Looking for long-term benefits. *Journal of Experimental Psychology: Applied, 2*(4), 330–342. https://doi.org/10.1037/1076-898X.2.4.330

Tocalli-Beller, A., & Swain, M. (2007). Riddles and puns in the ESL classroom: Adults talk to learn. In A. Mackey (Ed.), *Conversational interaction and second language acquisition: Empirical studies* (pp. 143–167). Oxford, UK: Oxford University.

Tyler, S.W., Hertel, P.T., McCallum, M.C., & Ellis, H.C. (1979). Cognitive effort and memory. *Journal of Experimental Psychology: Human Learning and Memory, 5*(6), 607–617. https://doi.org/10.1037/0278-7393.5.6.607

Urberti, H.Z., Scruggs, T.E., & Mastropieri, M.A. (2003). Keywords make the difference!: Mnemonic instruction in inclusive classrooms. *TEACHING Exceptional Children, 35*(3), 56–61. https://doi.org/10.1177/004005990303500308

Van Hell, J.G., & Candia Mahn, A.C. (1997). Keyword mnemonics versus rote rehearsal: Learning concrete and abstract foreign words by experienced and inexpereienced learners. *Language Learning, 47*(3), 507–546. https://doi.org/10.1111/0023-8333.00018

Van Hell, J.G., & de Groot, A.M.B. (1998). Conceptual representation in bilingual memory: Effects of concreteness and cognate status in word association. *Bilingualism: Language and Cognition, 1*(3), 193–211. https://doi.org/10.1017/S1366728998000352

Wang, A.Y., Thomas, M.H., & Ouellette, J.A. (1992). The keyword mnemonic and retention of second language vocabulary. *Journal of Educational Psychology, 84*(4), 520–528. https://doi.org/10.1037/0022-0663.84.4.520

Wang Ming-tzu, K., & Nation, P. (2004). Word meaning in academic English: Homography in the academic word list. *Applied Linguistics, 25*(3), 291–314. https://doi.org/10.1093/applin/25.3.291

Webb, S., & Nation, P. (2017). *How vocabulary is learned*. Oxford, UK: Oxford University.

Wei, Z. (2015). Does teaching mnemonics for vocabulary learning make a difference? Putting the keyword method and the word part technique to the test. *Language Teaching Research, 19*(1), 43–68. https://doi.org/10.1177/1362168814541734

West, M. (1953). *A general service list of English words*. London, UK: Longman, Green and Co.

Zou, D. (2017). Vocabulary acquisition through cloze exercises, sentence-writing and composition-writing: Extending the evaluation component of the involvement load hypothesis. *Language Teaching Research, 21*(1), 54–75. https://doi.org/10.1177/1362168816652418