TOWARDS AN INSTRUCTIONAL PROGRAMME FOR L2 VOCABULARY: CAN A STORY HELP?

Peter Prince, Laboratoire Parole & Langage, Université d'Aix-Marseille

This paper presents firstly a set of seven principles that lie behind the development of a vocabulary learning resource destined for French learners of English in higher education, and secondly an experiment investigating the use of narrative as a device to facilitate recall of target words. The seven principles can be combined in different ways, and it is suggested that all merit careful consideration when implementing a vocabulary instruction programme. The experiment compared a condition in which target words were embedded in a series of sentences within a narrative framework (story condition) to a condition in which similar sentences bore no connection to each other (unrelated condition). Results indicate that linking sentences containing target words within a narrative framework leads to better recall on an immediate post-test than when sentences are unrelated. The discussion section presents some theoretical and pedagogical implications of these results.

Keywords: Vocabulary, Computer-Assisted Language Learning

INTRODUCTION

In a review article on L2 vocabulary learning, Schmitt notes that many learners "fail to achieve even moderate vocabulary learning goals" (2008, p. 332). The present study evolved from the recognition of this all too familiar situation, and from an attempt to develop a vocabulary learning resource which might help to address it. Though I focus here on just one aspect, namely the use of a narrative framework, the resource as a whole incorporates several other features based on research findings. I therefore first present seven principles which lie behind its conception. These principles represent a personal compilation which overlaps with, but is not identical to other sets (e.g., Hunt & Beglar, 1998).

An experiment is then presented, dealing with the use of a specific type of context, namely a story, within which target words are presented. When stories are generated by learners, in a technique known as narrative chaining, they improve performance in short-term and working memory tasks (McNamara & Scott, 2001) and constitute a powerful mnemonic device for the retention of vocabulary (Bower & Clark, 1969). However, whilst mnemonic devices, such as the keyword method, can be effective (Pressley, Levin & Delaney, 1982; Avila & Sadoski, 1996; Hulstijn, 1997; Sagarra & Alba, 2006), learners seldom engage spontaneously in the more complex processing that the method entails compared to rote memorization (Hulstijn, 1997; Schmitt, 1997; Atay & Ozbulgan, 2007; Barcroft, 2009). The experiment therefore examines the effect of a teacher-provided—as opposed to learner-generated—story on the recall of L2 word forms and meanings.

VOCABULARY INSTRUCTION PRINCIPLES

Principle 1 is to target frequent words. Although the figures put forward vary, it has been estimated that a minimum of the 3000 most frequent word families needs to be known for learners to achieve some understanding of a text, a figure that could rise to as many as 8000 word families if they are to be comfortable with the demands of L2 reading or listening comprehension (Laufer, 1992; Nation, 2006; Schmitt, 2008; Webb & Rodgers, 2009a, 2009b; Laufer & Ravenhorst-Kalovski, 2010). If one is catering to beginners or to learners who despite years of study have not attained this level, the choice of frequency as a criterion for selecting target words is relatively uncontested.

Principle 2 calls for a combination of approaches. Rather than leave learners to acquire vocabulary randomly through reading, it has been suggested that a more systematic approach is needed (Nation, 2001; Hunt & Beglar, 2005, Schmitt, 2008). As Nation (2001, p. 232) points out, though, the two

approaches are not necessarily in opposition. Extensive reading helps learners to acquire words through incidental learning, while various forms of elaborative processing help them to consolidate specific target words.

Principle 3 is to provide opportunities for learners to engage in deep processing (Craik & Lockhart, 1972). Regarding L2 vocabulary, this involves at the least the mental effort of considering the semantic aspects of a word and linking it to other words. Testing the mental effort hypothesis, Hulstijn (1992) found that word meanings were better recalled when learners had to infer them through a multiple choice question than when they were given by means of a synonym. In a further development, the involvement load hypothesis (Hulstijn & Laufer, 2001; Laufer, 2001), proposes that one component of a learner's involvement in processing words is evaluation, in which the word is compared or combined with other words. Not all studies, however, have validated the hypothesis. Contrary to Hulstijn (1992), Mondria (2003) found that in intentional learning conditions, inferring word meanings did not lead to better retention rates than being given these meanings. In a study by Rodriguez and Sadoski (2000), inferring from context led to good immediate recall but in a delayed test retention fell off considerably. However, a combination of context and keyword proved to be highly effective. Other studies have found that, when time on task is taken into account, it may not always be the case that greater cognitive involvement leads to better retention (Webb, 2005; Keating, 2008). In sum, the findings are mixed: much depends on task characteristics, the degree of intentionality involved, and the type of comparison being made.

Principle 4 is to provide opportunities for encountering words in context. Here, the concept of context means the context of a single sentence as opposed to the continuous text advocated in Principle 2. Although inferring word meanings from this type of context can be unreliable (Laufer & Sim, 1985; Laufer, 1997; Frantzen 2003; Kaivanpanah & Alavi, 2008), sentences that are informative enough can be an aid to vocabulary improvement (Webb, 2008), providing a useful intermediary level between words with translations or definitions and longer texts (Laufer & Shmueli, 1997). The benefit of a context sentence may lie more in focusing learners' attention on a word's characteristics than on simply promoting retention (Seibert, 1930; Mondria & Wit-de Boer, 1991). In this respect, a single sentence may not always be more effective than a translation (Webb, 2007); however, the more translation equivalents behave differently in two languages, the more useful context is likely to be.

Principle 5 is to provide multiple encounters with target words. Estimations vary as to how many encounters are needed for word learning, but for a stable lexical representation to be formed, a reasonable aim is six to ten encounters (Zahar, Cobb, & Spada, 2001). This can be achieved by providing appropriate reading texts (Cobb, 2007), but also by ensuring that target words appear in subsequent context sentences.

Principle 6 is to provide translations. Nation (2001, pp. 298–299) summarises a number of studies that point to the effectiveness of learning words with their translations, either in lists or with flash cards. Learning words with their L1 translations remains a favoured strategy of L2 learners (Schmitt, 1997) and given the evidence, it might be wondered why one should bother with any other method. The reason, of course, is that although information other than simply word meanings may be gleaned from translations (Webb, 2007), they say less about the use of the word in L2 than L2 contexts do. Again though, there is no reason why the use of L2 context and L1 translation should be in opposition rather than complementary.

Principle 7 is to provide multiple aids to processing and recall. Presenting words in an enriched context, such that different modalities (auditory and visual) or different contextual clues (e.g., translations, L2 sentences, definitions) are processed, means that more cues are available when it comes to retrieving a memory trace for the purposes of recall. This principle applies as long as the elements present at encoding are additive, such that the extra processing involved increases the chance of recall rather than leading to cognitive overload (Chandler & Sweller, 1991; Plass, Chun, Mayer, & Leutner, 2003; Mayer & Moreno, 2003, Diao & Sweller, 2007). Dual modality presentation (auditory and visual) has been shown to be

effective for the recall of isolated L2 words (Nassaji, 2004) and for text comprehension, though reading only may be just as effective for word learning (Brown, Waring, & Donkaewbua, 2008). Within the visual modality, pictures can provide a strong cue for recall of associated stimuli (Paivio, 1969; Omaggio, 1979), both in a multimedia setting (Chun & Plass, 1996; Kost, Foss, & Lenzini, 1999; Jones & Plass, 2002; Jones, 2004; Yoshii, 2006; Sydorenko, 2010) and in a classroom (Snyder & Colon, 1988). Some studies have also found an advantage for video as opposed to still pictures (Hanley, Herron, & Cole, 1995; Al-Seghayer, 2001; but see Chun & Plass, 1996, for an opposite finding). Within the field of learning words as opposed to glossing text, it has been suggested that pictures are more effective than translations in establishing links to underlying concepts (Finkbeiner & Nicol, 2003; Comesaña, Perea, Pineiro, & Fraga, 2009; Barcroft, 2009; Tonzar, Lotto, & Job, 2009; but see Lotto & de Groot, 1998, for a learning advantage for translations over pictures, and Boers, Piquer Píriz, Stengers, & Eyckmans, 2009, for circumstances in which pictures were of no benefit).

Finally, in a different vein, but still to be regarded as a means to aid recall, is the use of additional links to supplement those provided by translations or context sentences. When such links are deliberately generated by learners, they can be classed as mnemonic devices. However, the material needed can also be provided by the teacher or CALL designer. The device then becomes part of the presentation context, its effectiveness depending on how willing or able the learner is to make use of it. A teacher-generated mnemonic has been shown to be more effective (Hall, Wilson, & Patterson, 1981) in some conditions, though not consistently so (Pressley, Levin, & Delaney, 1982). Whilst the majority of studies on mnemonics in vocabulary learning have focused on the keyword technique, narrative chaining, in which words to be learnt are linked in a story, is the focus of the experiment reported here. However, the aim is not to study the effectiveness of a learner-generated story as in Bower & Clark (1969), but of one provided by the teacher. The use of narrative has received very little attention in the research literature, yet as noted by Bolger & Zapata (2011), stories are widely used in school textbooks. Studying the effect of such a device is therefore of relevance to course designers.

METHODOLOGY

The Effect of Presenting Words in a Narrative Framework

It has been claimed that narrative is a fundamental aspect of cognition and that "we see ourselves, the rest of the world and our life story as a connected chain: a narrative" (Greenway, 2008, p.59). Despite its potential, however, there has been little research into the use of narrative as a device for learning vocabulary. Bower and Clark (1969) found that when participants constructed a meaningful story around the words to be remembered, recall was far better than in the control condition of normal study and rehearsal (93% versus 13%). Bolger and Zapata (2011) examined the effect of presenting semantically related vocabulary in stories. Grouping words according to their semantic relatedness may be a hindrance to learning rather than a help (Tinkham, 1993; Waring, 1997; Folse, 2004; Finkbeiner & Nicol, 2003; Erten & Tekin, 2008), but Bolger and Zapata (2011) found that the contextual support provided by the stories largely offset the disadvantage. Among the specific aspects of narratives that can facilitate recall are narrative structure (Horiba, van den Breuk, & Fletcher, 1993) and the number of connections between the different events within the narrative (Trabasso & van den Breuk, 1985). Unlike the confusion that may result from semantic relatedness, the establishment of associations at a thematic level (e.g., *haunted, moonlight, yell,* and *ghost* within a horror story framework) can be beneficial for vocabulary learning (Tinkham, 1997). Therefore the purpose of this experiment was to see whether a teacher-provided narrative could also help in the recall of L2 vocabulary.

This question was examined within a procedure that along with the use of context sentences (Principle 4) incorporated several of the features described in the other principles. Firstly, the target words were of a high enough frequency to be considered useful to know and as such figured in the vocabulary strand of several of the university's English courses (Principle 1). Secondly, the presentation of the sentences was

followed by an association test devised to encourage deeper processing of the target words (Principle 3). Thirdly, the context sentences containing the target words were accompanied by pictures (Principle 7). These measures were adopted in order to see whether a positive effect from the story would be obtained on top of any effect resulting from the presence of these features. As Chun and Plass (1996) have argued, it is important to examine language acquisition in a realistic environment, where many parameters may be present. Similarly, Bolger and Zapata (2011) stress the need to examine the relative importance of different factors in the learning of L2 vocabulary. The purpose of the present experiment, therefore, was not simply to see whether a story would help learners in recalling word forms and meanings, but to do so in a situation where other aids to learning were also present. This reflects the combinatory approach set out in the seven principles. Another, more practical reason for doing this is that a story per se does not convey word meanings. These can be inferred from context sentences and pictures, but a clearer meaning of a word is unlikely to be conveyed simply because the sentences are linked in a story. The effect of the story, if any, would be to facilitate the organisation of information in memory due to the extra linking opportunities it provides. Although this might help participants to retain word forms, it was necessary to provide additional information that would allow them to reach an understanding of word meanings. It was decided to use pictures for this rather than translations. The choice of words within a certain frequency range meant that some were more abstract than others, providing instances where the meaning to be inferred was not immediately obvious. As already noted, inferring may not in itself lead to better retention, but the aim was to see whether there would be any difference between the two conditions when meanings were not directly given.

The participants were not aware that their recall of the target words would be tested afterwards, but were told that the aim of the experiment was to test the effectiveness of the presentation conditions (picture plus sentence) by means of the association test. This raises the question of the extent to which any learning that occurred was incidental or intentional. Hulstijn (2001) suggests that the distinction between the two is not always easy to discern, and that what counts most is the quality of the mental processing. The two modes of learning can be seen as being on a continuum, with at one extreme the incidental learning that occurs in extensive reading, and at the other, the intentional learning that results from consciously committing information to memory. Clearly, the nature of the experiment meant that participants' attention was directed towards the target words, and they knew that those words would be the focus of the association test, so any learning that might occur cannot be incidental in the same way as it is in studies of vocabulary gains from reading of long texts (e.g., Horst, Cobb, & Meara, 1998). However the purpose of the experiment, as explained to participants, meant that although they were incited to understand the meanings of the target words, there was no incitement to commit the form-meaning link to memory for purposes of production. This did not preclude intentional learning, but restricted it to the recognition requirements of the immediately subsequent association test. Because the experiment examined the effect of a narrative framework on initial exposure to target words, rather than on long-term learning, the aim was thus to avoid the confounding factor of differing rehearsal strategies that could intervene before a recall test (Laufer & Hulstijn, 2001). The hypothesis was that recall in an immediate post–test would be better when context sentences were linked in a narrative framework than when they were unrelated.

Participants

Forty-eight participants performed the experiment in return for a certificate contributing to course credits. They were in their first or second year of Psychology at a French university. There were 42 female and 6 male participants, the average age being 20.6. They had been studying English for 8.5 years on average. They were randomly assigned to one of the two conditions (sentences connected in a story or sentences unrelated to each other). Their proficiency level, assessed by an in-house multiple choice exam, was approximately B1 on the Common European Framework of Reference. No participant had spent more than a month in an English-speaking country.
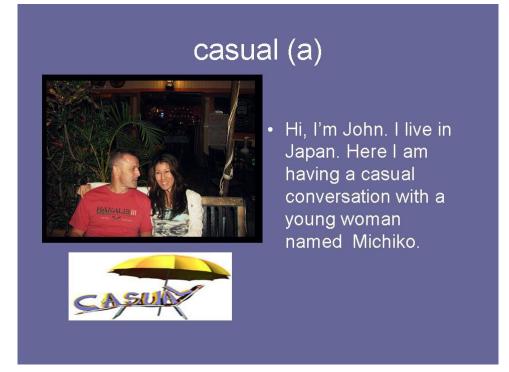
Materials

Sixteen semantically unrelated words were chosen, ranging in frequency rank from 1834 to 4159 in the lemmatised list of the 6318 most frequent words in the British National Corpus (Kilgariff, 1996), with a form frequency ranging from 15.6 per million to 50.1 per million (mean = 22.8, *SD* = 10.3). Two context sentences were composed for each word, one of the sentences being inserted into a narrative framework involving a character who lives in Japan (story condition), the other bearing no relation to other sentences (unrelated condition). The context sentences were similar in both conditions in terms of length (mean length: story condition = 14.8 words, *SD* = 2.4; unrelated condition =13.6 words, *SD* = 2.0), lexical density (mean number of content words other than target words: story condition = 4.4, *SD* = 1.2; unrelated condition = 4.9, *SD* = 1.0) and syntactic complexity (mean number of clauses: story condition = 2.7, *SD* = 0.9; unrelated condition = 2.3, *SD* = 0.6). The critical difference was that in the story condition they contained references to the characters and the setting. Two native English speakers judged the informativeness of the context sentences using Webb's (2008) four-scale classification. In all but two cases, sentences fell into the intermediate categories, whereby participants may gain partial knowledge of the target word's meaning (inter-rater reliability = .78). To complement the sentences, a picture was also provided in both conditions. For concrete words, the pictures gave a clear indication of meaning, but for abstract words they provided additional support for understanding and for retrieval from memory. All but five pictures were identical in the two conditions. Where pictures differed, it was in order to exclude any pictorial reference to Japan in the unrelated condition (see Figure 1 and Figure 2 for examples).
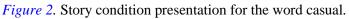

Figure 1. Unrelated condition presentation for the word *casual*.

Between the two conditions, none of the pictures differed completely in terms of composition. The informativeness of the sentence plus picture combination was assessed by a separate group of 24 French learners of English who were given the combinations with the target word missing in the sentences and asked to provide words in French that could fill the gap (Taylor, 1953). Answers were scored

0 (incorrect), 0.5 (partially correct) or 1 (correct) by two English speakers. Inter-rater reliability was .83, with differences resolved by discussion. Overall scores ranged from 0.13 (for the target word *mild)* to 0.96 (for *sword)*, with mean scores showing almost no difference between the two conditions (mean score: story condition = 0.51, *SD* = 0.29; unrelated condition = 0.49, *SD* = 0.29; *t* = 0.51, *p* = 0.6).

Figure 2. Story condition presentation for the word casual.

For each target word, an association test was devised, the task being to indicate which of two words presented along with the target word was closest in meaning to it. The choice was binary rather than multiple in order to keep the test as similar as possible between the two conditions. Although this meant participants had a good chance of guessing the correct answer, the test in itself was not the focus of the experiment and it was felt that the extra processing entailed by a multiple choice test might have led to extraneous differences between the two conditions. A feedback sentence was created for each item, comprising both the target word and the correct response to the question. For example, for the target word *thrust*, the choice was between (A) *push* and (B) *carry*, and the feedback sentence was *The gangster pushed him back and then thrust him against the wall.* The complete list of stimuli appears in the Appendix.

Procedure

Participants were informed that the experiment comprised two phases, one on paper and one on screen. They were then given the pre-test. In the pen and paper pre-test, participants were asked to provide a French translation for the English equivalents of 40 words in which the 16 target words were embedded. Participants were asked to indicate whether any of the words were familiar, and if possible to provide a translation in French. This procedure thus encompassed levels 1 to 4 of the Vocabulary Knowledge Scale (Paribakht & Wesche, 1997). Two raters scored the answers independently, the inter-rater reliability being 0.92, with differences resolved through discussion. Words were scored 1 for a correct translation and 0 when classified as unknown. Partially correct translations or words declared familiar but without translations were scored 0.5. As pointed out by Waring (2002), there is a chance that learners' knowledge could be overestimated this way: for example, some participants noted that they knew the word *thrust*, but

they gave the translation as *confiance* (trust), clearly a synform error (Laufer, 1997). Those who said they were familiar with *thrust* but gave no translation might have been doing the same. However, not to include the dimension of familiarity might mean participants' knowledge is underestimated, since the reactivation of lexical representations, however weak, differs from the establishment of new ones. Since there is no way of knowing if a synform error has occurred when translations are not given, scores were therefore also calculated without the familiarity level. Pre-test scores were naturally lower but uniformly so in both conditions (a mean loss of 0.69 in the story condition versus 0.66 in the sentence condition). The more conservative estimate of post-test gains was retained for purposes of analysis.

Scores at pre-test were not significantly different ($F(1, 47 = .17$, $p = .69$, $\eta^2 = .006$) between participants who subsequently saw words in the story condition (mean = 1.51, *SD* = 1.15) and those who saw the same words in the unrelated condition (mean = 1.6, *SD* = 1.33). After the pre-test, participants were placed in front of individual computers and informed that they would see 16 words from the list in a PowerPoint display in which each word was presented in a context sentence accompanied by a picture. They were told that the words would be presented in two blocks of eight, and after each block they would be given an association test. This test was presented on the screen but participants answered on paper by circling A or B for each item. They were instructed not to write anything else. They were not informed about any further phase to the experiment, nor that some of the sentences formed a story.

The material was counterbalanced such that participants saw all 16 words, half in the story condition and half in the unrelated condition. The presentation order was also counterbalanced, with half of the participants seeing the story condition followed by the unrelated condition and the other half seeing the blocks in the opposite order. The order of words within blocks was not randomised since the story condition imposed a fixed order. Each slide was presented for 15 seconds. Afterwards, participants filled in a sheet with biographical details for approximately three minutes. They were then informed of a third phase. Here they were asked to write on paper as many target words as possible in whatever order they recalled them. Participants were given four minutes for this. The sheet was removed and they were given another and asked to write down the target words again, but this time to provide wherever possible the L1 translations. This procedure differs from the more conventional method of testing for recall and knowledge of target items via cued recall, whether L1 to L2 or L2 to L1. The two different free recall tests were used in order to separate the two components, form and meaning, with a view to allowing participants to engage in a first free recall test without being encumbered by the search for meanings. This does not mean that they did not also recall meanings, but they were not obliged to do so. A point of interest here was whether the order of recall might differ between the two conditions due to the linear progression of the story. After the second free recall, the sheet was removed and participants were given a final sheet containing the L1 translations for the 16 words. They were asked to provide the L2 target words for the translations. Thus there were four scores altogether, not including the pre-test: (a) free recall 1 = free recall of L2 word forms without the need to provide L1 translations; (b) free recall 2 = free recall of L2 word forms plus the provision of L1 translations; (c) L2 to L1 = recall of word meanings as evidenced by the L1 translations of words provided in the second free recall, and (d) L1 to L2 = cued recall of L2 translations. The first three tests, with their emphasis on free recall, tap into the potential organising linear framework of the story; in the final test, L1 words were presented in semi-random order (12 different orders) so as not to replicate the initial presentation order. In all tests, participants were encouraged to write partial word forms, even if they could not remember the whole word. L2 words were scored using Barcroft's (2003) lexical production scoring protocol (LPSP). In this procedure, scores are assigned according to the number of correctly produced letters of a word: .25 for approximately a quarter of the word, .50 for a half, .75 for three quarters, and 1 for a word completely correct. Two raters scored the answers independently, the inter-rater reliability being 0.88, with differences resolved through discussion.

RESULTS

A two-way ANOVA was conducted with presentation condition (story or unrelated) and time (pre-test and post-test) as independent variables and post-test scores as dependent variables. At free recall 1, there was a significant effect of presentation ($F(1, 47 = 7.67)$, $p < .01$, $\eta^2 = .14$), with more words recalled in the story condition than in the unrelated condition. This advantage persisted at free recall 2, though not significantly so ($F(1, 47) = 3.69$, $p = .06$, $\eta^2 = .073$). In translation from L2 to L1, where the L2 words were those which the participants wrote down at free recall 2, there was a significant effect ($F(1, 47) = 7.51$, $p < .01$, $\eta^2 = .138$), with more words from the story condition translated correctly than from the unrelated condition. Finally, in L1-L2 translation there was no significant difference between the two conditions ($F(1, 47) = .47$, $p = .5$, $\eta^2 = .001$). The means for these results appear in table 1. There were significantly more words translated from L1 to L2 than from L2 to L1, ($F(1, 47) = 54.0$, $p < .0001$, $\eta^2 = .535$), and an interaction between direction of translation and presentation, with the gain in performance in the L1 to L2 direction compared to L2 to L1 being significantly higher for words seen in the unrelated condition that for those seen in the story condition ($F(1, 47) = 7.32$, $p < .01$, $\eta^2 = .135$).

Table 1. Mean scores (and gains) in each condition over the four measures

	Story condition		Unrelated condition	
	Mean (gain)	*SD*	Mean (gain)	*SD*
Pre-test	1.51	1.15	1.60	1.33
Free recall 1	3.96 (2.45)	2.10	3.04 (1.44)	1.72
Free recall 2	3.72 (2.21)	2.13	3.13 (1.53)	1.84
L2-L1	3.15 (1.64)	2.06	2.33 (0.73)	1.86
L1-L2	4.15 (2.64)	2.41	4.03 (2.43)	2.39

With regard to word form recall, the precision afforded by the LPSP allows partial word forms to be taken into account, but in fact when words were recalled, they were overwhelmingly fully correct. Of all the scores attributed, full points accounted for 85%. There was, however, considerable individual variation, with scores at free recall 1 ranging from 1.5 to 14.

DISCUSSION

The result of this experiment confirms and extends that of Bower and Clark (1969) relating to narrative chaining as a technique for learning vocabulary. The major difference was that the story condition in this experiment provided participants with a ready-made framework, thus removing the need for them to make it up themselves. The story allows for the linking of semantically unrelated words through a common context or framework and acts as an aid to recall.

Whilst presenting words in a story appears to be effective overall, it should be noted that the advantage for the story condition at free recall 1 did not persist to the same extent at free recall 2, which came immediately after. On average, 0.24 words in the story condition were forgotten between the two recalls while for the unrelated condition there was a gain of 0.09 words. A possible explanation is that at recall 1, the story may have facilitated recall of the target words from episodic memory, while at recall 2, the pictures and the association test may have come into play more, thus diminishing the relative effectiveness of the story. Partial evidence for this comes from the number of words recalled consecutively in the order in which they were presented. At recall 1 there were 39 instances of this in the story condition compared to 17 in the unrelated condition. At recall 2, these scores drop to 14 and 6

respectively. This suggests firstly that when unencumbered by the search for word meanings, participants made use of the story framework as an aid to recall of word forms, and secondly that when they searched their memory a second time, albeit immediately after, they made relatively more use of other cues as well. The experimental design adopted makes it difficult to tease out such effects, but clearly the pictures and feedback sentences contributed to both learning and recall. As they were available in both conditions, they do not constitute a confounding factor, but the findings suggest that the benefit of the story was due to an episodic memory trace quickly superseded by other cues.

Interestingly, the advantage of the story condition was re-established in L2-L1 translation. Here participants were asked to give translations of the words they had themselves written down at recall 2. They were therefore recalling meaning as well as form, and the narrative framework appears to have been helpful in this. One possible, though speculative, explanation is that the propositional content of the sentences is stored in memory in L1, and the story provides a framework allowing better recall of the content, which is then linked to the L2 words that the participant has remembered. In the unrelated condition, this did not result in a corresponding acquisition of word meanings, although performance improved at free recall 2. It should be noted, however, that the gains in the L2-L1 measure were the lowest of all, indicating that providing L1 translations at some point is not just useful but necessary. This can be seen in the final score, L1-L2 translation, where the gains were the highest overall and where there was no significant difference between the two conditions. Participants were able to use the pictures and the association test feedback here in order to link the L1 translations to the L2 forms whose meanings may not have been clear up to then.

Although the overall gains appear to be small, this is in line with other studies that have focused on incidental learning (in the sense that participants were unaware of the post-test). Hulstijn (1992) and Hulstijn, Hollander, and Greidanus (1996) report very low retention rates of word forms and meanings, albeit in conditions somewhat different from the present study in that participants read a complete text. However, target words in those studies were highlighted in various ways such that participants' attention was drawn to them. In the present study, target words were clearly signalled as such, but only three participants suspected they might be tested afterwards. The low rates of retention are therefore unsurprising, but they highlight the difference between retention derived from inferring as opposed to the larger gains to be obtained from memorising (Mondria, 2003). Thus, although the benefit of the story condition was significant, one may ask how meaningful this is when retention rates are so small. As already noted with regard to the pre-test scores, a conservative estimate of gains was used here, and if some words were falsely reported as familiar at pre-test, actual gains may have been slightly higher. Nonetheless, although being informed of a post-test does not necessarily lead to higher retention rates (Peters, 2007; Keating, 2008), it would be interesting to conduct a similar study in which participants are encouraged to commit form-meaning links to memory in order to see whether the story is also effective in helping learners to consciously organise the material to be learnt.

The absence of an intentional learning condition is thus the main limitation to this study, which gives only a partial idea of how the narrative framework operates. When used as a conscious learner-generated mnemonic device, as in Bower and Clark (1969), narrative chaining can be highly effective. The present study shows that a teacher-provided narrative can be an aid to recall when there is no intention to commit words to memory, but there is now a need to investigate a condition lying mid-way between the two, (i.e., when the story is provided and participants are encouraged to make use of it for long-term retention). This would require a delayed test of vocabulary retention as part of a study looking at the effect of consciously using the story as a means to organise target words in long-term memory. As Hulstijn (2001, p. 274) notes, a delayed post-test becomes appropriate when one is investigating the effect of various rehearsal strategies as opposed to the effect of initial exposure to target words. A delayed post-test could also, of course, investigate the effect of initial exposure provided that learners did not encounter the target words in the meantime. This may be possible with invented or low-frequency words but with the present

material could not be guaranteed. As pointed out by Bjork (1999), conditions that promote learning during a training phase may be ineffective in the longer term, and in the absence of a delayed post-test, the present study does not address this issue. However, in real learning conditions, the question is more whether learners engage in rehearsal or not, and if they do, whether organising the words to be learnt in a story can be of help.

Another limitation of this study is the absence among the post-tests of an exact replication of the pre-test conditions (i.e., cued L2-L1 recall), since there was thus no strict comparison between equivalent conditions. Thus, one cannot conclude from these results that the story condition is more effective in a purely receptive test. However, the difference in scores between pre-test and post-test reflects the ability to produce at post-test the form and/or meaning of a word that was not recognised at pre-test and thus indicates a gain in knowledge of that word. Furthermore, it could be argued that the free recall procedure adopted is more stringent than cued recall as it tests the effectiveness of the presentation conditions in triggering recall, first of target word forms and then of meanings. Although there were no doubt words that were not recalled in either condition but whose meanings might have been found in a receptive post-test, this does not invalidate the fact that the story appears to help learners generate cues internally for purposes of recall. This may be of benefit when no other cues are provided, as is often the case in written or oral expression, but it remains to be seen, of course, whether the effect reported here could be transferred to such tasks.

A final comment concerns to what extent the features proposed in the set of principles are additive or not. Demonstrating an effect in isolation is of value, but as noted by Bolger and Zapata (2011) the relative strength of that effect then needs to be assessed by comparing it to other effects. The present study sought to investigate the effect of the story within a presentation procedure that combined other features, but did not incorporate those features as experimental variables. Future research would need to address this issue if we are to gain a clearer understanding of how the different elements of an instructional package should be combined to ensure optimal learning. Think aloud protocols offer one way of sounding out participants' processing of input and thus gaining an insight into the depth to which the various elements of a multimedia presentation are processed (Lomicka, 1998; Bowles, 2004; Yanguas, 2009). From a theoretical standpoint a productive framework is offered by the generative theory of multimedia learning and cognitive load theory (Chandler & Sweller, 1991; Jones, 2004; Mayer & Moreno, 2003), but several questions remain. For example, order of presentation is one parameter that has received little attention, yet it clearly conditions the sort of processing that learners will engage in. In the present study context sentence and picture appeared simultaneously, but depending on the degree of explicitness of each, it is possible that deeper processing could be promoted by presenting them separately. Also, in the case of abstract words, it remains to be seen precisely what role pictures can play. Though they cannot convey the exact meaning(s) of an abstract word as well as translation(s), pictures relating to the sentence as a whole may nonetheless be helpful. While the result of the experiment indicates that the presence of a narrative framework was beneficial, the design did not allow the effect to be teased out with respect to the concrete-abstract dimension. Further research is needed to see how strong this effect is in conjunction with differing types of pictures.

CONCLUSION

The set of principles put forward in this article do not all command unanimous agreement; nevertheless, as they have all garnered sufficient empirical support when regarded as separate features, a consideration of their merits—when combined—should be of interest. The present article has sought to examine an approach to vocabulary learning that combines a rich contextual presentation with a systematic instructional programme.

One obvious pedagogical implication to this study is that course designers may be interested in making use of stories to enhance recall of specific sets of vocabulary items. The task of generating stories

comprising clear context sentences is time-consuming, but when spread between learners who then share the results using multimedia, a large number of words can be covered. Involving learners in the design of certain aspects of their own course can be both motivating (Dörnyei, 2001, p.104) and effective.

Nikolova (2002) found that students who helped to create annotations for target words learned them better than a control group with annotations already given, though not when time on task was taken into account. When I asked my students to compose stories in pairs, each pair using ten target words, considerable creativity was shown, even if the results had to be carefully checked and corrected. The stories were then redistributed, with the target words in bold type, and different pairs were asked to find pictures illustrating the context in which the words appeared. The final results were emailed to the group as a whole and formed the basis of a follow up production activity that was engaged in with much enthusiasm. Though little research has been conducted to assess the effect of stories on vocabulary learning in similarly authentic conditions, the present study suggests that they provide a potentially rich seam to tap.

APPENDIX. Sentences used in the experiment

Unrelated Condition Set 1

Subsidiary (n). She works for an international company which has subsidiaries all over the world.

Spin (v). His head is spinning because his teacher gave him so much work to do.

Thrust (v). The woman thrust a book into her friend's hand. "For you," she said.

Endorse (v). My colleagues at work endorsed my opinion that our boss is very competent.

Damp (adj). The autumn was beautiful, but the winter has been very cold and damp.

Sword (n). Last week, I saw an old sword in a museum.

Bid (v). To buy the picture, many collectors were at the sale to bid for it.

Yield (v). He would like a girlfriend but his attempts to find one have yielded no results.

Unrelated Condition Set 2

Casual (adj). The two young people sat down to have a casual conversation about their favourite films.

Tune (n). The Beatles are known for composing some very popular tunes.

Elbow (n). My brother took me by the elbow and led me over to the window.

Praise (v). He praised the woman's beauty and intelligence, and invited her to dinner at his house.

Trap (n). The animal they caught yesterday is being kept in a trap.

Cling (v). You have to cling hard to the rock if you want to get to the top.

Mild (adj). When he was told the police wanted to see him, he suffered a mild attack of panic.

Vanish (v). They stared at the space ship for as long as they could until it vanished.

Story Condition Set 1

Subsidiary (n). Hi! I'm John. I work for an international company which has subsidiaries all over the world.

Spin (v). My head is spinning because my company is sending me to Japan!

Thrust (n). My boss thrust a Japanese dictionary into my hand. 'Good luck!' she said.

Endorse (v). My colleagues endorsed my boss's opinion that I'm the man for the job.

Damp (a). The autumn in Japan is beautiful, but the winter can be very cold and damp.

Sword (n). My first month in Japan, I bought an old samurai sword.

Bid (v). To buy it, I had to bid for it at a private sale.

Yield (v). I would like a Japanese girlfriend, but so far my attempts to find one have yielded no results.

Story Condition Set 2

Casual (a). Hi, I'm John. I live in Japan. Here I am having a casual conversation with a young woman named Michiko.

Tune (n). Michiko and I danced to a new tune that was popular at the time.

Elbow (n). After dancing, Michiko took me by the elbow and led me over to the window.

Praise (v). I praised her beauty and intelligence, and asked her to be my girlfriend.

Trap (n). But later, when Michiko asked me to stay in Japan, I felt I was caught in a trap.

Cling (v). I accused Michiko of clinging to her Japanese traditions and not trying to understand my position.

Mild (adj). When Michiko said she was leaving me, I suffered a mild attack of panic.

Vanish (v). Shortly afterwards she vanished, and I was unable to find her anywhere. How stupid I am!

Association Test Stimuli

Subsidiary (n):   (a) attached   (b) isolated
Our subsidiaries all over the world are attached to our head office in Chicago.

Spin (v):   (a) grow   (b) turn
The dancer turned in slow circles at first, but then she began to spin quite quickly.

Thrust (v):   (a) push   (b) carry
The gangster pushed him back and then thrust him against the wall.

Endorse (v):   (a) support   (b) imagine
My friends always support me and endorse everything I say.

Damp (adj):   (a) hard   (b) wet
I put my wet pullover in the sun but after an hour it was still too damp to wear.

Sword (n):   (a) fight   (b) walk
If you get into a fight, a sword is not much use against a gun.

Bid (v):   (a) buy   (b) fall
My husband loves buying paintings - last year he bid a million dollars for an artist I'd never heard of.

Yield (v)   (a) produce   (b) follow
They knew the technique would produce what they needed but it yielded more than expected.

Casual (adj):   (a) informal   (b) formal
Our party will be very informal - you can wear casual clothes.

Tune (n):   (a) music   (b) sport
I love that music - it's such a beautiful tune.

Elbow (n):   (a) arm   (b) leg
Hey, mind your arm! You just put your elbow in my face!

Praise (v):   (a) criticise   (b) congratulate
His boss often praised him, and congratulated him on his results.

Trap (n)   (a) flower   (b) bird
The bird was caught in the trap and couldn't get out.

Cling (v)   (a) hold   (b) talk
The young child clings to its parents, holding their hand for comfort.

Mild (adj)   (a) rich   (b) small
It was a small shock so our surprise was only mild.

Vanish (v)   (a) disappear   (b) recognise
The magician's assistant vanished, disappearing in a cloud of smoke.

REFERENCES

Al-Seghayer, K. (2001). The effects of multimedia annotation modes on L2 vocabulary acquisition: A comparative study. *Language Learning & Technology, 5*(1), 202–232. Retrieved from http://llt.msu.edu/vol5num1/alseghayer/default.html

Atay, D., & Ozbulgan, C. (2007). Memory strategy instruction, contextual learning and ESP vocabulary recall. *English for Specific Purposes, 26*, 39–51.

Avila, E., & Sadoski, M. (1996). Exploring new applications of the keyword method to acquire English vocabulary. *Language Learning*, *46*, 379–395.

Barcroft, J. (2003). Effects of questions about word meaning during L2 lexical learning. *The Modern Language Journal*, *87*(4), 546–561.

Barcroft, J. (2009). Strategies and performance in intentional L2 vocabulary learning. *Language Awareness*, *18,*(1), 74–89.

Bjork, R. A. (1999). Assessing our own competence: Heuristics and illusions. In D. Gopher and A. Koriat (Eds.), Attention and peformance XVII. *Cognitive regulation of performance: Interaction of theory and application* (pp. 435–459). Cambridge, MA: MIT Press.

Boers, F., Piquer Píriz, A. M., Stengers, H., & Eyckmans, J. (2009). Does pictorial elucidation foster recollection of idioms? *Language Teaching Research, 13*(4), 367–382.

Bolger, P., & Zapata, G. (2011). Semantic categories and context in L2 vocabulary learning. *Language Learning, 61*(2), 614–646.

Bower, G. H., & Clark, M. C. (1969). Narrative stories as mediators for serial learning. *Psychonomic Science, 14*, 181–182.

Bowles, M. A. (2004). L2 glossing: To CALL or not to CALL. *Hispania*, *87*(3), 541–552.

Brown, R., Waring, R., & Donkaewbua, S. (2008). Incidental vocabulary acquisition from reading, reading-while-listening, and listening to stories. *Reading in a Foreign Language, 20*(2), 136–163.

Chandler, P., & Sweller, J. (1991). Cognitive load theory and the format of instruction. *Cognition and Instruction, 8,* 293–332.

Chun, D. M., & Plass, J. L. (1996). Effects of multimedia annotations on vocabulary acquisition. *The Modern Language Journal*, *80*(2), 183–198.

Cobb, T. (2007). Computing the vocabulary demands of L2 reading. *Language Learning & Technology, 11*(3), 38–63. Retrieved from http://llt.msu.edu/vol11num3/cobb/default.html

Comesaña, M., Perea, M., Piñeiro, A., & Fraga, I. (2009). Vocabulary teaching strategies and conceptual representations of words in L2 in children: Evidence with novice learners. *Journal of Experimental Child Psychology, 104*, 22–33.

Craik, F. I. M., & Lockhart, R. S. (1972). Levels of processing: A framework for memory research. *Journal of Verbal Learning and Verbal Behavior, 11*, 671–684.

Diao, Y., & Sweller, J. (2007). Redundancy in foreign language reading comprehension instruction: Concurrent written and spoken presentations. *Learning and Instruction, 17,* 78–88.

Dörnyei, Z. (2001). *Motivational strategies in the language classroom*. Cambridge: Cambridge University Press.

Erten, I. H., & Tekin, M. (2008). Effects on vocabulary acquisition of presenting new words in semantic sets versus semantically unrelated sets. *System, 36*, 407–422.

Finkbeiner, M., & Nicol, J. (2003). Semantic category effects in second language word learning. *Applied Psycholinguistics*, *24*(3), 369–383.

Folse, K. S. (2004). Myths about teaching and learning second language vocabulary: What recent research says. *TESL Reporter 37* (2) 1–13.

Frantzen, D. (2003). Factors affecting how second language Spanish students deriving meaning from context. *The Modern Language Journal, 87*, 168–199.

Greenway, S. (2008). *ID: The quest for identity in the 21st century*. London: Sceptre.

Hall, J. W., Wilson, K. P., & Patterson, R. J. (1981). Mnemonic techniques: Some limitations of the mnemonic keyword method for the study of foreign language vocabulary. *Journal of Educational Psychology, 73,* 345–357.

Hanley, J., Herron, C., & Cole, S. (1995). Using video as advance organizer to a written passage in the FLES classroom. *The Modern Language Journal, 79*(1), 57–66

Horiba, Y., van den Breuk, P., & Fletcher, C. R. (1993). Second language readers' memory for narrative texts: Evidence for structure–preserving top–down processing. *Language Learning*, *43*, 345–372.

Horst, M., Cobb, T., & Meara, P. (1998). Beyond A Clockwork Orange: Acquiring second language vocabulary through reading. *Reading in a Foreign Language, 11*, 207–223.

Hulstijn, J. H. (1992). Retention of inferred and given word meanings: Experiments in incidental vocabulary learning. In P. J. L. Arnaud & H. Bejoint (Eds.), *Vocabulary and applied linguistics* (pp. 113–125). London: Macmillan.

Hulstijn, J. H. (1997). Mnemonic methods in foreign language vocabulary learning: Theoretical considerations and pedagogical implications. In J. Coady & T. Huckin (Eds.), *Second language vocabulary acquisition* (pp. 203–224). Cambridge: Cambridge University Press.

Hulstijn, J. H. (2001). Intentional and incidental second–language vocabulary learning: A reappraisal of elaboration, rehearsal and automaticity. In P. Robinson (Ed.), *Cognition and second language instruction* (pp. 258–286). Cambridge: Cambridge University Press.

Hulstijn, J. H., Hollander, M., & Greidanus, T. (1996). Incidental vocabulary learning by advanced foreign language students: The influence of marginal glosses, dictionary use, and reoccurrence of unknown words. *The Modern Language Journal, 80*(3), 327–339.

Hulstijn, J. H., & Laufer, B. (2001). Some empirical evidence for the involvement load hypothesis in vocabulary acquisition. *Language Learning*, *51*(3), 539–558.

Hunt, A., & Beglar, D. (1998). Current research and practice in teaching vocabulary. *The Language Teacher*. Retrieved from http://www.jalt-publications.org/old_tlt/articles/1998/01/hunt

Hunt, A. & Beglar, D. (2005). A framework for developing EFL reading vocabulary. *Reading in a Foreign Language, 17*(1), 23–59.

Jones, L., & Plass, J. (2002). Supporting listening comprehension with multimedia annotations. *The Modern Language Journal, 86*(4), 546–561.

Jones, L. (2004). Testing L2 vocabulary recognition and recall using pictorial and written test items. *Language Learning & Technology, 8*(3), 122–143. Retrieved from http://llt.msu.edu/vol8num3/jones/default.html

Kaivanpanah, S., & Alavi, M. (2008). Deriving unknown word meaning from context: Is it reliable? *RELC Journal, 39*, 77–95.

Keating, G. (2008). Task effectiveness and word learning in a second language: The involvement load hypothesis on trial. *Language Teaching Research, 12*(3) 365–386.

Kilgariff, A. (1996). A lemmatised frequency list for the 6,318 words with more than 800 occurrences in the British National Corpus. http://www.kilgarriff.co.uk/BNClists/lemma.num

Kost, C., Foss, P., & Lenzini, J. (1999). Textual and pictorial gloss: Effectiveness on incidental vocabulary growth when reading in a foreign language. *Foreign Language Annals, 32*(1), 89–113.

Laufer, B. (1992). How much lexis is necessary for reading comprehension? In P. Arnaud & H. Béjoint (Eds.), *Vocabulary and applied linguistics* (pp. 126–132). London: Macmillan.

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: Huckin, T. (Eds.), *Second language vocabulary acquisition.* Cambridge: Cambridge University Press, 20–34.

Laufer, B. (2001). Reading, word-focused activities and incidental vocabulary acquisition in a second language. *Prospect, 16*(3), 44–54.

Laufer, B., & Hulstijn, J. H. (2001). Incidental vocabulary acquisition in a second language: The construct of task–induced involvement. *Applied Linguistics, 22*(1), 1–26.

Laufer, B., & Ravenhorst-Kalovski, G. C. (2010). Lexical threshold revisited: lexical text coverage, learners' vocabulary size and reading comprehension. *Reading in a Foreign Language 22,* 15–30.

Laufer, B., & Shmueli, K. (1997). Memorizing new words: Does teaching have anything to do with it? *RELC Journal, 28*(1), 89–108.

Laufer, B., & Sim, D. D. (1985). Taking the easy way out, non use and misuse of clues in EFL reading. *English Teaching Forum 23*(2), 7–10.

Lomicka, L. (1998). To gloss or not to gloss: An investigation of reading comprehension on-line. Language Learning & Technology, 1(2), 41–50. Retrieved from http://llt.msu.edu/vol1num2/article2/default.html

Lotto, L., & de Groot, A. M. B. (1998). Effects of learning method and word type on acquiring vocabulary in an unfamiliar language. *Language Learning*, *48*(1), 31–69.

Mayer, R. E., & Moreno, R. (2003). Nine ways to reduce cognitive load in multimedia learning. *Educational Psychologist, 38*(1) 43–52.

McNamara, D. S., & Scott, J. L. (2001). Working memory capacity and strategy use. *Memory and Cognition, 29*, 10–17.

Mondria, J. (2003). The effects of inferring, verifying, and memorizing on the retention of L2 word meanings: An experimental comparison of the "meaning–inferred method" and the "meaning–given method". *Studies in Second Language Acquisition*, *25*, 473–499.

Mondria, J., & Wit-de Boer, M. (1991). The effects of contextual richness on the guessability and the retention of words in a foreign language. *Applied Linguistics*, *12*, 249–267.

Nassaji, H. (2004). Input modality and remembering name–referent associations in vocabulary learning. *Canadian Journal of Applied Linguistics, 7*,(1), 39–55.

Nation, I. S. P. (2001). *Learning vocabulary in another language.* Cambridge: Cambridge University Press.

Nation, I. S. P. (2006). How large a vocabulary is needed for reading and listening? *Canadian Modern Language Review*, *63*(1), 59–82.

Nikolova, O. (2002). Effects of students' participation in authoring of multimedia materials on student acquisition of vocabulary. *Language Learning & Technology*, *6*(1), 100–112. Retrieved from http://llt.msu.edu/vol6num1/nikolova/default.html

Omaggio, A. C. (1979). Pictures and second language comprehension: Do they help? *Foreign Language Annals, 12*, 107–116.

Paivio, A. (1969). Mental imagery in associative learning and memory. *Psychological Review, 76*(3),241–263.

Paribakht, T. S., & Wesche, M. (1997). Vocabulary enhancement activities and reading for meaning in second language vocabulary development. In J. Coady & T. Huckin (Eds.), *Second language vocabulary acquisition* (pp. 174–200). Cambridge: Cambridge University Press.

Peters, E. (2007). Manipulating L2 learners' online dictionary use and its effect on L2 word retention. *Language Learning & Technology, 11*(2), 36–58. Retrieved from http://llt.msu.edu/vol11num2/peters/default.html

Plass, J. L., Chun, D., Mayer, R. E., & Leutner, D. (2003). Cognitive load in reading a foreign language text with multimedia aids and the influence of verbal and spatial abilities. *Computers in Human Behavior*, *19*, 221–243.

Pressley, M., Levin, J. R., & Delaney, H. D. (1982). The mnemonic keyword method. *Review of Educational Research, 52*, 61–91.

Rodriguez, M., & Sadoski, M. (2000). Effects of rote, context, keyword, and context/keyword methods on retention of vocabulary in EFL classrooms. *Language Learning*, *50*, 385–412.

Sagarra, N., & Alba, M. (2006). The key is in the keyword: L2 vocabulary learning methods with beginning learners of Spanish. *The Modern Language Journal, 90*(2), 228–243.

Schmitt, N. (1997). Vocabulary learning strategies. In N. Schmitt & M. McCarthy (Eds.), *Vocabulary: Description, acquisition and pedagogy* (pp. 199–227). Cambridge: Cambridge University Press.

Schmitt, N. (2008). Instructed second language vocabulary learning. *Language Teaching Research, 12*(3), 329–363.

Seibert, L. C. (1930). An experiment on the relative efficiency of studying French vocabulary in associated pairs versus studying French vocabulary in context. *Journal of Educational Psychology, 21*, 297–314.

Snyder, H., & Colon, I. (1988). Foreign language acquisition and audio–visual aids. *Foreign Language Annals, 21*(4), 343–384.

Sydorenko, T. (2010). Modality of input and vocabulary acquisition. *Language Learning & Technology, 14*(2), 50–73. Retrieved from http://llt.msu.edu/vol14num2/sydorenko.pdf

Taylor, W. L. (1953). "Cloze procedure": A new tool for measuring readability. *Journalism Quarterly, 30,* 415–33.

Tinkham, T. (1993). The effect of semantic clustering on the learning of second language vocabulary. *System, 21*, 371–380.

Tinkham, T. (1997). The effects of semantic and thematic clustering in the learning of second language vocabuary. *Second Language Research 13*(2), 138–63.

Tonzar, C., Lotto, L., & Job, R. (2009). L2 vocabulary acquisition in children: Effects of learning method and cognate status. *Language Learning*, *59*, 623–646.

Trabasso, T., & van den Broek, P. (1985). Causal thinking and the representation of narrative Events. *Journal of Memory and Language, 24*, 612–30.

Waring, R. (1997). The negative effect of learning words in semantic sets: A replication. *System 25*(2), 261–74.

Waring, R. (2002). Scales of vocabulary knowledge in second language vocabulary assessment. *Kiyo, The occasional papers of notre dame seishin univsersity*. Retrieved from http://www.doc88.com/p-51266788660.html

Webb, S. (2005). Receptive and productive vocabulary learning: The effects of reading and writing on word knowledge. *Studies in Second Language Acquisition, 27*, 33–52.

Webb, S. (2007). Learning word pairs and glossed sentences: The effects of a single sentence on vocabulary knowledge. *Language Teaching Research*, *11*, 63–81.

Webb, S. (2008). The effects of context on incidental vocabulary learning. *Reading in a Foreign Language, 20*(2), 232–245.

Webb, S., & Rodgers, M. P. H. (2009a). The vocabulary demands of television programs. *Language Learning*, *59*, 335–366.

Webb, S., & Rodgers, M. P. H. (2009b). The lexical coverage of movies. *Applied Linguistics*, *30*, 407–427.

Yanguas, I. (2009). Multimedia glosses and their effect on L2 text comprehension and vocabulary learning. *Language Learning & Technology, 13*(2), 48–67. Retrieved from http://llt.msu.edu/vol13num2/yanguas.pdf

Yoshii, M. (2006). L1 and L2 glosses: Their effects on incidental vocabulary learning. *Language Learning & Technology, 10*(3), 85–101. Retrieved from http://llt.msu.edu/vol10num3/yoshii/default.html

Zahar, R., Cobb, T., & Spada, N. (2001). Acquiring vocabulary through reading: Effects of frequency and contextual richness. *Canadian Modern Language Review, 57*(4), 541–572.