Improving Incidental L2 Vocabulary Learning
With Latent Semantic Analysis

Akira HAMADA
Graduate School, University of Tsukuba / The Japan Society for the Promotion of Science

Abstract

This study examined whether word-context semantic similarity computed by Latent Semantic Analysis (LSA) predicts the performance of incidental L2 vocabulary learning. LSA is relevant to the usage-based model of language learning, which premises that learners obtain lexical knowledge from the information about how words are used in context. In the experiment, 153 Japanese undergraduates were given 20 target words with contexts whose proposition had higher or lower semantic similarity to those words (HSS vs. LSS context) in the lexical inference and multiple-choice glosses tasks. A Vocabulary Knowledge Scale (VKS) test was used to assess incidental gains in word meaning and usage knowledge. The analyses of the task performances showed that the learners inferred the target word meanings from the HSS contexts more successfully than from the LSS ones, but obtained similar scores between the two context conditions in the multiple-choice glosses task. Nevertheless, the VKS results demonstrated that the HSS contexts greatly contributed to the incidental gains in word meaning and usage knowledge. Thus, LSA predicted the outcomes of incidental vocabulary learning, which indicates that the learners acquired lexical knowledge from usage-based contextual information.

1. Background

Successful vocabulary learning requires a good balance of various learning modes (Nation & Webb, 2011). Additionally, accounts of vocabulary acquisition have emphasized the importance of contextualized learning (e.g., Hamada, 2014; Hasegawa, 2012, 2013; Jiang, 2000; Webb, 2007) because word learning only from definitions often results in poorer performance on tasks using words in context, compared to contextualized learning (McKeown, 1985). In the study of contextualized learning, many researchers have largely focused on incidental learning, or by-products of reading activities (e.g., Fraser, 1999; Hulstijn, 1992; Wesche & Paribakht, 2010). Particularly, incidental learning is supported by the usage-based model, which argues that language can be learned from frequent input of information about how it is used in context (Ellis, 2002).

Given that context shows how words are used in it, maximizing its effectiveness in learning new words is important (Nation & Webb, 2011). In relation to this usage-based model, Landauer
and Dumais (1997) developed the algorithms of Latent Semantic Analysis (LSA), in which a computer learns word meaning from its coincidence frequency with other words used in a particular context (i.e., the degree of word-context semantic similarity) based on the statistical analyses of a large-scale corpus. If this word-context semantic similarity quantified by LSA has high validity for predicting learners’ word leaning performance, it can become a useful factor in creating informative contexts to improve incidental vocabulary learning. This study addressed this issue by expanding past research, which is reviewed below.

1.1 Strategies and Factors Affecting Incidental Vocabulary Learning

The acquisition of new words during reading begins with building a memory representation of these words through inference (Bolger, Balass, Landen, & Perfetti, 2008; Bordag, Kirschenbaum, Optiz, & Tschirner, 2014). Thus, deriving word meaning from context is the primary strategy in incidental vocabulary learning (e.g., Wesche & Paribakht, 2010). Although knowledge gained by inference is rarely full and complete, many prior studies showed that the memory representation constructed by inference is integrated into a mental lexicon as vocabulary knowledge (e.g., Bolger et al., 2008; Bordag et al., 2014; Hamada, 2011; Hulstijn, 1992; Wesche & Paribakht, 2010). However, if used alone, deriving meaning has not always been the most efficient strategy for second language (L2) learners to use. For example, Fraser (1999) reported that L2 learners failed in lexical inference when target words were not semantically constrained by contextual information and the learner ability to use contextual information was poor. In such cases, learners often ignore those unknown words without guessing (Fraser, 1999; Hulstijn, 1993), leading to poor performance in incidental vocabulary learning (Hulstijn, Hollander, & Greidanus, 1996).

The other strategies have been studied by Hulstijn and colleagues, who employed (a) multiple-choice glosses (MCG: Hulstijn, 1992; Hulstijn & Laufer, 2001) and (b) dictionary use (Hulstijn, 1993; Hulstijn et al., 1996). MCG promotes the cognitive processes of evaluation, or hypothesis testing of word meaning inferred from context (Hulstijn & Laufer, 2001). The effect of MCG was also explained by Rott (2005); this technique elaborates the inferred meaning of new words by directing learners’ attention to the word itself. Similarly, dictionary use helps such evaluation and elaboration. Hulstijn (1993) showed that L2 learners consulted a dictionary even after they inferred the meaning of target words. Moreover, evaluating and elaborating the meaning of unknown words represented in mind facilitated word meaning learning (e.g., Hulstijn et al., 1996).

In addition to the lexical processing strategies, other researchers have examined the linguistic factors affecting incidental vocabulary learning. It is beyond the scope of this paper to review all the possible factors, but it is important to recognize that the quality of context near a target word is one of the most influential factors for successful incidental learning of this word, because incidental learning occurs in the course of language comprehension in a meaningful context (e.g., Hulstijn, 1992). For example, contextual features such as contextual constraint and imageability helped learners represent word meanings in memory (Bolger et al., 2008; Hamada, 2013; Hasegawa, 2012,
Thus, these factors are relevant to inference activation, which builds a robust memory representation of words; therefore, readers can gain lexical knowledge with the help of context.

Although these contextual features should improve incidental word learning, it may not be a workable task for educators to create effective texts because collecting this kind of data requires a large number of respondents. In other words, teachers would have to collect experimental data for a particular context if they want to employ these context factors in creating a text for better word learning. For example, the strength of contextual constraint is calculated by the average score of a cloze task (Bolger et al., 2008; Hamada, 2013; Nation & Webb, 2011). Similarly, the imageability of linguistic items is based on human rating using a seven-point Likert scale (Hasegawa, 2012, 2013). Moreover, theoretically, both factors are related to reading comprehension processes (i.e., inference), but not to incidental word learning. In fact, studies of incidental learning have produced mixed results, although context quality was manipulated in terms of how easily a target word meaning can be inferred (Nation & Webb, 2011). These practical and theoretical problems highlight the need to apply a different methodology for creating more effective contexts.

1.2 Latent Semantic Analysis: Usage-Based Model for Word Learning

As the focus of this study is incidental word learning, the theory of how words are represented in context is relevant. Recent research has suggested that the meaning of a word depends on its co-occurrences with other words in a present context (i.e., usage-based model; Ellis, 2002; Inohara & Kusumi, 2012). This background theory affected the development of LSA as a computational tool for extracting and representing the meaning of words (Landauer & Dumais, 1997; Landauer, Foltz, & Laham, 1998). LSA is a simulation model of human language learning, whose most fundamental idea is that we learn language based on the co-occurrence between words and contexts (e.g., Inohara & Kusumi, 2012; Landauer et al., 1998). Empirical research in first language (L1) acquisition has supported this assumption by demonstrating that vocabulary knowledge is incidentally gained from the contexts to which L1 readers usually expose themselves (Inohara & Kusumi, 2012).

Based on the results of machine learning, LSA outputs the strength of semantic similarities between concepts described by words and contexts; it is represented as the cosine of the angle (hereafter, LSA value) formed by high-dimensional vectors (see Figure 1). The theoretical range of the LSA value is from −1.00 to 1.00, and semantic similarity strengthens as the value approximates 1.00 because the angle formed by two vectors approaches zero. Although there are many steps in calculating LSA values, this study introduces only the main two steps (for review, see Inohara & Kusumi, 2011, 2012; Landauer & Dumais, 1997; Landauer et al., 1998; Nahatame, 2012). First, a semantic space using a high-dimensional vector of each word meaning is created based on the information of weighted word frequency with a particular corpus. The second step is to evaluate the strength of semantic similarity between each word and its context. In LSA, large corpora are used as language input for learning word meaning; therefore, LSA values closely reflect representations of human language knowledge (Inohara & Kusumi, 2011; Landauer & Dumais, 1997).
Figure 1. A simplified vector space of semantic similarity computed by LSA (adapted from Hamada, 2014). Each arrow represents a high-dimensional vector of the corresponding word’s meaning.

Although Ellis (2002) claimed the possibility of application of LSA in L2 acquisition, the amount of past research is very limited, particularly in the area of incidental vocabulary learning. In the intentional learning mode, Hamada (2014) demonstrated that knowledge gains in both word meaning and usage were improved when Japanese learners of English as a foreign language (EFL) tried to associate a higher semantic similarity context (HSS; e.g., the LSA value is .69 in “The dog jumped up and licked his face”) with a target word meaning (italic), compared to using a lower semantic similarity context (LSS; e.g., the LSA value is .07 in “The girl craves expensive clothes and bags”). Given that performance on cloze tasks tends to be high when HSS contexts are used (Nahtatame, 2012), these findings suggest that Japanese EFL learners can derive word meaning from the information about the word’s usage in context.

The findings of past studies seem to support the picture of the process and outcome of learning L2 words suggested by the usage-based model (i.e., deriving and learning word meaning from the word-usage information in context). Nevertheless, the implication from the results of intentional learning might apply to incidental learning only in a limited manner (Nation & Webb, 2011). Because learners do not process input with conscious intention to learn it in incidental learning (Hulstijn, 2005), it is possible that the lack of attention to input reduces the knowledge gains that should be improved by HSS contexts. In contrast, another possibility is that HSS contexts improve incidental word learning because this learning mode requires involving contextual information in lexical processing, more than the intentional learning mode does (e.g., McKeown, 1985; Wesche & Paribakht, 2010). Namely, the context effects on word learning might increase more in the incidental mode than in the intentional one. In fact, Hamada (2014) demonstrated that the effects of HSS contexts on knowledge gains in word meaning and usage in intentional word learning were limited (effect sizes of $\eta^2 = .07$ and .01, respectively), although they were still superior to those of LSS contexts. And furthermore, the assumption of the language acquisition model suggested by LSA is essentially an incidental mode (e.g., Bolger et al., 2008). To elucidate these contrastive views, it is important to determine if LSA predicts the performance of incidental L2 vocabulary learning.

This study built a research design to compare the effects of semantic similarity (HSS vs. LSS) on incidental knowledge gains in word meaning and usage when learners engaged in lexical inference, MCG, and dictionary use tasks. It also aimed at examining whether the performance of
deriving word information from context differs according to LSA values because success in lexical inference is the prerequisite of incidental word learning. To assess the acquisition of two types of word knowledge, this study used an unannounced Vocabulary Knowledge Scale (VKS) test; thus, this study followed the definition of incidental learning as the learning mode “in which participants are not forewarned of an upcoming retention test for a particular type of information” (Hulstijn, 2005, p. 132). The experiment was designed to address the following three research questions (RQs): Does the higher semantic similarity between a target word and context computed by LSA improve the learners’ deriving unknown word meaning from context (RQ1), and the incidental learning of word meaning (RQ2) and usage (RQ3)?

2. Method

2.1 Participants

Participants were 153 Japanese EFL learners, who were all but one first-year undergraduates (104 females and 49 males). Their majors were child education, life science, physiotherapy, or nursing. They had studied English for more than six years in Japan. A reading section (k = 50) of a TOEIC Bridge® practice test (Educational Testing Service, 2007) and the 1,000 to 5,000 word level of the Mochizuki vocabulary size test version 3 (Aizawa & Mochizuki, 2010) were used to estimate their English reading proficiency. Given the results of the TOEIC Bridge® (M = 28.58, 95% CI [27.54, 29.61], SD = 6.45, Cronbach’s α = .84) and vocabulary size test (M = 2,754, 95% CI [2,654, 2,867], SD = 665, Cronbach’s α = .95), it was assumed that they were beginner-level EFL learners.

2.2 Materials

This study prepared two types of contexts including each target word: HSS vs. LSS contexts provided by LSA. Table 1 summarizes the characteristics of the contexts and target words.

Target words. A total of 20 target words were selected from past studies (Hamada, 2013; Webb, 2007; see Appendix). Half of them were embedded into HSS contexts (six nouns and four verbs), and the others into LSS contexts (seven nouns and three verbs). Based on Nation and Webb (2011), this study controlled as many factors of the target words critical for learning their meaning and use as possible between the HSS and LSS conditions (i.e., word length such as letters and syllables, word frequency, and word familiarity, reported in Table 1). In rating word frequency and familiarity, Japanese word properties (Amano & Kondo, 1999, 2000) were referred because word learning depends on whether the meaning of unknown words is well lexicalized in learners’ L1 (Wesche & Paribakht, 2010). All target nouns except headline were concrete, but its concept was assumed to be familiar to the participants. The target verbs had no novel or special meanings.

Contexts. The contexts corresponding to the target words were also adapted from Hamada (2013) and Webb (2007). Because the participants’ average vocabulary size was 2,754 words, each context was modified to consist of high-frequency basewords (levels 1 to 2) based on the JACET
8000 list of 8,000 basic words (JACET, 2003). Whereas some prior studies used discourse to examine incidental word learning (e.g., Hulstijn et al., 1996), this study adopted the sentence form because of the participants’ low English reading proficiency. The reason is that sentential information is a primary source of lexical inference, and the use of discourse knowledge requires higher L2 proficiency (e.g., Wesche & Paribakht, 2010). Hence, the sentence was used in this study to ensure that the participants could make better use of contextual information. Sentence length did not differ between HSS and LSS contexts (see Table 1).

Table 1

<table>
<thead>
<tr>
<th>Linguistic and Psycholinguistic Characteristics of Contexts and Target Words</th>
<th>HSS (k = 10)</th>
<th>LSS (k = 10)</th>
<th>Cohen’s (d)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Characters</td>
<td>(M)</td>
<td>(SD)</td>
<td>(M)</td>
</tr>
<tr>
<td>Contexts</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>LSA</td>
<td>.45</td>
<td>.12</td>
<td>.14</td>
</tr>
<tr>
<td>Words</td>
<td>8.90</td>
<td>2.28</td>
<td>9.80</td>
</tr>
<tr>
<td>Target words</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Letters</td>
<td>4.70</td>
<td>1.16</td>
<td>5.90</td>
</tr>
<tr>
<td>Syllables</td>
<td>1.30</td>
<td>0.48</td>
<td>1.80</td>
</tr>
<tr>
<td>Frequency</td>
<td>244.30</td>
<td>256.16</td>
<td>219.90</td>
</tr>
<tr>
<td>Familiarity</td>
<td>5.42</td>
<td>0.55</td>
<td>5.38</td>
</tr>
</tbody>
</table>

*Note.* The value of frequency means the number of times where the target words appear in the NTT database corpus \((N = 341,771)\). The range of word familiarity is 1.00 (very low) to 7.00 (very high).

The strength of semantic similarities between each target word and context meanings was computed using LSA (http://lsa.colorado.edu/), based on the semantic space of “General reading up to first year college” (i.e., the base corpus of LSA; see Dennis, 2007). A total of 10 contexts were determined to have a higher semantic similarity based on the median split on the LSA values. These HSS contexts had substantially higher semantic similarity than the LSS contexts, as shown in Table 1. This ensured that incidental learning outcomes were affected by the semantic similarity computed by LSA if any significant differences were found between the HSS and LSS contexts.

For the MCG task, the gloss options appeared in the margin of each context. Participants had three choices: the definition of the target words, the meaning of words phonologically similar to the target words, and the meaning of contextually unfit words. For example, in “The farmer milked the cattle,” two distractors of “cattle” were *yakan* [kettle] and *bokusou* [grass], respectively. Note that the results obtained from this task were not directly relevant to our research goals. Both MCG and dictionary use tasks were conducted to inform the participants of the correct meaning of all the target words because this study intended the LSS contexts to cause disadvantage in terms of lexical processing in the inference task. If only the inference task was used, different outcomes in incidental
word learning between the HSS and LSS contexts could be simply viewed as the degree of success in deriving target word meanings, but not as effects of word-context semantic similarity.

**Vocabulary Knowledge Scale.** The VKS test was used to evaluate the participants' incidental gains in word knowledge. The VKS is designed to assess knowledge of words, ranging from total unfamiliarity with a target word to the ability to use a target word in context (Wesche & Paribakht, 2010). Therefore, this test was appropriate to the goals of this study (i.e., incidental gains in word knowledge of semantics and usage).

### 2.3 Procedure

Before the experiment, the participants were notified of how the personal data collected would be used and informed consent was obtained. The experiment was conducted in two sessions within a 90-minute class period: (a) the incidental-learning phase (meaning identification tasks including lexical inference, MCG, and dictionary use) and (b) the testing phase.

In the lexical inference task, the 20 contexts were presented in a booklet, with the target words underlined. The participants derived their meanings from the contexts and wrote down their answers in Japanese (20 minutes). At that time, they circled any target words that they had already known prior to this task. The presentation order of the target words differed among participants using random function. In the next task, MCG, they were given the same contexts and target words with three gloss options, and asked to read the contexts again and select the best-fit meaning of each target word (15 minutes). After that, they consulted a dictionary to verify whether the meanings of the target words they selected were correct. Finally, the author confirmed that the participants appropriately completed these meaning identification tasks, and retrieved the worksheets (10 minutes). It should be noted that the participants' intention in these tasks was to practice lexical processing strategies, not to learn the words intentionally.

The participants received another worksheet for the VKS test, in which the target words were presented in an alphabetical order, and they were instructed to recall the target words. They were given as much time as they needed to complete the test (25 minutes). During the incidental-learning phase, the author did not warn them of the VKS test in advance. The participants had taken two profiling tests (i.e., TOEIC Bridge® and Mochizuki vocabulary size tests) in former class periods.

### 2.4 Scoring and Data Analysis

Prior to assessing the task performances and the VKS test, every target word that any participant reported to have known was excluded from the analysis. Strict scoring was employed to evaluate the performance of the inference and MCG tasks; one credit was given when the inferred and selected meaning was correct, as shown in Appendix. The internal consistency of each task was enough (Cronbach's α = .85 for the inference; .76 for the MCG). A two-way analysis of variance (ANOVA) was performed for the mean score rate; it included Context (HSS and LSS) and Task (Inference and MCG) as within-participants variables.
The scoring of the VKS test was made in accordance with Wesche and Paribakht (2010). When the participants reported that they knew a target word meaning, one credit was given if they properly produced its definition (i.e., VKS3-4). When the participants wrote any sentences using a target word, the author determined if its usage was appropriate (i.e., VKS5). Because the participants’ English proficiency was at the beginner level, a credit was given when the target word usage was appropriate even if the whole sentence was not grammatical (e.g., “This puppy are very cute”). Thus, the criterion was whether the target word was used correctly in terms of its part of speech, inflection, semantics, and syntax. Any spelling mistakes were also disregarded. Finally, when a participant wrote a meaning or sentence with a target word but could not complete it accurately, it was treated as if they gained no knowledge of that word’s meaning or usage (i.e., VKS2). Cronbach’s \( \alpha \) was .82 for assessing the gains in word meaning and .85 for word usage.

To examine the knowledge gains in word meaning and usage, VKS scores were interpreted as follows: (a) VKS1-2, in which the participants could not recall the meaning of a target word or use it, was regarded as no gains in word meaning or usage knowledge; (b) VKS3-5 means that they succeeded in integrating the semantic knowledge of a target word into their mental lexicon; therefore, their mean rate was regarded as meaning knowledge gain; and similarly, (c) the mean rate of VKS5 was seen as usage knowledge gain. A two-way ANOVA was conducted; Context (HSS and LSS) and Knowledge (Meaning and Usage) were within-participants variables.

An alpha level of .05 was consistently used in this study. In all subsequent analyses of any interactions and multiple comparisons, Bonferroni adjustment was applied. This study reported ANOVA results with effect sizes and 95% confidence intervals (CI) of mean differences (\( M_{\text{diff}} \)) to interpret the degree to which the Context factor affects incidental L2 vocabulary learning (for review, see Mizumoto & Takeuchi, 2008).

3. Results and Discussion

3.1 Meaning Identification Tasks

Table 2 shows the descriptive statistics of learner performances for the inference and MCG tasks. Considering the usage-based model (i.e., learners derive the lexical properties of words from context), the LSA values should be related to the lexical inference performance. The most important result of the two-way ANOVA was a significant interaction of Context × Task, \( F(1, 152) = 156.03, p < .001, \eta^2 = .05 \). The main effects of Context, \( F(1, 152) = 147.02, p < .001, \eta^2 = .07 \), and Task, \( F(1, 152) = 943.19, p < .001, \eta^2 = .49 \), were also significant. A subsequent analysis was implemented to examine the Context effects on each task performance.

Inference task. Figure 2 summarizes the data distribution of the two tasks’ scores. The Context factor greatly affected the inference performance, \( F(1, 152) = 232.85, p < .001, \eta^2 = .61 \). It showed that the participants could more accurately infer the meaning of the target words from HSS than LSS contexts, and the mean difference was large (\( M_{\text{diff}} = .33, 95\% \text{ CI}[.29, .37] \)).
Table 2

<table>
<thead>
<tr>
<th>Task</th>
<th>HSS context</th>
<th></th>
<th>LSS context</th>
<th></th>
<th>Mean difference</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>M</td>
<td>95% CI</td>
<td>SD</td>
<td>M</td>
<td>95% CI</td>
</tr>
<tr>
<td>Inference</td>
<td>.49</td>
<td>[.45,.54]</td>
<td>.27</td>
<td>.16</td>
<td>[.13,.19]</td>
</tr>
<tr>
<td>MCG</td>
<td>.82</td>
<td>[.79,.85]</td>
<td>.19</td>
<td>.79</td>
<td>[.76,.82]</td>
</tr>
</tbody>
</table>

![Figure 2. Box plot of learner performances (N = 153) for the inference and MCG tasks reported in Table 2. The boxes represent the interquartile range with the median value; the whiskers represent the maximum and minimum values of each condition.](image)

In relation to RQ1, the experiment produced results consistent with the usage-based model. When the target words were presented in the HSS contexts, the performance of L2 lexical inference was greatly promoted. Considering that the HSS contexts evaluated by LSA reflected the high typicality of word meaning in a context (Inohara & Kusumi, 2011, 2012; Landauer & Dumais, 1997; Landauer et al., 1998), this finding indicates that L2 learners extracted word meaning from such word-usage information. For example, the target word *lick* was used in the HSS context like “the dog jumped up and did something to her owner’s face.” In this case, it is reasonable to suppose that the participants found the most plausible meaning of the target word by thinking about how it was used to convey a coherent message; they relied on their prior experiences of word usage in contexts similar to this one.

The effects of contextual features on meaning identification during reading had been established by prior studies. Bolger et al. (2008) and Hamada (2013) showed that high contextual constraint, as measured by a cloze test, promoted the activation of lexical inference. As LSA values are likely to be correlated with the cloze probability provided by lower-proficiency L2 learners (Nahatame, 2012), these two factors may be inseparably related in a natural text. Although it is difficult to distinguish their validity in terms of improving the performance of identifying the meaning of unknown words from context, it can be concluded that LSA is available to predict L2 learners’ performance in lexical inference.

**Multiple-choice glosses task.** In the MCG task, the mean score rates substantially increased, compared to the inference task, by .33 (95% CI [.29,.37]) in the HSS contexts, $F(1, 152) = 267.27, p < .001, \eta^2 = .64$, and by .63 (95% CI [.59,.67]) in the LSS contexts, $F(1, 152) = 1034.18, p < .001,
\[ \eta^2 = .87 \]. Importantly, the mean score rates did not differ between the HSS and LSS contexts, \( F(1, 152) = 1.85, p = .176, \eta^2 = .01, M_{\text{diff}} = .03 \) (95% CI [-.02, .06]).

Prior research often suggests that MCG led to better performance in meaning identification than lexical inference (Hulstijn, 1992; Hulstijn & Laufer, 2001; Rott, 2005). The results of this study showed that, consistent with past studies, meaning identification was improved by this task. Interestingly, the effects of word-context semantic similarity disappeared; however, there are no studies that examined context effects on MCG performance, to my knowledge. In this study, the participants were asked to select the most plausible gloss, distinguishing it from two distractors. Because the distractors were completely unrelated to the contextual propositions regardless of the difference in the LSA values (see Section 2.2), the insignificant differences in the performance of MCG can be attributed to the fact that the distractors were very easy to reject in either type of context.

As explained in Section 2.2, the participants engaged in the MCG and dictionary use tasks to verify the correct meanings of the target words. Obtaining the high and similar scores in the MCG task between HSS and LSS contexts supported the claim that the participants could process all the target words regardless of the context conditions. After the MCG task, they used a dictionary and checked the target word meanings. These ensure that VKS results do not simply reflect the degree of success in lexical inference but the incidental gains in word knowledge affected by the Context factor.

### 3.2 Incidental Word Learning

Table 3 provides the descriptive statistics for the VKS test. Our interest was in the effect of word-context similarity on the incidental gains in word meaning and usage knowledge, and the Context effect was large, \( F(1, 152) = 316.08, p < .001, \eta^2 = .17 \). Moreover, the knowledge gains depended on its type, \( F(1, 152) = 190.93, p < .001, \eta^2 = .18 \), and the two-factor interaction was also significant, \( F(1, 152) = 9.50, p = .002, \eta^2 < .01 \). To answer RQs 2 and 3, this study reported Context effects on incidental learning of word meaning and usage respectively.

<table>
<thead>
<tr>
<th>Knowledge</th>
<th>HSS context</th>
<th>LSS context</th>
<th>Mean difference</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>( M )</td>
<td>95% CI</td>
<td>( SD )</td>
</tr>
<tr>
<td>Meaning</td>
<td>.83</td>
<td>[.80, .86]</td>
<td>.18</td>
</tr>
<tr>
<td>Usage</td>
<td>.50</td>
<td>[.44, .55]</td>
<td>.34</td>
</tr>
</tbody>
</table>

**Incidental gains in word meaning.** Figure 3 displays the context effects on the incidental gains in word meaning knowledge. A post hoc analysis showed that word meaning learning was promoted better in the HSS contexts than the LSS ones, \( F(1, 152) = 232.85, p < .001, \eta^2 = .61 \), and this Context effect was large, \( M_{\text{diff}} = .33 \) (95% CI [.29, .37]).
This result supported the assumption of RQ2, demonstrating that the participants retained the word meanings represented in memory from the HSS contexts better than those from the LSS contexts. Although they knew all the target word meanings regardless of the context conditions thanks to the MCG and dictionary use tasks, a remarkable difference in the incidental gains in word meaning knowledge was found between the HSS and LSS contexts. Similar effects of word-context semantic similarity have been obtained in the intentional learning mode (Hamada, 2014). Thus, this finding is consistent with prior L2 vocabulary learning research at the point that a context factor will appear regardless of the learning modes (Hasegawa, 2012, 2013; Jiang, 2000; McKeown, 1985).

More importantly, the LSA values proved to predict the outcomes of incidental learning of L2 word meaning. This indicates that L2 learners obtain this type of knowledge by experiencing word usage in context, as predicted by the assumption of LSA (e.g., Landauer et al., 1998). Although the participants had an opportunity to derive the target word meanings even from the LSS context in the MCG task, the results show that the HSS contexts helped the participants establish a robust memory representation of the target words. The large difference in word meaning learning can be attributed to making lexical inference based on contextual information. Whereas in the intentional learning using definitions, in which lexical inference is not necessary, the Context effects were medium ($\eta^2 = .07$ in Hamada, 2014), they turned out to be very large in this study. Thus, the findings of this study demonstrated that constructed memory of a word learned from a meaningful context is incidentally integrated into the mental lexicon, as suggested by prior research (Bolger et al., 2008; Bordag et al., 2014; Hamada, 2011; Hulstijn, 1992; McKeown, 1985; Wesche & Paribakht, 2010).

When the possible meaning of unknown words is semantically constrained by context, incidental learning of word meaning also occurs (Bolger et al., 2008). However, there is a substantial difference between word-context semantic similarity and contextual constraint in terms of how the context features can be quantified. The strength of contextual constraint, evaluated by a cloze test, will differ according to any individual differences among learners (e.g., L2 proficiency), and Nation and Webb (2011) have claimed that this technique vitiated ecological validity, relating to how well the findings from experiments reflect the real world. In contrast, LSA represents real-world data based on a corpus linguistic method (Landauer & Dumais, 1997; Landauer et al., 1998). Although this study did not establish which factors are more relevant, the use of natural language data like a
large-scale corpus should be important in discussing the process of L2 vocabulary learning based on the usage-based model and empirical data.

**Incidental gains in word usage.** The context effects on the incidental gains in word usage knowledge can also be visualized in Figure 3. A subsequent analysis demonstrated that incidental learning of word usage was promoted better in the HSS contexts than in the LSS ones, $F(1, 152) = 194.97, p < .001, \eta^2 = .56$, and the Context effect was large, $M_{diff} = .26$ (95% CI [.22, .30]). The other result was that the participants had more difficulty learning word usage than word meanings, regardless of the context conditions: HSS, $F(1, 152) = 171.19, p < .001, \eta^2 = .53$, $M_{diff} = .33$ (95% CI [.29, .37]); LSS, $F(1, 152) = 128.07, p < .001, \eta^2 = .46$, $M_{diff} = .26$ (95% CI [.22, .30]).

Acquisition of word knowledge is an incremental process; partial knowledge of a word grows into full knowledge through multiple exposures (e.g., Bolger et al., 2008; Jiang, 2000; Nation & Webb, 2011). This perspective was replicated by the significant difference in the knowledge gains between word meaning and usage in this study. This result is consistent with past studies, which showed that knowledge gains in word usage were followed by those in word meaning (Bolger et al., 2008; Hamada, 2014; McKeown, 1985; Wesche & Paribakht, 2010).

The more valuable finding is that Context effects were very large in incidental learning of word usage (RQ3). The grand average of the VKS5 rates (7.4 out of 20 target words; see Table 3) showed that the use of various lexical processing strategies, including evaluation and elaboration, led to a certain degree of knowledge gain in word usage. Furthermore, the HSS contexts still promoted incidental learning much better than the LSS contexts. Whereas learning word meanings depends on the quality of context in the intentional learning mode (Hasegawa, 2013), it has null or small effects on learning word usage ($\eta^2 = .01$ in Hamada, 2014; see also Webb, 2007). However, in incidental vocabulary learning, in which learners have to make better use of context, the finding showed that the quality of context greatly affected the incidental gains in word usage knowledge. Such large effects of context are consistent with the assumption of the usage-based model of language learning (e.g., Ellis, 2002). Taken together, the results suggest that when L2 learners engage in cognitive processes to identify the meaning of new words from the information about how to use them in context, this information source (i.e., word usage) can also be integrated into L2 learners’ mental lexicon simultaneously.

**4. Conclusions**

This study demonstrated that the application of LSA improved the effectiveness of incidental L2 vocabulary learning. The general findings showed that the word-context semantic similarity computed by LSA should be considered an essential factor to examine the quality of contexts for purposes of incidental word learning. The answers to the three RQs supported some aspects of the usage-based model of language learning because (a) the Japanese EFL learners could glean lexical properties from the HSS contexts better than from the LSS contexts (RQ1), and (b) they incidentally
gained knowledge of word meaning and usage based on the higher semantic similarity between a target word and context (RQs 2 and 3).

Although the current study expanded past findings of L2 vocabulary learning research, it still has a few limitations. Given that incidental vocabulary learning is an incremental process and it is influenced by various factors (e.g., Bolger et al., 2008; Jiang, 2000), it is important to investigate its interaction with any possible factors, such as frequency of word encounter and individual differences in learners. Frequency of input should especially be considered because the learning source of LSA is also massive input of language (i.e., large-scale corpora), following the usage-based model (e.g., Landauer et al., 1998). Future research, therefore, must focus on the time-dependent change of incidental word learning from multiple exposures to target words (Nation & Webb, 2011). This point is also critical to revealing how lexical knowledge gained from context is integrated into long-term memory because this study could not use a delayed vocabulary test.

This study collected data on word learning performance from practical tasks in a regular class; the results may provide some pedagogical implications. Although all the learners in this study engaged in the same lexical processing strategies, the outcomes of incidental learning differed according to the quality of context. This suggests that LSA can be a useful tool in improving word-learning materials. If an original context does not provide the lexical meaning properties of a target word, it can be modified to better help incidental learning of the target word. Next, LSA can be applied to the modified version of the context to evaluate its informativeness. According to Inohara and Kusumi (2011), the strength of semantic similarity between a context and a target word grows if the context is provided with words that are semantically related to the target word. Because access to the LSA web site is open (Dennis, 2007), teachers can easily compare the quality of the original contexts with revised ones. As the performance of L2 learners in this study was congruent with the prediction of LSA, more studies are needed to consider natural language data and computational tools. Such approaches allow us to improve the effectiveness of incidental L2 vocabulary learning and provide valuable insights into discussing the theory of L2 vocabulary learning.

Acknowledgements

This study was partly supported by Grant No. 25487 from the Japan Society for the Promotion of Science. I would like to express my deep gratitude to Professor Yuji Ushiro, his colleagues, and three anonymous reviewers for their constructive comments to revise an earlier version of this article.

References

The imagery theories of English language learning.

Bilingualism: Language and Cognition. Advance online publication doi:10.1017/S136672891400078


shinmitsudo [NTT database series “lexical properties of Japanese word” vol.1 word familiarity]. Tokyo, Japan: Sanseido.


## Appendix

**Target Words With L1 Equivalents, Contexts, and LSA Values Used in the Experiment**

<table>
<thead>
<tr>
<th>Target word</th>
<th>Part of speech</th>
<th>L1 equivalent</th>
<th>Context</th>
<th>LSA value</th>
</tr>
</thead>
<tbody>
<tr>
<td>HSS condition</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>lick&lt;sup&gt;b&lt;/sup&gt;</td>
<td>verb</td>
<td>なめる</td>
<td>The dog jumped up and licked his face.</td>
<td>.69</td>
</tr>
<tr>
<td>puppy&lt;sup&gt;a&lt;/sup&gt;</td>
<td>noun</td>
<td>子犬</td>
<td>Her little puppy grew up to be a big dog.</td>
<td>.60</td>
</tr>
<tr>
<td>monarch&lt;sup&gt;a&lt;/sup&gt;</td>
<td>noun</td>
<td>君主</td>
<td>Elizabeth II of England is the most famous monarch in Europe.</td>
<td>.48</td>
</tr>
<tr>
<td>sob&lt;sup&gt;b&lt;/sup&gt;</td>
<td>verb</td>
<td>むせび泣く</td>
<td>She stopped crying in a big voice and began to sob.</td>
<td>.48</td>
</tr>
<tr>
<td>spear&lt;sup&gt;b&lt;/sup&gt;</td>
<td>noun</td>
<td>槍</td>
<td>He was killed with the long hunting spear.</td>
<td>.46</td>
</tr>
<tr>
<td>cattle&lt;sup&gt;a&lt;/sup&gt;</td>
<td>noun</td>
<td>家畜用の牛</td>
<td>The farmer milked the cattle.</td>
<td>.41</td>
</tr>
<tr>
<td>doze&lt;sup&gt;b&lt;/sup&gt;</td>
<td>verb</td>
<td>うとうとする</td>
<td>She closed her eyes and dozed.</td>
<td>.37</td>
</tr>
<tr>
<td>mourn&lt;sup&gt;b&lt;/sup&gt;</td>
<td>verb</td>
<td>嘆く</td>
<td>They continue to mourn for years after the death of their friend.</td>
<td>.36</td>
</tr>
<tr>
<td>mane&lt;sup&gt;a&lt;/sup&gt;</td>
<td>noun</td>
<td>長髪</td>
<td>The pretty dancer wears a flower in her golden mane.</td>
<td>.32</td>
</tr>
<tr>
<td>reef&lt;sup&gt;b&lt;/sup&gt;</td>
<td>noun</td>
<td>(サンゴ)礁</td>
<td>The small boat went south around the reef.</td>
<td>.30</td>
</tr>
<tr>
<td>LSS condition</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>pier&lt;sup&gt;b&lt;/sup&gt;</td>
<td>noun</td>
<td>埠頭</td>
<td>My brother and I were at the end of the pier and trying to catch fish.</td>
<td>.23</td>
</tr>
<tr>
<td>vehicle&lt;sup&gt;a&lt;/sup&gt;</td>
<td>noun</td>
<td>乗り物</td>
<td>George taught his son to drive the vehicle.</td>
<td>.22</td>
</tr>
<tr>
<td>boulder&lt;sup&gt;b&lt;/sup&gt;</td>
<td>noun</td>
<td>巨石</td>
<td>The boulder was as large as a small house.</td>
<td>.21</td>
</tr>
<tr>
<td>mako&lt;sup&gt;a&lt;/sup&gt;</td>
<td>noun</td>
<td>サメ</td>
<td>The surfers were attacked by a dangerous mako in the sea.</td>
<td>.17</td>
</tr>
<tr>
<td>recluse&lt;sup&gt;b&lt;/sup&gt;</td>
<td>noun</td>
<td>世捨て人</td>
<td>He was a recluse and never came to the town.</td>
<td>.14</td>
</tr>
<tr>
<td>terminus&lt;sup&gt;a&lt;/sup&gt;</td>
<td>noun</td>
<td>終着駅</td>
<td>The train always arrives on time at the terminus in Tokyo.</td>
<td>.10</td>
</tr>
<tr>
<td>headline&lt;sup&gt;a&lt;/sup&gt;</td>
<td>noun</td>
<td>重要ニュース</td>
<td>My husband always watches the seven o’clock headline in the evening.</td>
<td>.09</td>
</tr>
<tr>
<td>pawn&lt;sup&gt;b&lt;/sup&gt;</td>
<td>verb</td>
<td>賭に入れる</td>
<td>He pawned his watch to buy some new clothes.</td>
<td>.08</td>
</tr>
<tr>
<td>crave&lt;sup&gt;b&lt;/sup&gt;</td>
<td>verb</td>
<td>欲しめる</td>
<td>The girl craves expensive clothes and bags.</td>
<td>.07</td>
</tr>
<tr>
<td>abhor&lt;sup&gt;b&lt;/sup&gt;</td>
<td>verb</td>
<td>ひどく嫌い</td>
<td>We really abhor his English class.</td>
<td>.06</td>
</tr>
</tbody>
</table>

*Note. Contexts<sup>a</sup> were from Hamada (2013), and contexts<sup>b</sup> were from Webb (2007).*