What Effects Do Text-based Tasks Promoting Generation Have on Incidental Vocabulary Acquisition?

ANGELA JOE
Victoria University of Wellington

A large proportion of vocabulary is acquired incidentally from written contexts. However, in text-based studies promoting generative processing, it is not clear if, or to what extent, generation influences incidental vocabulary learning. This study examined the effects of text-based tasks and background knowledge (prior vocabulary knowledge and a disposition to use generative learning tactics when tackling new vocabulary) on incidental vocabulary acquisition. Forty-eight adult ESL learners were randomly assigned to one of three treatments: (a) reading and retelling a text with explicit generative training and without access to the text during recall, (b) reading and retelling a text without explicit generative training but with access to the text during recall, and (c) neither reading nor retelling a text. All subjects sat a pre-test (individual interviews and a read and retell task) and post-tests (individual interviews and two multiple-choice tests) designed to tap partial vocabulary knowledge gains. Results indicated that the process of reading and retelling a text promotes incidental vocabulary learning and that generative processing enhances vocabulary learning with greater levels of generative processing leading to greater vocabulary gains for unknown words.

1 INTRODUCTION

Incidental or indirect vocabulary learning, vocabulary learning which occurs without the specific intent to focus on vocabulary, has been shown to be an effective way of learning word meanings from context (Jenkins et al. 1984, Saragi et al. 1978, Nagy et al. 1985, Day et al. 1991). Several researchers (e.g., Craik and Tulving 1975, Baddeley 1990, Schmidt 1990, Ellis 1991) have suggested that the way in which learners process material influences incidental learning. They stress the importance of ‘noticing’, or attending, in second language learning and maintain that incidental learning can result from task demands which cause learners to focus attention on specific features of input which are crucial for learning.

Few studies, however, have examined the relationship between tasks which engage learners in generative processing (Wittrock 1974) and incidental vocabulary acquisition. Generative tasks require learners to process information at semantic levels and to integrate new information with acquired knowledge. Such tasks demanding ‘deeper’ processing are thought to result in
more effective learning. This study applies the generative learning theory and the depth of processing theory to incidental vocabulary learning. Of particular interest is what learners do in the process of recall to enhance vocabulary learning. By investigating the types of tasks which can lead learners to engage in generative or deeper processing we may be able to promote better tasks for vocabulary learning.

Depth of processing model

The depth of processing model (Craik and Lockhart 1972) has been applied to several vocabulary studies (e.g., Brown and Perry, Jr 1991, Stahl and Fairbanks 1986, Stahl and Clark 1987) as the theoretical basis for explaining how different levels of cognitive processing are engaged in as a result of the demands of a task. One criticism of the levels approach is the difficulty in defining the 'depth' of processing (Craik and Tulving 1975, Baddeley 1990). In this study, the depth of generative processing, pertaining to the learning of vocabulary, was operationalized by the number of novel constructions associated with a target word and the number of semantic elaborations.

Generative model

The generative model is an adaptation of the depth of processing model which focuses specifically on developing the semantic processing level. It promotes the integration of new and known information. Generative models of learning were first introduced by Wittrock (1974) and Slamecka and Graf (1978). The generative learning model posits that learning is a process which involves actively transferring, interpreting, and constructing meaning for unfamiliar concepts, information, and events according to one's prior knowledge, experience, abilities, attitudes, and background (Wittrock 1974).

The underlying assumption behind the generative model is that generative processing, generation or elaboration (Craik and Tulving 1975) leads to improved retention by learners actively generating their own creative versions of language in response to target items read in a text, for example, reformulating in their own words the meaning of a word and enriching and embellishing aspects of the target item with related existing knowledge. This process connects new information with existing information and enriches new items with what is already known.

Several laboratory experiments examining the learning of words from lists (e.g., Hirshman and Bjork 1988, McDaniel et al. 1988, McDaniel and Waddill 1990) support a positive 'generation effect' (Slamecka and Graf 1978) on learning. These studies have shown that subjects who generate words remember them better than subjects who merely read because generation requires the construction of more elaborate connections between new and old information.
Generative tasks

Fincher-Kiefer et al. (1988) maintain that tasks which require reading and recall facilitate deeper cognitive processing and the integration of a learner’s domain knowledge, or background knowledge of a topic, with that from the text. Reading without subsequent recall does not demand the retrieval of information from the text or the structuring of information based on prior knowledge (Fincher-Kiefer et al. 1988).

Ellis (1991) also makes a case for demanding recall in tasks to facilitate the process of second language acquisition. Furthermore, he emphasizes the importance of modifying output and making texts comprehensible to the reader. He argues that tasks which require learners to modify their initial output, such as clarifying and rephrasing speech, help learners to compare new and known information. They can create new hypotheses about the language based on the gaps that they notice between new input and acquired knowledge.

Another factor which may indicate that generative processes are occurring is when learners retrieve target words during recall. Learners who expend effort linking the form of the word with its related concepts have a greater chance of learning the word. Baddeley (1990, 156) states that the act of successfully recalling an item increases the chance that that item will be remembered. This is not simply because it acts as another learning trial, since recalling the item leads to better retention than presenting it again, it appears that the retrieval route to that item is in some way strengthened by being successfully used.

Generation and vocabulary learning

Several studies (e.g., Marks et al. 1975, Stahl and Fairbanks 1986, Stahl and Clark 1987, Hall 1991, Newton 1993) investigating the moderating effect of generative processing on vocabulary learning reported findings which show that generation enhances the acquisition of vocabulary learning.

Of particular interest is the work of Stahl and Clark who used a semantic mapping task, group discussion, and a reading passage to teach fifth graders key science concepts. They demonstrated that the covert process of actively generating new ways to use target words based on existing knowledge and new information enhances vocabulary learning. The critical factor affecting vocabulary learning was cognitive involvement in discussion, regardless of whether learners were active participants in discussion, covertly formulating a response to an answer, or just listeners. A possible reason for this may be attributed to the task conditions. Irrespective of whether one actively listens, overtly contributes, or covertly rehearses a response to discussion, pooled knowledge from group discussion may provide sufficient input for at least incomplete vocabulary knowledge to be established. In addition, the requirement to complete a semantic mapping task and read a text may
encourage more connections to be made between known and new information. Presumably too, greater exposure to target items and greater opportunities to activate relevant prior knowledge on the topic enhanced learning.

Despite Stahl and Clark's findings that overt responses are not necessary to promote vocabulary learning, it is hypothesized that learners who overtly generate target items will learn those items better than learners who do not. This is based on Baddeley and Ellis' work stressing the importance of retrieval and modifications to speech respectively.

The present study attempted to investigate overt generation of vocabulary items in text-based tasks and its effect on incidental vocabulary acquisition. The research questions were the following:

1. Does observable generative processing affect the learning of vocabulary?
2. Will participation in tasks requiring deep level cognitive processing lead to more generative use?
3. Does the degree of generation affect the degree of learning vocabulary?

2 THE STUDY

Subjects

The subjects were 48 adult volunteers learning English as a second language in non-credit bearing courses in Wellington, New Zealand. Forty-one subjects were completing a three-month intensive English proficiency course and the other subjects were participating in an ESL programme in a technical institute. Three quarters of the participants were Asian and the other quarter was comprised mostly of Samoans. The sample included 15 males and 36 females, ranging from 19 to 46 years old. Subjects were placed in lower-intermediate to advanced classes according to placement scores upon entry into the programmes. The battery of tests included dictation, cloze, and vocabulary tests. Their vocabulary scores ranged from 23-95 out of a possible total of 102 with 86 indicating mastery of the first, second, and third thousand general words of English (Nation 1984). A quarter of the learners scored less than 51, four learners scored 86 or above and approximately two-thirds scored between 51 and 86.

Design

Three matched groups—experimental (Group 1), comparison (Group 2) and control (Group 3)—were used in this study. Each group consisted of 16 subjects who did not know the real purpose of the study and were informed that the study was about reading tasks and language learning. They were differentiated in terms of background knowledge. The background knowledge factor was a composite variable of existing vocabulary knowledge, language proficiency, and a disposition to use generative learning tactics when
Pre-test scores from a self-report interview adapted from Wesche and Panbakht (1993) were used to assess both learners' background knowledge of words associated with the topic 'pain' and their knowledge of nine words within the first and second thousand general word list (Nation 1984). Language proficiency was judged impressionistically during the interview and a pre-test read and retell task. The read and retell task was of a similar design to the task used in the main study and a generative scale was used to gauge learners' use of generative tactics during the retelling.

Two levels of the background knowledge factor were distinguished. Those subjects above the group median on the composite variable were categorized high and those below the median, low. Subjects were randomly assigned to form three separate groups. Evenly distributing subjects into high and low background knowledge bands across each of the three groups made it possible to apportion variability in vocabulary learning to both technique and prior learning. Five instructors were trained in using the main study and test procedures.

3 MATERIALS

An expository text, 'Frontiers of Pain Control', was adapted for use in the task (Appendix A). The topic, pain, was chosen because the underlying concept was one which learners were likely to be familiar with and have had some experience and knowledge of.

Vocabulary

With the exception of the target words, the vocabulary was kept within a first and second thousand vocabulary list (Nation 1984). The aim of controlling the vocabulary was to bring the majority of the words within the learners' experience without oversimplifying the text. In addition, reducing the number of unknown words in the text frees up the amount of cognitive space required to attend to the message. Tasks demanding greater mental effort or more decision making use up more processing space. By enabling learners to process known information effortlessly, more effort could be spent focusing on crucial concepts and items pre-selected by the teacher (Nation 1990).

Calculating the number of target words in the text

Based on Liu and Nation (1985) and Laufer's (1989) suggestions that about 95 per cent\(^2\) of words in the text should be known, 12 target items were chosen from the 338-word text 'Frontiers of Pain Control'.
Selecting target words

Three measures of word difficulty were considered. First, Nation's (1984) vocabulary lists were consulted to determine the difficulty level of words in the text. All the words in the text beyond the first 2,000 word list were considered for target words. Second, three participants from pilot studies identified what they considered to be the 12 most difficult and important words in the passage. Third, three experienced teachers of English as a second language identified 12 words in the text according to two criteria: the high level of difficulty and high information value. Those items contained within the text which were beyond the first and second 1,000 word list and not identified as target words were replaced with more familiar words or phrases within the second thousand words.

4. MEASURES

Three measures were used to assess the effects of the different instructional techniques on vocabulary knowledge: a self-report interview procedure, multiple-choice questions, and a generative scale. The interview and the multiple-choice tests measured the product of learning, and the generative scale was used to measure the possible cause of vocabulary gains made in the process of learning. Because the present study aimed to detect any partial gains in vocabulary knowledge, it was imperative that the measures employed were sensitive to shifts in degrees of knowledge and that different degrees of word knowledge could be tapped with each type of measure. Each of the three measures enabled learners to demonstrate different aspects of word knowledge. These measures are outlined below.

Self-report interview

The interview procedure and self-report descriptions were adapted from Wesche and Paribakht's (1993) 'Vocabulary Knowledge Scale' (VKS). The interview involved the learners reporting knowledge of a word using the VKS (Table 1) and allowed learners to express their depth of word knowledge using whatever contexts they chose. Subsequently, instructors elicited extra information from the learners to clarify or develop previous responses. Subjects were credited with a score from one to six for each target word. Interpretations for each score are given in Table 2.

The scale of generativeness

According to Wittrock (1974, 1991), generative processing leads to better recall. However, it is not clear to what extent generation or elaboration affects how well a word is learned. Consequently, a holistic scale was constructed to investigate the extent, if any, to which a learner's observable level of generation in a retelling task affected vocabulary learning. Retelling key ideas...
Table 1 Vocabulary Knowledge Scale  Self-report categories

<table>
<thead>
<tr>
<th>Score</th>
<th>Category</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>I haven't seen this word before</td>
</tr>
<tr>
<td>2</td>
<td>I have seen this word before, but I don't know what it means</td>
</tr>
<tr>
<td>3</td>
<td>I haven't seen this word before, but I think</td>
</tr>
<tr>
<td>4</td>
<td>I have seen this word before and I think</td>
</tr>
<tr>
<td>5</td>
<td>I know this word  It means</td>
</tr>
<tr>
<td>6</td>
<td>I can use this word in a sentence</td>
</tr>
</tbody>
</table>

(Adapted from Wesche and Panbakht 1993)

from the text about pain required learners to retrieve target word forms and to demonstrate semantic and syntactic knowledge within a prescribed context. The theoretical basis for the scale of generativeness comes from Craik and Lockhart's (1972) depth of processing theory, Craik and Tulving's (1975) elaboration theory, Wittrock's (1991) generative learning model and research on the dimensions of vocabulary knowledge (Richards 1976, Ringbom 1987, Nation 1990) The scale assessed both the degree of generation evident in the retelling overall and the degree of generation demonstrated for one item As learners progressed to higher levels of generation they were required to construct more connections between data-driven information, previously acquired knowledge and experience beyond the text Inter-rater reliability.

Table 2 Demonstrated word knowledge

<table>
<thead>
<tr>
<th>Score</th>
<th>Interpretation</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>The word is not familiar at all</td>
</tr>
<tr>
<td>2</td>
<td>The word is familiar but the meaning is not known</td>
</tr>
<tr>
<td></td>
<td>An affix is familiar but the base and general meaning is unknown</td>
</tr>
<tr>
<td>3</td>
<td>The word is familiar, an association or general meaning is given</td>
</tr>
<tr>
<td>4</td>
<td>The correct synonym, paraphrase or translation is given</td>
</tr>
<tr>
<td></td>
<td>The word is used accurately and appropriately in a sentence, but only the general meaning is supplied</td>
</tr>
<tr>
<td>5</td>
<td>The word is used with semantic appropriateness in a sentence</td>
</tr>
<tr>
<td>6</td>
<td>The word is used with semantic appropriateness and grammatical accuracy in a sentence (also uses acceptable collocations)</td>
</tr>
</tbody>
</table>

(Adapted from Wesche and Panbakht 1993)
measured by the items marked identically by the researcher and inter-rater, was 83 per cent

Levels

Each level on the four point scale is presented below. The examples provided are taken from subjects recalling the meaning of the word *acupuncture* after reading the text *Frontiers of Pain Control*.

0—No generation  Verbatim repetition from the original text. There appears to be no effort to connect the target word with its specific meaning, its associated concepts, appropriate contexts or other known words, either within or beyond the text.

In this method, needle are inserted into at particular points on the body’s surface to relieve pain in other parts of the body.

1—Low generation  Learners demonstrate very little effort to connect the target word with its specific meaning, its associated concepts, appropriate contexts or other known words either within or beyond the text. Learners provide very general definitions for target items which do not go beyond what is read in the text. Learners focus on a word’s structural features, such as form, rather than on meaning. Learners attempt to integrate prior knowledge or experience with words and their concepts but cannot draw the relationships together.

The pain is so terrible things, so you have to go to acupuncture. Acupuncture is very useful treatment for chronic pain, so I don’t understand acupuncture.

2—Reasonable generation  There appears to be some effort to connect the target word with its specific meaning, its associated concepts, appropriate contexts or other known words within and beyond the text. Learners use their own words to integrate key concepts with personal experience and existing knowledge. However, some key concepts or important details are inaccurate or not very specific. Learners need to clarify meaning with more elaboration.

They point the needle in the sore area and they replace it, decrease the pain.

3—High generation  There appears to be considerable effort to connect a target word with its specific meaning, its associated concepts, appropriate contexts and other known words within and beyond the text. Learners interrelate information from different parts of the text to verify their predictions, they use paraphrasing, examples, analogy, synonyms and elaboration. Learners integrate key words and concepts from the text with personal experience or existing knowledge.

Maybe you know, it is really from China. They use needles and you
have special spots just for the pain on the body, and if you put the
needle to the spot you will get cured

Multiple-choice tests
The multiple-choice tests provided opportunities for subjects to demonstrate
their receptive word knowledge which they might not have been able to
express in the other tests Forrethermore, the multiple-choice tests verified
learners' knowledge of the words reported to be familiar or known in the
interview
Two sets of distractors, differing in form, were constructed for the same
set of stems to differentiate between learners who knew the finer shade of
meaning for a word from learners who only had a vague sense of the
word An example of the two levels of difficulty is shown in Table 3 Each
of the twelve items in both the easy and difficult versions of the test
consisted of a stem, a key, three distractors and a 'don't know' option
Moreover, the multiple-choice items in the easy and difficult test were
ordered differently

Easy version
The easy version of the test was sensitive to partial word knowledge It aimed
to maximize opportunities for the learners to demonstrate anything they
knew about a word The correct answers for each of the twelve items provided
a very general description of the word, hence vague knowledge of the broad
semantic category to which the word belonged would have been sufficient for
learners to select the correct answer Knowledge of a word's part of speech
could also have provided adequate clues Each of the distractors was
semantically distant from the stem, but still within the same topic domain

Difficult version
In contrast, the difficult version of the multiple-choice test aimed to
discriminate between learners with a greater depth of knowledge about a
target word and those with only a receptive or partial knowledge of the word
Each of the distractors was phrased using the same part of speech and was
semantically close to the target item Selecting the correct answer required
specific knowledge of a word's meaning within a specific context since very
fine distinctions of meaning separated each of the distractors Each correct
response for the easy version of the test was given one mark and correct
responses for the difficult test were given two marks

5 PROCEDURE
The study was implemented over three stages with both Group 1
(experimental) and Group 2 (comparison) participating in the following
Table 3 Example of two levels of multiple choice items

Easy

tolerance means
(a) to put yourself in danger
(b) a fit and healthy lifestyle
(c) the ability to accept something
(d) to have particular beliefs
(e) don’t know

Difficult

tolerance means
(a) the inability to accept drugs
(b) the ability to give someone drugs
(c) when your body has no bad reactions to a drug
(d) the ability of knowing when you do or do not need to use drugs
(e) don’t know

1 Pre-test (self-report interview and a read and retell task).
2 Treatment, in the form of a practice read and retell task and the main read and retell task. The nature of the treatment differed between the two groups, as described below.
3 Post-tests (self-report interview, easy multiple-choice, difficult multiple-choice)

Group 3 (control) only performed the pre-test and post-tests and continued with their regular class programme without receiving any treatment.

Pre-test

The pre-test consisted of a self-report interview on learners’ vocabulary knowledge of 28 words and a read and retell task of similar design to the main task. The 28 words consisted of 12 target words from the expository text, three other words from the text related to pain, four words beyond the text from the domain of pain, and nine words from the first and second 1,000 general vocabulary list (Nation 1984). The target words were not obviously noticeable and constituted 42 per cent of the words in the list. Learners’ knowledge of domain and target words was assessed using the scores from the pre-test interview. Learners were instructed not to use a dictionary to look up any words they had encountered in the pre-test. They individually performed the read and retell task with their instructors in separate rooms using think-aloud procedures.
Treatment

After a break of at least two days, subjects in Groups 1 and 2 participated in separate group practice sessions carried out by the researcher. Learners in Group 2 practised a fifteen-minute read and retell task similar to the main task. To begin, learners activated background knowledge by predicting the answers to two questions based on the text and read five cue questions intended to guide reading and retelling. They then read and retold the key concepts from the passage using think aloud phrases, such as ‘I’m not sure but I think ...’ to explain concepts they were either unsure of or unfamiliar with. If necessary, learners referred to the reading passage.

Learners in Group 1 received the same treatment as those of Group 2 with the following exceptions: subjects in Group 1 did not have access to the input text during the retelling of the key points and they received explicit instruction on generative learning strategies. Specifically, learners were instructed in the following ways: (a) to recall prior experience and knowledge to make sense of unfamiliar concepts or words in the text, that is, learners were told to add their own examples, experience, and knowledge to information from the text, and to offer personal opinions and comments, (b) to paraphrase, use synonyms, examples, or analogy, and (c) to discuss why some examples of learners’ generative strategies were better than others.

At least a day later, individual subjects in Group 1 and 2 worked with their assigned instructors and completed the main read and retell task. Instructors read aloud ‘pre-reading’ and ‘while reading’ questions before learners silently read the text ‘Frontiers of Pain Control’ (see Appendix A), then, immediately after reading, they were instructed to reread the questions. Subsequently, the subjects proceeded with the retelling. Comparison group subjects had the option of referring to the text when necessary whereas the experimental group subjects did not.

Post-tests

All three groups completed the post-tests. The pre-test and post-test interview procedures were identical but the post-test list contained only the 12 target words from the text. The sequencing of the post-tests consisted of the self-report interview, the easy multiple-choice test and finally, the difficult multiple-choice test. Throughout stages one and three all subjects were audio-recorded.

Data analysis

To compare the effect of the two ways of performing a read and retell task, on the one hand, with no treatment at all, a 3x2x2 factorial analysis of variance (ANOVA) was used, with instructional technique (instruction, without text, no instruction, with text, control), background knowledge (high versus low),
and display of vocabulary learning as factors (interview and multiple-choice).
Each group's vocabulary performance was based on a post-test interview and
the multiple-choice tests. The two groups participating in the read and retell
tasks were expected to outperform the group that did not do the task. A
moderating factor, background knowledge, was expected to influence
vocabulary learning, irrespective of treatment.

A 2x2x3 ANOVA compared two ways to deliver the instructional technique,
and employed an additional dependent variable—vocabulary used in the free
recall of text. The group that received explicit instruction on generative tactics
and retold the passage without the aid of the text was expected to show
greater vocabulary gains than the group that did not receive explicit
instruction but had the text available to them during retelling. In addition,
high background learners were expected to attain higher vocabulary scores
than low background learners and learners demonstrating high generation in
recall were expected to make greater word knowledge gains than learners
demonstrating low generation.

A hierarchical repeated measures design was selected for both ANOVAs
with an alpha level of .05 for the 3x2x2 ANOVA and a level of .01 for the
2x2x3 ANOVA. The 2x2x3 ANOVA had a higher alpha level to account for
some data in the 3x2x2 ANOVA being included in the analysis.

6 RESULTS
3x2x2 ANOVA
A statistically significant effect was found for treatment and background
knowledge with levels \( p < .001 \) for \( F (2, 42) = .1733 \) and \( F (1, 42) = 38.36 \)
respectively. An \textit{a posteriori} test of significance (Winer 1971: 564) revealed
that both the experimental and comparison groups performed significantly
better than the control group, but did not differ significantly from each other.
No statistically significant interaction effects were found for final scores taken
from the three dependent measures in Table 4. In other words, participation
in read and retell tasks as well as high background knowledge were associated
with substantial vocabulary gain. Predictably, learners who did not perform
the task, or had low background knowledge, made fewer word knowledge

2x2x3 ANOVA
The results of the second ANOVA revealed no statistically significant
differences between the two forms of the instructional technique, but again,
there was a significant effect for background knowledge \( F (1, 28) = 35.0, \)
\( p < .001 \)
Table 4 Means and (standard deviations) for the three dependent measures expressed in raw scores

<table>
<thead>
<tr>
<th>Group</th>
<th>Background</th>
<th>*Multiple choice</th>
<th>#Interview</th>
<th>*Recall</th>
</tr>
</thead>
<tbody>
<tr>
<td>Experimental</td>
<td>high</td>
<td>28.5 (3.16)</td>
<td>53.4 (9.8)</td>
<td>18.0 (3.3)</td>
</tr>
<tr>
<td></td>
<td>low</td>
<td>16.1 (5.84)</td>
<td>37.8 (10.2)</td>
<td>9.6 (6.2)</td>
</tr>
<tr>
<td>Comparison</td>
<td>high</td>
<td>27.2 (5.84)</td>
<td>54.8 (8.6)</td>
<td>19.9 (4.1)</td>
</tr>
<tr>
<td></td>
<td>low</td>
<td>19.8 (5.65)</td>
<td>37.5 (6.3)</td>
<td>8.6 (5.9)</td>
</tr>
<tr>
<td>Control</td>
<td>high</td>
<td>17.7 (7.47)</td>
<td>39.3 (10.4)</td>
<td>–</td>
</tr>
<tr>
<td></td>
<td>low</td>
<td>11.5 (7.19)</td>
<td>22.5 (2.1)</td>
<td>–</td>
</tr>
</tbody>
</table>

*Total possible 36
#Total possible 72

Generation and retrieval of items in protocols

The target items used with greater frequency in the protocols were important for comprehension of the text and were learned better than less salient target items.

The subjects who saw the text during recall used the words much more.

Table 5 Comparison of item retrieval verbatim use versus generation

<table>
<thead>
<tr>
<th>Target word</th>
<th>Comparison</th>
<th>Experimental</th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Verbatim</td>
<td>Generated</td>
<td>Verbatim</td>
<td>Generated</td>
<td>Total</td>
<td></td>
</tr>
<tr>
<td>Chronic</td>
<td>14</td>
<td>48</td>
<td>3</td>
<td>42</td>
<td>107</td>
<td></td>
</tr>
<tr>
<td>Acupuncture</td>
<td>14</td>
<td>24</td>
<td>4</td>
<td>30</td>
<td>72</td>
<td></td>
</tr>
<tr>
<td>Narcotics</td>
<td>13</td>
<td>34</td>
<td>–</td>
<td>16</td>
<td>63</td>
<td></td>
</tr>
<tr>
<td>Placebo</td>
<td>15</td>
<td>21</td>
<td>1</td>
<td>23</td>
<td>60</td>
<td></td>
</tr>
<tr>
<td>Electrical-stimulation</td>
<td>6</td>
<td>23</td>
<td>1</td>
<td>14</td>
<td>44</td>
<td></td>
</tr>
<tr>
<td>Electrode</td>
<td>7</td>
<td>9</td>
<td>–</td>
<td>8</td>
<td>24</td>
<td></td>
</tr>
<tr>
<td>Severe</td>
<td>4</td>
<td>5</td>
<td>–</td>
<td>6</td>
<td>15</td>
<td></td>
</tr>
<tr>
<td>Psychological</td>
<td>1</td>
<td>6</td>
<td>–</td>
<td>1</td>
<td>8</td>
<td></td>
</tr>
<tr>
<td>Addictive</td>
<td>1</td>
<td>3</td>
<td>–</td>
<td>2</td>
<td>6</td>
<td></td>
</tr>
<tr>
<td>Incapacitated</td>
<td>4</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>4</td>
<td></td>
</tr>
<tr>
<td>Tolerance</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td></td>
</tr>
<tr>
<td>Vicinity</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td></td>
</tr>
</tbody>
</table>
frequently than the experimental subjects. However, further classifying the repeated items into (a) words that were generated such as

I'm not sure of the name of the method exactly but one of them is to put some needles at the point of pain the point they have painful and the needle can release the painful and the name of the method I said acupuncture

or (b) words that were merely repeated verbatim, for instance, 'Chronics pain is sometimes so severe that some people are incapacitated yeah incapacitated by it', revealed a different picture. Although subjects who saw the text during recall repeated the form of the word without generation more frequently, the frequency with which they generated target items was similar in number to those shown by the experimental group. Perhaps, the more often a word is retrieved, the greater the opportunities for that word to be generated.

The design of the task may have led some comparison learners to reproduce information in the text verbatim, rather than to reconstruct it, although the instruction to paraphrase the retelling was intended to avoid this.

In their study of the interactive effects on the importance of text-based and task-based information, Schraw et al. (1993 654) noted that paraphrase, as opposed to verbatim recall, constituted approximately 84 per cent of text segments in a read and recall task. This appears to support the view that learners interpret the recall task as reconstructive rather than reproductive. Further research using longer texts and evaluative comprehension questions where learners must elaborate on and integrate new information is needed to validate this point.

**Generation facilitates vocabulary learning**

Figure 1 summarizes the experimental and comparison group's gain scores attained between the pre-test and post-test interview as a function of vocabulary knowledge. Scores of six in both the pre- and post-test were omitted from the graph because a ceiling effect was attained and these subjects could not gain further credit for their knowledge of a word. The gain scores shown in Figure 1 are separated according to whether or not target items had been generated during the retelling of key ideas from the text. The graph shows that it was possible for many words that were not generated in the retelling task to increase by one between the pre-test and post-test. However, as greater gains were made (gains of 2, 3, 4, and 5) the number of non-generated items steadily decreased.

**Depth of generation affects learning**

Figures 2 to 4 show the number of target items in the comparison and experimental groups which increased in score between the pre- and post-test interview. Post-test scores are apportioned according to the demonstrated
level of generation (none, low, reasonable, or high) for target words used in the read and retell task.

Figure 2 presents words which were reported as unrecognizable (1) in the pre-test. It indicates that the likelihood of making large vocabulary gains, such as defining a word precisely (4), using a word with appropriate meaning in a sentence (5) and in a grammatically accurate way (6), is much greater when a reasonable or high level of generation is shown. In contrast, relatively less cognitive effort was required to acquire a vague sense of a previously unknown word (3) such as recognizing word form, having an awareness of a word, or being able to express the general meaning. This analysis is important because it indicates that for reportedly unknown words, generative processing
Figure 3: Comparison and experimental subjects' demonstrated level of generation and increases in word knowledge for words reported as seen or heard but unknown in the pre-test interview (VKS score of 2)

Figure 4: Comparison and experimental subjects' demonstrated level of generation and increases in word knowledge for words reported as familiar in the pre-test interview (VKS score of 3)

is associated with the ability to use words appropriately and accurately in context.

Figures 3 and 4 show the post-test scores for words identified in the pre-test as seen or heard (2), or familiar, i.e. a general meaning is given (3). In contrast to words reported as unknown, a clear patterning of greater vocabulary knowledge gains was not borne out for words that were somewhat familiar prior to the read and retell task or generated during recall.

Taken together, these graphs suggest that generative processing helps learners to move toward the ability to define words and use target words productively in context. This is particularly so for words that are unknown at the outset.
7 DISCUSSION

The result of the 3x2x2 ANOVA suggests that read and retell tasks stimulate learners to process written information in a way which facilitates vocabulary learning. The background knowledge effect indicates that learner differences independently moderate the extent of vocabulary learning. As one would expect, learners who already possess high levels of general vocabulary, who have acquired world and linguistic experiences and are reasonably proficient speakers in the L2 can expect to acquire a greater depth and breadth of word knowledge (see Anderson and Freebody 1981).

The 2x2x3 ANOVA showed that there was no significant difference between the experimental and the comparison groups across any of the dependent measures. This result can be accounted for in three ways:

(a) the experimental group had an increased learning burden,
(b) the design of the comparison task induced generative processing and
(c) training and the presence or absence of text were unimportant variables.

Increased learning burden

Increased task demands and a higher learning burden imposed on the experimental group may have contributed to their failure to outperform the comparison group. The time needed to decode the text, select and retain critical words and concepts, as well as to comprehend and generate new and relevant information may have been insufficient, particularly for learners with less general or domain-specific vocabulary knowledge (Shefelbine 1990, Spiro 1980).

Furthermore, recall with the aid of the text may have allowed subjects in the comparison group to allot more time to generate ways of using the target form meaningfully without the added burden of memorizing the target word forms.

Generative task

Many comparison group subjects employed generative strategies with or without recourse to the text even though they did not receive any explicit generative instruction. It appears that involvement in the read and retell task led the learners to generate unfamiliar words. Presumably, these learners adopted tactics that allowed them to focus selectively on information that related directly to the demands and goals of the task (Schraw et al. 1993).

Instructional effects

These findings emphasize the importance of student factors in determining the extent of vocabulary learning. More importantly, these results suggest that prior instruction and presence or absence of text during recall did not produce measurable differences in vocabulary learning. It appeared to make no
difference whether learners in the experimental groups received prior instruction on generative strategies or whether they recalled the passage with or without the aid of the text. The task itself seems to have stimulated the learners to engage in cognitive processing during recall and led the learners to perform similarly in the vocabulary tests. This would suggest that the demands of the task were a much more effective teaching tool than pre-instruction. Alternatively, the lack of difference in vocabulary performance between the two treatment groups may be attributed to insufficient training of learners in the experimental group. The one-off practice task may have been inadequate for such learners to develop confidence or improve their existing skills in using generative tactics. This would apply particularly to learners whose style it is to be less verbal. Moreover, benefits from the practice session may have been attenuated by the delay between the practice and the main task.

**Generation and retrieval in recall**

Target words that were retrieved and generated during recall were learned better than words whose general properties were discussed but not explicitly linked with the word. Moreover, reasonable and high generation subjects compared general and specific aspects of the item with other related words within or beyond the text. This supports Wittrock's (1981), Baddeley's (1990), and Craik and Tulving's (1975) claims that constructing semantic links of this kind lead to better memory and learning of encoded words. Tasks which encourage learners to retrieve the target form during recall and to provide an original sentence using the target word will give learners extra opportunities to practice using the word to-be-learned.

The finding that greater levels of generation led to greater gains for completely unknown words was not a pattern found for partially known words. An increase in word knowledge for unknown words may be attributed to greater cognitive effort due to item novelty (Loewenthal 1971, Nagy et al, 1985, Shore and Durso, 1990). Quite possibly, students may have underestimated or misreported their vocabulary knowledge in the pre-test because of lack of time in the interview or lack of contextual clues available.

**8 CONCLUSION**

In the present study, the comparison and experimental treatment required learners to focus primarily on reading and recalling the main points of the passage without the specific intention to learn vocabulary.

Taken together with previous studies of incidental vocabulary learning from reading, this study suggests that engagement in tasks requiring reading and recall without explicitly focusing on vocabulary can facilitate the acquisition of vocabulary. What is less clear is the contribution of the reading text (or task) without the retelling.

The significant background knowledge effect indicates that individual
student factors, such as prior vocabulary knowledge and generation, to a large extent determine the vocabulary gains which will be made.

The main findings regarding generation and vocabulary learning are that (a) generative processing appears to enhance vocabulary learning, with higher levels of generation producing greater gains for previously unknown words and (b) greater use and retrieval of the target form in recall is likely to strengthen the learning pathway and may create favourable conditions for generation.

(Revised version received October 1997)

APPENDIX A: THE TEXT USED IN THE TASK

Frontiers of pain control

One of the greatest human problems is chronic pain—continuing, often severe pain caused by such disorders as lower-back problems and cancer. Chronic pain is sometimes so severe that people are incapacitated by it, the pain stops people from doing things they usually do and doctors are at a loss in treating it effectively. In recent years, however, researchers have been developing some revolutionary treatments for chronic pain. Four of these treatments will be discussed.

Acupuncture

Acupuncture is now widely used for the treatment of chronic pain. In this method, needles are put into particular points on the body's surface to relieve pain in other parts of the body.

Electrical stimulation

Another approach to the control of chronic pain is the use of electrical nerve stimulation machines. A simple battery within the machine produces an electric current which passes through one or more pairs of electrodes. First, the electrodes are placed on the skin of the patient. They are put on the approximate area of the pain and the electricity is felt in this vicinity. Next, when the electricity flows through the electrodes, the patient feels as if sharp points are sticking into the vicinity of the pain. For the best response the sharp feeling mixes in with the pain and replaces it.

Placebos

Placebos aren't really drugs at all but rather sugar pills that contain no active chemical substances. Sometimes patients suffering from severe pain can be helped by such 'drugs.' The surprising effectiveness of placebos is proof of the psychological aspects of pain. If the person believes the treatment will work, it just might.

Narcotics

Narcotic drugs, such as Morphine, have long been used as painkillers. However, patients develop increasing tolerance for these drugs, which means that larger and
larger amounts are needed to be effective. In addition, a heavy dependence on large amounts of narcotics can make them highly addictive, and in rare cases they can cause death. Scientists have been continually trying to develop drugs that will relieve pain more safely.

Source: Adapted from Curne, P and E Cray (1987) *Strictly Academic* New York Newbury House pp 30–31

ACKNOWLEDGEMENTS

This research is based on a master's thesis completed at Victoria University of Wellington. I would like to thank my supervisor, Paul Nation, and Jim Dickie for their valuable suggestions and insightful comments.

NOTES

1. In this study, disposition to use generative tactics was included as part of the background knowledge factor because drawing links exemplifying, and paraphrasing reflect one's strategic ability to exploit available resources to construct meaning (see Baddeley (1990) and Wittrock (1974)).

2. As one reviewer suggested, we need to be cautious not to apply the 95 per cent figure too rigidly when estimating the extent of text coverage required for inferencing unknown words, as other factors such as background knowledge come into play.

REFERENCES

Anderson, R C and P Freebody 1981 *Vocabulary knowledge* in J T Guthrie (ed) *Comprehension and Teaching* Delaware IRA

Baddeley, A D 1990 *Human Memory Theory and Practice* Hove Lawrence Erlbaum Associates


Craik, F I M and R S Lockhart 1972 'Levels of processing A framework for memory research' *Journal of Verbal Learning and Verbal Behaviour* 11 671–84

Craik, F I M and E Tulving 1975 'Depth of processing and the retention of words in episodic memory' *Journal of Experimental Psychology General* 104/3 268–94

Curne, P and E Cray 1987 *Strictly Academic* New York Newbury House

Day, R R, C Omura, and M Hiramatsu 1991 'Incidental EFL vocabulary learning and reading' *Reading in a Foreign Language* 7/2 541–51


Fincher-Kiefer, R, T A Post, T R Greene, and J F Voss 1988 'On the role of prior knowledge and task demands in the processing of text' *Journal of Memory and Language* 27/4 416–29

Hall, S 1991 *The Effects of Split Information Tasks on the Acquisition of Mathematics Vocabulary* Unpublished MA thesis Victoria University of Wellington

Hirshman, E and R A Bjork 1988 'The generation effect Support for a two-factor theory' *Journal of Experimental Psychology: Learning, Memory and Cognition* 14/3 484–94


Laufer, B 1986 'Possible changes in attitude towards vocabulary acquisition research' *IRAL* 24/1 67–75

Laufer, B 1989 'What percentage of text-lexis
is essential for comprehension?' in C. Lauren and M. Nordman (eds) Special Language From Humans Thinking to Thinking Machines Clevedon Multilingual Matters

Loewenthal, K. 1971 ‘A study of imperfectly acquired vocabulary’ British Journal of Psychology 62(2) 225–33


Marks, C., M. C. Wittrock, and M. Doctorow. 1974 ‘Reading as a generative process’ Journal of Educational Psychology 67(4) 484–9


Nation, I. S. P. 1990 Teaching and Learning Vocabulary New York: Newbury House


Richards, J. 1976 ‘The role of vocabulary teaching’ TESOL Quarterly 10(1) 77–89


Saragi, T., I. S. P. Nation, and G. F. Melster. 1978 ‘Vocabulary learning and reading’ System 6(2) 72–8

Schraw, G., S. E. Wade, and C. A. Kardash. 1993 ‘Interactive effects of text-based and task-based importance on learning from text’ Journal of Educational Psychology 85 652–61


Shefelbine, J. L. 1990 ‘Student factors related to variability in learning word meanings from context’ Journal of Reading Behavior 22(1) 71–97

Shore, W. J. and F. T. Durso. 1990 ‘Partial knowledge in vocabulary acquisition: General constraints and specific detail’ Journal of Educational Psychology 82(2) 315–18


Wescz, M. and T. S. Paribakht. 1993 ‘Assessing vocabulary knowledge: Depth versus breadth’ (draft manuscript) Second Language Institute, University of Ottawa


Wittrock, M. C. 1974 ‘Learning as a generative process’ Educational Psychologist 9(1) 87–95


Wittrock, M. C. 1991 ‘Generative teaching of comprehension’ Elementary School Journal 92(2) 169–84