Learning and Retaining Specialized Vocabulary From Textbook Reading: Comparison of Learning Outcomes Through L1 and L2

DANA GABLASOVA
Lancaster University
Department of Linguistics and English Language
Lancaster LA1 4YW
United Kingdom
Email: d.gablasova@lancaster.ac.uk

This study investigated the acquisition of specialized vocabulary from L1 and L2 textbook reading by 64 Slovak high school students who were intermediate or advanced users of English. The students were divided into two groups: One group read the academic texts in their L1, the other group in their L2. In a posttest and a delayed posttest, they were asked to orally recall the meanings of 12 technical words that appeared in the texts. The word meanings recalled by the students immediately after reading and 1 week later were examined in terms of their breadth and depth. Results showed that although the L2-instructed students acquired the meanings of the specialized vocabulary items to a considerable degree, they still differed significantly from their L1-instructed counterparts in several respects: They could recall fewer word meanings after the reading; they acquired the words to a lesser depth; and after a week, their knowledge of the words faded more rapidly than that of the L1-instructed participants. The significance of the findings for L2 vocabulary acquisition and bilingual education is discussed.

Keywords: incidental vocabulary acquisition; vocabulary learning; specialized vocabulary; reading to learn; L1 and L2 users

KNOWLEDGE OF WORDS UNDOUBTEDLY plays an important role in the educational development of students. Research evidence shows that greater vocabulary size and depth of lexical knowledge are positively related to academic achievement (Laufer, 1997a; Saville–Troike, 1984; Verhallen & Schoonen, 1993). While most of this research has been concerned with general or academic vocabulary, knowledge of specialized technical words is likely to have an even stronger impact on the academic development of students.

Technical vocabulary and knowledge of the subject are closely linked in several ways. Studying an academic subject is a process in which the students' knowledge of a certain field is broadened or deepened while they simultaneously acquire words for the new concepts they learn (Mohan & van Naerssen, 1997). As a result, learning the meanings of new specialized words is an integral part of learning new subject knowledge (Bravo & Cervetti, 2009; Woodward–Kron, 2008). Moreover, the appropriate use of technical words can be also an indicator of the depth of

subject knowledge (Drum & Konopak, 1987) and can thus serve as a basis for the assessment of subject competence. Studying the acquisition of specialized vocabulary can therefore provide valuable insight into students' lexical and more general language learning as well as their subject matter development.

However, despite its prominent role in the development of subject knowledge, the acquisition of technical vocabulary from academic learning has not yet received much attention in the literature. This study therefore sets out to contribute to filling this gap in our understanding of more advanced vocabulary learning in two areas: (a) the learning of L2 specialized words in the course of a meaning-oriented activity, that is, *reading to learn*, and (b) subject learning through the medium of an additional language, common to students in bilingual programmes such as Content and Language Integrated Learning (CLIL). Specifically, the study compares the outcomes of lexical learning by L2 users and L1 users, thereby highlighting the similarities and differences in the content and lexical knowledge acquired through the first or an additional language.

The study focuses on high school students who study content through the medium of an additional language. This group is of interest both cognitively and linguistically, as students work with specialized topics and language in the course of their schooling, and are also expected, more substantially than students in lower grades, to develop their use of appropriate technical vocabulary. Nevertheless, their subject-specific language development has not received adequate attention (cf. Cumming, 2008). Research is especially scarce with respect to students' ability to communicate their subject knowledge orally (Menyuk & Brisk, 2005), although this is a common medium for sharing and assessing content knowledge in the classroom at the upper-secondary level (cf. Dalton–Puffer, 2007).

VOCABULARY ACQUISITION THROUGH L1 AND L2

It is often assumed that the acquisition of L2 vocabulary in an EFL setting is a process qualitatively different from L1 acquisition for two reasons: (a) a lack of sufficient, contextualized input in L2 and (b) the presence of an established semantic/conceptual system in students' L1 through which all new lexical learning is mediated (Jiang, 2000). In the process of new semantic and lexical acquisition, this usually leads to the mapping of a new L2 form onto an already existing L1 concept and the word meaning associated with it.

While this may be true of much of L2 lexical learning, it is not the case with more advanced learners who often learn new information about the world along with learning an additional language (Hague, 1987). This applies especially to students in L2-medium education who simultaneously acquire new subject knowledge (e.g., new concepts) and the language to communicate this knowledge. Academic learning is a natural source of new vocabulary gains (Mohan & van Naerssen, 1997) for students who are studying through their L1 as well as for those studying through their L2.

Despite the importance of an L1 baseline for a better understanding of L2 language development, comparative studies of L1 and L2 vocabulary acquisition (e.g., Hague, 1987; Meara, 1988; Stoller & Grabe, 1995) are still scarce. There are even fewer studies that compared learning through the two languages directly. These studies focused on differences in lexical inferencing (Wesche & Paribakht, 2010), input through glosses in L1 and L2 (Yoshii, 2006), and technical vocabulary acquisition (Haynes & Baker, 1993; Lessard–Clouston, 2006).

In SLA research, L1 data can provide an important baseline for interpreting the outcomes of L2 learning (Foster & Tavakoli, 2009). In a study with implications for pedagogical practice, it can also help to focus teachers' attention on the gaps between the performance of L1 and L2 users and ensure that any difficulties experienced by bilingually educated students are addressed.

Academic Reading as a Source of Vocabulary Learning

There is an extensive literature on various aspects of incidental lexical learning from reading (see for example de Bot, Paribakht, & Wesche, 1997; Huckin & Coady, 1999; Huckin, Haynes, & Coady, 1993; Paribakht & Wesche, 1999). However, little work has addressed learning of specialized words in the course of learning an academic subject, despite the fact that reading often plays a central role in acquiring both subject knowledge and language. As Huckin and Bloch (1993) pointed out, "reading is the primary means by which academic knowledge is transmitted, and it is also a useful secondary source for information that might be missed in a class discussion or a lecture" (p. 154). Written sources have arguably an even more important role in learning through an additional, not yet

fully mastered language. They provide support for knowledge (and language) development by offering students further input if they do not understand the lecture or class explanation (Burger, Wesche, & Migneron, 1997; Johnson, 1997).

Among studies on learning of technical words that involved reading, Parry (1991, 1993) provided a detailed description of the L2 learning process and outcomes, and Lessard–Clouston (2006) and Haynes and Baker (1993) employed a comparative approach with the language of reading/learning (L1 or L2) as the independent variable. This design allowed them to examine to what extent the observed development in L2 learners is typical of learning vocabulary in general, or is unique to learning vocabulary through L2. Not surprisingly, a difference between students who were acquiring subject-specific vocabulary through their L1 and through their L2 was found with respect to (a) the breadth of the vocabulary acquired (Haynes & Baker, 1993), with the L2-instructed students acquiring fewer words and (b) the depth of lexical knowledge, with the L2-instructed students developing a less complete understanding of the technical words (Haynes & Baker, 1993; Lessard–Clouston, 2006).

Learning Technical Words from Lexical Familiarization in Textbooks

Learning technical words in the course of academic study is to some extent a specific form of incidental vocabulary learning. First, it often involves a specific type of reading, which has been referred to as reading to learn (Grabe, 2009) and which, as a form of reading for meaning, has been found to result in greater retention of the information from the text than reading for fun or general comprehension (Swanborn & de Glopper, 2002). Second, a crucial factor in readers' ability to learn the meanings of new words is the presence of information in the text from which their meanings can be inferred (Beck, McKeown, & McCaslin, 1983). Writers of textbooks recognize the importance of specialized vocabulary acquisition for subject learning. As technical words are usually low frequency vocabulary, textbook writers use different *lexical familiarization* techniques in the texts to maximize the chances that a technical term, often unfamiliar to students beforehand, will be understood and learned. Lexical familiarization has been defined "as a contextual aid, intentionally and explicitly provided by the author when writing for a specific readership. The writer's intention is to help

[readers], by providing [them] with sufficient familiarity with the new word, as employed in its context, so that the reader can continue reading with understanding" (Bramki & Williams, 1984, p. 170).

Definitions are a common means of lexical familiarization in educational contexts, particularly in textbooks and lectures (Bramki & Williams, 1984; Flowerdew, 1992; Haynes & Baker, 1993; Lessard–Clouston, 2009; Nation, 2001). They form an integral part of a text and its meaning and provide an appropriate amount of semantic and syntactic information about a specific technical word. This enables the reader to construct the meaning of the word at a depth appropriate for the particular text. In this way, the reader receives sufficient information about the word and can avoid a situation in which a lack of contextual support results in incorrect guesses and a wrong meaning being inferred (Huckin & Bloch, 1993; Stein, 1993).

Learning Specialized Vocabulary in a CLIL Context

Gaining a good command of subject-related vocabulary in L2 is an important issue in educational contexts that combine subject and language learning, such as CLIL (Content and Language Integrated Learning) programmes. CLIL is an approach to language learning that stresses naturalness in the language acquisition process such that language is both socially and contextually embedded. It involves creating a context in which language is both the medium of communication as well as the object of learning. As a result, CLIL is a dual-focused approach in which language learning and content learning are of equal importance (Baetens Beardsmore, 2009; Coyle, Hood, & Marsh, 2010) and the curriculum taught through the L2 is parallel or relatively equal to that taught through the medium of the L1 in mainstream classes. This is in contrast to some bilingual programmes whose primary aim is to develop general L2 proficiency of students (Lasagabaster & Serra, 2010). To assist the students in learning academic subjects through an additional language, it is necessary to identify areas that may be difficult for them and to offer well-designed pedagogical support to overcome the increased learning and cognitive load. This study seeks to identify these potential problem areas with respect to the acquisition of specialized lexis.

RESEARCH QUESTIONS

Learning the meanings of new specialized words is an integral part of learning new subject

knowledge and is acknowledged as such by textbook writers. Lexical familiarization is a common way of communicating the meaning of new words to students. On the one hand, the learning of new technical terms by L1- and L2-speaking students is similar in that both learn new concepts and new words from a rich context. On the other hand, learning in the L1 or the L2 are different experiences, and it is necessary to understand the differences in order to address the possible difficulties faced by bilingually educated students. The present study did so by comparing the learning outcomes of L1- and L2-instructed students in two respects: *breadth* (the number of new entries in students' lexicon) and *depth* (the quality of these new entries). The study also investigated these two dimensions with respect to the retention of the new knowledge, with the aim of providing a holistic picture of the learning of new words through L1 and L2. The following research questions were addressed:

RQ1. Breadth of lexical knowledge: How many word meanings can L1- and L2-instructed students recall (a) in a test immediately following the instruction and (b) after a week?

RQ2. Depth of lexical knowledge: If a form–meaning relationship was established, how many meaning components can be recalled by students in each group (a) in a test immediately following the instruction and (b) after a week?

METHOD

Participants

Participants in this study were 64 Slovak high school students aged 17–20 years. They were drawn from two schools that have implemented a CLIL bilingual programme aimed at additive bilingualism. These schools teach part of their curriculum via the medium of the students' second language (English). Participants were recruited from among students in the last two years of high school (year 4 and 5) and thus had spent at least 3.5 years at the school. As a result, they had extensive experience with several aspects of studying through an additional language, such as reading academic texts, speaking on academic topics, and learning content knowledge in English.

Participants were divided into two groups (both $N = 32$), one of which received input in their L1 (Slovak) and the other in their L2

(English). The groups were balanced for gender and proficiency. All participants were at least intermediate-level users of English as established by three tests: two measures of vocabulary size at 5k and 5–10k level (X Lex and Y Lex) (Meara, 2005) and one measure of general productive proficiency (C-test). An independent samples *t*-test did not reveal any statistical differences between the proficiency of the two groups of participants.

Materials

Two academic-type texts were used in the study. The first was a history text describing the changes in the Maori lifestyle from the arrival of Polynesians in New Zealand until today. The second was a geography text describing the living conditions and farming in the South Island High Country in New Zealand. The two texts were developed for this study using two authentic sources[1] that were modified and to which target words (TWs) and/or their lexical familiarization were added if they were not already included. Two language versions of the texts were prepared: Slovak and English. The texts were developed in English and then translated into Slovak by the researcher and checked by two native speakers of Slovak; both versions were approximately 800 words in length. Every effort was made to make the texts as fully comparable as possible, but due to typological difference between Slovak and English some translation effects could not be ruled out (this material is provided as Supporting Information Appendix S1 in the online version of the article).

Lexical analysis[2] showed that almost 90 % of the words in the texts were from the 5,000 most frequent words in English. Based on the estimated vocabulary size of the participants (X Lex: $M = 4048.9$, $SD = 564.6$, Y Lex: $M = 2132.6$, $SD = 648.7$), it can be assumed that the texts were moderately difficult to read for the students. Four high school students who did not take part in the study but were asked to read the texts reported them as being easy to moderately difficult to read.

In the study, participants were asked to read the two texts and listen to a recorded reading of these texts. While reading was expected to be the main source of learning the TWs' meaning, listening to the texts (while being able to follow the text) was introduced in order to familiarize the students with TW pronunciation. This was deemed important as participants' knowledge of TWs was elicited orally in a test. Listening to the texts was also introduced for methodological reasons

in order to ensure that participants attended to the whole of the text.

Recordings of the English texts were made by a native English speaker (female), a teacher at one of the schools in which the study was carried out. The Slovak versions were recorded by a female native speaker of Slovak.

Target Words

The TWs were selected according to the following criteria: (a) All were technical terms related to the topics (and main ideas) of the two texts; (b) the words (and their meanings) were new to participants, both in their L1 and L2; (c) all words were nouns, a common characteristic of technical words; (d) half of the words were concrete, the other half abstract nouns; and (e) the words were similar in Slovak and English form and pronunciation, thus allowing for direct comparison of the acquisition of the TWs through the two languages.

The two groups of students were tested on the same set of words, so that differences in vocabulary learning could, with some confidence, be attributed to the language of learning rather than to individual properties of the TWs. Also, as the languages of testing were both English and Slovak, it was necessary to ensure that students would be able to recognize the TW in both languages. Table 1 shows the orthographic form and pronunciation of the 12 TWs in English and Slovak, using IPA conventions. Given the previously stated criteria, it was not possible to control

other aspects of learnability of the words, such as their length and other properties that may affect retention (cf. de Groot & Keijzer, 2000). To preserve the naturalness of the text, the number of occurrences of the TWs in the text also was not controlled. Individual TWs appeared in the text one to three times but each appearance was restricted to a specific passage and every effort was made to tie each TW to a main idea in the text. Table 2 shows an example of lexical familiarization (definitions or explanations) with which the target words appeared.

Pretest

Prior knowledge of the TWs and the texts' content was pretested using (a) a set of open questions related to the main ideas in the texts (cf. Zaki & Ellis, 1999) and (b) a vocabulary list. Apart from the 12 TWs, the vocabulary list contained another 56 lower frequency words that appeared in the texts, placed in alphabetical order, so as not to guide attention to the TWs. In the list, participants were asked to indicate their familiarity with the words by ticking one of the following boxes (adapted from Paribakht & Wesche, 1997): (1) I have never seen/heard the word; (2) I have seen/heard the word before but don't know what it means; (3) I know what the word means. For practical reasons, the pretest was completed as a pen and paper test; it established that the participants did not have any knowledge of the TWs or the topics that the texts described.

TABLE 1
List of Twelve Target Words in English and in Slovak (Orthographic Shape and IPA Transcription)

Word in English	Pronunciation (IPA)	Word in Slovak	Pronunciation (IPA)
ampelography	æmpɪˈlɒgrəfi	ampelografia	ˈampɛlografia
diastrophism	dʌɪˈæstrəfɪz(ə)m	diastrofizmus	ˈdɪastrofɪzmus
ecocentrism	iːkəʊˈsɛntrɪz(ə)m	ekocentrizmus	ˈɛkotsɛntrɪzmus
	ɛkəʊˈsɛntrɪz(ə)m		
kumara	ˈkuːmərə	kumara	ˈkumara
	ˈkuːmʌrʌ		
moa	ˈməʊə	moa	ˈmoa
moko	ˈməʊkəʊ	moko	ˈmoko
pa	ˈpɑː	pa	ˈpa
perendale	ˈpɛrəndɛɪl	perendale	ˈpɛrɛndɛɪl
RCD	ɑːsiːˈdiː	RCD	ɛrˈtsɛːˈdɛː
terroir	tɛˈwɑː	terroir	ˈtɛrua:
			ˈtɛrua:r
transhumance	trænzˈhjuːməns	transhumancia	ˈtranshumantsia
	trɑːnzˈhjuːməns		
whanau	ˈfɑːnaʊ	whanau	ˈfɑːnaʊ

TABLE 2
Examples of Lexical Familiarization Used in the Texts

Target Word	Lexical Familiarization
Transhumance	Transhumance—the seasonal movement (before winter) of stock from exposed, high mountain slopes to the more sheltered foothills and river flats. This avoids large stock losses due to the bitter cold of winter.
Moko	Moko—this is body and face marking signifying the person's rank in society. Moko is a type of tattoo. Men generally received moko on their faces, buttocks, and thighs. Women usually wore moko on their lips and chins.

Posttest

The computer-administered posttest consisted of 36 questions about the two texts: 12 questions asking for the meanings of the TWs and 24 distractor items asking about other information contained in the texts. Each question about the TW had the same format: "What is X?" The definition task was selected to measure the knowledge of the TWs' meanings as it can provide rich insight into participants' understanding of the meaning of TWs (Parry, 1991; Verhallen & Schoonen, 1993).

Participants received the questions in two blocks: One consisted of 18 questions in Slovak, the other comprised 18 questions in English. Each block contained six target words. The blocks of questions in a particular language were counterbalanced across participants in each group so that half of the participants in each group received the English part of the test first, whereas the other half started with the Slovak part of the test. The questions and the language in which they were asked were kept constant, that is, the question about "ampelography" was always asked in Slovak. Participants were instructed to answer in the language in which the question was presented. Half of the questions were thus answered in the L1 and the other half in the L2.

Delayed Posttest

A delayed posttest was used to test participants' knowledge of the target words one week after reading and listening to the texts. This enabled measurement of longer-term retention of the word meanings. The delayed posttest consisted of 36 questions that had already appeared in the posttest. The questions were presented in the same manner as in the posttest. Although students were not presented with any new information or feedback on their definitions, merely recalling information constitutes an instance of additional learning. As a result of this procedure, some retesting effects could not be avoided.

The Language of Data Elicitation in the Posttest and Delayed Posttest

Studying L2-mediated learning involves choosing the language through which to assess participants' knowledge. The problems associated with choosing either L1 or L2 to assess L2-instructed participants were previously pointed out in research that compared outcomes of reading in L1 and L2 (cf. Chen & Donin, 1997; Lee, 1986) as well as in the studies on assessment of L2-educated students (e.g., Short, 1993). As this research evidence suggested that each language might elicit a different response (in terms of the amount and type of information), this study adopted an approach in which both languages were used to obtain responses from the participants (i.e., participants answered half of the questions in their L1 and the other half in their L2). However, as the effect of the language of response on students' performance is not among the questions addressed in this article, the responses in L1 and L2 are always presented together.

Procedure

This article presents data that were collected as part of a larger study on the quality and quantity of lexical learning from textbook reading. The whole study consisted of a pretest, a reading and listening session, a posttest, an extended definition task, a word association task, and a delayed posttest. In this article, only the data from the posttest and the delayed posttest are discussed.

In all sessions, participants were tested individually. At the beginning of the reading session, participants were told to pay attention to the meaning of the text. After the reading, they would be asked questions about information contained in the text. They began by completing the pretest, followed by the learning session, which took approximately 30 minutes and consisted of reading and listening to the two texts. In this session, participants were first given 10 minutes to

read the first text and then listened to it (while being able to follow the text). The same process was repeated with the second text.

This was followed by a 2-minute nonverbal distractor task, after which the participants were given a posttest. The questions in the posttest were computer-administered. Next, participants completed another computer-administered session, in which they provided word associations for the target words. This was followed by an extended definition task, where more information about the meaning of the target words was elicited from participants in a face-to-face interview with the researcher. One week later, participants were given the delayed posttest.

In the computer-administered tests, the questions appeared on the screen and participants responded orally. On a separate occasion, usually preceding the learning session, participants completed the proficiency tests used in the study.

Data Coding and Analysis

Breadth of Lexical Knowledge.   In order to answer the first research question about word meaning recall immediately after the reading and with one week delay, participants' responses were divided dichotomously into (a) those responses that showed at least minimal evidence of learning of the TW meaning (i.e., they contained at least one correct meaning component) and (b) those which did not show any lexical learning of this kind (e.g., participants answered "I don't know" or gave an incorrect meaning of the TW). Participants' answers were scored as 1 (evidence of learning) or 0 (no evidence of learning). The scores for the 12 TWs were added up and compared using an independent samples *t*-test.

Depth of Lexical Knowledge.   The second research question examined the depth of word meanings acquired by participants. As indicated, participants' responses had already been classified into those that showed evidence of learning of new words and those that did not. The first group of responses was analysed further to determine depth of word knowledge, that is, the amount of semantic information participants could recall about each word. Following the tradition of research using definitions (cf. Davidson, Kline, & Snow, 1986; Verhallen & Schoonen, 1993), depth of knowledge was operationalised in terms of the number of meaning components.

First, all possible meaning components in the elicited word meanings were identified based on participants' responses and the information in the texts. Only the components considered obligatory for defining the word and distinguishing it from different words in the same semantic field were taken into consideration. These meaning components were selected according to the most frequent components provided by the participants and by consulting the *Oxford English Dictionary* (Simpson, 1989) and *Te Ara: The Encyclopedia of New Zealand* (Phillips, 2005). In this way, two to five core components were identified for each target word (Nippold et al., 1999).

Second, in order to compare the depth of the word meanings retained by the two groups, a mean standardised score of core meaning components was calculated for each participant. To obtain this score, a percentage of the core components contained in participants' definitions was calculated (e.g., if a participant's definition of a TW contained three out of four possible core meaning components, the score would be 75%). The scores of participants in the two groups were compared using a *t*-test.

Statistical Analyses.   As individual variation is an important factor in vocabulary learning, the independent samples *t*-test was used to compare the learning outcomes of the two groups of participants. The *t*-test establishes whether the difference between the compared groups is larger than the differences between individual participants. Cohen's *d* was calculated for statistically significant results to establish the effect size.

RESULTS

Breadth of Lexical Knowledge

Table 3 shows the results for both the posttest and the delayed posttest (both groups $N = 32$). As can be seen from the table, L1-instructed students were able to recall more word meanings after reading the two texts. This finding was of a medium effect size ($d = .60$). The delayed posttest showed that after a week, the L1-instructed participants outperformed the L2-instructed participants on form–meaning connections. Moreover, while the performance of the L1-instructed group as observed in the posttest decreased only a little, the L2-instructed participants recalled considerably fewer word meanings than in the posttest. This difference had a large effect size ($d = .98$). This finding indicates a different rate of retention by the L2-instructed participants from that of the L1-instructed participants. It should also be noted that the performance of the L1-instructed group was more homogenous than

TABLE 3
Between-Group Comparison of the Number of Form–Meaning Connections

| | L1-instructed | | L2-instructed | | | | | |
	Mean	SD	Mean	SD	t	df	Sig.	d
Posttest	9.56	2.03	8.00	3.16	2.353	62	.022*	.60
Delayed posttest	9.25	2.05	6.81	3.23	3.608	54.481	.001**	.98

$^*p < .05$, $^{**}p < .001$

that of the L2-instructed participants as indicated by the smaller standard deviations.

Figure 1 further explored the differences between the two instructional groups according to the individual words. As can be seen from the figure, in the L1-instructed group more than 80% of the participants were able to recall 7 of the 12 target words, and another 4 words were recalled by more than 65% of participants in this group. Only one word, *diastrophism*, was recalled by fewer than half of the participants. By contrast, in the L2-instructed group, only 3 of the words were recalled by more than 80% of students, and 5 words fell in the range of 60 to 80%. Four words were recalled by fewer than 60% of participants in this group, with *perendale* being recalled by only one quarter of the L2-instructed participants. A similar pattern was observed in the delayed posttest (Figure 2). The results in the delayed posttest reflected those from the posttest in two respects: First, the L2-instructed participants recalled fewer TWs than their L1-instructed counterparts and, second, the recall rate of individual words was similar in the posttest and the delayed posttest; that is, the words that had been easier to recall in the posttest were also

those more likely to be recalled in the delayed posttest.

Analysis of the incorrect answers also revealed that on several occasions participants formed an incorrect connection between a word's meaning and form; in other words, a definition of a particular TW was given when a form of another TW was presented. For example, this was the case when a student answered the question "what is a *kumara*?" with "it is a tattoo," which is in fact a definition of *moko*. L1-instructed participants demonstrated incorrect form–meaning connections in the posttest 11 times (2.6% of all answers), while the L2-instructed participants did so 18 times (4.1%). In the delayed posttest, the L1-instructed participants made 12 errors (3.1%) while the L2-instructed participants made 27 errors (7%) of this kind. Thus, on these occasions, the L2-instructed participants showed that the relevant information from the text was comprehended and available to them after the reading, but there was no link between the target word form and its meaning.

Visual inspection of the data suggested that longer words tended to be more difficult for the students to learn. To test whether this was indeed

FIGURE 1
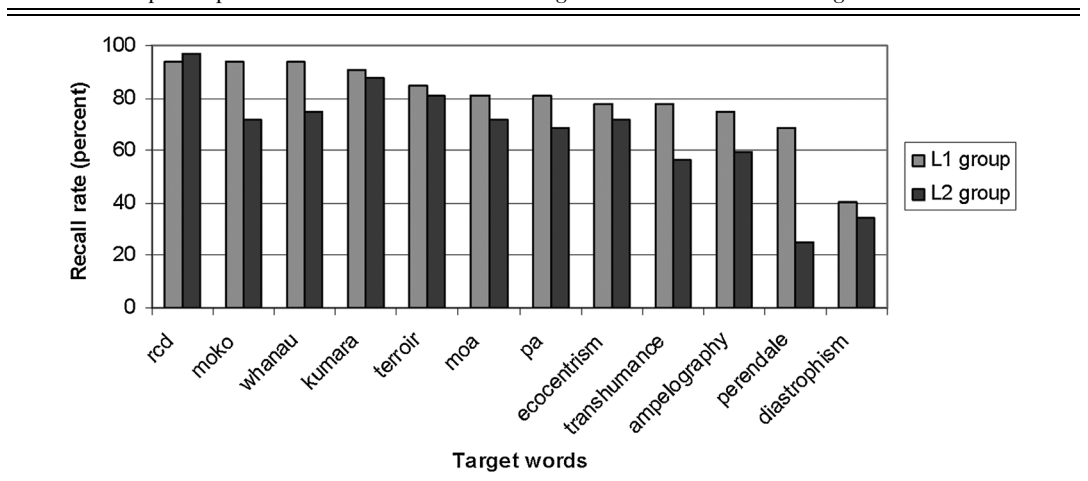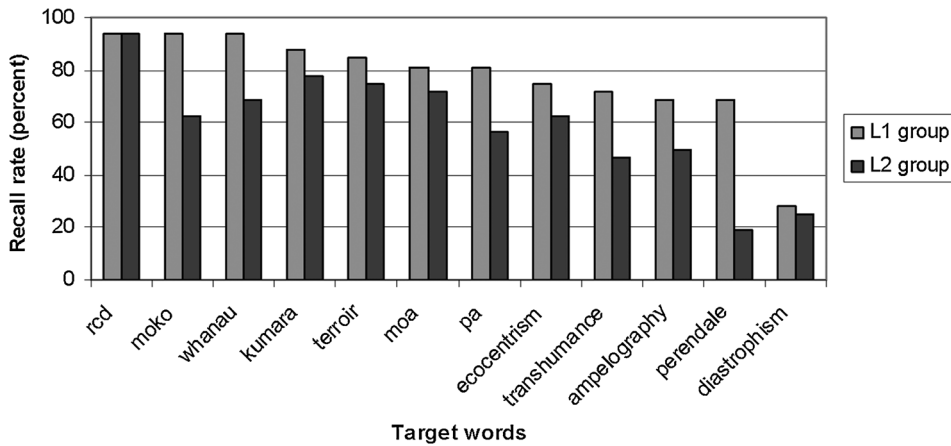Between-Group Comparison of Recalled Form–Meaning Connections for Each Target Word in the Posttest

FIGURE 2

Between-Group Comparison of Recalled Form–Meaning Connections for Each TW in the Delayed Posttest

the case, the length of TWs (calculated in letters as the number of syllables of some TWs differed in Slovak and in English) was correlated with the scores. The results showed that there was a moderately strong negative relationship between word length and learnability for both L1-instructed ($r = -.628$, $p < .05$) and L2-instructed participants ($r = -.586$, $p < .05$).

Depth of Lexical Knowledge

Table 4 displays the results of the *t*-test comparing the number of core meaning components provided by students in the two groups. As shown in the table, the number of core meaning components produced by the L1-instructed group in the posttest was higher than that produced by the L2-instructed group by approximately 5%. The finding was of a medium effect size ($d = .64$). No such difference was found in the delayed posttest. Here, if the L2-instructed participants were able to recall a particular TW, they were also able to provide the same number of meaning components as the L1-instructed participants. Similar to the results reported in Table 3, the performance of L2-instructed partic-

ipants was more heterogeneous than that of the L1-instructed students as documented by the higher standard deviations.

In order to establish whether there was a relationship between the initial degree of completeness and elaboration of the word meanings and participants' ability to retain and recall the word meanings after a week, a correlation between the total number of recalled word meanings in the delayed posttest and the mean number of core meaning components in the posttest was calculated. The results showed a moderate positive relationship between the target variables for the L1-instructed group ($r = .394$, $p < .05$) and somewhat stronger relationship for the L2-instructed group ($r = .545$, $p < .01$).

DISCUSSION

This study sought to describe the difference between knowledge of new specialized words gained through L1 and L2. It focused on three dimensions of word knowledge: (a) the ability of students to broaden their lexicon with new entries, (b) the quality (depth) of the new entries, and (c) the ability of students to remember new

TABLE 4

Between-Group Comparison of the Number of Core Meaning Components

	L1-instructed		L2-instructed					
	Mean (N)	*SD*	*Mean (N)*	*SD*	*t*	*df*	*Sig.*	*D*
Posttest	72.83 (32)	7.90	67.06 (32)	11.98	2.198	47.727	.033*	0.64
Delayed posttest	71.22 (32)	7.54	68.89 (29)	11.01	.969	59	.336	–

*$p < .05$

words after a week. The findings from this study indicate that the ability of proficient L2 users to learn new content and lexical knowledge through L2 is to some extent similar to that of L1-instructed students, yet it differs in some important respects.

Breadth of Acquired Lexical Knowledge

With respect to the number of words learned, there seem to be two possible reasons for the fact that L2-instructed participants established fewer initial connections between words and their meanings than their L1-instructed counterparts. It is possible that they either failed altogether to establish the link between the TW form and its meaning or that this connection was formed but was not strong enough and thus faded from memory. Both noticing a link between a word form and its meaning and retaining it are prerequisites of lexical learning (Gass, 1988; VanPatten et al., 2004).

With respect to noticing, according to Van-Patten et al. (2004), an initial connection is established "when a learner somehow cognitively registers a form, a meaning, and the fact that the form encodes the meaning in some way" (p. 5). As the purpose of placing lexical familiarization in a text is to make the meaning of the word form explicit to the reader, noticing should be a straightforward process. However, as observed by Flick and Anderson (1980), when looking at an implicitly expressed relationship between a word and its definition, readers sometimes remain unaware of this connection and are likely to treat the definition as a new and unrelated piece of information (cf. Stein, 1993). The failure to form the connection may have been directly language-related in that L2 reading difficulties—such as insufficient vocabulary size or insufficiently automatic word recognition (Grabe, 2009; Haynes & Baker, 1993; Huckin & Coady, 1999)—could result in some L2-instructed participants failing to understand the link between a word form and its definition.

Alternatively, the additional effort associated with reading in a nonnative language, even one mastered to a high degree, could result in increased processing demands on working memory, which might have interfered with the noticing, formation, and retention of the association between the TW and its definition. As Kellerman and Bialystok (1997) state, "in any cognitive activity, we are able to attend to only some selected portion of available information at any given time. Situations invariably present more information than it is possible to process, and cognition involves continual selection from that pool of information" (p. 33). This appears as a likely explanation, as the lower ability of the L2-instructed participants to recall the word meanings was at least in some cases not related to problems with noticing or comprehension of the definitions in the text. This could be seen in the responses where evidently some information about the target word meaning was picked up and recalled by participants, but was not linked with the correct word form (i.e., a form was associated with an incorrect meaning).

However, despite the lower number of new word–meaning connections established by the L2-instructed group, it appears that similar strategies of reading and learning of the word meanings were at play in both groups of students. This is indicated by the fact that there was considerable overlap between the two groups in terms of which words were easy or difficult to learn. Table 5 shows the top 4 and bottom 4 words from the two groups.

The learning pattern of the L2-instructed group followed the pattern of the L1-instructed group, except that the L2-instructed students acquired knowledge to a lesser degree. It seems that for both groups, the most difficult words to acquire were mostly longer and/or abstract words while the words easiest to acquire were the shorter and/or concrete words. This is in line with the findings that length and abstractness of a word affect its retention (Laufer, 1997b; Nagy, Anderson, & Herman, 1987). However, it should also be

TABLE 5
The Least and Most Difficult Words to Acquire by Instructional Group

Posttest & delayed posttest	L1-instructed group	L2-instructed group
Least acquired	diastrophism, perendale, ampelography, transhumance, ecocentrism*	perendale, diastrophism, transhumance, ampelography
Most acquired	rcd, moko, whanau, kumara	rcd, kumara, terroir, whanau

transhumance and *ecocentrism* had the same recall rate

noted that, whereas learning of most of the TWs showed a similar pattern in both groups, some words proved significantly more difficult for the students working with the texts in their L2 than for those reading them in their L1. Such words were, for example, *moko,* and, in particular, *perendale* (rate of recall differed by more than 40%). We may speculate that specific properties of these words or the context in which they were embedded made them less accessible to L2-reading students.

Overall, findings from this study show that some aspects of lexical learning through L2 are quite predictable on the basis of L1 performance (e.g., words that were difficult to learn for L1 users were also more difficult for L2 users). However, for some target words, the learning outcomes of L2 users differed significantly from those of L1 users (see Figure 1). These findings stress the complexity of learning by advanced users of L2. This complexity is also reflected in the diversity of findings from previous research: While Haynes and Baker (1993) found a relatively large difference between their native and nonnative speakers, this was not confirmed by Lessard–Clouston (2006), who, by contrast, reported no difference in the number of technical words learned by L1 and L2 users. These results imply that contextual factors (e.g., a specific L1–L2 pairing or the number of encounters with the target vocabulary) as well as learner characteristics (e.g., L2 proficiency) may play a considerable role in determining the learning rate of a particular group of L2 users.

Depth of Acquired Lexical Knowledge

The results showed that the participants in the two groups differed in the depth of their understanding of TW meanings in the posttest, with the answers of the L2-instructed participants showing a more limited conceptual knowledge of the technical terms. The different range of word knowledge recalled by participants can be seen in the following definitions of *moko* from student test responses (for the original definition of *moko* see Table 2):

(1) a moko is type of tattoo
(2) moko is type of tattoo typical for Maori
(3) moko is a special type of tattoo that Maoris wear, men wear it on their thighs or on their faces and buttocks and women wear it on their lips and chins
(4) a moko is a type of tattoo that that signifies the position in the society of Maoris

The study found that the L2-instructed students systematically recalled fewer meaning components than the L1-instructed participants. Thus, for example, these students provided more general answers similar to Examples (1) and (2) while L1-instructed participants provided more answers that showed more precise definitions of the terms such as examples (3) and (4). Nevertheless, as this study looked only at expressible word knowledge (i.e., the knowledge of the word that students are able to verbalise), it is possible that L2-instructed students also gained some receptive knowledge of the words but were not able to express it fully. However, in the learning process, expressible word knowledge is often crucial as it is used for assessing students' knowledge and might have a direct influence on what the teacher decides to do next in the class.

The fact that the L2-instructed participants recalled fewer meaning components of the target words in the posttest can be explained in part by increased cognitive demands of reading through a nonnative language proposed previously. In addition, knowledge of particular lexis might also play a role in the lower number of meaning components recalled by the L2-instructed group. This is suggested by the fact that, upon further examination of the meaning components provided by the L2-instructed group, some of the components were found to be systematically omitted. These omissions appear related to the low frequency of words used in communicating these meaning components.

One example of such a systematic omission was found in the students' definitions of *transhumance.* Its definition as it appeared in the English version of the geography text can be seen in Table 2. Analysis of participants' responses revealed that *stock* appeared to have been the problematic component for the L2-instructed group in the definition of this word. In the posttest, out of 20 students who were able to give a definition only 14 referred to "farm animals" in some form (some narrowed the concept to "sheep" only; some extended it to "animals" in general). Two other students omitted the component altogether, as illustrated in example (5), and four others replaced it with an incorrect one, usually "people", as in example (6).

(5) It's moving from mountainous areas to more to the places where the rivers are especially before winter.
(6) It is a seasonal movement of the people from High Country.

This was not the case with the L1-instructed group where, out of 29 definitional attempts, 28 contained the component referring to "animals" in some way and only one response contained an incorrect component. No participant in this group who provided a definition omitted the component. The difference between the number of participants in the two groups who were able to provide the component was statistically significant (Log Likelihood = 4.76, $p < .05$).[3]

It seems that familiarity with a particular vocabulary item (or lack thereof) contributed to the lower number of meaning components and thus to a lower degree of elaboration of the word meanings among the L2-instructed participants. As a consequence of the omitted components, the representation of the words' meanings by L2-instructed learners was not as precise as that of their L1-instructed counterparts. This is an issue that has been raised in research on differences between the information gained from L1 and L2 reading (Bernhardt, 2010; Grabe, 2009; Haynes & Baker, 1993). That the word meaning acquisition process is a gradual one regardless of the language in which it takes place is an accepted fact among vocabulary researchers (Nation, 2001). However, the findings from this study suggest that the knowledge gained through L2 is likely to develop more incrementally than that acquired through L1.

Retention of Word Meanings

In addition to previous studies on the learning of technical vocabulary, which focused on the breadth and depth of lexical knowledge, this article addressed the important dimension of retention. After all, committing information to memory is a very important aspect of subject learning. Findings showed that the knowledge of the L2-instructed participants was less permanent. If the L1-instructed participants learned the word meanings, their knowledge suffered nearly no attrition in the span of a week; this was not the case with the L2-instructed students. Further analysis revealed that this could be partly attributed to a difference in the initial quality of acquired knowledge. This finding is in line with studies that provided evidence that a certain degree of robustness or strength of the initial association between a word form and its meaning is necessary (if no further input is available) for committing information to long-term memory (Baddeley, 1997; VanPatten et al., 2004). Also, the more elaborate the initial input is, the more links may be expected to be formed with knowledge already stored in the long-term memory or the mental lexicon, which can have a positive impact on retention of the new knowledge. The link between the initial depth of the word meanings and their retention is further supported by the fact that in the delayed posttest no difference was found between the number of meaning components recalled by the two groups. It appears likely that the target words remembered with a lower number of meaning components faded from memory after a week, leaving only those TWs that had a certain level of initial elaboration.

Durability, while perhaps a less tangible and visible facet of lexical knowledge than size and depth, is nevertheless important. The difference in the durability of the knowledge of the two groups of students suggests that more targeted pedagogical approaches might be required in L2-medium content classes to enhance the retention of learned information. As the retention of new TWs seems partly related to the initial quality of the word knowledge, pedagogical interventions should focus on the period when the form–meaning connections are formed or shortly after. However, as this study measured retention only after a week, further research is necessary to more fully understand the impact of the time factor in combination with language of instruction on the ability of students to remember newly acquired technical words.

Implications for Specialized Vocabulary in CLIL Classrooms

The main aim of this study was to raise awareness of the challenges that L2-instructed students face when learning subject-specific vocabulary from textbook reading. This was achieved by comparing the learning outcomes of L1- and L2-reading students. The study highlighted several areas where L2-instructed students appeared at a disadvantage and which therefore deserve further pedagogical attention in order to improve the learning experience of students who study content through their nonnative language.

First, the L2-instructed students were able to learn and recall fewer word–meaning connections after reading the subject-related texts. Moreover, the words that they learned were not acquired to the same depth as they were by the L1-instructed students. These findings are consistent with the trend found in the two previous studies that focused on learning specialized vocabulary through L1 and L2 (Haynes & Baker, 1993; Lessard–Clouston, 2006). This growing body of evidence suggests that the development

of specialized language in L2-medium content learning should receive heightened attention given the equal focus in CLIL programmes on curricular learning as well as general L2 proficiency. As was stressed before, specialized vocabulary is an essential part of subject learning (e.g., Woodward–Kron, 2008), and a relatively developed and precise knowledge of subject-related concepts is expected at a pre-university and university level of education. In the results, however, the knowledge of L2-reading students was more general and superficial than that of the L1 baseline. Moreover, it appears that the problem with insufficient depth of learning is not limited to textbook learning only. Results from Lessard–Clouston's (2006) long-term study demonstrated that the gap between native and nonnative speakers also is not easily rectified in learning situations that include a face-to-face component. On the contrary, after a semester at a university, the gap between native and nonnative students in Lessard–Clouston's study widened further.

The findings of the current study suggest that students' ability to learn new word meanings was affected by problems with noticing the links between the words as well as understanding and/or retaining the information from the definitions. To encourage learning outcomes similar to those of mainstream students (i.e., native speakers of the instructional language), it may be useful to employ additional pedagogical techniques that support effective vocabulary learning from content area reading. However, as Flanigan and Greenwood (2007) warn, the recommendations from vocabulary specialists are often too general to be of direct use in content area vocabulary teaching. Instead, specific strategies for content vocabulary instruction that can be directly incorporated into the content classroom activities are needed.

It is due to this specific need that recommendations formulated in work by Hirsh and Coxhead (2009), Harmon, Hedrick, and Wood (2005), and Bravo and Cervetti (2009) are especially useful for developing conceptual knowledge of core subject vocabulary. These techniques are concerned with fostering both the connections between the words and their meanings as well as in-depth understanding and expressible knowledge of words. As was also argued in this study, retention of newly learned words may be related to the initial amount of knowledge of these words. Thus, enhancing the depth of vocabulary knowledge can also positively influence the durability of the learned information. The techniques described

in the studies just mentioned draw on well-established tenets of vocabulary learning, such as multiple and varied exposure to the target lexis, enhanced engagement of students with the words, and guiding students' attention explicitly to word meanings. However, these studies are of special relevance for content vocabulary instruction, as they translate the recommendations of vocabulary experts to the reality of content classrooms. As these techniques were often developed for mainstream content classes, they are meaning-focused in character, which is a trait that lies at the heart of CLIL pedagogy.

Finally, it should be pointed out that, whereas there were some similarities in learning of the two groups, the nature of the knowledge of the L2-instructed students was not fully predictable on the basis of the performance of L1-instructed students. The performance of L2-instructed students was considerably more heterogeneous as indicated by the standard deviation, which, in the case of meaning components, was twice as large as that of the L1-instructed group. A similar observation was made in research on the learning outcomes of bilingually educated students whose performance showed a greater range than that predicted on the basis of the results of students from mainstream monolingual classes (Turnbull, Lapkin, & Hart, 2001).

As a result, specific approaches to teaching and assessment are needed to support the learning experience and educational development of L2-instructed students (de Graaff et al., 2007). As part of these approaches, it appears important also to decide what performance standards the CLIL students are expected to reach. For example, it may be useful for teachers to consciously reflect on the degree of understanding of core subject vocabulary that they want their students to master (Flanigan & Greenwood, 2007) and to adjust their teaching and assessment practices according to these aims.

Limitations and Future Research

Evidence from both experimental and classroom studies is needed before potential intervention studies can be designed and carried out. This study adopted a controlled experimental design in order to provide detailed analysis of the differences between students using L1 and L2 in their learning. Whereas this design allowed a close investigation of different cognitive aspects of learning, it also had several limitations. First, the study focused on one dimension of vocabulary

knowledge—namely word meaning—and studied it from three perspectives (breadth, depth, and retention). There are, however, other dimensions of word knowledge, both productive and receptive (Nation, 2001), that could be explored.

Second, the study tested learners' knowledge orally. Future research should employ a written testing component as well, as this may affect learners' ability to express the knowledge they have (e.g., they may revisit and edit their answers or they may experience less time pressure). Finally, this study focused on contrasting the performance of two groups of students. As a result, individual differences in students' performance (suggested by standard deviations) were not explored in greater depth.

CONCLUSION

By comparing the performance of students learning through their first and second language, this study identified learning areas that may be more challenging for L2-instructed students. It is perhaps not surprising that students working in their second language, even if they are relatively advanced, still fall short of the average performance of students learning through L1. This study, however, went further than just restating this fact and explored the processes that lie behind this observation. The close analysis of the learning outcomes allowed us to identify with greater precision the places where pedagogical intervention may be most effective. Also, the study focused on different aspects of learning that can assist teachers with targeting these problem areas with more specific strategies. Overall, the study demonstrated the complexity of technical vocabulary acquisition and the need for a systematic exploration of this area of vocabulary that lies at the intersection of language and subject learning.

ACKNOWLEDGMENTS

NOTES

[1] The primary source for the history text was *A Concise New Zealand History* (Wikibooks, 2007). The source for the geography text was *Year 12 Geography Study Guide: NCEA Level 2* (Billing et al., 2008).

[2] The frequency of the words in the texts was measured by the RANGE programme (Heatley, Nation, & Coxhead, 2002), which calculates the percentage of words in a text according to frequency bands based on BNC wordlists.

[3] Log Likelihood rather than chi-square was used here as more than 20% of cells in the contingency table had frequency lower than five which can affect the reliability of the chi-square statistic.

REFERENCES

Baddeley, A. (1997). *Human memory: Theory and practice.* Needham Heights, MA: Allyn & Bacon.

Baetens Beardsmore, H. (2009). Language promotion by European supra-national institutions. In O. García (Ed.), *Bilingual education in the 21st century: A global perspective* (pp. 197–217). Malden, MA: Wiley–Blackwell.

Beck, I. L., McKeown, M. G., & McCaslin, E. S. (1983). Vocabulary development: All contexts are not created equal. *Elementary School Journal, 83,* 177–181.

Bernhardt, E. B. (2010). *Understanding advanced second-language reading.* New York: Taylor & Francis.

Billing, A., Broad, T., Carter, K., Coombe, P., Maguire, G., Skinner, B., & Stirling, S. (2008). *Year 12 geography study guide: NCEA level 2.* Auckland, New Zealand: ESA Publications.

Bramki, D., & Williams, R. (1984). Lexical familiarization in economics text books. *Reading in a Foreign Language, 2,* 169–181.

Bravo, M. A., & Cervetti, A. (2009). Teaching vocabulary through text and experience in content areas. In M. F. Graves (Ed.), *Essential readings on vocabulary instruction* (pp. 141–152). Newark, NJ: International Reading Association.

Burger, S., Wesche, M., & Migneron, M. (1997). "Late, late immersion": Discipline-based second language teaching at the University of Ottawa. In R. K. Johnson & M. Swain (Eds.), *Immersion education: International perspectives* (pp. 65–84). New York: Cambridge University Press.

Chen, Q., & Donin, J. (1997). Discourse processing of first and second language biology texts: Effects of language proficiency and domain-specific knowledge. *Modern Language Journal, 81,* 209–227.

Coyle, D., Hood, P., & Marsh, D. (2010). *CLIL: Content and language integrated learning.* New York: Cambridge University Press.

Cumming, A. (2008). Foreword. In D. Albrechtsen, K. Haastrup, & B. Henriksen (Eds.), *Vocabulary and writing in a first and second language: Processes and development* (pp. xiii–xvii). Basingstoke, UK: Palgrave Macmillan.

Dalton–Puffer, C. (2007). *Discourse in Content and Language Integrated Learning (CLIL) classrooms.* Philadelphia/Amsterdam: John Benjamins.

Davidson, R. G., Kline, S. B., & Snow, C. E. (1986). Definitions and definite noun phrases: Indicators of chidren's decontextualized language skills. *Journal of Research in Childhood Education, 1,* 37–47.

de Bot, K., Paribakht, T. S., & Wesche, M. B. (1997). Toward a lexical processing model for the study of second language vocabulary acquisition: Evidence from ESL reading. *Studies in Second Language Acquisition, 19,* 309–329.

de Graaff, R., Koopman, G. J., Anikina, Y., & Westhoff, G. (2007). An observation tool for effective L2 pedagogy in content and language integrated learning (CLIL). *International Journal of Bilingual Education and Bilingualism, 10,* 603–624.

de Groot, A., & Keijzer, R. (2000). What is hard to learn is easy to forget: The roles of word concreteness, cognate status, and word frequency in foreign-language vocabulary learning and forgetting. *Language Learning, 50,* 1–56.

Drum, P. A., & Konopak, B. C. (1987). Learning word meanings from written context. In M. G. McKeown & M. E. Curtis (Eds.), *The nature of vocabulary acquisition* (pp. 73–87). London: Lawrence Erlbaum.

Flanigan, K., & Greenwood, S. C. (2007). Effective content vocabulary instruction in the middle: Matching students, purposes, words, and strategies. *Journal of Adolescent & Adult Literacy, 51,* 226–238.

Flick, W. C., & Anderson, J. I. (1980). Rhetorical difficulty in scientific English: A study in reading comprehension. *TESOL Quarterly, 14,* 345–351.

Flowerdew, J. (1992). Definitions in science lectures. *Applied Linguistics, 13,* 202–221.

Foster, P., & Tavakoli, P. (2009). Native speakers and task performance: Comparing effects on complexity, fluency, and lexical diversity. *Language Learning, 59,* 866–896.

Gass, S. M. (1988). Integrating research areas: A framework for second language studies. *Applied Linguistics, 9,* 198–217.

Grabe, W. (2009). *Reading in a second language: Moving from theory to practice.* New York: Cambridge University Press.

Hague, S. A. (1987). Vocabulary instruction: What L2 can learn from L1. *Foreign Language Annals, 20,* 217–225.

Harmon, J. M., Hedrick, W. B., & Wood, K. D. (2005). Research on vocabulary instruction in the content areas: Implications for struggling readers. *Reading & Writing Quarterly, 21,* 261–280.

Haynes, M., & Baker, I. (1993). American and Chinese readers learning from lexical familiarization in English texts. In T. Huckin, M. Haynes, & J. Coady (Eds.), *Second language reading and vocabulary learning* (pp. 130–152). Norwood, NJ: Ablex.

Heatley, A., Nation, I. S. P., & Coxhead, A. (2002). RANGE and FREQUENCY programs. Accessed 11 April 2014 at http://www.victoria.ac.nz/lals/about/staff/paul-nation

Hirsh, D., & Coxhead, A. (2009). Ten ways of focusing on science-specific vocabulary in EAP. *English Australia Journal, 25,* 5–16.

Huckin, T., & Bloch, J. (1993). Strategies for inferring word-meanings in context: A cognitive model. In T. Huckin, M. Haynes, & J. Coady (Eds.), *Second language reading and vocabulary learning* (pp. 153–176). Norwood, NJ: Ablex.

Huckin, T., & Coady, J. (1999). Incidental vocabulary acquisition in a second language. *Studies in Second Language Acquisition, 21,* 181–193.

Huckin, T., Haynes, M., Coady, J. (Eds.). (1993). *Second language reading and vocabulary learning.* Norwood, NJ: Ablex.

Jiang, N. (2000). Lexical representation and development in a second language. *Applied Linguistics, 21,* 47–77.

Johnson, R. K. (1997). The Hong Kong education system: Late immersion under stress. In R. K. Johnson & M. Swain (Eds.), *Immersion education: International perspectives* (pp. 171–189). Cambridge: Cambridge University Press.

Kellerman, E., & Bialystok, E. (1997). On psychological plausibility in the study of communication strategies. In G. Kasper & E. Kellerman (Eds.), *Communication strategies. Psycholinguistic and sociolinguistic perspectives* (pp. 31–48). London: Longman.

Lasagabaster, D., & Serra, C. (2010). Immersion and CLIL in English: More differences than similarities. *ELT Journal, 64,* 367–375.

Laufer, B. (1997a). The lexical plight in second language reading: Words you don't know, words you think you know and words you can't guess. In J. Coady & T. Huckin (Eds.), *Second language vocabulary acquisition: A rationale for pedagogy* (pp. 20–52). Cambridge: Cambridge University Press.

Laufer, B. (1997b). What's in a word that makes it hard or easy: Some intralexical factors that affect the learning of words. In N. Schmitt & M. McCarthy (Eds.), *Vocabulary: Description, acquisition and pedagogy* (pp. 140–155). Cambridge: Cambridge University Press.

Lee, J. F. (1986). On the use of the recall task to measure L2 reading comprehension. *Studies in Second Language Acquisition, 8,* 201–211.

Lessard–Clouston, M. (2006). Breadth and depth specialized vocabulary learning in theology among native and non-native English speakers. *The Canadian Modern Language Review/La Revue Canadienne des Langues Vivantes, 63,* 175–198.

Lessard–Clouston, M. (2009). Definitions in theology lectures: Implications for vocabulary learning. *Asian English for Specific Purposes Journal, 5,* 1, 7–22.

Meara, P. (1988). *Learning words in an L1 and an L2.* Unpublished discussion paper. Accessed 11 April 2014 at http://www.lognostics.co.uk/vlibrary/meara1987.pdf

Meara, P. (2005). Designing vocabulary tests for English, Spanish and other languages. In C. S. Butler, M. de

los Ángeles Gómez–Gonzáles, & S. Doval–Suárez (Eds.), *The dynamics of language use: Functional and contrastive perspectives* (pp. 271–285). Philadelphia/Amsterdam: John Benjamins.

Menyuk, P., & Brisk, M. E. (2005). *Language development and education: Children with varying language experiences.* New York: Palgrave.

Mohan, B., & van Naerssen, M. (1997). Understanding cause–effect: Learning through language. *Forum, 35,* 4, 22–29.

Nagy, W. E., Anderson, R. C., & Herman, P. A. (1987). Learning word meanings from context during normal reading. *American Educational Research Journal, 24,* 237–270.

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

Nippold, M. A., Hegel, S. L., Sohlberg, M. K. M., & Schwarz, I. E. (1999). Defining abstract entities: Development in pre-adolescents, adolescents, and young adults. *Journal of Speech, Language, and Hearing Research, 42,* 473–481.

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

Paribakht, T. S., & Wesche, M. (1999). Reading and "incidental" L2 vocabulary acquisition. *Studies in Second Language Acquisition, 21,* 195–224.

Parry, K. (1991). Building a vocabulary through academic reading. *TESOL Quarterly, 25,* 629–653.

Parry, K. (1993). Too many words: Learning the vocabulary of an academic subject. In T. Huckin, M. Haynes, & J. Coady (Eds.), *Second language reading and vocabulary learning* (pp. 109–129). Norwood, NJ: Ablex.

Phillips, J. (Ed.). (2005). *Te Ara: The encyclopedia of New Zealand.* Accessed 29 April 2014 at http://www.teara.govt.nz

Saville–Troike, M. (1984). What really matters in second language learning for academic achievement? *TESOL Quarterly, 18,* 199–219.

Short, D. J. (1993). Assessing integrated language and content instruction. *TESOL Quarterly, 27,* 627–656.

Simpson, J. (Ed.). (1989). *The Oxford English dictionary.* Oxford: Oxford University Press.

Stein, M. J. (1993). The healthy inadequacy of contextual definition. In T. Huckin, M. Haynes, & J. Coady (Eds.), *Second language reading and vocabulary learning* (pp. 203–212). Norwood, NJ: Ablex.

Stoller, F., & Grabe, W. (1995). Implications for L2 vocabulary acquisition and instruction from L1 vocabulary research. In T. Huckin, M. Haynes, & J. Coady (Eds.), *Second language reading and vocabulary learning* (pp. 24–45). Norwood, NJ: Ablex.

Swanborn, M. S. L., & de Glopper, K. (2002). Impact of reading purpose on incidental word learning from context. *Language Learning, 52,* 95–117.

Turnbull, M., Lapkin, S., & Hart, D. (2001). Grade 3 immersion students' performance in literacy and mathematics: Province-wide results from Ontario (1998-99). *The Canadian Modern Language Review/ La Revue Canadienne des Langues Vivantes, 58,* 9–26.

VanPatten, B., Williams, J., Rott, S., & Overstreet, M. (Eds.). (2004). *Form–meaning connections in second language acquisitions.* Mahwah, NJ: Lawrence Erlbaum.

Verhallen, M., & Schoonen, R. (1993). Lexical knowledge of monolingual and bilingual children. *Applied Linguistics, 14,* 344–363.

Wesche, M., & Paribakht, T. S. (2010). *Lexical inferencing in first and second language: Cross-linguistic dimensions.* Bristol, UK: Multilingual Matters.

Wikibooks. (2007). *A concise New Zealand history* Accessed 29 April 2014 at http://en.wikibooks.org/wiki/New_Zealand_History

Woodward–Kron, R. (2008). More than just jargon—The nature and role of specialist language in learning disciplinary knowledge. *Journal of English for Academic Purposes, 7,* 234–249.

Yoshii, M. (2006). L1 and L2 glosses: Their effects on incidental vocabulary learning. *Language Learning and Technology, 10,* 85–101.

Zaki, H., & Ellis, R. (1999). Learning vocabulary through interacting with a written text. In R. Ellis (Ed.), *Learning a second language through intraction* (pp. 151–169). Philadelphia/Amsterdam: John Benjamins.

SUPPORTING INFORMATION

Additional Supporting Information may be found in the online version of this article at the publisher's Web site.