The Weight of Involvement Load in College Level Reading and Vocabulary Tasks

Vanessa Beal

A Thesis

in

The Department

of

Education

TESL Centre

Presented in Partial Fulfilment of the Requirements
for the Degree of Master of Arts (Applied Linguistics) at
Concordia University
Montreal, Quebec, Canada

December 2007

Your file Votre référence
ISBN: 978-0-494-40797-4
Our file Notre référence
ISBN: 978-0-494-40797-4

ABSTRACT

The Weight of Involvement Load in College Level Reading and Vocabulary Tasks

Vanessa Beal

This study provides some evidence in support of the Involvement Load Hypothesis proposed by Laufer & Hulstijn (2001) which claims that tasks inducing high levels of *need*, *search*, and *evaluation* of meanings of unfamiliar words are most effective for vocabulary acquisition. Immediate and delayed retention of ten unfamiliar words taken from a short story reading text were investigated in learning conditions of varying "Involvement Loads". These were: low, glossary provided; moderate, multiple choice glossary; high, dictionary based sentence production; and control, reading only. Participants were 118 students from seven intact ESL college level classes at a Montréal area college. On an immediate retention test, there was a statistically significant difference between the means from both the multiple choice glossary task and dictionary based sentence production task and the control group. For delayed retention, there was a statistically significant difference only between the means from the multiple choice glossary task and the control group results.

With a view to addressing unanswered questions in previous Involvement Load research regarding how learners undertake reading comprehension and vocabulary tasks, the present study included a qualitative aspect in which ten participants provided concurrent think aloud protocols whilst completing vocabulary reading tasks of moderate and high involvement loads. Think aloud protocols provided valuable insights into both search and

evaluation strategies used in lexical intervention tasks; furthermore it was clear that metalinguistic knowledge of parts of speech and etymology were enabling strategies. The protocols also highlighted the relative weight of '*need*' in calculations of task Involvement Load as participants indicated a clear preference for the structure and format of multiple choice glossary tasks.

ACKNOWLEDGEMENTS

Writing this thesis has really been a learning process and many people have contributed in a variety of ways. First and foremost I would like to express my sincere gratitude to my thesis supervisor, Marlise Horst. A vocabulary course began a discussion which led to a pilot project, and finally a thesis. All along the way I have benefited from her insightful feedback and encouragement. This has made the process of writing and revising most enjoyable.

Thank you to my committee members Tom Cobb, Pavel Trofimovich, and Joanna White for your careful reading of my proposal and informative guidance about conducting the study. I'd also like to thank Beth Gatbonton for encouraging research by raising so many 'burning questions' right from my first day of the M.A. and Randall Halter for his clear explanations to tricky questions.

I'd like to thank Joan Thompson and Steven Sharpe for their enormous help in conducting this study. I would like to thank all my friends and colleagues who have helped and encouraged me along the way.

Last but not least, I'd like to thank my family, particularly my children, Félix and Zoë, for enduring a sometimes tired and impatient mother. Most of all, I must emphasise my gratitude to my husband, Jean Frédéric, for his love and support, along with all those hours spent at the computer helping me understand my data and attempting to instil rudimentary notions of statistics. Without you all, life would be dreadfully dull!

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

CHAPTER 1: INTRODUCTION

Learning second or foreign language vocabulary necessitates both considerable exposure to the target language and mental effort. As both a language teacher and language learner, I can appreciate the frustration felt when conversation fails because we do not have adequate words, or when looking at a word and feeling that we have seen it many times before, but were unable to assign a meaning. Drawing on both anecdotal and empirical evidence, Nation (1990) suggests that lexical errors are very important and disruptive to communication; learners certainly perceive a limited vocabulary as an important hurdle for their progression. Among others, Nation advocates a systematic and principled approach to teaching and learning vocabulary knowledge, both by teachers and learners, using a wide variety of methods to maximise the vocabulary acquisition process (Nation, 1990, 2001). This study explores the efficacy of a few of these methods in an ESL reading context.

In a review of psychological literature on attention, Schmidt states that, "Nothing is free ... in order to acquire vocabulary one must attend to both word form (pronunciation, spelling) and to whatever clues are available in input that can lead to identification of meaning" (Schmidt, 2001, p.30), concluding that language learners who take a passive approach are likely to be slow and unsuccessful learners.

Intuitively most people would agree with this. Indeed we do not learn to ski without paying attention and we do not begin on vertical mogul filled slopes; there is a gradual progression from the nursery hill to more challenging slopes, gradually integrating information that our instructors have repeated numerous times, and only when we have practiced enough does the information we were told (bend your knees!) seem to

make sense. There is a gradual increase in knowledge going from a general idea of the concept of sliding down a mountain on narrow planks to a more precise notion of the relationships of the physics of movement. Learning vocabulary in a second language (L2) classroom follows a similar path; networks of meanings are created as learners gradually 'get' words and fit them into their ever evolving systems of lexical knowledge. Many learners of a second or foreign language feel, understandably, overwhelmed by the burden of vocabulary learning and wonder how they will ever succeed in the formidable task of learning thousands of words effectively.

Nation states that vocabulary knowledge enables language use, language use enables the increase of vocabulary knowledge, knowledge of the world enables the increase of vocabulary knowledge and language use and so on (1993, cited in Schmitt and McCarthy, 1997); learners have difficulty finding a way to get onto this circle. It is generally understood in second language acquisition (SLA) research that reading promotes second language skills (e.g. Krashen, 1989), which raises questions about how many words learners need to know in order to comprehend a text and begin learning the new words they encounter in it.

Of various estimates of vocabulary sizes, the most common figure is that native speakers have a vocabulary of around 20,000 word families (Goulden, Nation, & Read, 1990). Evidence from various studies (Hirsh & Nation, 1992; Liu Na & Nation, 1985) indicates that a vocabulary of the 3000 most frequent words provides approximately 95% coverage ("Percentage of total text covered by a word or set of words," Schmitt & McCarthy, 1997, p.328) of a text, after which it is implied that learners can use context clues to deduce meanings of the remaining unfamiliar words.

Language teachers also face challenges; they are uncertain as to how they can guide students in acquiring masses of lexical items whilst maintaining interesting communicative classroom environments that are not focused on learning word lists. As a teacher of college (CEGEP) level francophone, English as second language (ESL) learners, I can clearly see the difficulties students face with limited lexicons. There is a common perception that explicit vocabulary instruction is not needed; reading exercises tend to focus on reading strategies, implying that vocabulary will be 'picked up' incidentally. Relatively few words are explicitly taught; therefore most vocabulary that my students acquire must occur incidentally through reading and other L2 exposure. However, despite generally having had English second language instruction since grade four (age nine), many of the students in the intermediate level college courses struggle with both reading and production exercises in English, which suggests that the incidental approach to vocabulary is not working.

The main goal of this thesis is to explore how second language reading activities can facilitate new word learning and how teachers can best structure their reading classes to help students learn and retain new vocabulary.

The next chapter begins by presenting two theoretical positions relevant to vocabulary learning, implicit and explicit learning, and considering how they contribute to a better understanding of vocabulary acquisition and the learning activities that facilitate it. Then I will review research that looks at vocabulary gains derived implicitly from extensive reading and compare this with studies researching explicit instruction and vocabulary gains. The literature review then moves on to consider how reading-based

tasks can be designed for effective vocabulary learning with particular focus on the contribution of the Involvement Load Hypothesis (Laufer & Hulstijn, 2001).

CHAPTER 2: LITERATURE REVIEW

Theories regarding implicit and explicit learning contribute to a better understanding of vocabulary acquisition and effective language learning activities; these will be considered first.

Implicit and Explicit Learning

The concepts of 'awareness' or 'noticing' come out of cognitive psychology and serve as defining features of the implicit-explicit distinction which is widely discussed in SLA literature (see Ellis, 1994; Krashen,1989). Krashen's (1989) input hypothesis (IH) is based on the assumption that language is acquired implicitly by understanding messages. Krashen argues that competence in spelling and vocabulary is best attained by comprehensible input in the form of reading: spelling and vocabulary are acquired naturalistically in the same way as the rest of language. Krashen postulates the presence of a Language Acquisition Device (LAD) through which language is subconsciously acquired: conscious focus is on the message (meaning) not the form; in this way language (and vocabulary) acquisition is implicit. According to Krashen, language learners simply need significant exposure to comprehensible input: learners must read in order to acquire vocabulary. Krashen (1989) relegates the role of explicit, learned competence to that of a 'monitor' for output production. However, this perspective has been seriously questioned by theorists who ascribe a more important role to explicit processes.

Ellis (1994) has addressed the relative contributions of both implicit and explicit learning and knowledge in numerous studies. According to Ellis, implicit learning is, "The acquisition of knowledge about the underlying structure of a complex stimulus environment by a process which takes place naturally, simply and without conscious

operations" (Ellis, 1994, p. 214); explicit learning, in contrast, involves conscious operations. In his framework, implicit learning involves attention to the stimulus, but does not involve other conscious operations. Explicit learning is more conscious; the learner makes and tests hypotheses in a search for a structure. Explicit learning can involve a search or an application of rules; Ellis also refers to a "mediational" aspect which involves mapping or linking of knowledge of the word form to knowledge of the meaning of the word. Ellis claims that some aspects in vocabulary learning, phonology for example, are learned implicitly through repeated exposure, whereas other aspects, such as processing at the semantic and conceptual level to make form-meaning connections involve the meanings of words being learned explicitly. Ellis (1997) argues that acquisition is strongly affected by the quality of the mental processing and can be speeded up by making the underlying patterns more salient as a result of explicit instruction or processes of 'awareness' raising.

Depth of Processing

The depth of processing hypothesis and many theories which emerged from it deal with explicit learning. In the field of educational psychology, Craik and Lockhart's (1972) seminal paper proposes that

> ....human memory be viewed as the record of mental operations carried out primarily for the purposes of perception and comprehension. The operations are performed at various levels of processing where 'deeper' levels involve greater semantic analysis and are associated with longer-lasting memory traces (p.671).

It is suggested that memory trace can best be described in terms of 'depth of processing' or degree of stimulus elaboration. Craik and Lockhart postulate a continuum of

processing varying from shallow sensory analyses to deeper, semantic analyses. The argument is that the deeper the processing or semantic encoding involved, the more durable the learning will be.

Applied to the field of second language vocabulary acquisition, Schouten-Van Parrenen's review of the literature (1995, cited in Folse, 2004) reveals three possible explanations for word retention being related to depth of processing: *elaboration, distinctiveness,* and *difficulty*. In concrete terms, an exercise to fill in the blanks with target words requires more from the learner (i.e. elaboration) than just noticing the target words in a text. Distinctiveness refers to the target word's particularities in terms of semantics, orthography or pronunciation. Jacoby, Craik, and Begg (1979, cited in Laufer, 2005) argue that words which are guessed with some difficulty will be remembered better since difficulty leads to processing effort, which subsequently creates a more distinctive 'memory trace'. Hulstijn (2001) also attributes long-term retention of words to more elaborate processing. The level of elaboration increases as learners are able to pay more attention to more aspects of word meaning, such as morpho-phonological, orthographic, semantic and pragmatic features, as well as inter-word relations.

Input Processing Models

Drawing on similar notions, but in reference to grammar, VanPatten and Cadierno (1993) discuss the role of explicit instruction in L2 classrooms in terms of the input processing involved. VanPatten (1996) proposes that learners process input for its meaning before they process it for form. Input processing concerns the conversion of input to intake using the strategies and mechanisms that promote form-meaning connections during comprehension; intake is then restructured into the developing

linguistic system. As the learner processes input, the 'internal processor' must attend to how the content is encoded linguistically so that acquisition can occur. Sharwood Smith (1981) also distinguished between comprehension and acquisition; similar to VanPatten (1996), he suggests that for acquisition to occur, the learner needs to attend to the linguistic features of the input, rather than just the message.

In a similar vein, Schmidt (1994) proposes a crucial construct of 'noticing' to account for the way in which only input that is noticed becomes available for intake and effective processing. In contrast to Krashen (1989), Schmidt (1994) claims that a level of awareness is important before material can be incorporated into the learners' developing interlanguage system. In his model of 'noticing' Schmidt discusses the significance of instruction and task effects alongside salience and frequency of input to develop a paradigm. This model attempts to show how long-term and short-term (working) memory interact with 'noticing' input so as to process output (Skehan, 1998).

Output Hypothesis

Developing from discussion about the role of 'noticing' in explicit learning is Merrill Swain's (1995) Output Hypothesis. Swain examines the difference between production and comprehension, the difference being that output (production) pushes learners to process language more deeply (with more mental effort) than input (comprehension). Output may be a way to make learners stretch their interlanguage and become actively responsible for their learning. The Output Hypothesis proposes firstly that output promotes 'noticing' a gap between what learners want to say and what they can say; this may trigger cognitive processes as learners bring their attention to items they need to know in the L2. Secondly output functions to test hypotheses about language;

output can result in feedback which can lead learners to 'reprocess' their output. The third function of output is metalinguistic, to enable learners to reflect on their use of the L2. Output, it is claimed, will push learners to move from the semantic processing of comprehension towards the complete grammatical processing which is necessary for accurate production (Skehan, 1998). Output thus has a potentially significant role in the development of interlanguage and specifically, here, in lexical development.

Form-Focused Instruction

Another area of theoretical interest for explicit instruction is the construct of form-focused instruction. Long (1991, discussed in Doughty & Williams, 1998) distinguishes between 'Focus on FormS' and 'Focus on Form'. The former (FonFS) characterises earlier synthetic approaches to language teaching where specific features based on a linguistic syllabus are systematically treated and the primary objective is the linguistic structure that is being targeted (e.g. past tenses). In focus-on-form (FonF) instruction, the focal point of attention is on meaning or communication. "Focus-on-form often consists of an occasional shift of attention to linguistic code features by the teacher and/or one or more students triggered by perceived problems with comprehension or production" (Long & Robinson, 1998, p. 23). In this way the learner's attention is drawn to a linguistic form as it is required in communication. Focus-on-form comprises incidental FonF (unfocused, spontaneous tasks) and planned FonF (focused, predetermined tasks) (Ellis, et al. 2002). Ellis (2001) argues that the distinctions relate to how students view themselves and the language: with focus on formS, students view themselves as learners of a language with the language as the object of study (i.e. words to learn). In focus-on-form, learners view themselves as language users and language (i.e. vocabulary acquisition) is a

communicative tool. The rationale behind focus-on-form instruction relates to the necessity of learners to notice or attend to form (Swain, 1995) and to limited processing capacity (VanPatten, 1996) whereby learners have difficulty dealing simultaneously with meaning and form and will prioritize meaning.

The current theoretical consensus thus suggests that vocabulary learning needs to be explicit and focused, rather than implicit. The implication of this for efficient classroom based vocabulary learning through reading is that drawing specific attention to words in reading class and instruction (focus on form) is likely to be more effective than just reading. Research that explores this proposition is reviewed in the next sections, beginning with an overview of studies that explore the effects of learning vocabulary incidentally through extensive reading and continuing with studies showing the effect of different types of explicit instruction on reading and vocabulary learning.

The 'Reading Only' Approach

Extensive Reading

There is a considerable body of research which looks at the vocabulary gains derived implicitly from extensive reading. Grabe and Stoller (2002) define extensive reading as, "Reading in which learners read large quantities of material that are within their linguistic competence" (p. 259). It is a concept which is applied both in the fields of L1 and L2 reading programmes. Day and Bamford (1998) expand this definition and highlight the importance of reading for pleasure in conditions where learners select reading material that is interesting and appropriate for their level. Reading authentic material underpins communicative language teaching. Krashen (1989) repeatedly stated the importance of reading as comprehensible input for vocabulary acquisition in both

naturalistic and instructed contexts: ideally this involves huge amounts of reading, so that there are possibilities of repeated exposures to words. Whilst considerable evidence from L1 studies suggests that extensive reading for meaning leads to vocabulary acquisition over time, evidence from L2 studies indicates that although vocabulary acquisition does occur incidentally through exposure to language input, its efficiency remains debatable. There is a general consensus in SLA literature that the minimum number of contextual reading encounters for a new word to be learned well enough for the reader to be able to match the item with a meaning is around ten (Saragi, Nation, & Meister, 1978); their study documented native speakers' acquisition of "nadsat" (slang words of Russian origin) words from Anthony Burgess' *A Clockwork Orange*.

Pitts, White, and Krashen (1989) replicated the Saragi et al. study with adult ESL students reading two chapters of *A Clockwork Orange*. Learners were subsequently tested for their understanding of nadsat words. Results revealed small but significant incidental acquisition of nadsat words relative to the control groups who had not read the chapters. The researchers claim this shows that L2 learners can acquire vocabulary by reading. Similarly, Horst, Cobb and Meara (1998), using a simplified version of Thomas Hardy's *The Mayor of Casterbridge* to investigate incidental L2 vocabulary acquisition, found that "...small but substantial amounts of incidental vocabulary learning can occur as a result of reading" (p. 214), but they also conclude that, "The power of incidental L2 vocabulary learning may have been overestimated" (p. 220). As this research shows, incidental exposure to language, through extensive reading, does not result in large amounts of vocabulary acquisition and retention.

A Default Hypothesis

Laufer (2005) explains the fault of the 'default hypothesis' whereby it is believed that vocabulary is acquired incidentally by reading. She argues that none of the lexical assumptions from L1 studies regarding skills such as noticing, guessing, repetition and retention, and extrapolation can be assumed in L2 learning. Evidence of this position can be seen in Laufer and Yano's study (2001) showing that not only do L2 learners not necessarily 'notice' unfamiliar words in a reading task, but they overestimate their understanding of unfamiliar items by as much as 60%. The authors explain this in terms of the low level input in instructed language learning compared to naturalistic acquisition; many words are somewhat familiar, but not yet firmly established in the mental lexicon. In addition learners who have understood the overall idea of the reading do not pay attention to the precise meaning of individual words.

Laufer (2005) argues that reading alone does not suffice for lexical development; finding that not only do learners fail to notice unfamiliar words, but when they do notice them, they frequently cannot infer meaning as context clues may be stated in words that are just as unfamiliar. Research in the area of the known-word density necessary to infer new word meanings from context highlights the learners' predicament. This research shows that L2 readers whose known-word coverage of a particular text is less than 95% will have difficulty trying to infer the meanings of the remaining words (Laufer, 1997; Hirsh & Nation, 1992). In the L2 classroom setting it is particularly unrealistic to assume that there will be 95% coverage of reading texts for the majority of students as there is a general requirement for texts read intensively in class to be authentic, challenging, and

interesting. It could be argued that students would not be motivated to read texts that offer high known word density because they perceive them as too easy.

The degree of difficulty (real and perceived) influences the 'depth of processing' learners use to deal with vocabulary tasks. Mondria and Wit-de-Boer (1991) compared guessing and retention of words in what were labelled 'pregnant', 'moderately pregnant' and 'non pregnant' sentence contexts whereby pregnant was defined as offering an abundance of clues that support the meaning. They found that words appearing in 'pregnant' contexts were guessed correctly with ease, but were not remembered well, whereas words whose meaning it was relatively difficult to guess from a sentence context tended to be better recalled than words from simpler contexts.

The experiments in L2 vocabulary learning reviewed above have shown that incidental vocabulary learning as a result of exposure to input is very low. Learning may be inaccurate if contexts are not well understood, but new words may not be retained or even noticed if contexts are readily comprehensible. The next section will consider empirical evidence showing that reading plus explicit instruction is more effective for vocabulary acquisition than reading alone.

Reading PLUS Instruction

Mental Processing

The argument that reading-only approaches are insufficient is bolstered by many studies that show the benefits of explicit vocabulary instruction. In a meta-analysis, Stahl and Fairbanks (1986) analysed the effectiveness of vocabulary instruction methods and concluded that explicit, cognitive strategies, such as taking notes and creating personal

ownership and structure for new material, were effective for the task of vocabulary acquisition.

Studies indicate that lexical retention is in part determined by the way in which that information has been processed; more mental elaboration will result in higher retention scores than will less elaboration. Increasingly research has focused on showing measurable benefits of explicit conditions for vocabulary instruction. Studies (e.g. Paribakht & Wesche, 1993) seem to point to the usefulness of explicit instruction for lexical development with intensive reading texts as learners cannot realistically be expected to 'pick up' words or guess word meanings correctly from all contexts.

An early study exploring these ideas was done by Hulstijn (1992). He initially found that if L2 learners read a text for comprehension of its content, and not with the intention of expanding their L2 vocabulary, they are more likely to remember the form and meaning of an unknown word in the text when they have inferred its meaning by themselves (high mental effort) than when words that were explained by a synonym (low mental effort). Following on from this, Hulstijn, Hollander and Greidanus (1996) compared the effects of giving the meaning of unknown words in a reading comprehension task in two ways: by means of marginal glosses and by allowing participants to use dictionaries. They discovered that the overall retention from marginal glosses was greater than that from dictionary use partly because participants in the dictionary condition often did not use the dictionary at all! However, when they did use the dictionary, they remembered the words looked up as well as, or better than, participants who had used the marginal glosses, presumably because using a dictionary

requires more effort. Thus, it is suggested that more successful retention is related to the greater amounts of mental processing involved in learning a word and its meaning.

Further evidence of the relationship between greater mental processing and better vocabulary learning comes from Hulstijn and Laufer (1998, cited in Hulstijn & Trompetter, 1998), who compared incidental L2 vocabulary learning in reading and writing conditions. In the reading condition, participants were given marginal glosses to explain the meaning of ten low frequency words appearing in the reading text; the more demanding writing condition consisted of a list of the same ten target words (with L1 translation and a sample sentence) and participants were required to write a text using the ten words. The findings showed that the writers retained more words, suggesting that word processing for writing (the task demanding most effort) leads to better retention than the processes for reading.

Hulstijn and Trompetter's (1998) study also speaks to the importance of effort involved in elaborative learning. They used computer-assisted reading and writing tasks to compare incidental vocabulary learning of French L2 for Dutch L1 learners under a reading and a writing condition. Of the 110 participants, only 70 looked up more than ten words while performing the reading or writing task. There was not a significantly better retention in either condition; however Hulstijn and Trompetter (1998) suggest that this is accounted for by the fact that participants of this study processed the lexical information they looked up only very briefly and in a superficial way, requiring minimum amounts of mental elaboration.

Paribahkt and Wesche (1993, 1996, & 1997) have similarly argued that engaging in "greater mental effort" is essential for more efficient learning and retention of words.

In 1993 and 1996, Paribakht and Wesche attempted to categorize reading related vocabulary exercise types in textbooks to explore the role of various vocabulary instruction techniques based on reading texts for vocabulary learning by university ESL students. A *reading plus* treatment was designed through examination of L2 vocabulary teaching textbooks. The researchers hypothesized a hierarchy of mental processing activity and grouped exercises into five distinct categories listed here from least to most demanding: selective attention (e.g. boldfacing, signaling target words in texts); recognition (e.g. matching target words with definitions); manipulation (e.g. using stems and affixations); interpretation (e.g. multiple choice cloze, classifications, and guessing from context), production (e.g. open cloze, labeling pictures) requiring recall and reconstruction of words.

The Vocabulary Knowledge Scale

To render operational their hierarchy of mental processing, Paribakht and Wesche (1993) developed the Vocabulary Knowledge Scale (VKS), which is a 5-point scale, involving self-report and performance items to track learners' acquisition of specific words in their written form. The self-report categories in this scale are shown in Figure 1.

I:	I don't remember having seen this word before;
II:	I have seen this word, but I don't know what it means;
III:	I have seen this word before, and I think it means _____ (translation or synonym);
IV:	I know this word. It means_____ (translation or synonym);
V:	I can use this word in a sentence: _____ (write a sentence; if you do this section make sure that you also complete IV).

Figure 1. VKS scale (Paribakht & Wesche, 1993)

In Paribakht's & Wesche's subsequent studies on theme-based readings, students exposed to the *reading plus* treatment, in which tasks were designed to increase the salience and cognitive processing of targeted words in the texts, made significantly higher gains quantitatively and qualitatively than those exposed to the *reading only* treatment where participants read the core reading text plus two thematically related texts.

Enhanced Readings

In recent years there have been numerous studies considering the cognitive benefits of 'enhanced' readings (glosses, dictionary use, computer assisted) and supplementary vocabulary tasks. For instance, Rott (2005) investigated which reading interventions facilitated creation of lexical form-meaning connections (FMC) that would be robust and retained over time. L2 learners read a text enhanced with either multiple choice glosses (MCG) or single translation glosses (STG). In both treatments the target words occurred three more times after the glossed occurrence. The results suggested that MCGs may lead to stronger, more developed form meaning connections than STGs. Strengthened form meaning connections seemed related to the integration of multiple

metacognitive and semantic elaborative resources: the repeated search and assessment of individual word meanings as well as recursive reading strategies. Readers had equal comprehension of the main ideas regardless of gloss type, but multiple choice gloss readers showed a tendency to comprehend more supporting ideas. This may have been because they were more engaged in the reading interaction: the think aloud protocols used in this study showed how the participants consistently checked the meanings that they had selected in the MCG and thus at the end of the reading had a strong sense of the appropriate form meaning connection in four different contexts. The study supports the argument that higher involvement and mental effort required for inferring and hypothesis–testing of word meaning leads to better word retention (Hulstijn, 1992). It also highlights the usefulness of think aloud protocols in understanding how learners make use of resources such as glosses.

Planned Lexical Instruction

Laufer (2006) developed the notion of explicit instruction and vocabulary learning further and compared the effectiveness of Focus-on-Form (FonF; where communication of meaning remains the main focus) and Focus-on-FormS (FonFS; where the language item is isolated and linguistic feature is the focus) with six classes of high school English foreign language students in Israel. Laufer (2006) argues that there is a place for FonFS in vocabulary learning as, "Meaning focused learning does not necessarily lead to a satisfactory vocabulary development" (p.152). In her study students in the FonF condition read a text containing the target words and answered comprehension questions. Students in the FonFS condition studied the target words as discrete items with their meanings and examples of usage. The FonFS condition yielded significantly higher

results than FonF (72% as opposed to 47% of the word meanings were retained). Laufer advocates FonFS on theoretical grounds, "Lexical competence is a combination of different aspects of vocabulary knowledge, together with vocabulary use, speed of access, and strategic competence." In planned lexical instruction, both FonF and FonFS have a role. "Planned lexical instruction ensures noticing, provides correct lexical information, and creates opportunities for forming and expanding knowledge through a variety of word focused activities " Laufer (2005, p. 311).

In summary, the research discussed in this section shows that recognition of the need for planned vocabulary instruction goes some way towards addressing the need for something more than an implicit 'pick up' of new words; however, empirical studies in explicit learning and knowledge continue to face the difficulty of operationalizing notions of 'depth of processing' and assessing the appropriateness of particular tasks for vocabulary learning objectives. Laufer and Hulstijn (2001) advanced the Involvement Load Hypothesis to compensate for the lack of operationalizable definitions in depth of processing theory (Craik & Lockhart, 1972) to define depth and to determine which language learning processing tasks were deeper or more superficial. This research is discussed in the next section.

The Development of the Involvement Load Hypothesis

Laufer and Hulstijn proposed a motivational-cognitive construct of involvement, comprising three basic components: *need, search,* and *evaluation* which they believed could be readily operationalized in order to address this problem. Each of these will be discussed in turn.

The *need* component refers to the motivational, non-cognitive dimension of involvement. A considerable amount of research has been directed towards understanding the role of affective variables, in explaining individual differences in second or foreign language learning. In particular, motivation has been given a great deal of attention. Pioneers in this area were Gardner and Lambert (1959), who postulated the existence of two classes of orientations (goals): *integrative* in which the learner wants to learn the L2 in order to interact, identify, and possibly join the L2 community and *instrumental* in which the learner is learning the L2 for a practical goal such as a job, credit, or social status. Theories of an intrinsic and extrinsic motivation continuum (Dörnyei, 1990) have largely replaced the integrative-instrumental dichotomy. In this new paradigm, intrinsic motivation refers to motivation to engage in an activity because it is enjoyable and satisfying to do. Extrinsically motivated behaviours are actions carried out to achieve some instrumental end, such as a reward or to avoid punishment.

In terms of L1 reading motivation, Guthrie (2001) defines *engaged reading* as a merger of motivation and thoughtfulness, and engaged readers as mastery oriented, intrinsically motivated, and possessing self-efficacy. Engaged readers, who are intrinsically motivated to read for knowledge and enjoyment, focus on meaning and avoid distractions. Self-efficacy is another aspect of reading motivation: this is an ability to organize and execute courses of action required to attain designated types of performances: students with high self-efficacy perceive difficult reading tasks (i.e. vocabulary processing) as challenging and work diligently to master them. Motivation processes are the foundation for coordinating cognitive goals and strategies in reading.

Students with high intrinsic motivation, a task orientation, and high self-efficacy are relatively active readers and high achievers (Guthrie, et al. 2000). It can be assumed that their motivation to read would correlate to some degree with a motivation to acquire vocabulary competence. In the Involvement Load Hypothesis, Laufer and Hulstijn (2001) define two degrees of prominence for *need*, moderate and strong, in terms of the intrinsic and extrinsic motivation distinction. Thus, their conceptualisation of need goes some way towards incorporating the motivation theories outlined above.

Search and Evaluation

Search and *evaluation* are the two cognitive or information processing elements of involvement; they are conditional upon the allocation of learner attention to form-meaning relationships (Schmidt, 1994 cited in Hulstijn & Laufer, 2001). For Laufer and Hulstijn (2001) *search* is the attempt to find the meaning of an unknown L2 word or to find the L2 word to express a concept (i.e. translation from L1 to L2) by consulting a dictionary or authority such as a teacher. In their framework, the term *evaluation* refers to the assessment of an appropriate meaning or use of a given word as is prescribed by a specific context. For example, in a writing task in which an L1 word is looked up in a dictionary and three L2 alternatives are given; the learner will need to make decisions to select the most suitable term based on a criterion of semantic and formal appropriateness of the word and its context. Laufer and Hulstijn (2001) propose that if the evaluation entails recognising differences between words, or differences between several senses of a word in a given context the evaluation is 'moderate'. By comparison, if evaluation necessitates making a decision about additional words which will combine with the new word in an original sentence or text it is referred to as 'strong' evaluation. According to

Laufer and Hulstijn (2001), all of the three involvement factors may not be at work simultaneously during a reading task. The combination of these three factors with their degrees of prominence comprises the Involvement Load of a task.

Empirical Studies of Involvement Load

Laufer and Hulstijn (2001) proposed an "involvement index", which assigns numerical weights in which the absence of a factor equals 0, a moderate presence equals 1, and a strong presence of a factor equals 2. The basic claim of the Involvement Load Hypothesis is that "retention of unfamiliar words is, generally, conditional upon the degree of involvement in processing these words" (Laufer & Hulstijn, 2001). The researchers acknowledge that the three involvement factors (need, search, and evaluation) may not share equal weight in determining task effectiveness with regard to learning. In particular, they suggest that the 'weight' of search may be lower, less significant, than that of need and evaluation. Similarly the researchers do not distinguish between the effect of input tasks and that of output tasks with identical involvement indices. The Involvement Load Hypothesis does, however, present a research paradigm through which to manipulate factors in evaluating different tasks, so as to explore the relationship between retention and aspects of deep processing with varying combinations.

In order to test their hypothesis specifically, Hulstijn and Laufer (2001) conducted two parallel experiments with young adult EFL learners in Israel and the Netherlands to test the assumption that tasks with a higher Involvement Load would be more effective for vocabulary retention than tasks with a lower Involvement Load. Ten low frequency target words (TWs) and expressions were selected from a text and three tasks with different Involvement Loads administered to different groups of students. Task 1

consisted of a reading comprehension with marginal glosses; its involvement index was 1 because there was moderate need (it was teacher /task imposed) but neither search (as the glosses were provided) nor evaluation (the appropriate meaning for the context was provided). Task 2 consisted of a reading comprehension plus 'fill in the gaps'. The same text and questions were used as in Task 1, but the ten target words were deleted leaving gaps; the target words and five distracters with L1 translations and L2 explanations were listed on a separate sheet. The participants had to read the text and fill in the ten gaps with words from the list and answer comprehension questions. Task 2 induced moderate need (as it was teacher/ task imposed), no search (meanings for target words in L1 and L2 were provided) and moderate evaluation (because the context was provided); therefore its involvement index was 2. The final task consisted of writing a composition and incorporating the ten target words; they were instructed to write a letter in which "you must use the following ten words". Task 3 again induces a moderate need, no search (as meanings were provided), and a strong evaluation (because the words had to be used appropriately in an original context). Its involvement index was 3. The tasks thus varied in terms of the degree of evaluation involved.

Drawing on the Involvement Load Hypothesis, Hulstijn and Laufer hypothesised that the retention scores would be highest in Task 3 with the highest involvement index. The results partially fulfilled the hypothesis in that the mean scores from the composition group were significantly different from the marginal gloss and fill-in-the blank groups; however there was no significant difference between the means of these two groups and the researchers indicate that the superiority of the results from the composition task may be due to time taken and to learners having been required to stretch their linguistic

resources (thereby supporting Swains' Output Hypothesis, 1995, discussed above, which claims that learners need tasks that will push output and trigger mental processes of 'noticing' and analysis).

The Involvement Load Hypothesis has definite strengths: it presents a concise framework to analyse tasks and interpret varying levels of retention. It is also well grounded in cognitive psychology and supported by empirical studies suggesting that when learners have to infer or induce a solution to a problem they will invest more mental effort than when the solution is given and that information that has been attained with more mental effort can later be better retrieved and recalled than information that has been attained with less mental effort (Hulstijn, 1992). However, similar to depth of processing theory, it remains difficult to operationalize in practice. Language learning is an internal process: it is difficult to evaluate which processes work, for whom, and under which circumstances. There is clearly a need for a better understanding of how the theory works in real L2 classrooms. In addition to environmental factors such as programme constraints, or learners' perception of the utility of the L2, individual factors such as motivation, aptitude, and learning styles make it difficult to generalise and to identify the most effective vocabulary acquisition tasks.

It was with these issues in mind that I ran a pilot project, replicating to some degree the Hulstijn and Laufer (2001) study with young adults learning English in a second language context in Québec, Canada. The project applied the Involvement Load Hypothesis to assess incidental vocabulary from reading comprehension tasks with college level students in a Montreal area college. Before presenting the main study, I will briefly present the results of the pilot project.

CHAPTER 3: PILOT PROJECT

The pilot project was conducted with three intact classes (65 participants) of high intermediate students during the winter 2006 semester, at a Montréal area college. Three short story reading texts were selected: *The Lottery* by Shirley Jackson; *The Sniper* by Liam O'Flaherty; and *Lamb to the Slaughter* by Roald Dahl. A counterbalanced research design was used, meaning that the students in each of the three groups were exposed at different times to each of the three conditions; as such they served as each others' control groups. Similar to the Hulstijn and Laufer (2001) study, ten low-frequency words were selected from each reading text. In all three conditions students needed to find the word in its context, pay attention to it by writing out the complete sentence from the text; all also had to answer comprehension questions and discuss the story in class. Students also completed one of three post-reading tasks that were designed with different (increasing) Involvement Loads: A) examining a glossary of target words with definitions in English, B) completing multiple choice glosses (with synonyms/definitions in English), C) providing dictionary definitions in English or synonyms for the target words and using the words in original sentences.

A vocabulary awareness pretest indicated that I could be fairly confident that the thirty target words (TWs) selected (ten from each short story) were not known prior to the treatments. In each of the three sessions, students from each group read one of the three short stories as part of their English course such that each group were exposed to all of the three treatments. The experimental treatments were assigned as homework; students also prepared for a quiz for the following class which tested knowledge of the ten target words along with reading comprehension and grammar. These quizzes involved

matching the TW with definitions provided. Performance on this testing (Posttest 1) was compared to students' pretest knowledge of the targets. The second test (Posttest 2) took place two weeks after completion of the tasks related to the last of the three experimental readings. It consisted of a complete list of the thirty TWs and a multiple choice task of selecting the correct meaning from three distracters.

Results showed that vocabulary acquisition did take place: gains consistently occurred in all three participant groups for all three texts, regardless of the study condition. This is not entirely surprising since previous research has shown that incidental acquisition of vocabulary will occur from reading short stories with a reading comprehension focus. Results in the three study conditions are shown in Table 1. The mean retention scores and standard deviations of the students for both the immediate (T1) and delayed (T2) posttests are indicated.

Table 1.
Number of Participants, Mean Retention Scores (%), and Standard Deviations in Immediate and Delayed Posttests in Pilot Project.

	Immediate (T1)			Delayed (T2)		
	N	Mean	SD	N	Mean	SD
Task A: Glossary	60	61.5	2.9	65	75.2	1.96
Task B: Multiple choice	62	71.1	2.5	65	70.5	2.43
Task C: Definitions	59	68.3	2.8	65	73.1	2.38

Performance means for the tasks with higher involvement levels (Tasks B and C) were higher than the means for the task with low involvement (Task A) at the immediate (T1) posttest. Although there was no significant difference between task type

$F(1.\ 93) = 14.\ 86, p = 0.147$ in the immediate test (T1) or in the delayed (T2) test $F(0.\ 66) = 3.\ 43, p = 0.515$, the immediate posttest (T1) results indicated a trend that 'more involvement' did indeed equate with higher scores.

The pilot study highlighted a number of issues that motivated the main study. It was concluded that a greater control over factors such as pretest vocabulary knowledge and timing of the testing is needed to clarify the effect of the tasks. The influence of time was particularly evident in this study: students performed better on the tasks associated with the third, most recently read text. In addition, the delayed test (T2) was conducted two weeks after reading the third text, three weeks after the second text and four weeks after the first which meant that some of the words from earlier in the term were harder to retrieve. More important though was the familiarity that had crept into the project. By the time that students were reading the third text, *Lamb to the Slaughter*, they had become very familiar with the routine of having reading comprehension questions and a vocabulary task followed by a quiz in the subsequent class which included the vocabulary items. Consequently, results may merely reflect a test effect, rather than the effect of differential task Involvement Loads.

In addition to addressing the problem of recency effects and test effects, there were a number of other concerns. Firstly, there was no other instructional attention given to words in the pilot study; however, certain words seemed to have been 'picked up' and used spontaneously more than others. At least ten students used the TW *remorse* in a completely unrelated listening comprehension written summary task. Research supplemented with a qualitative approach may be more able to gain insight into 'depth of knowledge' and thus evaluate whether some words are more straightforward to learn than

others. Secondly, the results of the pilot project also indicated that a better portrait of participants' prior vocabulary knowledge needs to be provided: one of the classes involved appeared on the whole to be weaker and less motivated than the others. For this reason, participants in the main study completed more extensive pretesting including the Vocabulary Levels Test (Schmitt, 2000) in addition to a vocabulary awareness pretest to ensure that the target words selected were unknown and that the participating groups were comparable. Additional unanswered questions that arose from the pilot study pertained to students' perception of the words and their usefulness, interest in the readings, and desire to complete the tasks. One of the concerns with the Involvement Load construct is the relative weighting of need, search, and evaluation when assigning task load; the inclusion of a qualitative, talk aloud, introspective aspect in the main study aimed to explore learners' perceptions and strategies when doing the vocabulary tasks.

In addition, the pilot study revealed the need for more sensitive testing to try and identify whether students exposed to the more involved tasks perform better in terms of depth of knowledge with the target vocabulary items.

Most importantly in terms of research design, administration of the pilot project showed the difficulties of isolating exposure to treatment and ensuring the effects of tasks were assessed in a valid manner. Students became used to the tests that followed the reading and vocabulary tasks and as the weeks went by, I could no longer be certain that they had not ignored the task instructions and simply studied the words. Over time there was also more likelihood of them comparing notes with students in other groups who were exposed to the other treatments.

For this reason, for the present study, the three-way counterbalanced research design was abandoned in favour of a more conventional methods-comparison study using a single reading text. One intact class was also included as a control group exposed to the reading and comprehension exercises. Two intact classes were exposed to treatment A (glossary); two other classes were exposed to treatment B (multiple choice glossary), and two intact classes were exposed to treatment C (dictionary-production). Using one text to compare the methods also permitted a greater control of variables.

In summary, the motivation for this thesis arises from the review of the literature and the results of the pilot project's manipulation of Hulstijn and Laufer's constructs of *search* and *evaluation* that constitute parts of a task's measurable Involvement Load. The next chapter describes the research questions and hypotheses behind the study.

CHAPTER 4: RESEARCH QUESTIONS

The goal of this study was to investigate the effect of Involvement Load on the retention of low-frequency English single-word vocabulary items in reading comprehension tasks by young adult, high intermediate English L2 learners in Québec. In addition to testing the Involvement Load hypothesis in an ESL setting rather than the EFL context of Hulstijn and Laufer's research, the study is designed to explore in greater depth how learners approach second language reading comprehension and vocabulary tasks of differing involvement levels and suggest why some tasks may be more effective than others.

The research questions addressed are:

a. Do learners of English acquire knowledge of new vocabulary items (incidentally) when reading a short story with a focus on reading comprehension?
b. Which types of vocabulary exercises accompanying reading texts are more effective for short term vocabulary acquisition? For long term vocabulary acquisition?
c. What do think aloud protocols reveal about the activities learners engage in while they complete two different tasks: 1) completing multiple choice gloss tasks and 2) writing sentences based on dictionary look-ups?

Drawing on the conclusions of the literature review and the pilot study, the following hypotheses have been formed:

1. Some incidental acquisition of vocabulary will occur from reading short stories with a reading comprehension focus;

2. Tasks with a higher Involvement Load will result in higher levels of both short term and long term retention.

In terms of explaining the hypothesized quantitative results, it is believed that oral think aloud protocols will reveal that:

3. Sentence production tasks that include dictionary look-ups will engage learners in a greater number and variety of cognitively demanding activities than those required to complete the multiple choice glossary task.

CHAPTER 5: METHOD

Participants

The participants of this study were seven intact classes of high intermediate ESL students at a Montréal area, Francophone College. This study occurred during the winter 2007 session. All were in sections of 604-101, which is the first of two required ESL courses taught at the institution. From a total of 167 eligible students, 155 (93%) initially participated, that is, they filled out the consent form and completed all the pretesting (vocabulary level and vocabulary awareness tests described below). Of these, 118 (71%) students completed all the tests; data from these 118 were included in the main analyses.

I taught three of the classes involved in the study and four other classes were taught by another experienced, very helpful, teacher in the language department. The students ranged in age from 17- 20 years. All participants had completed both elementary and secondary school in French, and learned English through regular Québec ESL school programmes and exposure to a wide range of English language media. Whilst the students may have had varied educational experiences in English as a second language, the 167 participants in intact classes can be considered to be at a similar level of ESL competency as they are placed in courses on the basis of a combination of their secondary V ESL exam results and a locally administered ESL placement test. The participants are in the first year of their college programme and are taking the first of two compulsory English second language courses. The ESL courses mix students from all programmes, so generally each class of approximately 26 is formed of students, male and female, from pre-university and technical (vocational) programmes. The curriculum is based on communicative teaching methodology but incorporates some form-focused language

teaching for grammar and lexical instruction. Students use bilingual and monolingual dictionaries, although most lexical work requires that students provide definitions in English. Techniques for vocabulary acquisition are generally only mentioned in passing and vary according to teacher and class.

So as to check that the intact groups were indeed comparable, all the participants completed the Vocabulary Levels Test (Schmitt, 2000). The Vocabulary Levels Test (VLT) was originally created by Nation (1990) and provides an estimate of receptive vocabulary size. It is easy to administer and evaluate and only takes about 30 minutes to complete. The test samples words from the 2000 to 10,000 most frequent words of English. In this study, participants took Schmitt's (2000) expanded version of the VLT. The receptive knowledge test used in the study samples 60 words from each of the three frequency level bands: 1001-2000, 2001-3000, and 3001-5000; these are presented in order of decreasing frequency which corresponds to increasing difficulty. The 60 words at each level are presented in 10 clusters of 6 words and 3 definitions. These are shown in Figure 2.

Choose the right word to go with each meaning. Write the number of that word next to its meaning.

1. business
2. clock ____6____ part of a house
3. horse ____3____ animal with four legs
4. pencil ____4____ something used for writing
5. shoe
6. wall

Figure 2. Sample VLT cluster

As shown, the participants write the number of the appropriate word beside the corresponding definition. A score out of 30 is given for each level, and this score is used to estimate the proportion of words that a learner can be expected to know from a particular frequency level. The results from the VLT testing in the participant groups are shown in Table 2.

Table 2.
Group Mean Scores for VLT.

	Number of test-takers	VLT level 2000 score (mean out of 30)	VLT level 3000 score (mean out of 30)	VLT level 5000 score (mean out of 30)
Participants in Task A	42	24.52	24.95	21.33
Participants in Task B	43	23.98	24.07	20.44
Participants in Task C	47	24.53	23.64	20.60
Participants in Control group	23	24.78	24.78	21.57
Total	155			

The scores from the VLT were submitted to an ANOVA in which independent variables (task type; group) were analysed on the basis of dependent variable (scores). No significant difference was found between the participants in different groups $F,(3,117) = 0.41$, $p = 0.74$ at the 5000 level band. Thus, the groups were comparable in terms of vocabulary knowledge.

For the qualitative component of the research, ten participants who had been part of the quantitative study provided concurrent think-aloud protocols which were audio-

taped and analysed. The participants were volunteers, selected mainly on the basis of convenience: they had completed in-class assignments and were available to stay for an extra 20 minutes after class.

Procedure

Much SLA research makes reference to the importance of classroom motivation and learners' interest in the material. The text chosen for the study was partly selected on the basis of the results of the pilot study in which three short stories were used. Participants in the pilot study expressed a clear interest and preference for this story above others; more of the participants completed all of the exercises, and classroom discussion after the reading was animated.

The reading text used was *Lamb to the Slaughter* by Roald Dahl, a story in which a betrayed wife murders her husband with a frozen leg of lamb and then serves the murder weapon with a selection of vegetables to the investigating police officers. The story is available online; it is frequently used with ESL students at the college level in Québec and is considered appropriate both in terms of language and interest level. The story can be found in Appendix F.

The Experimental Treatments

In the present study, three tasks that varied in level of motivational cognitive involvement (see earlier discussion of Involvement Load) were implemented as follows. Students in all three conditions and the control group read the experimental text at home. Participants answered reading comprehension questions and in the subsequent class they discussed the story in small groups and then individually completed the tasks. Each task involved initially highlighting the sentence in the text where the target word appears; the

tasks differed in the various conditions in the ways shown below. The task work sheets can be found in Appendix B.

Task A, "Glossary Provided": Students engage in the highlighting activity using a worksheet where the target words (TW) are presented with a glossary of definitions in English.

For example:

placid (adj): calm in nature or appearance; tending not to become excited or disturbed.

My adaptation of Hulstijn and Laufer's scheme can be seen in Table 3, in which the Involvement Load for doing the task under these fairly undemanding conditions totals **2**. In all three conditions, motivation (need) is considered moderate and assigned a value of 1 because it is teacher imposed. Search is considered moderate in task A and is assigned a value of 1 because students have to search for the contextual use of the word. Evaluation is assigned a value of 0; students are given the definitions and do not need to assess appropriate form meaning relationships.

Task B, "Multiple Choice Glossary": Students need to select the correct meaning for target words from a multiple-choice glossary of synonyms / definitions in English.

For example,

placid (adj): calm in nature or appearance / clear understanding / becoming excited

The Involvement Load total for task B is **3**. Search is moderate and assigned a value of 1 because students have to check the given meanings against the contextual use of the word. Evaluation is assigned a value of 1; it is moderate because students have to evaluate which of the definitions in the multiple choice glossary is the most appropriate

for the particular context. Motivation (need) is considered moderate and assigned a value of 1 because it is teacher imposed.

Task C, "Sentence Production": For the most involved task, which has a total value of **5**, students need to provide definitions in English or synonyms for the target words and use the words in written original sentences.

For example, the task for the word *placid* reads as follows:

placid (adj):

a. *Find the context for the word taken from the story.*
b. *Provide a dictionary definition / synonym.*
c. *Use the word in a sentence to show that you understand its meaning.*

Search is strong and assigned a value of 2 because students have to find the contextual use of the word and utilise a dictionary. Evaluation is assigned a value of 2; it is strong because students have to use the dictionary correctly and evaluate which of the definitions is the most appropriate for the particular context. Again, motivation (need) is considered moderate and assigned a value of 1 because it is teacher imposed.

Control group: The control group has no vocabulary task assigned. They read the same short story, answered comprehension questions, and participated in small group discussions.

The tasks and their Involvement Load weightings are summarized in Table 3.

Table 3.
Classification of Experimental Tasks According to Involvement Load Criteria from Hulstijn and Laufer (2001)

Task	Status of Target words	Involvement factors			Involvement Index
		need	search	evaluation	
A. Reading comprehension; attention to context of TW.	glossary	+	+	-	2
B. Reading comprehension; attention to context of TW; select meaning from multiple choice distracters.	multiple choice glossary (listed with distracters)	+	+	+	3
C. Reading comprehension; attention to context of TW; define TW in English and write sentence.	listed only	+	++	++	5

The study attempts to address the imbalance of time taken for output in the previous Involvement Load research by making the time taken for each task reasonably comparable. Although output is involved for Task C, it is not a composition (like in the Laufer & Hulstijn, 2001 study) and whilst students are required to process the meaning and compose an original sentence, they do not have to plan a complete text. Nonetheless, time on task was different; Tasks A, B, and C took about 35-40, 45-60 and 50-60 minutes respectively. There was also more interaction between students using dictionaries in Task C; this will be discussed later.

Participants who completed Tasks B (multiple choice glossaries) and C (sentence production) received feedback regarding the correct definitions for the target words as they are used in the text. This was done orally and students wrote the corrections onto their task sheets. In order to check that the tasks had been completed properly, the participants' worksheets were collected by the teachers and returned at the end of the experimental period. The week after the completion of tasks, the ten target words were incorporated into a quiz of reading comprehension and grammar. This immediate posttest (T1) consisted of matching the target word with an appropriate definition in English. The second, delayed posttest (T2) adopted the self- report and performance elements of the Vocabulary Knowledge Scale (VKS) (Paribakht & Wesche, 1993). Due to the research goal of examining *incidental* vocabulary acquisition, neither the tasks nor tests were announced to students ahead of time. The tests and analyses for the target words will be discussed in the section below.

Quantitative Measures

An electronic version of the short story was submitted to lexical frequency profiling (www.lextutor.ca) to select target words that were both important to the comprehension of the story and 'infrequent' (this is operationalised as 'off-list', that is, words not on the General Service Lists of the most common words in English or on Coxhead's (2000) list of words that occur frequently in academic reading). In this way it could be assumed the target words were not already known. Starting with 28 'infrequent' words from the short story, I created a vocabulary awareness pretest following a model used by Horst (2005). This awareness pretest consisted of a checklist of target word items to which the students responded whether or not they *know* the word, or are uncertain.

Indicate **YES** if you are **sure** you know the meaning of the word. **NS** if you have an idea about the meaning, but you are **not sure**; **NO** if you **do not know** the word.

	YES	NS	NO
Placid	X		
Swaying		X	
Bewildered		X	

Figure 3. Example of vocabulary awareness pretest

This vocabulary awareness test was administered in the first week of the course immediately after the VLT, and took participants about ten minutes to complete. Frequencies of responses for each vocabulary item for the vocabulary awareness pretest were generated: the ten target words selected had at least 60% responses "no" or "not sure" ensuring that they were not known by a large majority of the students prior to the treatments. Responses to the awareness tests were analysed, the reading text was assigned in the subsequent class (week 2) and the following week (week 3) participants completed the treatment tasks outlined above.

The following week (week 4), the immediate posttest (T1) was administered; two weeks later, in week 6, students completed the delayed posttest (T2).

Immediate Posttest (T1)

This test consisted of matching the target word with the appropriate English definition. There were 11 definitions provided and 10 target words. The test was included in a weekly quiz testing homework grammar and reading comprehension. Participants did not know in advance that they would be tested on the words from the previous week's task as the researcher's goal was to examine incidental vocabulary learning. The complete test took about 20 minutes to complete. The test can be seen in Appendix C.

Delayed Posttest (T2)

The subsequent delayed test (T2) took place two weeks (week 6) after the immediate posttest (T1); in other words it was administered three weeks after the in-class completion of tasks. It followed the format of self-report and performance of the VKS (Paribakht & Wesche, 1993). Thus it evaluates the quality of knowledge for a particular word by combining receptive (i.e. recognizing a word) and productive (i.e. writing an appropriate definition and using the word in a semantically and grammatically accurate sentence) abilities. The goal was to identify whether students exposed to the more involved tasks would perform better in terms of retention and depth of knowledge of target words.

Figures 4 and 5 outline the adaptation of the VKS scale utilised along with the categories for scoring. This streamlined version was created to make the testing less time consuming and maximise the response rate. The original VKS categories I and II (which corresponded to "I don't remember having seen this word before" and "I have seen this word, but I don't know what it means") were condensed into "I don't know what this word means" because having come across words in the reading text, all the participants had definitely seen the words recently in the awareness pretest. I also wanted to push them to attempt to give a definition.

I	I don't know what this word means.
II	I am not sure. I think it means _____ (give the meaning in English, French or your language)
III	I know this word. It means_____ (translation or synonym) and I can use this word in a sentence: _____ (write a sentence).

Figure 4. Adaptation of the VKS scale (Paribakht & Wesche, 1993) for posttest 2 (T2).

Self-report categories	Possible scores	Meaning of scores
I	1	The meaning of the word is not known
II, III	2	The word is familiar, but an incorrect meaning is given
III	3	A correct meaning (synonym or translation) is given only
III	4	The word is correctly defined and used with semantic appropriateness in a sentence
III	5	The word is correctly defined and used with both semantic appropriateness and grammatical accuracy in a sentence

Figure 5. VKS scoring categories and the meaning of scores posttest 2 (T2).

The ten target words were randomly presented and test-takers indicated their self-perceived knowledge of each item. The researcher then evaluated the level of their knowledge based on their performance in each category. For example, a test-taker may indicate III: "I know this word", but then give an incorrect definition, the score will then be 2. Due to practical constraints, I did most of the VKS ratings; however, there was a

degree of triangulation of assessment as in situations where it was difficult to assign a score, I worked with the other teacher involved. She scored a batch and then we discussed 'hard to assign' sentences until we reached agreement. Figure 6 shows examples of some of the responses from the VKS test (T2).

Score	Example
1	The meaning of the word is not known (no attempt made).
2	The word is familiar, but an incorrect meaning is given: (shove) to shovel something (placid) a place that is hit by acid rain (congeal) frozen (bewildered) *être sauvage* (to be wild)
3	A correct meaning (synonym or translation) is given: (congeal) something that passes from liquid to solid (placid) have a calm expression
4	The word is used with semantic appropriateness in a sentence (swaying) She hit him he swaying. (frantic) At job I am frantic.
5	The word is used with semantic appropriateness and grammatical accuracy in a sentence (congeal) There is congealed blood on his head. (shove) I shoved the man under my bed.

Figure 6. Examples of VKS scoring (T2).

Words assigned a score of three or above were considered to have been *acquired.* There was considerable variation in the amount of time participants took to complete the VKS test, which will be discussed in the results section. It was also more complex to score than T1 as the rater needed to determine the range of semantic and grammatical appropriateness accepted for each score. This was difficult in cases such as *He is very*

frantic where a minimal amount of words are used as opposed to more extensive examples, such as *He late he drive frantically*. My feeling is that the person that gave the latter response has more understanding of the word, but technically the former gets a higher score because the sentence is grammatically correct. In such situations, the other teacher and I checked for 'completeness' of the meanings given for each item then tried to agree on a score. In this instance, both *He is very frantic* and *He late he drive frantically* were assigned scores of five because the former was accompanied by the appropriate synonym *anxious* (the participant seemed to prefer minimal language responses!) and the latter despite lacking the third person singular's', which is a very common error for these learners, makes an attempt to elaborate and explain why someone might be *frantic*.

Qualitative Measures

In order to address the research question about the activities learners engage in while they complete two different vocabulary-reading tasks, ten participants from the investigator's classes who had participated in the quantitative study (having completed either Task A, B or C) were invited to participate in the qualitative, introspective aspect of the study. As mentioned, Laufer and Hulsijn's (2001) study of task involvement lacks this crucial component; they do not explore how learners perceived the various tasks, nor do they investigate the kinds of strategies learners use in executing them. Ten participants in this study provided concurrent think-aloud protocols (or CVPs, concurrent verbal protocols), in which they verbalized what they were thinking and doing while doing the task. These participants ranged from two very keen, highly motivated 'good' students to one who was struggling to get through the course and saw his participation as an

opportunity to have one-on-one time with the teacher. However, the participants were selected mainly on the basis of convenience: they had completed in-class assignments and were available to stay for an extra 20 minutes after class. During this time the participants completed vocabulary exercises of differing Involvement Loads related to a short story text that had been assigned to the whole class as reading material and which had been used successfully in the pilot study; the text used was *The Sniper* by Liam O'Flaherty. The vocabulary exercises assigned were categorised as tasks B (multiple choice glossary), and C (dictionary look up; production). The participants each did four or five examples of both types of tasks. The tasks used can be found in Appendix E.

Introspection is the process of individuals observing and reflecting on their thoughts, feelings, motives, reasoning processes, and mental states. It is one of very few data collection methods available for going beyond observable behaviour and attempting to access the underlying mental processes that determine that behaviour (Nunan, 1994; Wesche & Paribakht, 2000). Ericsson and Simon (1993) define concurrent think aloud as the process whereby participants verbalise their thoughts as new information enters; they verbalise thoughts without analysis or decisions. Polio (2007, October) addressing the issues pertaining to the use of CVPs in L2 reading and writing research noted that, in terms of validity, think aloud protocols are subject to questions of *veridicality*- the extent to which the verbal report is a true and complete representation of thinking and of *reactivity*- whether the act of thinking aloud changes the process or outcome of the task being studied. Ericsson and Simon (1993) make the claim that, if done correctly, concurrent think aloud protocols do not interfere with cognitive processes as long as participants are simply reporting contents from short term memory. The use of such

methodology is obviously limited by the participants' ability to observe and report their mental processes accurately, their willingness to do so, or desire to please the researcher (Ericsson & Simon, 1993). Since the validity and reliability of introspection is known to be improved through training participants, there was a practice session immediately before the think-aloud procedure. Following the recommendations for think-aloud protocols outlined by Macaro (2001) in Mackey and Gass (2005), the participants were instructed to explain, in either English or in French (their L1), as best as they could and with as much detail as possible what they are thinking and doing **while** doing the vocabulary reading tasks. The investigators modelled 'think- aloud' with another text and exercise before participants began the tasks. The investigators made a conscious effort to encourage 'talk' by using phrases such as *"What are you doing at the moment? And how did you select that response?"* rather than phrases such as *"That's good"* which encourage closure rather than further introspection. Resisting this *teacherly* tendency to provide encouragement was particularly challenging. The think-aloud protocols were complemented by open ended interviews which questioned participants about their approaches to the tasks and to vocabulary and language learning in general. Five think aloud protocols were conducted by the college's language assistant and the other five by the main investigator either in the classroom or the language laboratory; each think aloud interview took about 20 minutes to complete. The think-aloud protocols were taped, coded and analysed. Mackey and Gass (2005) describe coding as an attempt to organise data "into a manageable, easily understandable, and analyzable base of information" (p.221). In order to analyse the qualitative introspective data, the researcher listened to two of the ten audio-taped protocols closely to create a taxonomy of activities, or sub-

tasks, related to the Involvement Load Hypothesis undertaken in the two task conditions. Examples from this taxonomy include: *Searches for meaning in text: uses context; searches for meaning in English/ French dictionary; substitutes possible synonyms or words in French.*

The resulting checklist was used to analyse the rest of the recorded protocols, and was adjusted as necessary during the process. Frequencies of the various behaviours were tallied and categorised to indicate possible patterns and characteristic strategies relating to reading comprehension vocabulary tasks.

The goal behind this combination of quantitative and qualitative methods is to gain insights into what college level ESL learners think and why they approach tasks in particular ways. This information combined with results regarding the effectiveness of different tasks will enable teachers to approach vocabulary acquisition tasks through reading comprehension more systematically.

CHAPTER 6: RESULTS

Quantitative results

RQ 1: Incidental Vocabulary Acquisition

Answering the first research question pertaining to incidental acquisition regardless of learning condition involved assigning each target word a score based on the extent to which it was known on each of the three word knowledge measures (pretest, posttest 1, posttest 2). These percentages were based on numbers of participants who knew each of the words and were examined to see if the expected patterns of increase occurred. Words for which there was a 'yes' rating (pretest), a correct definition match (T1), or a score on the VKS (T2) were counted as 'known' in this initial comparison. Table 4 shows the percentages for each target word (TW) for two main groups of participants: the control group, who simply read the experimental text, and learners in the experimental conditions, who read the text and completed additional vocabulary tasks (which varied in levels of Involvement Load).

The figures in the columns presenting immediate (T1) and delayed (T2) average scores show a mixed pattern in the control group (see left half of Table 4). On the immediate posttest, knowledge percentages are higher than the pretest percentages for only five of the ten words; two appear to have remained at about the same level, and there appears to be a decline in knowledge of three. A clear case of decline is *bewildered* which 33.3% of the control group participants claimed to 'know' at the pretest, but which was assigned a correct definition by only 6.7% of them at both immediate and delayed posttests. Comparisons must be made with caution though, firstly because the control group is small (n = 15). Also, since the first (awareness) test does not involve

performance it is plausible that participants over-estimated their knowledge. Nonetheless, the percentages give a rough indication of patterns in the data.

Table 4.
Comparison of Control (Reading Only) Group and Experimental Groups' Knowledge of Each of the 10 Target Words (Percentages).

	Reading only control group (n=15)			Experimental groups (n=103)		
Target words	Pretest rated 'Yes'	T1 matched to correct definition	T2 assigned VKS score of 3-5	Pretest rated 'Yes'	T1 matched to correct definition	T2 assigned VKS score of 3-5
stride	6.7	13.3	0.0	4.9	61.2	30.1
blissful	33.5	33.3	33.3	4.9	47.6	20.4
frowning	6.7	33.3	20.0	9.7	32.0	15.5
congeal	13.5	13.3	33.3	16.5	87.4	35.0
bewildered	33.3	6.7	6.7	9.7	36.9	10.7
swaying	13.3	26.7	13.3	14.6	56.3	35.0
frantic	13.3	40.0	26.7	14.6	42.7	19.4
placid	33.3	40.0	33.3	28.2	52.4	40.8
blunt	26.7	20.0	6.7	25.2	52.4	16.5
shove	53.3	0.0	0.0	34.0	50.5	18.4

The consistent pattern of increases from pretest to immediate posttest in the experimental groups (see right half of Table 4) suggest that some amount of Involvement Load, or focus on a vocabulary task (regardless of type), increases the likelihood of learning. Percentages in the rightmost column of Table 4 point to long term retention, but at lower levels than at the immediate posttest. For example, for the first target word in the list *stride*, only 6.7% of the control group and 4.9% of participants in the experimental (task) groups reported that 'yes' they knew the meaning at pretesting. In the immediate

posttest (T1) 13.3% of the control group participants and 61.2% of the experimental group participants matched the correct definition to the target word; at delayed testing (T2) none of the control group participants and 30.1% received a VKS score of three (correct definition) and above.

The averages for short term retention are considerably higher than those for long term retention; however in interpreting the patterns in Table 4, it is important to bear in mind that the test measures are not the same: the first posttest (T1) involved recognition (matching the target words with definitions), the delayed VKS posttest (T2) involved production (recalling definitions and, if possible, writing sentence examples). This goes some way to explain why an item like *shove* (last on the list in Table 4) was known to none of the control group participants at either T1 or T2 despite having been rated 'yes' by 53.3% of the control group at the pretest. For the same item, whilst 50.5% of the experimental groups matched the TW to a correct definition at the immediate posttest (T1), only 18.4% of these participants produced a correct definition at the delayed test (T2).

The overall pattern in the data for the control group provides weak support to confirm the first hypothesis: there appears to have been some incidental acquisition of new vocabulary through reading a short story with a reading comprehension focus, but gains were not consistent. The consistent pattern of increase in the experimental groups taken together suggests that for both short term and longer term retention, some amount of lexical focus, or Involvement Load, is necessary. However, with only 15 participants in the control group, the conclusions that can be drawn are limited. Further studies of this

question should compare a larger number of 'reading only' participants to experimental groups.

RQ 2: Task and Short and Long Term Retention

In order to answer the second research question pertaining to task effect and immediate and delayed retention, analysis of variance (ANOVA) procedures were conducted. The mean retention scores and standard deviations of the participant scores (N=118) for the pretest and both the immediate (T1) and delayed (T2) posttests in the different learning conditions were calculated. Table 5 shows these results and the results are represented graphically in Figure 7.

Table 5.
Number of participants, mean retention scores (%), and standard deviations in immediate and delayed posttests.

		Awareness pretest		Immediate (T1)		Delayed (T2)	
	Participants (N= 118)	Mean (%)	SD	Mean (%)	SD	Mean (%)	SD
Control: Reading only	15	30.67	17.41	28.66	14.57	31.33	8.34
Task A: Glossary	35	24.71	18.75	47.42	25.36	33.77	9.75
Task B: Multiple choice	36	30.28	19.12	58.61*	27.37	41.16*	16.30
Task C: Definitions	32	32.50	23.04	49.38*	30.47	37.12	14.94

* Statistically significant contrasts to control group mean, Tukey multiple comparisons test, $p < 0.05$

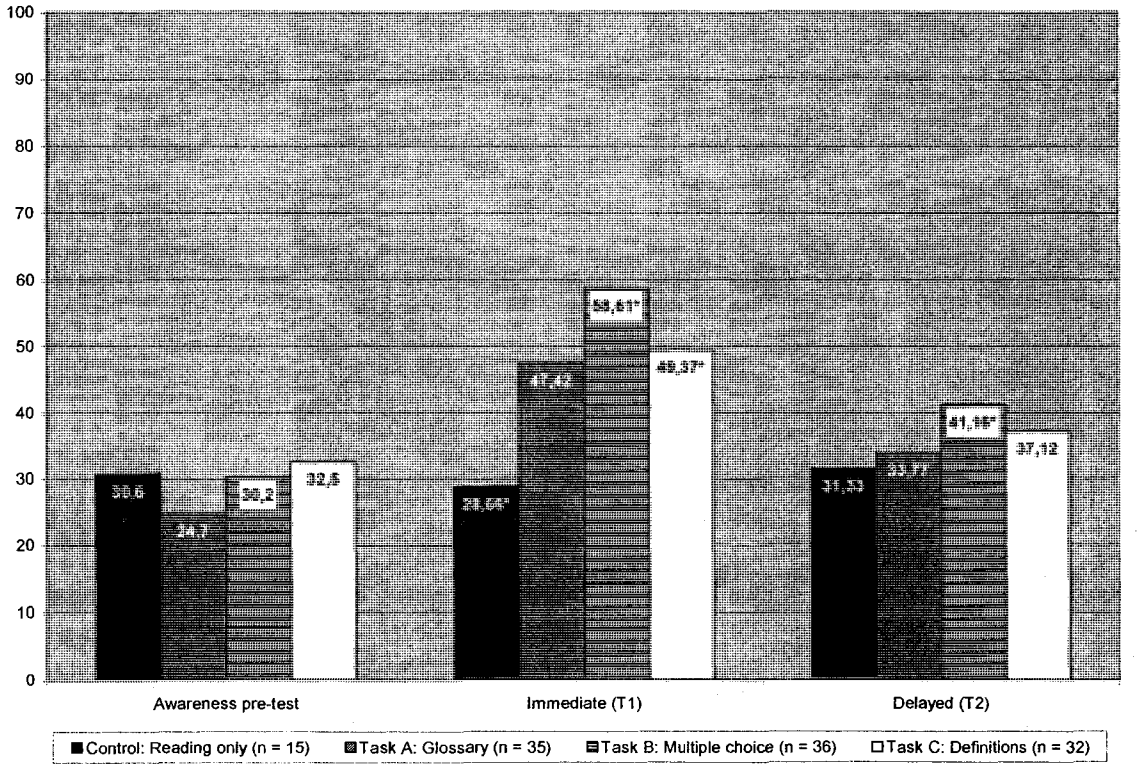

Figure 7. Mean retention test scores shown according to task

As Figure 7 shows, the means for all participants except those in the control group are higher at the immediate posttest (T1) than at the pretest. Contrary to the expected pattern of least to most weighted Involvement Load being mirrored in the distribution of participants' scores, the means for Task B (multiple choice glossary) are higher than those for Task C (dictionary look up and sentence production) at both T1 and T2.

The scores were submitted to analysis of variance (ANOVA) with task as the between subject factor (gloss, multiple choice gloss, dictionary/sentence production, control). The results are shown in Table 6. A Tukey HSD posthoc multiple comparisons test was selected as an option in the ANOVA. At the target word awareness pretest, there

are no statistically significant differences between participant groups $F(3,117) = 0.94$, $p = 0.42$

The ANOVA for the immediate posttest (T1) indicated a statistically significant task effect $F(3,117) = 4.58$, $p = .005$. The post hoc Tukey HSD tests revealed that at the immediate posttest (T1) the mean scores of both the Task B and Task C groups differed significantly from the control (reading only) group, but that the means of the tasks with different Involvement Load levels (Tasks A, B, and C) did not significantly differ from each other. Since the mean scores of both Task B and Task C groups test as significantly different from the mean scores of the control group, this confirms the hypothesis that a task with a higher Involvement Load will result in higher short term retention. The tasks with more involvement are associated with higher levels of short term retention than the control group's task of reading only accompanied by comprehension questions. It can be claimed that the second hypothesis is partly confirmed: tasks that engage learners in 'moderate' or 'high' need, search and evaluation activities using either multiple choice glossary or dictionary sentence production result in higher levels of short term retention than the minimal task of simply reading the text for comprehension (control group). Table 6 shows the ANOVA output for differences between the tasks for both immediate and delayed tests.

Table 6.
One-way ANOVA Testing the Difference in Means Between Tasks

	Source	*SS*	*df*	*MS*	*F*	*p*
Immediate (T1)	Between	9618.01	3	3206.00	4.58	.005
	Error	79859.96	114	700.53		
Delayed (T2)	Between	1455.59	3	485.20	2.71	.048
	Error	202422.00	114	179.14		

In terms of long term retention, the average scores at T2 are lower than T1 mean scores in all of the experimental groups, as might be expected. As can be seen in Table 6, there is a statistically significant difference in the data between the means, $F(3,117) = 2.71$, $p = .048$. Again, a Tukey HSD posthoc multiple comparisons test was selected as an option in the ANOVA. The post-hoc Tukey test involving multiple comparisons indicated that there was a significant difference once again between the control group and Task B means. No other significant differences were identified. As the test measures utilised were not the same, it is not possible to compare the differences over time; however, the same pattern seen in the T1 results is clearly present. The students who completed the multiple choice glossary task (Task B) have a higher mean score than those in the control group who completed only the reading with accompanying comprehension exercises. Although the difference is not significant, the Task B mean is again higher than that of Task C where Involvement Load was the highest. Overall, participants who did one of the two tasks with higher Involvement Loads (B or C) obtained higher averages than those with the least involved task (Task A) and the reading only control group. This pattern suggests that Involvement Load may be a relevant factor in longer term retention of

lexical items. But contrary to expectation, Task B, which was not the most heavily weighted task in the Involvement Load scheme, proved to be the most effective in promoting both short and long term retention of vocabulary.

Qualitative Results

RQ 3: Learners' Activities during Tasks

The third hypothesis regarding the qualitative aspect of the study assumed that oral think aloud protocols would reveal dictionary look-up sentence production tasks to engage learners in a greater number and variety of cognitively demanding activities than the multiple choice gloss task. The protocols were expected to show the great variety of elements involved in learners' approaches to the tasks. The oral think aloud also worked as a spring board for conversation during the open-ended interview in which participants discussed the vocabulary learning tasks. Based on the constructs of the Involvement Load Hypothesis a checklist of activities undertaken in the two task conditions was created as outlined in the methodology section. Frequencies of the various behaviours were tallied and categorised to indicate possible patterns and characteristic strategies relating to reading comprehension vocabulary tasks. Although no clear patterns emerged, both tasks elicited a variety of useful strategies. These can be categorised according to participants' references to the main components of the Involvement Load: *need*, *search* and *evaluation* as shown in Table 7.

Table 7.
Categories of Strategies Revealed in Think Aloud Protocols.

Category	Number of instances	Strategies	Examples
Search: Re-reads definitions and distracters	9	- puzzle solves- eliminates words it can not be - repeats sentence out-loud; substitutes possible definitions - uses statement or question intonation	"The other choices don't make sense…looking at the sentence".
Search: Looks for meaning in sentence context	6	- checks sentence context - needs the context because words have many meanings	"I read the text, I look at the words before I read the text...each time I see a word I look in the text to see what the meaning is and if I can't find it I look it up in a dictionary, but I do them one at a time as I see them…I finish the paragraph. "
Search: Dictionary look up a. French b. English	 a.) 6 b.) 1	- Order: uses French –English dictionary then context OR vice versa. - Systematic approach	"I look at the context, and then use a bilingual dictionary. I use an English dictionary if I'm not certain. "
Evaluation: Identifies part of speech (sentence level)	6	- checks part of speech	"*tattered* shawl…it must be an adjective...I don't know what is shawl…I find *tatter*. " "I'm looking in the paragraph I'm looking in the dictionary because in the sentence it is an adjective and I could just erase it." "*Spasmodically* is an adverb with the verb."
Evaluation: Etymology of word; identification of cognates (word level)	7	- breaks down the parts of the words	"*spasm*…it is the same word in French. " "The word is about the same word as in French. I will write the sentence I think of in French then I will switch to English. " "*gutter*…I think it is like *gouttière*. "

Evaluation: Checks 'fit' in context; substitution of possible synonyms	4	- substitutes alternative synonyms	"In the multiple-choice I read aloud and try and fit the other definitions." "In the paragraph *gutter* is an *endroit*...so it is in the street." "the other choices don't make sense..just looking at the sentence. "
Affective: Perception of exercise- Ease	6	- perceives task as easier - expresses preference	"Multiple choice exercise is easier – not easy to learn the words- it is easier to know what the word is with the text, the meaning... so we don't look and we don't search as much as we do for the other task" (Prefers Task B)... "substitute the definitions offered in the context. Sometimes I select the wrong dictionary definition...sometimes I look up the same words again and again. " "Multiple choice tasks are easier- you can be more certain you are right. Dictionary look up is better for learning because you need to find the good definition- you cannot guess. "
Affective: Perception of exercise- Usefulness	4	- believes dictionary look-up promotes longer term learning	"It is not difficult if I look at the dictionary –I prefer the multiple choice because I can figure out the meaning...Don't have choice so have to find the right meaning. To learn *long time* figure it out by looking at the text and searching in the dictionary if you find it by searching you remember more. "
Affective: Confidence; certainty	4	- lacks confidence- uses question intonation to have guesses confirmed - needs dictionary	

Beginning with the cognitive elements *search* and *evaluation*, using the dictionary to search for meanings and writing a definition and an original sentence requires the learner to both search and evaluate to judge for appropriateness; as shown in the top row of the table, students often read the definitions and tried to substitute words provided in the dictionary definitions in the context. In writing their sentences they were pushed to look at the words semantically and grammatically and consider the word as part of a whole system (i.e. to evaluate). Thus to some extent, the results from the think aloud protocols confirmed the expectation that dictionary look-up sentence writing tasks would engage learners in a greater number and variety of cognitively demanding activities than the multiple choice gloss task.

In the multiple choice glossary (Task B) students frequently searched the context for meaning; there were six instances of reading and re-reading the context sentences along with the distracters and nine instances of searching using puzzle-solving strategies of eliminating distracters by figuring out which ones no longer seemed possible. In contrast, the dictionary look-up sentence production tasks prompted students to refer to a French –English dictionary six times and to a monolingual English dictionary once. The students drew less on puzzle solving and more on evaluation strategies which meant that they looked at the part of speech of the target word and its role in the sentence; for example, "*Tattered shawl...it must be an adjective...I don't know what is shawl...I find tatter.* " Seven students used evaluation strategies where they looked at the etymology, broke the word into parts and looked for cognates, for example, "G*utter...I think it is like gouttière'* and *'Spasm* (part of TW *spasmodically*)*...it is the same word in French.*"

Moreover, the retention results that favour Task B should be considered in the light of students' perception of the ease and usefulness of the exercise for their vocabulary acquisition. Remarks such as, *"Sometimes I select the wrong dictionary definition...sometimes I look up the same words again and again,"* and *"Multiple choice tasks are easier - you can be more certain you are right,"* indicate a clear preference for the multiple choice task. The following statements show that learners have insights into the value of cognitive involvement and favour the multiple-choice task for this reason:

"It is not difficult if [I] look at the dictionary –I prefer the multiple choices because I can figure out the meaning."

"...Don't have choice so have to find the right meaning."

Another student verbalised what seemed to be belief in a '2 for the price of 1' strategy inherent in the multiple choice gloss:

"It is more useful to have the choices in English- I learn two ways to say the same thing to better understand the words - I have two choices to use if I want to write a text." This strategy of having "one word matched to another word" may be a clue to explain the higher mean associated with Task B.

Involvement Load also includes an affective component and in light of this, it is interesting that learner observations point to an affective advantage for Task B. There were six mentions that the multiple choice glossary exercise (Task B) was easier and that they could be confident that it would give them a correct answer, or an appropriate synonym. More than half of the participants were comfortable and more used to doing such multiple choice vocabulary exercises; four participants mentioned insecurity related to dictionary look-up exercises. These students showed awareness of the variety of

meanings that a word could have depending on the context, and two indicated that despite thinking that doing more to get at a meaning would translate into better retention, they were often uncertain about having identified the correct definition in dictionary look-up exercises. Although perceptive of the intricacies of language, they were unsure of their abilities to select the correct meanings.

Perhaps most significantly, the think aloud protocols raise further questions regarding the role and weight of the affective-motivational component of the Involvement Load construct. It is possible that the creators have overlooked the importance of learners' familiarity with a task type and the extent to which learners' perceive the task as manageable.

In summary, the quantitative results indicate that for immediate retention (T1) mean scores on the 'moderate' weighted multiple choice glossary task (Task B) and the 'high' weighted dictionary and sentence writing task (Task C) differed significantly from the means in the reading only control group. Tests for delayed retention (T2) followed a similar pattern to that of T1 but only Task B proved to be associated with a long term retention advantage over the control group. The data from the qualitative aspect of the study suggest that in addition to 'search' and 'evaluation' strategies, preference, or familiarity with Task B type exercises may offer an explanation for this result. These results are briefly summarized in Table 8 below and will be discussed in the next chapter.

Table 8.
Summary of Results

	Hypothesis	Findings
1	Some incidental acquisition of vocabulary will occur from reading short stories with a reading comprehension focus regardless of task	• Partially confirmed
2	Tasks with a higher Involvement Load will result in higher levels of both short term and long term retention.	• Partially confirmed • Participants with moderate Involvement Load task (Task B) outperform control group in both T1 and T2 • Significant difference between reading only control and Task B and Task C at T1 • Significant difference between reading only control and Task B at T2
3	Sentence production tasks that include dictionary look-ups will engage learners in a greater number and variety of cognitively demanding activities than those required to complete the multiple choice glossary task.	• Not confirmed • Task B: promotes context searching; puzzle solving • Task C: promotes searching parts of speech; etymology • Relative weight of motivational and affective IL component of *need* requires adjustment: preference, confidence, perceived usefulness need to be taken into account.

CHAPTER 7: DISCUSSION

This study set out to investigate how levels of Involvement Load in lexical intervention tasks affect vocabulary acquisition and retention. The results from the quantitative and qualitative data will first be discussed in relation to the research questions and hypotheses guiding the study.

Hypothesis 1: Incidental Acquisition

The focus of the first research question was to assess whether reading a short story text with a focus on reading comprehension, rather than vocabulary, would result in *incidental acquisition* (defined by Hulstijn, 2001, as learning words as a by-product of any activity that is not explicitly geared to lexical learning). The quantitative data go some way towards confirming this hypothesis that some acquisition of vocabulary will occur from reading short stories with a reading comprehension focus. There is some evidence to suggest that even some very infrequent words, with single textual exposure will be picked up incidentally through reading comprehension tasks. This was possibly the case with the target word *swaying* which in the story describes the husband's actions after he has been hit over the head. In the reading control group 13% 'knew' the word at the pretest and in the immediate retention test, 26% matched the word to a correct definition. As the control group was very small (n=15), this is a considerable gain and suggests it would be interesting to assess word type and study incidental acquisition of other 'graphic' words in reading comprehension texts.

Hypothesis 2: Involvement Load and Retention

The focus of the second research question was to measure whether reading comprehension vocabulary tasks triggering higher levels of *need*, *search* and *evaluation*

would be more effective for vocabulary acquisition and retention of the ten unfamiliar target words. Patterns in word knowledge scores in reading only and reading-plus-task groups suggested that motivationally and cognitively demanding tasks enhance word learning and retention. This was substantiated in a closer analysis of immediate and delayed retention scores. At T1 (immediate posttest) mean performance on the multiple choice glossary task (Task B), which had an Involvement Load of 3, and the dictionary with sentence production task (Task C), with an Involvement Load of 5, were found to differ significantly from the mean scores in the reading only control group. At T2 (delayed posttest) the data followed a similar pattern to that of T1 with statistically significant difference in mean scores between Task B and the control group.

Contrary to expectations, the group mean for the task that had the greatest Involvement Load, dictionary look-up and sentence production with an Involvement Load of 5 (Task C), was not significantly higher than other group means on the measures of long term retention. But, since the multiple choice glossary task (Task B) had a greater Involvement Load value than the control condition, the T2 results provide limited support for the idea that weight of involvement load affects both short term and longer term retention. These results, complemented with insights drawn from the think aloud protocols, suggest that Involvement Load does plays a role in how well lexical items are retained.

At the outset it had been expected that the weight of Involvement Load for the tasks would be the determining factor for immediate and delayed retention score differences. The study did not find significant differences between the types of experimental task. In the analysis, the tasks with relatively high Involvement weights

(Tasks B and C) differ from the reading-only control group, but not from each other. However, both short and particularly long term results point to a distinct advantage for Task B over Task C.

There are a number of possible explanations why the more demanding dictionary work in Task C did not result in the highest scores. Two considerations are the type of test and the timing of its administration. It had been expected that the VKS (T2) would indicate that those participants with the most involved task would score higher on this depth of knowledge measure; however this does not seem to have been the case. Participants who completed the multiple choice glossary task (Task B) continued to score higher than the other groups. The Task C participants had been obliged to provide dictionary definitions and utilise the words in original sentences; thus the task had provided a basis for productive word knowledge that the other tasks did not favour. My feeling was that they should have been better at providing sentences that showed their understanding of how the words fit into the general schema than students who did other tasks. However, the VKS consists of an 'easy' self report option (to indicate basic knowledge) and a more arduous production task (to indicate 'deeper' knowledge).The astute student who doesn't want to spend an extra ten minutes in the classroom quickly realises that the test is more rapidly completed by reporting only a little knowledge rather than attempting to formulate informative sentences. This was very likely a factor with my group of Task C participants whose class met late in the day, from 3-6pm. It would be interesting to use a different test – possibly a cloze or simple composing type exercise that would tap into this assumed receptive and productive knowledge more effectively.

Hypothesis 3: Think Aloud Protocols

Another question that this study raised was why the multiple choice glossary task (Involvement load 3) led to higher mean retention scores than the dictionary look-up with sentence production task with the highest Involvement Load (Involvement load 5). The think aloud protocols offer insights that may explain this finding. In the think aloud protocols, there was a recurring perception among the participants that the multiple choice glossary task was both easier and more straightforward. Participants felt more confident when they had a correct 'teacher provided' response to learn than when the response was more open-ended and they were led to seek out a variety of different dictionary definitions and diverse ways of expressing a similar meaning. Whilst the search and evaluation involved in the dictionary-production task may enrich the students' lexicons in the long term, the results of this study imply that students at this level want *one word = another word* equivalents. The almost exclusive use of bilingual dictionaries revealed in the think aloud protocols suggests that the participants prefer to define new words with words in their L1; after this they are willing to work with synonyms and definitions in the target language providing the task *"gives the right meaning"* rather than pushes them to find what might well prove to be "*the wrong dictionary definition.*"

This is an interesting finding, because as a teacher I see student opportunities to read the dictionary as exposure to more input with the advantage that they may 'pick-up' or understand other words through reading the definitions and using the words in a sentence. Teachers tend to see consulting a dictionary as a learning experience. The participants in the qualitative aspect of this study seemed to have understood a belief held by many teachers that, "the more you have to do to get at the meaning, the better the

chance that you'll learn the word" along with the notion that they need to be 'pushed' to learn the words. However the data indicate that dictionary consultation may not be as beneficial for students at this level, particularly with words of low frequency where there is only one instance of textual exposure. As can be seen in Table 5, in the experimental treatment groups, the mean retention scores for Task B (multiple choice glossary) are considerably higher than the mean scores for both Task A (glossary given) and Task C (dictionary look-up and sentence production) at the immediate and delayed posttests. Thus it appears that the more constrained, but moderately demanding task (Task B) is the most appropriate pedagogical solution for learners in this context.

The findings that favour Task B support other studies in this area. For instance, Rott (2005) showed that reading accompanied by answering multiple choice glosses (MCG) facilitated stronger, more developed form meaning connections than single translation glosses (STG). This was related to the learners' involvement in the task. The target words occurred four times in the reading text and think-aloud protocols showed how repeated search and assessment of individual word meanings resulted in a strong sense of the appropriate form meaning connection in four different contexts.

Some further evidence of the possible benefit of multiple choice glossaries comes from Martinez (2007, October) who presented research that tested the Involvement Load Hypothesis with six target words in three conditions: single gloss, fill-in-the-blanks, and multiple choice glossary. Like Rott (2005) this study found that longer term retention was higher for MCG than single gloss and think aloud protocols revealed inference and hypothesis testing to contribute positively to retention, suggesting the importance of a certain degree of engagement, or involvement, with the material. However, contrary to

Hulstijn (1992) and Rott (2005), she found little evidence for the effectiveness of the most involvement weighted condition (multiple choice glosses) compared to the moderately involvement weighted condition of fill-in-the-blanks. However, one of the limitations of the Martinez study is that participants were not given feedback regarding correct responses in the MCG, so it is feasible that they had difficulty creating form-meaning connections if they remained uncertain whether they were learning the 'right' meanings.

The experiment reported here makes two interesting contributions to the investigation of the "involvement index" proposed by Laufer and Hulstijn (2001). Although they assign numerical weights in which the absence of a factor equals 0, a moderate presence equals 1, and a strong presence of a factor equals 2, Laufer and Hulstijn (2001) acknowledge that the three involvement factors (need, search, and evaluation) may not share equal weight in determining task effectiveness with regard to learning. My experiment shows that this is indeed an issue. The data from this study suggest that the *need* factor should be considered in its widest sense; it is important that elements such as the student's perception of ease, utility, and confidence be factored into the calculation of *need* in a task's Involvement Load. Future studies of Involvement Load would do well to find ways to assign numerical values to these task factors.

Furthermore, previous research into the Involvement Load Hypothesis has focused mainly on the relative weights of evaluation and search. A strength of this study was the inclusion of think aloud protocols. These revealed the many interesting ways in which students actually tackle vocabulary assignments. The study provides evidence of the importance of both evaluation and search strategies. In addition, for participants who

provided think aloud protocols, it was clear that metalinguistic knowledge of parts of speech and etymology enabled students to meet the challenges rather than become confused by the tasks. Further research in this field should continue to include think aloud protocols in order to identify key attitudes and strategies that facilitate vocabulary learning.

CHAPTER 8: LIMITATIONS

Over the course of conducting this study a number of limitations became apparent. These related to time with the reading text, administration of the tasks, difficulty of target words and the practicality of research in real classroom settings.

Time with Text

Firstly, although the difference of time with each experimental task was taken into consideration and considered to be minimal, one of the limitations of the study is that the difference between the mean retention scores for the control group and those for all the experimental groups could have been related to the amount of time participants spent working with the text. In doing the focused vocabulary tasks, participants in all the experimental groups spent more time with the short story text than the reading only control group, who simply answered comprehension questions and participated in class discussions.

Task Administration

Secondly, another element to consider is how the tasks were completed. The results from the qualitative aspect of the study suggest that the amount of contextual checking is in fact quite limited; some participants may have done the tasks almost as lexical assignments independent from reading and understanding the text. All the participants had to find the target word context and highlight the sentence, but they may simply have scanned the text to recognise the word. They were not required to say the word aloud or to show that they could see how it 'fit' in the sentence; the tasks were administered to the whole class and the way in which each individual worked with the reading text and target words was not controlled. In relation to this, when administering

the tasks I noticed that the majority of participants doing the multiple choice glossary task (Task B) in the main study worked individually. Although there were no specific instructions to work silently, the task seemed to lend itself more to silent individual work. In contrast and perhaps because so many students had come to class without dictionaries, the dictionary look-up and sentence production tasks produced more collaborations and exchanges of ideas. Students certainly addressed elements other than the definitions of target words in completing the tasks; for example, they incorporated work on subject verb agreement and verb tenses in their sentence production. In the Task C class that I taught I noted students working in teams. One worked with the dictionary while the other looked for the context and they then 'created' the sentences together. This collaborative aspect, whilst enabling students to practice diverse elements of English may in actual fact have reduced the degree of Involvement Load of the task. In a controlled laboratory experiment it would be feasible to insist on silent individual work for all tasks; however in a real classroom setting the results reflect the real, and unpredictable, nature of classroom interaction.

Target Word Difficulty

Thirdly, as mentioned earlier in the discussion, the study also revealed that there were some target words, such as *placid, swaying, congeal* (refer to Table 4) with higher mean retention scores than others at the immediate and particularly at delayed posttesting. This suggests that despite the pretesting, the words are not 'equal' – *placid* and *congeal* have similar French cognates (*placide* and *congeler*). In order to give a clearer picture of 'new' word learning, future investigations of Involvement Load would do well to select words that are not cognates. The word *swaying* evoked a very visual,

essential image in the short story (the husband sways after being hit on the head by the frozen leg of lamb) which may explain why performance on this item is higher. The selection of particular target words greatly affects a study; this underlines the importance of considering parts of speech, word class (abstract or concrete) along with their importance to the meaning of the story.

Real Classroom Research

Finally, one of the methodological difficulties in the study was the use of real, intact classes and different instruments for testing. As the study was exploring "incidental acquisition", the tests were not announced ahead of time. This is consistent with Hulstijn's (2001) contrast of an incidental learning situation to one that is intentional; in incidental learning, individuals process new information without the intention to commit this information to memory. Thus incidental vocabulary learning is the learning of words as a by-product of any activity that is not explicitly geared to lexical learning. In this study, in the absence of prior warning about upcoming tests, students made no effort to be present on the days when test measures were administered. As a result the number of participants was significantly reduced from an initial 167 to 118 due to absences for one or more of the tests. Had all 167 been present at all of the data collection points, differences between the task conditions might have been clearer. The results would also have been "tidier" had the same test been administered to the classes at three time intervals rather than utilising three different test measures. However, this would probably have created a test effect, with participants very likely becoming aware of a forthcoming vocabulary test. Despite the limitations outlined above, this last factor was avoided by the current study's design.

CHAPTER 9: IMPLICATIONS AND CONCLUSION

Many readers might find the statement, *"The more work you do, the more you will learn,"* to be ridiculously self-evident. While I was conducting this study it became quite the amusement among colleagues who summed up my research as a project to investigate whether our students would indeed learn if the teachers made them work. It would have been a nice result to be able to report that YES-- the higher the Involvement Load the higher the amount of language acquisition. However, this study has underlined that it is not such a straightforward matter.

Nonetheless, the results of this study have at least four clear implications for teaching practice.

Efficiency of Multiple Choice Glossary

First, multiple choice glossary (Task B) type activities appear to be effective and in some contexts more efficient than dictionary consultation and sentence production (Task C) type activities. Teachers can prepare focused vocabulary worksheets with multiple choice glossary activities that accompany reading texts. With feedback on the correct answers given by the teacher in class, these provide clear synonyms or definitions with which learners feel confident. The think aloud protocols from this study suggest, firstly, that learners at this level are familiar or comfortable doing this type of task and are in some ways building word maps or networks of language by seeing how some words, as opposed to others, work as synonyms in different contexts. All of this, however, is reliant on prompt and reliable feedback from the teacher to ensure that the "word map" theories being created are useful and accurate. The study gives reason to

question the efficiency of more time consuming dictionary work or output tasks such as sentence writing in classroom settings where time is a major constraint.

Word Frequency and Task

Secondly, teachers should choose words judiciously when they design Task B activities because not all words benefit from this type of treatment. For instance the cognates found in this reading text proved to be fairly easy to learn for the participants of this study. Teachers need to make decisions about the relative importance and frequency of a word in making decisions about tasks. In this study for example, in T1 immediate retention tests, even the lowest Involvement Load weighted task (Task A: glossary provided) made greater retention gains than the reading only control group, the mean score for Task A was 47% and whereas for the Control group's was 28%. This could be interpreted as support for increasing exposure to readings and providing accompanying glossaries; however, returning to both the literature review and the results in this study, I would argue that 'depth of processing' or Involvement Load will promote the stronger word map connections that are conducive to longer term retention.

Furthermore, more frequent words will probably benefit from more exposure (input) so perhaps less Involvement Load is necessary, whereas less frequent words, such as those found in academic English (the Academic Word List) or specialised lexicons used in courses of English for specific purposes, would benefit from tasks with higher Involvement Loads as repeated exposures to the words will be less available. Hulstijn and Laufer (2001) discuss whether it is the task type or the number of times that a learner is exposed to an item that is most important in vocabulary learning. Hulstijn (2003) argues that empirical studies showing low retention rates in incidental vocabulary learning

highlight the need for repeated exposures to make a sufficiently deep imprint in memory, so that the word is available for retrieval in the long term. In this study, the low scores for long term retention (T2), regardless of levels of task involvement load, tend to suggest that a one off exposure to the word is not sufficient for long term retention.

Task Trade-Off

Thirdly, teachers should be aware of the trade-off that comes with Task B. With moderate Involvement Load Type B activities, there appears to be a gain in efficient learning, but at the possible expense of collaborative interaction. It would seem that the Involvement Load of a task is reduced if learners collaborate and share the components of the task; however, with interaction comes real language use in the form of organising the task, negotiating meanings, and arriving at appropriate usages. This may also contribute to leaving an imprint of words in learners' memories and facilitating longer term retention in ways this study was not able to assess.

Quality and Quantity

The fourth implication that becomes clear is that working with the word once is simply not enough, whatever the activity. In this study, the low scores for long term retention (T2), regardless of levels of task involvement load, tend to suggest that a one off exposure to the word is not sufficient for long term retention. Therefore, teachers should as much as possible build in repetitions for exposure to the targeted words. For example, they can do this by focusing on the same set of words in both reading and listening activities or by designing productive activities where students are pushed to incorporate new vocabulary.

Finally, students must be encouraged to read in *quantity* as exposure to words must be repeated to permit learners to construct, test out, and refine theories of form-meaning relationships. Assigned texts must be of an appropriate level, with the percentage of unknown words appropriate to the proportion of known items. Reading must be interesting and challenging; learners must not be blocked from fluent textual understanding because they are stuck at the word level, checking each word in a dictionary or gloss. However, the teacher cannot do everything; students also need to take some responsibility: in order to learn they need motivation to read *more* and apply search and evaluation strategies to work at developing their word maps or vocabulary knowledge.

In conclusion, vocabulary acquisition tasks are subject to the challenge of finding the right balance. For teachers working within the constraints of classrooms, curricula and fixed amounts of time with students, the question really becomes one of how much attention it is possible to give to new words within these constraints. In the quest for effective vocabulary learning tasks, I believe it is important to consider language as a whole, and not isolate words from their natural settings too much. The focus should remain on using the words to comprehend text and to communicate; as Rott (2005, p.4) emphasises, "...word interventions should not only promote word learning, but also text comprehension, or at least they should not interfere with the comprehension process." In this study think aloud participants believed in the utility of Task B, they felt confident that they were getting something 'right' to help them to understand the reading text and to build their vocabulary. It would seem that part of the key to L2 reading comprehension

and vocabulary acquisition is to offer students a mix of motivating, interesting, challenging texts of appropriate levels and exercises such as multiple choice glossaries with moderate to high degrees of task processing involvement. This helps ensure that learners enjoy reading and also engage in learning tasks that are suited to their individual needs.

Lexical intervention tasks, such as the multiple choice glossary, that trigger *need*, *search*, and *evaluation* activities promote the creation of networks of meanings as learners proceed in understanding words and fitting them into their ever evolving systems of lexical knowledge. Such tasks may also be associated with better reading comprehension if readers are involved in more attentive focused reading. Assessment of both reading comprehension and vocabulary task Involvement Load was beyond the scope of this research; however, this study contributes to the understanding of vocabulary acquisition and retention by combining insights into task effectiveness from both quantitative and qualitative perspectives. Future research should examine the role of different learner strategies mentioned here in relation to particular vocabulary tasks and consider the relative efficiency of receptive (input) versus productive (output) tasks as well as the balance of individual versus collaborative work in vocabulary building tasks.

This study discovered that the multiple choice glossary task with a moderate Involvement Load was the most effective for longer term retention of words with these learners in this particular context. Future research should address the questions raised here regarding the effectiveness of particular tasks in specific learning contexts.

REFERENCES

Coxhead, A. (2000). A new academic word list. *TESOL Quarterly, 34*, 213-238.

Craik, F., & Lockhart, R. (1972). Levels of processing: a framework for memory research. *Journal of Verbal Learning and Verbal Behaviour, 11*, 671-684.

Day, R., & Bamford, J. (1998). *Extensive reading in the second language classroom*. Cambridge: Cambridge University Press.

Dörnyei, Z. (1990). Conceptualizing motivation in foreign language learning. *Language Learning, (40)*1, 45-78.

Doughty, C., & Williams, J. (1998). *Focus on form in classroom second language acquisition*. Cambridge: Cambridge University Press.

Ellis, N.C. (1994). *Implicit and explicit learning of languages*. London: Academic Press.

Ellis, N.C. (1997). In N. Schmitt & M. McCarthy (Eds.), *Vocabulary: Description, acquisition and pedagogy.* Cambridge: Cambridge University Press.

Ellis, R. (Ed.). (2001). Form-focused instruction and second language learning. *Language Learning, 51.*

Ellis, R., Basturkmen, H., & Loewen, S. (2002). Doing focus on form. *System, 30*, 419-432.

Ericsson, K.A., & Simon, H.A. (1993). *Protocol Analysis: verbal reports as data.* Cambridge, USA: Massachusetts Institute of Technology Press.

Folse, K.S. (2004). *Vocabulary myths: Applying second language research to classroom teaching*. Michigan, USA: University of Michigan Press.

Gardner, R.C., & Lambert, W.E. (1959). Motivational variables in second language acquisition. *Canadian Journal of Psychology, 13*, 266-272.

Goulden, R., Nation, P., & Read, J. (1990). How large can a receptive vocabulary be? *Applied Linguistics, 11, (4),* 341-363.

Grabe, W., & Stoller, F.L. (2002). *Teaching and researching reading*. London: Pearson Education.

Guthrie, J.T., Wigfield, A., & Von Secker, C. (2000). Effects of integrated instruction on motivation and strategy use in reading. *Journal of Educational Psychology, 92*, 331-341.

Guthrie, J.T. (2001). Contexts for engagement and motivation in reading. *Reading Online, 4(8)*. Available at http://www.readingonline.org/articles/art_index.asp?HREF=/articles/handbook/guthrie/index.html

Hirsh, D., & Nation, P. (1992). What vocabulary size is needed to read unsimplified texts for pleasure? *Reading in a Foreign Language, 8* (2), 689-696.

Horst, M. (2005). Learning L2 vocabulary through extensive reading: A measurement study. *Canadian Modern Language Review, 61*, 355-382.

Horst, M. (in press). Developing definitional vocabulary knowledge and lexical access speed through extensive reading. In Z. H. Han & N. J. Anderson (Eds.), *Learning to read and reading to learn: On the dual process of L2 reading instruction (working title).* Alexandria VA: TESOL.

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

Hulstijn, J.H. (1992). Retention of inferred and given word meanings: experiments in incidental vocabulary learning. In: P.J.L. Arnaud & H. Béjoint (Eds.), *Vocabulary and Applied Linguistics*. (pp. 113-125). London: Macmillan.

Hulstijn, J.H. (1993). When do foreign-language readers look up the meaning of unfamiliar words? The influence of task and learner variables. *The Modern Language Journal, 77(ii)*, 139-147.

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

Hulstijn, J.H. (2003). Incidental and intentional learning. In C. Doughty & M.H. Long (Eds.), *Handbook of second language acquisition.* Malden, MA: Blackwell.

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

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

Hulstijn, J. H., & Trompetter, P. (1998). Incidental learning of second language vocabulary in computer-assisted reading and writing tasks. In D. Albrechtsen, B. Hendriksen, I. Mees, & E. Poulsen (Eds.), Perspectives on foreign and second language pedagogy (pp. 191-200). Odense, Denmark: Odense University Press.

Krashen, S. (1989). We acquire vocabulary and spelling by reading: additional evidence for the input hypothesis. *The Modern Language Journal, 73*, 440-464.

Laufer, B. (1997). What's in a word that makes it hard or easy? Intralexical factors affecting the difficulty of vocabulary acquisition. In N. Schmitt & M. McCarthy (Eds.), *Vocabulary: description, acquisition and pedagogy.* (pp. 140-155). Cambridge: Cambridge University Press.

Laufer, B. (2005). Instructed second language learning: the fault with the 'default hypothesis.' In A. Housen & M. Pierrard (Eds.), *Investigations in instructed second language acquisition* (pp. 311-329). Berlin: Mouton de Gruyter.

Laufer, B. (2006). Comparing focus on form and focus on formS in second language vocabulary learning. *The Canadian Modern Language Review, 63*, 149-166.

Laufer, B., & Hill, M. (2000). What lexical information do L2 learners select in a call dictionary and how does it affect word retention? *Language Learning & Technology, Vol. 3*(2), 58-76.

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

Laufer, B., & Yano, Y. (2001). Understanding unfamiliar words in a text: do L2 learners understand how much they don't understand? *Reading in a Foreign Language 13*, 549-566.

Laufer, B., Elder, C., Hill, K., & Congdon, P. (2004). Size and strength: do we need both to measure vocabulary knowledge? *Language Testing, 21*(2), 202-226.

Liu, N., & Nation, P. (1985). Factors affecting guessing vocabulary in context. *RELC Journal 16*(1), 33-42.

Long, M., & Robinson, P. (1998). Focus on form: Theory, research and practice in C. Doughty & J. Williams (Eds.), *Focus on form in classroom second language acquisition* (pp.15-41). Cambridge: Cambridge University Press.

Mackey, A., & Gass, S. (2005). *Second language learning: Methodology and design.* Mahwah, New Jersey: Lawrence Erlbaum.

Martinez, A. (2007, October). *Depth of processing, type of task and type of item in incidental L2 vocabulary learning*. Paper presented at the annual conference of the Second Language Research Forum. Champaign Urbana, IL.

Mondria, J., & Wit-de-Boer, M. (1991). The effects of contextual richness on the guessability and retention of words in a foreign language. *Applied Linguistics*, *12*(3), 249-267.

Nation, I.S.P. (1990). *Teaching and learning vocabulary*. New York: Newbury House.

Nation, I.S.P., & Waring, R. (1997). In N. Schmitt & M. McCarthy (Eds.), *Vocabulary: description, acquisition and pedagogy.* (pp.6-19). Cambridge: Cambridge University Press.

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

Nunan, D. (1992). *Research methods in language learning*. Cambridge: Cambridge University Press.

Paribakht, T.S., & Wesche, M. (1993). Reading comprehension and second language development in a comprehension-based ESL programme. *TESL Canada Journal, 11,* 9-27.

Paribakht, T.S., & Wesche, M. (1996). Enhancing vocabulary acquisition through reading: A hierarchy of text-related exercise types. *Canadian Modern Language Review, 52* (2), 155-178.

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

Pitts, M., White, H., & Krashen, S. (1989). Acquiring second language vocabulary through reading: A replication of the clockwork orange study using second language acquirers. *Reading in a Foreign Language, 5* (2), 271-275.

Polio, C. (2007, October). *The use of think-aloud protocols for understanding reading and writing processes in a second language*. Paper presented at the monthly meeting of the Concordia University Applied Linguistics and Psychology Research Group, Montreal.

Robinson, P. (2001). *Cognition and second language instruction*. Cambridge: Cambridge University Press.

Rott, S. (2005). Processing glosses: a qualitative exploration of how form-meaning connections are established and strengthened. *Reading in a Foreign Language, Vol. 17,* (2). Retrieved October 28, 2005, from http://nflrc.hawaii.educ/rfl/October2005/rott/rott.html

Saragi,T., Nation, P., & Meister, F. (1978). Vocabulary learning and reading. *System, 6*, 72-78.

Schmidt, R. (1994). Implicit learning and the cognitive unconscious: Of artificial grammars and SLA. In N.C. Ellis (Ed.), *Implicit and explicit learning of languages* (pp.165-210). London: Academic Press.

Schmidt, R. (2001). Attention. In P. Robinson (Ed.), *Cognition and second language instruction.* Cambridge: Cambridge University Press.

Schmitt, N. (2000). *Vocabulary in language teaching.* Cambridge: Cambridge University Press.

Schmitt, N., & McCarthy, M. (1997). *Vocabulary: description, acquisition and pedagogy.* Cambridge: Cambridge University Press.

Sharwood Smith, M. (1981). Consciousness raising and the second language learner. *Applied Linguistics, 2,* 159-168.

Skehan, P. (1998). *A cognitive approach to language learning.* New York: Oxford University Press.

Stahl, A., & Fairbanks, M. (1986). The effects of vocabulary instruction: a model-based meta-analysis. *Review of Educational Research, 56*, 72-110.

Swain, M. (1995). Three functions of output in second language learning. In G. Cook & B. Seidlhofer (Eds.), *Principle & practice in applied linguistics* (pp. 125-144). New York: Oxford University Press.

VanPatten, B. (1996). *Input processing and grammar instruction: Theory and research.* Norwood, NJ: Ablex.

VanPatten, B., & Cadierno, T. (1993). Explicit instruction and input processing. *Studies in Second Language Acquisition, 15*, 225-243.

Waring, R., & Nation, P. (2004). Second language reading and incidental vocabulary learning. *Angles on the English-speaking world*, 4, 11-22.

Wesche, M., & Paribakht, T.S. (2000). Reading-based exercises in second language vocabulary learning: an introspective study. *The Modern Language Journal, 84*, (2), 196-213.

APPENDIX A: Consent form to participate in research

Formulaire de consentement à la participation à une recherche

Par la présente, je déclare consentir à participer à un projet de recherche mené par Mme Vanessa Beal _____ et Marlise Horst marlise@education.concordia.ca du département d'éducation de l'Université Concordia.

A. BUT DE LA RECHERCHE

On m'a informé du but de la recherche, soit d'évaluer la relation entre les capacités de lecture et l'acquisition de l'anglais langue seconde chez les étudiants du niveau collégial.

B. PROCÉDURES

Ce projet de recherche se déroulera au Collège Edouard Montpetit durant le cours 604-101. La participation à cette étude consistera en la lecture de courts textes et la réalisation de courts travaux de compréhension et de vocabulaire faisant partie du cours habituel. L'ensemble des travaux seront intégrés au cours 604-101 et n'entraîneront pas de temps supplémentaires de la part des participants. La participation à cette étude n'aura aucune influence sur les résultats du cours 604-101. Tous les travaux recueillis dans le cadre de cette étude seront divulgués de façon anonyme.

C. CONDITIONS DE PARTICIPATION

La participation à cette étude n'entraîne aucun risque pour le participant. La participation à cette étude est réalisée dans le cadre d'un programme de formation universitaire.

- Je comprends que je peux retirer mon consentement et interrompre ma participation à tout moment, sans conséquences négatives.
- Je comprends que ma participation à cette étude est confidentielle (c.-à-d. que la chercheuse ne révélera pas mon identité)
- Je comprends que les données de cette étude puissent être publiées ou présentées à un colloque scientifique.
- J'atteste qu'on m'a remis une copie du présent formulaire de consentement.

J'AI LU ATTENTIVEMENT CE QUI PRÉCÈDE ET JE COMPRENDS LA NATURE DE L'ENTENTE. JE CONSENS LIBREMENT ET VOLONTAIREMENT À PRTICIPER À CETTE ETUDE.

NOM : (caractères d'imprimerie) : _____

SIGNATURE : _____

DATE : _____ 2007

Si vous avez des questions en ce qui a trait à vos droits en tant que participants à la présente étude, veuillez communiquer avec Adela Reid, Agente d'éthique en recherche/conformité, Université Concordia, au (514) 848-2424, poste 7481, ou par courriel a l'adresse areid@alcor.concordia.ca

APPENDIX B: Sample experimental tasks

Task A

Lamb to the Slaughter, by Roald Dahl

Vocabulary:
Find the context for the following words taken from the story. Copy out the sentence where the word appears and check that you understand its meaning.

sniper (n): hidden shooter; someone who shoots from a concealed position
e.g.: On a rooftop near O'Connell Bridge, a Republican sniper lay watching.

placid (adj): calm in nature or appearance; tending not to become excited or disturbed

gravel (n): small stones for paths or driveway

blissful (adj): perfectly happy, serene (usually because unaware of something)

stride (n / v): long, brisk steps

bewildered (adj): extremely confused, puzzled

frowning (adj): facial expression of displeasure or concentration by wrinkling the brow

swaying (v): swing back and forth; stagger from side to side

shove (v): push something or someone forcefully or in a rude / careless way

frantic (adj): out of control emotionally; hysterical; beside yourself

congeal (v): become thick and solid; cause a liquid to solidify

blunt (adj): not sharp; dull

Task B

Lamb to the Slaughter, by Roald Dahl

Vocabulary:
Which is the correct meaning for the following words taken from the story?
Check the context, underline the BEST response.
Copy out the sentence where the word appears.

sniper (n): concealed child / person on show / hidden shooter
e.g.: On a rooftop near O'Connell Bridge, a Republican sniper lay watching.

placid (adj): calm in nature or appearance / clear understanding / becoming excited

gravel (n): large rocks / small stones for driveway / cement entry

blissful (adj): unhappy / perfectly happy, serene / delectable

stride (n / v): walking on tiptoe / leisurely stroll / brisk steps

bewildered (adj): confused, puzzled / wild behaviour / bemused

frowning (adj): facial expression of displeasure or concentration / expression of contentedness / feeling of freezing

swaying (v): movement up and down / swing back and forth / balancing on one leg

shove (v): grasp or grab / push something or someone forcefully / to do carelessly

frantic (adj): out of control emotionally; hysterical / rapid; fast / disorganised

congeal (v): melt; become liquid / cause a liquid to solidify / freeze

blunt (adj): not sharp; dull / hard texture / pointed end

Task C

Lamb to the Slaughter, by Roald Dahl

Vocabulary:

A) Find the context for the following words taken from the story.
B) Provide a dictionary definition / synonym.
C) Use the word in a sentence to show that you understand its meaning.

e.g. sniper (n): A. *On a rooftop near O'Connell Bridge, a Republican sniper lay watching.*

B. A hidden shooter; someone who shoots from a concealed position
C. In USA, many civilians were killed or injured by a drive-by sniper.

placid (adj):
gravel (n):
blissful (adj):
stride (n / v):
bewildered (adj):
frowning (adj):
swaying (v):
shove (v):
frantic (adj):
congeal (v):
blunt (adj):

Comprehension Questions: (accompanying all tasks)
Answer the following questions with complete sentences and be prepared to discuss the answers in class.

1. What is the setting (time and place it occurred, both general and specific) of the story?
2. What is the mood like at the beginning of the story? Choose some words or phrases that help create the mood. What effect do these words have on the reader?
3. What is the story actually about? What do you think the author wanted to achieve?
4. Who are the main characters in the story? Describe them. Find some quotations that show what they are like. Which characters do you like or dislike?
5. Is there suspense in the story? When does it begin? What adds to the suspense?
6. What did you enjoy most or least about the story?
7. What is the significance of the title?

APPENDIX C: Sample of immediate posttest

Posttest 1: QUIZ

A. Reading Comprehension:
What happened in Lamb to the Slaughter? In five (5) complete sentences using different verbs in the past tense tell the main events of the story in chronological order.

B. Vocabulary: Match the correct definitions with the vocabulary items. (5)

1. placid (adj):
2. blissful (adj):
3. stride (n / v):
4. bewildered (adj):
5. frowning (adj):
6. swaying (v):
7. shove (v):
8. frantic (adj):
9. congeal (v):
10. blunt (adj):

a. small stones for paths or driveway
b. calm in nature or appearance; tending not to become excited or disturbed
c. not sharp; dull
d. extremely confused, puzzled
e. become thick and solid; cause a liquid to solidify
f. long, brisk steps
g. facial expression of displeasure or concentration by wrinkling the brow
h. perfectly happy, serene (usually because unaware of something)
i. out of control emotionally; hysterical; beside yourself
j. push something or someone forcefully or in a rude / careless way
k. swing back and forth; stagger from side to side

C. Grammar: Select the appropriate form of the verb in the simple past or past progressive. (10)
George _____*was working*____ (work) at the front desk yesterday morning and _____ (1. look) forward to a calm and relaxing morning. As he _____ (2. help) the customer who had just arrived, the phone _____ (3. ring). George politely _____ (4. interrupt) the customer and _____ (5. answer) the phone. It _____ (6. be) an angry client who _____ (7. want) to complain about a new computer programme he had bought and which _____ (8. work, NEG).
George asked the caller, "Where _____ (9. buy, you- QUESTION) the programme?" Then he _____ (10. go) back to finish helping the other customer.

APPENDIX D: Sample of delayed posttest

Posttest 2 (example of VKS test)

What do you know about these words? Please circle 1, 2, or 3 and complete.

placid

1. I don't know what this word means.
2. I am not sure. I think it means .....
(Give the meaning in English, French or your language)

3. I know this word. It means ..... and I can use it in a sentence. (Write the sentence)

.....

blissful

1. I don't know what this word means.

2. I am not sure. I think it means .....
(Give the meaning in English, French or your language)

3. I know this word. It means ..... and I can use it in a sentence. (Write the sentence)

.....

APPENDIX E: Sample of think aloud protocol tasks

Think Aloud Protocol Tasks: *The Sniper*

TASK 1:

- **D) Find the context for the following words taken from the story. Copy out the sentence where the word appears.**
- **E) Provide a dictionary definition / synonym.**
- **F) Use the word in a sentence to show that you understand its meaning.**

***e.g. sniper* (n)**: A. *On a rooftop near O'Connell Bridge, a Republican sniper lay watching.*
B. A hidden shooter; someone who shoots from a concealed position
C. In USA, many civilians were killed or injured by a drive-by sniper.

tattered (adj):
lodged (v):
gutter (n):
remorse (n):
gibber (v):
reckless (adj):

TASK 2:
Which is the correct meaning for the following words taken from the story?
Check the context, underline the BEST response.
Copy out the sentence where the word appears.

twilight (n): sun is below the horizon / bright sunshine / midday heat

---

beleaguered (adj): besieged or under attack / lost / involved

---

spasmodically (adv): intelligently / intermittently / paradoxically

---

gleam (n): momentary indication of something / shining surface / brightness

---

parapet (n): protective wall or ramparts / fortress / parachute

---

ooze (v): to drink alcohol / slow, gradual leakage or flow / to spurt out

---

APPENDIX F: Short Story: Lamb to the Slaughter, Roald Dahl

The room was warm and clean, the curtains drawn, the two table lamps alight-hers and the one by the empty chair opposite. On the sideboard behind her, two tall glasses, soda water, whiskey. Fresh ice cubes in the Thermos bucket.

Mary Maloney was waiting for her husband to come him from work.

Now and again she would glance up at the clock, but without anxiety, merely to please herself with the thought that each minute gone by made it nearer the time when he would come. There was a slow smiling air about her, and about everything she did. The drop of a head as she bent over her sewing was curiously tranquil. Her skin -for this was her sixth month with child-had acquired a wonderful translucent quality, the mouth was soft, and the eyes, with their new placid look, seemed larger darker than before. When the clock said ten minutes to five, she began to listen, and a few moments later, punctually as always, she heard the tires on the gravel outside, and the car door slamming, the footsteps passing the window, the key turning in the lock. She laid aside her sewing, stood up, and went forward to kiss him as he came in.

"Hullo darling," she said.

"Hullo darling," he answered.

She took his coat and hung it in the closet. Then she walked over and made the drinks, a strongish one for him, a weak one for herself; and soon she was back again in her chair with the sewing, and he in the other, opposite, holding the tall glass with both hands, rocking it so the ice cubes tinkled against the side.

For her, this was always a blissful time of day. She knew he didn't want to speak much until the first drink was finished, and she, on her side, was content to sit quietly, enjoying his company after the long hours alone in the house. She loved to luxuriate in the presence of this man, and to feel-almost as a sunbather feels the sun-that warm male glow that came out of him to her when they were alone together. She loved him for the way he sat loosely in a chair, for the way he came in a door, or moved slowly across the room with long strides. She loved intent, far look in his eyes when they rested in her, the funny shape of the mouth, and especially the way he remained silent about his tiredness, sitting still with himself until the whiskey had taken some of it away.

"Tired darling?"

"Yes," he said. "I'm tired," And as he spoke, he did an unusual thing. He lifted his glass and drained it in one swallow although there was still half of it, at least half of it left. She wasn't really watching him, but she knew what he had done because she heard the ice cubes falling back against the bottom of the empty glass when he lowered his arm. He paused a moment, leaning forward in the chair, then he got up and went slowly over to fetch himself another.

"I'll get it!" she cried, jumping up.

"Sit down," he said.

When he came back, she noticed that the new drink was dark amber with the quantity of whiskey in it.

"Darling, shall I get your slippers?"

"No."

She watched him as he began to sip the dark yellow drink, and she could see little oily swirls in the liquid because it was so strong.

"I think it's a shame," she said, "that when a policeman gets to be as senior as you, they keep him walking about on his feet all day long."

He didn't answer, so she bent her head again and went on with her sewing; bet each time he lifted the drink to his lips, she heard the ice cubes clinking against the side of the glass.

"Darling," she said. "Would you like me to get you some cheese? I haven't made any supper because it's Thursday."

"No," he said.

"If you're too tired to eat out," she went on, "it's still not too late. There's plenty of meat and stuff in the freezer, and you can have it right here and not even move out of the chair."

Her eyes waited on him for an answer, a smile, a little nod, but he made no sign.

"Anyway," she went on, "I'll get you some cheese and crackers first."

"I don't want it," he said.

She moved uneasily in her chair, the large eyes still watching his face. "But you must eat! I'll fix it anyway, and then you can have it or not, as you like."

She stood up and placed her sewing on the table by the lamp.

"Sit down," he said. "Just for a minute, sit down."

It wasn't till then that she began to get frightened.

"Go on," he said. "Sit down."

She lowered herself back slowly into the chair, watching him all the time with those large, bewildered eyes. He had finished the second drink and was staring down into the glass, frowning.

"Listen," he said. "I've got something to tell you."

"What is it, darling? What's the matter?"

He had now become absolutely motionless, and he kept his head down so that the light from the lamp beside him fell across the upper part of his face, leaving the chin

and mouth in shadow. She noticed there was a little muscle moving near the corner of his left eye.

"This is going to be a bit of a shock to you, I'm afraid," he said. "But I've thought about it a good deal and I've decided the only thing to do is tell you right away. I hope you won't blame me too much."

And he told her. It didn't take long, four or five minutes at most, and she say very still through it all, watching him with a kind of dazed horror as he went further and further away from her with each word.

"So there it is," he added. "And I know it's kind of a bad time to be telling you, bet there simply wasn't any other way. Of course I'll give you money and see you're looked after. But there needn't really be any fuss. I hope not anyway. It wouldn't be very good for my job."

Her first instinct was not to believe any of it, to reject it all. It occurred to her that perhaps he hadn't even spoken, that she herself had imagined the whole thing. Maybe, if she went about her business and acted as though she hadn't been listening, then later, when she sort of woke up again, she might find none of it had ever happened.

"I'll get the supper," she managed to whisper, and this time he didn't stop her.

When she walked across the room she couldn't feel her feet touching the floor. She couldn't feel anything at all- except a slight nausea and a desire to vomit. Everything was automatic now-down the steps to the cellar, the light switch, the deep freeze, the hand inside the cabinet taking hold of the first object it met. She lifted it out, and looked at it. It was wrapped in paper, so she took off the paper and looked at it again.

A leg of lamb.

All right then, they would have lamb for supper. She carried it upstairs, holding the thin bone-end of it with both her hands, and as she went through the living-room, she saw him standing over by the window with his back to her, and she stopped.

"For God's sake," he said, hearing her, but not turning round. "Don't make supper for me. I'm going out."

At that point, Mary Maloney simply walked up behind him and without any pause she swung the big frozen leg of lamb high in the air and brought it down as hard as she could on the back of his head.

She might just as well have hit him with a steel club.

She stepped back a pace, waiting, and the funny thing was that he remained standing there for at least four or five seconds, gently swaying. Then he crashed to the carpet.

The violence of the crash, the noise, the small table overturning, helped bring her out of the shock. She came out slowly, feeling cold and surprised, and she stood for a while blinking at the body, still holding the ridiculous piece of meat tight with both

hands.

All right, she told herself. So I've killed him.

It was extraordinary, now, how clear her mind became all of a sudden. She began thinking very fast. As the wife of a detective, she knew quite well what the penalty would be. That was fine. It made no difference to her. In fact, it would be a relief. On the other hand, what about the child? What were the laws about murderers with unborn children? Did they kill then both-mother and child? Or did they wait until the tenth month? What did they do?

Mary Maloney didn't know. And she certainly wasn't prepared to take a chance.

She carried the meat into the kitchen, placed it in a pan, turned the oven on high, and shoved it inside. Then she washed her hands and ran upstairs to the bedroom. She sat down before the mirror, tidied her hair, touched up her lips and face. She tried a smile. It came out rather peculiar. She tried again.

"Hullo Sam," she said brightly, aloud.

The voice sounded peculiar too.

"I want some potatoes please, Sam. Yes, and I think a can of peas."

That was better. Both the smile and the voice were coming out better now. She rehearsed it several times more. Then she ran downstairs, took her coat, went out the back door, down the garden, into the street.

It wasn't six o'clock yet and the lights were still on in the grocery shop.

"Hullo Sam," she said brightly, smiling at the man behind the counter.

"Why, good evening, Mrs. Maloney. How're you?"

"I want some potatoes please, Sam. Yes, and I think a can of peas."

The man turned and reached up behind him on the shelf for the peas.

"Patrick's decided he's tired and doesn't want to eat out tonight," she told him. "We usually go out Thursdays, you know, and now he's caught me without any vegetables in the house."

"Then how about meat, Mrs. Maloney?"

"No, I've got meat, thanks. I got a nice leg of lamb from the freezer."

"Oh."

"I don't know much like cooking it frozen, Sam, but I'm taking a chance on it this time. You think it'll be all right?"

"Personally," the grocer said, "I don't believe it makes any difference. You want these Idaho potatoes?"

"Oh yes, that'll be fine. Two of those."

"Anything else?" The grocer cocked his head on one side, looking at her pleasantly. "How about afterwards? What you going to give him for afterwards?"

"Well-what would you suggest, Sam?"

The man glanced around his shop. "How about a nice big slice of cheesecake? I know he likes that."

"Perfect," she said. "He loves it."

And when it was all wrapped and she had paid, she put on her brightest smile and said, "Thank you, Sam. Goodnight."

"Goodnight, Mrs. Maloney. And thank you."

And now, she told herself as she hurried back, all she was doing now, she was returning home to her husband and he was waiting for his supper; and she must cook it good, and make it as tasty as possible because the poor man was tired; and if, when she entered the house, she happened to find anything unusual, or tragic, or terrible, then naturally it would be a shock and she'd become frantic with grief and horror. Mind you, she wasn't expecting to find anything. She was just going home with the vegetables. Mrs. Patrick Maloney going home with the vegetables on Thursday evening to cook supper for her husband.

That's the way, she told herself. Do everything right and natural. Keep things absolutely natural and there'll be no need for any acting at all.

Therefore, when she entered the kitchen by the back door, she was humming a little tune to herself and smiling.

"Patrick!" she called. "How are you, darling?"

She put the parcel down on the table and went through into the living room; and when she saw him lying there on the floor with his legs doubled up and one arm twisted back underneath his body, it really was rather a shock. All the old love and longing for him welled up inside her, and she ran over to him, knelt down beside him, and began to cry her heart out. It was easy. No acting was necessary.

A few minutes later she got up and went to the phone. She knew the number of the police station, and when the man at the other end answered, she cried to him, "Quick! Come quick! Patrick's dead!"

"Who's speaking?"

"Mrs. Maloney. Mrs. Patrick Maloney."

"You mean Patrick Maloney's dead?"

"I think so," she sobbed. "He's lying on the floor and I think he's dead."

"Be right over," the man said.

The car came very quickly, and when she opened the front door, two policemen walked in. She know them both-she know nearly all the man at that precinct-and she fell right into a chair, then went over to join the other one, who was called O'Malley, kneeling by the body.

"Is he dead?" she cried.

"I'm afraid he is. What happened?"

Briefly, she told her story about going out to the grocer and coming back to find him on the floor. While she was talking, crying and talking, Noonan discovered a small patch of congealed blood on the dead man's head. He showed it to O'Malley who got up at once and hurried to the phone.

Soon, other men began to come into the house. First a doctor, then two detectives, one of whom she know by name. Later, a police photographer arrived and took pictures, and a man who know about fingerprints. There was a great deal of whispering and muttering beside the corpse, and the detectives kept asking her a lot of questions. But they always treated her kindly. She told her story again, this time right from the beginning, when Patrick had come in, and she was sewing, and he was tired, so tired he hadn't wanted to go out for supper. She told how she'd put the meat in the oven- "it's there now, cooking"- and how she'd slopped out to the grocer for vegetables, and come back to find him lying on the floor.

"Which grocer?" one of the detectives asked.

She told him, and he turned and whispered something to the other detective who immediately went outside into the street.

In fifteen minutes he was back with a page of notes, and there was more whispering, and through her sobbing she heard a few of the whispered phrases-"...acted quite normal...very cheerful...wanted to give him a good supper... peas...cheesecake...impossible that she..."

After a while, the photographer and the doctor departed and two other men came in and took the corpse away on a stretcher. Then the fingerprint man went away. The two detectives remained, and so did the two policemen. They were exceptionally nice to her, and Jack Noonan asked if she wouldn't rather go somewhere else, to her sister's house perhaps, or to his own wife who would take care of her and put her up for the night.

No, she said. She didn't feel she could move even a yard at the moment. Would they mind awfully of she stayed just where she was until she felt better. She didn't feel too good at the moment, she really didn't.

Then hadn't she better lie down on the bed? Jack Noonan asked.

No, she said. She'd like to stay right where she was, in this chair. A little later, perhaps, when she felt better, she would move.

So they left her there while they went about their business, searching the house.

Occasionally on of the detectives asked her another question. Sometimes Jack Noonan spoke at her gently as he passed by. Her husband, he told her, had been killed by a blow on the back of the head administered with a heavy blunt instrument, almost certainly a large piece of metal. They were looking for the weapon. The murderer may have taken it with him, but on the other hand he may have thrown it away or hidden it somewhere on the premises.

"It's the old story," he said. "Get the weapon, and you've got the man."

Later, one of the detectives came up and sat beside her. Did she know, he asked, of anything in the house that could've been used as the weapon? Would she mind having a look around to see if anything was missing-a very big spanner, for example, or a heavy metal vase.

They didn't have any heavy metal vases, she said.

"Or a big spanner?"

She didn't think they had a big spanner. But there might be some things like that in the garage.

The search went on. She knew that there were other policemen in the garden all around the house. She could hear their footsteps on the gravel outside, and sometimes she saw a flash of a torch through a chink in the curtains. It began to get late, nearly nine she noticed by the clock on the mantle. The four men searching the rooms seemed to be growing weary, a trifle exasperated.

"Jack," she said, the next tome Sergeant Noonan went by. "Would you mind giving me a drink?"

"Sure I'll give you a drink. You mean this whiskey?"

"Yes please. But just a small one. It might make me feel better."

He handed her the glass.

"Why don't you have one yourself," she said. "You must be awfully tired. Please do. You've been very good to me."

"Well," he answered. "It's not strictly allowed, but I might take just a drop to keep me going."

One by one the others came in and were persuaded to take a little nip of whiskey. They stood around rather awkwardly with the drinks in their hands, uncomfortable in her presence, trying to say consoling things to her. Sergeant Noonan wandered into the kitchen, come out quickly and said, "Look, Mrs. Maloney. You know that oven of yours is still on, and the meat still inside."

"Oh dear me!" she cried. "So it is!"

"I better turn it off for you, hadn't I?"

"Will you do that, Jack. Thank you so much."

When the sergeant returned the second time, she looked at him with her large, dark tearful eyes. "Jack Noonan," she said.

"Yes?"

"Would you do me a small favor-you and these others?"

"We can try, Mrs. Maloney."

"Well," she said. "Here you all are, and good friends of dear Patrick's too, and helping to catch the man who killed him. You must be terrible hungry by now because it's long past your suppertime, and I know Patrick would never forgive me, God bless his soul, if I allowed you to remain in his house without offering you decent hospitality. Why don't you eat up that lamb that's in the oven? It'll be cooked just right by now."

"Wouldn't dream of it," Sergeant Noonan said.

"Please," she begged. "Please eat it. Personally I couldn't touch a thing, certainly not what's been in the house when he was here. But it's all right for you. It'd be a favor to me if you'd eat it up. Then you can go on with your work again afterwards."

There was a good deal of hesitating among the four policemen, but they were clearly hungry, and in the end they were persuaded to go into the kitchen and help themselves. The woman stayed where she was, listening to them speaking among themselves, their voices thick and sloppy because their mouths were full of meat.

"Have some more, Charlie?"

"No. Better not finish it."

"She wants us to finish it. She said so. Be doing her a favor."

"Okay then. Give me some more."

"That's the hell of a big club the guy must've used to hit poor Patrick," one of them was saying. "The doc says his skull was smashed all to pieces just like from a sledgehammer."

"That's why it ought to be easy to find."

"Exactly what I say."

"Whoever done it, they're not going to be carrying a thing like that around with them longer than they need."

One of them belched.

"Personally, I think it's right here on the premises."
"Probably right under our very noses. What you think, Jack?"
And in the other room, Mary Maloney began to giggle.