Frontier Words in the L2 Mental Lexicon

Alla Zareva

Ohio University

Lexical knowledge development is subject to partial increments which suggests that between the two extremes of familiar and unfamiliar words, there is an intermediate level of partially familiar vocabulary (frontier words) that, however, has received little attention in L2 research. Thus, the present study set out to examine the partial word knowledge of native speakers, L2 advanced and intermediate learners of English with regard to three features from Richards' (1976) taxonomy of aspects describing what knowing a word entails. To capture partial familiarity, the participants completed in writing a test containing low and mid frequency content words accompanied by a word knowledge scale. The analysis showed that there were three distinctive patterns of partial familiarity but their distribution across the three groups was quite different, which indicated that partial knowledge was linked to different word features across the three proficiency levels.

There is no doubt that the mental lexicon is highly organized. The fact that an educated native speaker (NS) of English knows several thousand words—the more conservative estimates range from 14,000 to 20,000 words (e.g., D'Anna, Zechmeister, & Hall, 1991; Goulden, Nation, & Read, 1990; Nusbaum, Pisoni, & Davis, 1984; Zareva, Schwanenflugel, & Nikolova, 2005; Zechmeister, Chronis, Cull, D'Anna, & Healy, 1995), while the more liberal ones suggest 50,000 words and more (e.g., Aitchison, 1987)—which he/she can access very fast (in 200 ms) upon recognition says a lot to that effect. In addition, most NSs are impressively fast at finding the words they need in speech production—for example, in prepared discourse, NSs of English maintain a rate of about 150 words per minute which suggests that it roughly takes one third of a second for a NS to retrieve a word he/she needs to use in a sequence of words. Of course, these are only average estimates which, nonetheless, are quite useful in helping us get a better idea about the size and the nature of the lexical task in second language (L2) learning.

L2 users are usually thought to have much smaller vocabulary sizes than native speakers which claim, however, to my knowledge, has not been supported by empirical findings, neither does it help us decide how to distribute the lexical learning effort accordingly across differing proficiencies. In this regard, it is more realistic to argue that L2 vocabulary size crucially depends on learners' proficiency level, rather than lump all L2 speaking individuals together in one unitary, yet shapeless, group. Indeed, in comparing three groups of participants at different levels of proficiency—NSs of English, L2 advanced, and intermediate learners of English—Zareva et al. (2005) found that size-wise, there were no meaningful differences between the NS (range 8,500 – 20,700 words) and advanced L2 group (range 7,900 – 20,700 words), while on average the vocabulary size of the L2 intermediate participants (range 4,900 – 14,800 words) was way below the NSs' and the advanced L2 learners' sizes alike. What seems to be evident, then, is that given that at the higher levels of

proficiency the L2 lexicon closely approximates to NSs' lexical size and the demands for it to be functional are equally pressing, it has to be highly organized as well. In this regard, there have been several attempts to set up a framework of meaningful criteria that account for the organization of lexical knowledge, where the notion of knowing a word has been operationalized with respect to a number of features related to a word's form, meaning, function, and features it projects onto the syntactic level. In reviewing some of these frameworks, however, Meara (1996) has rightly noted that most of the researchers, who have tried to specify the components of lexical knowledge (e.g., Gass & Selinker, 2001; Nation, 1990; 2001; Singleton, 1999; Wesche & Paribakht, 1996) have more or less re-iterated Richards' (1976) taxonomy of seven aspects of word knowledge which he identified as follows:

1) to know the probability of encountering a word in speech or writing;
2) to know the limitations of word use according to function and situation;
3) to know its syntactic properties;
4) to know the word's underlying form as well as its derivations;
5) to know the associations between the word and other words in the language;
6) to know the semantic value of the word;
7) to know many of the meanings associated with the word.

Regardless of the framework one may choose to adhere to, the fact that for an adult language user word knowledge is best described in terms of a set of features that capture what it means to know a word suggests that knowing a word is not an either-or state of affairs. Rather, it should be conceptualized along a continuum where, at the one extreme, we have complete familiarity with words, i.e. knowledge of all seven aspects (in Richard's terms) for every word that falls there and, at the other extreme, we would have unfamiliar words. Thus, assuming degrees of knowing or not knowing a word is an approach that recognizes both the cumulative nature of the process of lexical learning as well as the graduated nature of word knowledge. In other words, such an approach assumes that, between the two extremes, there are intermediate levels of word knowledge where individuals are not certain about some aspects of a word's meaning, even though they may have encountered that word before (Dale, 1965). To capture those intermediate levels of knowing a word and find out what information is available from familiar words that is not completely available from the partially familiar or the unfamiliar ones, first language (L1) researchers (e.g., Dale, 1965; Shore & Durso, 1990; Durso & Shore, 1991) began using a scale of word familiarity. One of the most commonly used scales is the one proposed by Dale (1965), who distinguished between 4 stages of knowing a word:

Stage 1: "I never saw it before".

Stage 2: "I've heard of it, but I don't know what it means."

Stage 3: "I recognize it in context – it has something to do with …."

Stage 4: "I know it." (Dale 1965, p. 398).

Similar scales are also frequently used in L2 studies (e.g., Hulstijn & Laufer, 2001; Hulstijn, Hollander, & Greidanus, 1996; Knight, 1994; Paribakht & Wesche, 1993; Wolter 2001; Zareva, 2007; Zareva et al. 2005) to account for familiarity effects, especially when words from a range of frequency bands are used as stimulus words (SWs). Thus, it would be safe to say that, in the last couple of decades, conceptualization of word familiarity in L2

research shifted from ***assumptions*** about familiarity to ***obtaining evidence*** of familiarity, based on participants' ratings and verification of knowledge of words from a range of frequency bands. However, very little research has been done on L2 users' partially familiar words with the aim of identifying the features that are most frequently absent from their word knowledge as related to their language proficiency. Having better understanding of the status of those words in L2 learners' minds would help us put stronger emphases in our vocabulary teaching practices on aspects that L2 learners themselves may unknowingly overlook and de-emphasize others that may be self-evident.

The following study was designed to compare two groups of adult L2 learners of English at two levels of proficiency–L2 advanced and intermediate learners–with NSs of English on their frontier words, to use Trembly's term (1966 cited in Durso & Shore, 1991). Trembly (1966) explained that frontier words, or partially familiar words, are the ones that exist 'in the frontier region between the point where every word is known and the point where no words are known" (p. 229). Thus, those are words with a somewhat intermediate status in one's lexicon of which status language users may not even be fully aware. For the purposes of this study I adopt the term *frontier words* for words that were judged as familiar by the participants but could not be acceptably explained. To capture partial word familiarity, I used a vocabulary knowledge scale on which the participants rated their word knowledge as unfamiliar ("I have not seen this word before and I don't know what it means") or familiar with two degrees of certainty—one, indicating a lower degree of certainty ("I think this word means …") and another one, indicating a high degree of certainty ("I know that this word means ... "). In both cases, the participants had to verify their knowledge with the SW by giving an explanation or synonym that reflected both the lexical class (LC) and, at least, one of the meanings of the test items. This in turn allowed for identification of their frontier words.

It has to be pointed out here that the test design did not allow for examining all seven aspects of word knowledge as specified by Richards (1976), but it captured meaningfully three of them, i.e.: 1) knowledge of a word's underlying form; 2) knowledge of its syntactic properties; and 3) knowledge of the semantic value of the word. As important as all seven features are, the three chosen for investigation in this study are the ones that can be successfully examined in a single word task, which in itself is a rigid measure of word knowledge. It was also of interest to find out whether differences among the three groups, if any, could point to aspects we as instructors may need to stress on more in vocabulary teaching so that the partially familiar words move faster from the state of being frontier to familiar.

In light of these questions, the following null hypotheses were tested in the study:

The first hypothesis tested the assumption that there were no differences among the three groups in the proportion of words they identify as unfamiliar.

The second one hypothesized no differences among the three groups in the proportion of their partially familiar words.

A closer inspection of the category of frontier words revealed that this is not a unitary category and that it can be subdivided into three other subcategories reflecting patterns of partial familiarity. Thus, the third hypothesis tested the assumption of no differences among the participants in the proportion of words they judged as familiar but, in actuality, they had no knowledge of.

The second pattern of frontier words contained words for which the participants did not have a good sense of LC, yet revealed some knowledge of their meaning; thus, hypothesis four tested how the three groups compared on the proportion of such words.

The last pattern of partially familiar vocabulary reflected participants' knowledge of the LC of some SWs, but no knowledge of their meaning; hence, the fifth hypothesis examined the assumption of no difference among the participants in the proportion of words for which they knew the LC but not the meaning.

Method

Participants

Altogether, 120 students—both NSs and L2 learners of English—participated in the study. The NSs (n = 40) were undergraduate students of different majors at two U.S. universities. The L2 learners of English (n = 80) were NSs of seventeen different languages. All non-native speakers (NNSs) had learned English in their respective native countries through formal instruction ($M_{Advanced}$ = 9.9 and $M_{Intermediate}$ = 8.5 years of instruction) and had lived in the U.S.A. or other English speaking countries for an average of 1.8 years. The L2 participants were placed in two proficiency groups—an advanced learner group (n = 40) and an intermediate learner group (n = 40)—based their *TOEFL* scores. All *TOEFL* tests were taken in controlled settings. The participants in the L2 advanced group had an official mean score of 584 (range 550 – 630). The participants placed in the intermediate learner group had a mean score of 475 (range 427– 500).

Materials and procedure

The test insrument contained 25 SWs selected from a larger sample. The test items were from three lexical categories—nouns (n = 15), verbs (n = 5), and adjectives (n = 5)—and each lexical category contained low and mid frequency words (see Appendix A). A guiding principle in the SW selection was to have words that are grammatically unambiguous—that is, the words in the sample could belong to just one grammatical class. The frequency of each item was identified by its u-value taken from *The educator's word frequency guide* (Zeno, Ivens, Millard, & Duvvuri, 1995).

The participants completed a vocabulary test in writing. Each SW was accompanied by a word familiarity scale on which they rated their word knowledge on a scale of three options corresponding to three degrees of word familiarity. Choosing the first option ("[I] I have not seen this word before and I don't know what it means") did not require from them any further explanation. A choice of one of the two other options (" [II] I think this word means …" or "[III] I know that this word means ...") required from them to verify their familiarity by writing down a synonym or brief explanation of the word.

Coding of responses and WA classification procedure

The data were analyzed in several steps. First, to find out how the participants compared on unfamiliar words, the number of items they self-reported as unfamiliar in option I ("I have not seen this word before and I don't know what it means") was counted and then transformed into a proportion. Next, the unacceptable responses from either option II ("I think this word means …") or III ("I know this word means …") were counted and turned into proportions, indicating the percentage of words the participants thought they knew but in reality they did not quite know. These responses formed the proportion of frontier words. The decision about response acceptability was based on criteria used in lexicography for defining words by part of speech, as outlined by Landau (2001) (for more detail, see Zareva et al.,

2005). The category of partial knowledge was not uniform so, to refine it, the unacceptable responses were further classified into three sub-categories, i.e.:

1) Responses which showed that the participant had no knowledge of the SW and/or had completely misinterpreted its meaning.

2) Responses that did not account for the LC of the SW, yet revealed some knowledge of the meaning of the word.

3) Responses that reflected knowledge of the LC of the SW, but failed to reflect correctly its meaning.

Results

A series of One-way ANOVAs was conducted to examine the relationship between language proficiency and participants' partial word knowledge. The independent variable was language proficiency with three levels (NSs, L2 advanced, and intermediate learners). The dependent variables were: (1) proportion of words the participants self-reported as unfamiliar; (2) overall proportion of participants' frontier words; (3) proportion of responses which showed that the participants had no knowledge of some SWs and/or had completely misinterpreted their meaning; (4) proportion of responses that did not account for the LC of some SWs but revealed some knowledge of their meaning; and (5) proportion of responses that reflected knowledge of the LC of some SWs but did not reflect correctly their meaning.

Proficiency level showed significant main effects on several of the dependant variables (see Table 1 for means and standard deviations). First, there were significant differences among the three groups of participants on their proportion of self-reported unfamiliar words, $F(2, 117) = 15.830$, $p < .001$, $\omega^2 = .20$, which indicated that the proportion of unfamiliar vocabulary can largely be explained by level of proficiency. The three groups also differed in the overall proportion of partially familiar words, i.e. words they believed they knew and attempted to explain, but the explanation revealed partial knowledge of those words, $F(2, 117) = 7, 438$, $p < .01$, $\omega^2 = .10$. When the category of partially familiar vocabulary was broken down into the three subcategories discussed earlier, there were significant differences among the three groups in the proportion of responses which revealed participants' lack of knowledge of some SWs and/or some misinterpretation of their meaning, $F(2, 117) = 7, 435$, $p < .01$, $\omega^2 = .10$, and in the proportion of responses that reflected knowledge of the LC of some SWs but no adequate knowledge of their meaning, $F(2, 117) = 8, 083$, $p < .01$, $\omega^2 = .12$. However, the differences in participants' proportion of responses that did not account for the LC but revealed some knowledge of the meaning of the SWs turned out nonsignificant ($p = .706$). Thus, the results indicated that, overall, the participants had a stronger sense of LC than meaning of their partially familiar vocabulary regardless of their proficiency level.

Table 1

*Means and Standard Deviations for the Group of NSs (n=40), L2 Advanced (n=40), and L2 Intermediate (*n*=40) Learners of English*

	Mean			*SD*		
	NSs	*L2 Adv.*	*L2 Intrm.*	*NSs*	*L2 Adv.*	*L2 Intrm.*
% Self-reported unfamiliar words	50.8	55.4	72.3	21.3	15.3	16.8
% Total frontier words:	9.6	19.2	25.0	9.1	18.3	23.8
% defined words but no knowledge	5.3	6.3	16.1	6.7	12.2	19.5
% wrong LC but close meaning	3.1	3.0	2.0	5.3	6.0	7.6
% correct LC but wrong meaning	1.5	9.6	7.0	3.1	10.8	11.2

To find out where exactly the differences among the three groups were, post hoc comparisons were conducted. Based on Bonferroni rationale to control for Type I error, all post hoc comparisons were adjusted to .05 alpha level of significance. The comparisons of participants' proportion of self-reported unfamiliar words revealed significant differences between the L2 intermediate learners and the native speaking group (95% CI = 11.726, 31.2741; $p < .05$) as well as the L2 advanced participants (95% CI = 7.126, 26.674; $p < .05$) indicating a much higher proportion of unfamiliar vocabulary among the intermediate learners than the NSs or the advanced learners. Similar differences were observed with the comparison of their defined words that actually showed lack of familiarity (95% CI $_{NSs}$ = 3.298, 18.303 and 95% CI $_{L2\ Adv.}$ = 2.247, 17.252; $p < .05$) (see Figure 1).

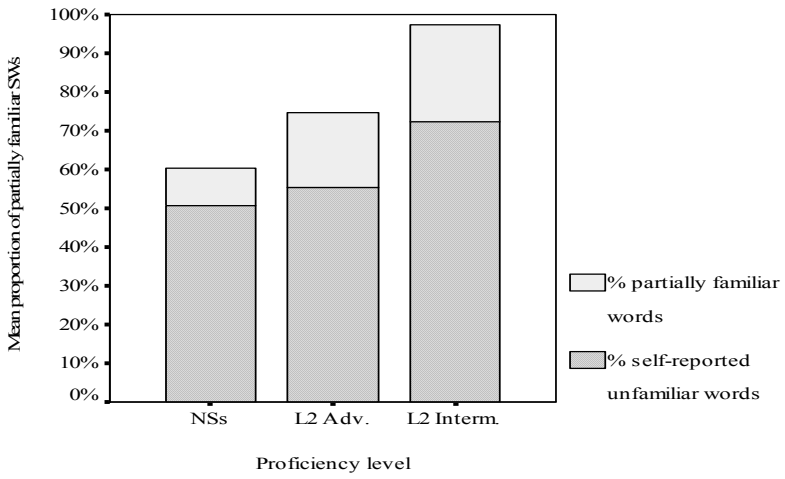

Figure 1. Proportion of self-reported unfamiliar and partially familiar words.

The intermediate learners also differed significantly from the NSs in the overall proportion of partially familiar words (95% CI = 5.618, 25.383; $p < .05$) and in the proportion of responses that showed LC knowledge of the SWs but no adequate knowledge of their meaning (95% CI = .439, 10.411; $p < .05$). Quite surprisingly, the advanced learners also differed from the NSs in that respect (95% CI = 3.113, 13.086; $p < .05$), but I will elaborate on this finding in the discussion section. Figure 2 shows the three different types of familiarity misconceptions across the three groups, of which only the proportion of responses that failed to account for the LC but revealed some knowledge of the meaning of the SWs was nonsignificant.

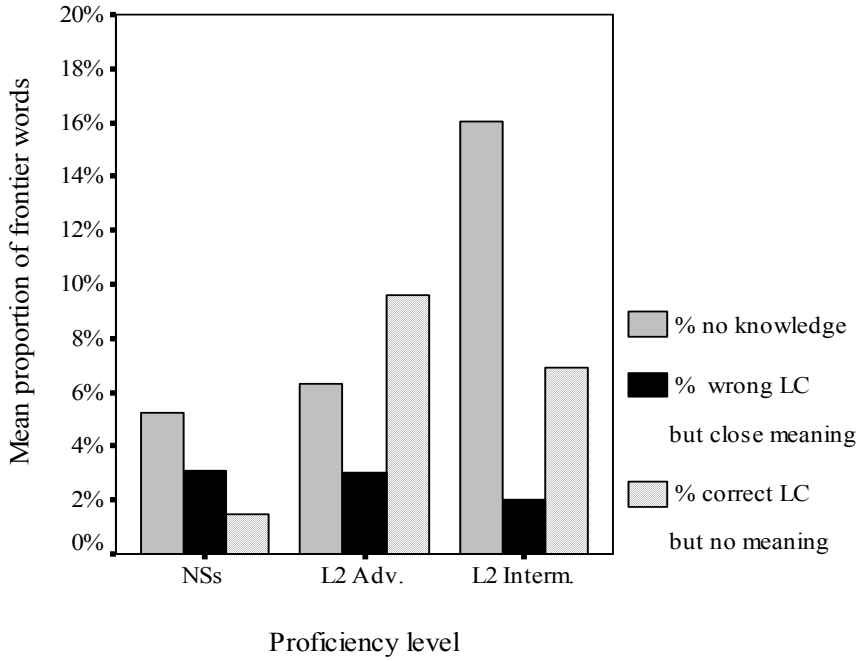

Figure 2. Effects of proficiency level on participants' familiarity misconceptions.

Discussion

As noted elsewhere in the literature, the development of vocabulary knowledge is subject to partial increments and, be it in L1 or L2, there are words which we are certain that we know (and we are able to prove it) as well as words we know that we do not know. In-between these two extremes, there is an intermediate level of partial familiarity (some words being more familiar than others) which has not been clearly operationalized. If we consider Richards' (1976) seven features describing what knowing a word entails, not knowing one or more of the features would render a word partially familiar. Thus, words that fall into this intermediate level of word knowledge may range from "I know there is such a word, but I don't know what it means", in Dale's (1965) terms, to words that have been encountered in some form for which, for example, a person may not have a good sense of their frequency of distribution (feature 1 in Richards' framework), yet know the rest of the aspects. There might be even words there which one believes he/she knows but, in actuality, he/she does not. Thus, it is clear that partial knowledge is not just an intermediate level, but a level that is also arranged on a continuum. While it is not possible to say which aspects are more important than others, it is possible to say that knowledge of same aspects can be tested at a single word level, while knowledge of others can only be tested at a syntactic, discourse, or pragmatic level.

The present experiment was set to examine comparatively the partial word knowledge of three groups of adults at different levels of proficiency—NSs, L2 advanced and intermediate learners of English—with regards to three word features, namely knowledge of a word's underlying form, its syntactic properties, and its semantic value. Those features were chosen because they could successfully be examined by means of a single word task, in which the participants had to verify their knowledge of the test items without supportive context. The test items were selected from three lexical categories (nouns, verbs, and adjectives) and each lexical category contained only low and mid frequency words. Five hypotheses were tested in the study and I shall elaborate on the findings regarding each of them in the paragraphs below.

The first hypothesis tested the assumption that there were no differences among the three groups in the proportion of words they identify as unfamiliar and, not surprisingly, the hypothesis was rejected. The L2 intermediate learners self-reported a much greater percentage of unfamiliar words than the NSs and the advanced learners—that is, the L2 intermediate participants identified 72% of the low and mid frequency words as unfamiliar, while the NSs did so for 51% of those words and the advanced learners—for 55%. After the participants identified the words they considered unfamiliar, they had to verify their knowledge of the words they judged as familiar, which in turn provided evidence for their partial familiarity with some of the items they thought they knew. Thus, the second hypothesis tested whether, overall, the participants differed in the proportion of their frontier words and, indeed, the results indicated that, in addition to the unfamiliar vocabulary, the L2 intermediate learners had another 25% of frontier words and in that they significantly differed from the NSs (9.5%), but not from the advanced learners (19%). This overall difference prompted further investigation into the nature of participants' partial lexical knowledge since, after analyzing the explanations and synonyms they provided, it was possible to identify distinctive patterns of misinterpretations of some SWs. The first pattern of misinterpretations contained words the participants thought they knew but their explanations or synonyms showed that they either completely misinterpreted the meaning of the stimulus or did not know any of its meanings. In this regard, Laufer (1997) has rightly noted that, in a reading comprehension task, some words may not even be recognized by readers as unfamiliar and be completely misinterpreted because of their "deceptive transparency" (p.25). She further explains that such cases usually include: 1) words with deceptive morphological structure (e.g., *drawback* can easily be misinterpreted as "something that draws you back" if the consisting morphemes are interpreted literary); 2) idioms whose deceptive transparency is of different nature, yet in some ways similar to the words with deceptive morphological structure (e.g., *look up a word,* which if interpreted literary does not make much sense); 3) false friends—i.e. L2 words that resemble L1 words in form, yet have a completely different meaning (e.g., *ten* in English—the numeral "10" and "ten" in Bulgarian, meaning "complexion"); 4) polysemous words, with whose multiple meanings an L2 learner may not be familiar (e.g., *transport* as a verb can mean "to transfer from one place to another" but it can also mean "to carry away with strong and pleasant emotion") and 5) synforms, which include pairs of words with similar forms but different meaning (e.g., *economic/economical*, *then/than*, etc.).

Indeed, some of these cases could easily be recognized among the responses the participants in this study gave, more so in the intermediate and advanced learner data than among the NSs' responses. The most common types of misinterpretations included synforms. For example, the intermediate students commonly misinterpreted the SW *prefect* as "perfect", the advanced learners—*parable* as "parabola", and the NSs—*virtuosity* as "virtuousness". In addition to the misinterpretations, there was a good deal of responses that showed no

knowledge whatsoever of the prompt word—for instance, in the NSs' responses, *prefect* was explained as "utopia", *abattoir* as "to strike repeatedly", etc; the L2 advanced learners explained *bracelet* as "to hold", *bursar* as "cover", *forgo* as "false", etc.; and the intermediate students had *amoral* explained as "good things", *concede* as "forwarded", *custodian* as "daily", etc. Overall, this category accounted for about 5% of the NSs' frontier words, 6% of the advanced learners', and 16% of the intermediate learners' partially familiar words.

What I find disturbing here is the overall instability of intermediate learners' lexical knowledge of mid and low frequency words. If we add the percentage of words they judged as unfamiliar (72%) to their partially familiar vocabulary (25%), we get the striking 97% of low and mid frequency vocabulary intermediate learners are uncertain about, of which there was evidence they actually did not know 88% (72% + 16% frontier words that may actually be unfamiliar). This figure is quite problematic in light of the fact that the intermediate learners were participants with a mean paper-based TOEFL score of 475 who have studied English for an average of 8.5 years and who were planning to move to degree granting programs at U.S. universities in the very near future. Yet, those learners perhaps have only about 12% of low and mid frequency sight vocabulary, of which approximately 7% consists of vaguely familiar ones. Going into English-based academic programs with such a state of lexical knowledge may turn out to be a set up for the students (as well as for the programs) since their vocabulary knowledge will very likely hold them back in their academic comprehension and production. For instance, most studies show that the threshold for reading comprehension is largely lexical (Laufer, 1997) and while text coverage in academic texts by the first most frequent 2,000 words of English is still high (about 78%), the role of mid and low frequency vocabulary (including academic, technical, specialized vocabulary, etc.) significantly increases from 10% in the conversational register to 22% in academic texts (Nation, 2001). Thus, the perceived importance of mid and lower frequency vocabulary and learners' efforts to improve their knowledge of those words should reflect the demands that their academic studies will put on them in the future.

The second pattern of frontier words contained words for which the participants did not show a good sense of LC, yet revealed some knowledge of their meaning. The results indicated that the cumulative effect of this sub-category of partially familiar words was very small (NSs – 3%, L2 advanced learners – 3%, intermediate group – 2%) and the differences among the three groups were not meaningful.

The last pattern of partially known vocabulary contained responses that reflected knowledge of the LC of some SWs, but no knowledge of their meaning and this category sharply distinguished NSs from NNSs. Quite surprisingly, the advanced learners' dominated the pattern with almost 10% of their responses falling in this sub-category, followed by the intermediate participants—with approximately 7%, while the NSs had only 1.5% of such responses. This again proves that of all three aspects of word knowledge tested in this study, knowledge of the LC of words is the aspect that seems to be the least problematic feature for adults at the higher level of proficiency. However, knowledge of the semantic value of words showed to be the greatest obstacle, particularly for the advanced learners. Responses in the advanced group data that were classified in this category revealed a different degree of deviation from the actual meaning of the SWs, which indicated a good sense of LC but a vague to non-existent knowledge of the meaning. With the intermediate learners, the examples did not contain vagueness, but rather a clear indication of non-existent semantic knowledge.

Overall, if we look at the three patterns of partial familiarity (see Figure 2), the distribution across the three groups is quite different. The NSs' pattern is dominated by words

they considered familiar and attempted to explain but the explanations showed they had no knowledge of these words, followed by words whose correct LC was not encoded in the explanation, and finally—a very small proportion of words whose grammatical class was accounted for but the meaning did not seem to be known. The greatest difference between the NS group and the advanced learner group, as shown on the diagram, is in the much higher proportion of words whose meaning was vaguely familiar or non-existent to the advanced learners. By and large, it can be reasonably hypothesized that these 10% of mid and low frequency vocabulary would be the most difficult category to move from the status of being frontier to a status of familiar since it is highly possible that the learners themselves are not aware of their not knowing the meanings of these words. Thus, even if they encounter them in context, unless the context is constrained, they will most likely be not able to successfully guess their meaning. The same holds true for the intermediate learners, though their greatest problem with frontier words is with the ones they think they know but they do not. This category was not further analyzed to single out the misinterpretations from the responses that show unfamiliarity with more than 1 feature, but the fact that this sub-category may account for as high as 16% of the intermediate learners' frontier words is already disturbing. When we add to those 16 % the 7% of words whose meaning the intermediate students do not know, the task of changing the status of these two sub-categories (23% altogether) from frontier to familiar will require a lot of effort.

In conclusion, it should be pointed out that a test format which comes with a scale for rating lexical knowledge and requires from test-takers to provide evidence for their word familiarity allows for identification of 10% to 25% of partially familiar vocabulary which a check-list test format, for example, would not capture. It is also important to note that the proportion of frontier words increases with a decrease in proficiency and it can be as high as 25% for L2 intermediate learners who, in addition to having much lower vocabulary sizes than advanced learners or NSs, probably, do not have sufficient experience with mid and low frequency words. The analysis of the category of frontier words revealed that there were three distinctive patterns that account for partial familiarity; however, the distribution of these patterns differed across the three proficiency groups. The most general conclusion to be drawn from this line of investigation is that both the advanced and intermediate learners need to work on their lexicons but from somewhat different angles. The advanced learners need to improve on words they consider familiar and have a good sense of what LC they belong to but not their meaning. The intermediate learners need to work on that aspect too, in addition to attending to the soaring 25% of words they thought they knew but their explanations showed that they either completely misinterpreted their meaning or they had no knowledge of any of the words' features.

On a final note, even though testing words in isolation has been most often criticized for depriving test takers from context that may disambiguate the meaning of words, studies in reading comprehension show that 'unless the context is very constrained, which is a relatively rare occurrence, or unless there is a relationship with a known word identifiable on the basis of form and supported by context, there is little chance of guessing the correct meaning" (Kelly, 1990, p. 203). Thus, knowing as many features of a word as possible out of context will unquestionably result in higher degree of L2 comprehension and production. In this regard, working with dictionary definitions may remedy some of the problems L2 intermediate and advance learners experience with mid and low frequency frontier words since a monolingual dictionary definition usually imparts specific information about a word and some general constraint information. Thus, at minimum, it will supply unambiguous information about a word's underlying form and its variants, its grammatical class, and semantic values. The lexical task for L2 learners at the higher levels of proficiency

(intermediate and above) should be seen not only in increasing their vocabulary size and improving its functionality but also in moving words from the status of frontier to familiar. This, in turn, will improve their lexical size, connectivity and functionality of the lexicon—a cycle which comes with more benefits than drawbacks. Likewise, our task as instructors should be to raise learners' awareness of the disadvantages of partially familiar vocabulary (since for the most part learners are usually not aware of their partial familiarity with some words), so that they themselves develop sensitivity to identifying their frontier words and moving them to a stable status of familiarity.

References

Aitchison, J. (1987). *Words in the mind. An introduction to the mental lexicon*. New York: Basil Blackwell.

D'Anna, C. A., Zechmeister, E. B., & Hall, J. W. (1991). Toward a meaningful definition of vocabulary size. *Journal of Reading Behavior, 23*, 109–122.

Dale, E. (1965). Vocabulary measurement: Techniques and major findings. *Elementary English, 42*, 395–401.

Durso, F. T. & Shore, W. J. (1991). Partial knowledge of word meaning. *Journal of Experimental Psychology, 2*, 190–202.

Gass, S. M., & Selinker, L. (2001). *Second language acquisition: An introductory course.* Mahwan, NJ: Lawrence Earlbaum.

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

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

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. *Modern Language Journal, 80*, 327–339.

Kelly, P. (1990) Guessing: No substitute for systematic learning of lexis. *System, 18,* 199–207.

Knight, S. (1994). Dictionary use while reading: The effects on comprehension and vocabulary acquisition for students of different verbal abilities. *Modern Language Journal, 78*, 285–299.

Landau, S. I. (2001). *Dictionaries: The art and craft of lexicography*. Cambridge University Press.

Laufer, B. (1997). The lexical plight in second language acquisition: Words you don't know, words you think you know, and words you can't guess. In J. Coady and T. Huckin (Eds.), *Second language vocabulary acquisition* (pp. 20–34). Cambridge: Cambridge University Press.

Longman dictionary of English language and culture. (1992). Essex, England: Longman.

McNiel, D. (1966). A study of word association. *Journal of Verbal Learning and Learning behavior, 5*, 548–557.

McNiel, D. (1970). *The acquisition of language*. New York: Harper & Row.

Meara, P. (1978). Learners' associations in French. *Interlanguage Studies Bulletin (3),* 192–211.

Meara, P. (1996). The dimensions of lexical competence. In G. Brown, K. Malmkjæ, and J. Williams (Eds.), *Competence and performance in language learning* (pp. 35–53). Cambridge: Cambridge University Press.

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

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

Nusbaum, H. C., Pisoni, D. B., & Davis, C. K. (1984). Sizing up the Hoosier mental lexicon. Measuring the familiarity of 20,000 words. (*Research on Speech Production Report No. 10*). Bloomington, IN: Indiana University.

Paribakht, T., & Wesche, M. (1993). The relationship between reading comprehension and second language development in a comprehension-based ESL program. *TESL Canada Journal, 11*, 9–29.

Piper, T. H., & Leicester, P. F. (1980). Word association behavior as an indicator of English language proficiency. (ERIC Document Reproduction Service No. ED 227 651).

Richards, J. C. (1976). The role of vocabulary teaching. *TESOL Quarterly, 10*, 77–89.

Shore, W. J. & Durso, F T. (1990). Partial knowledge in vocabulary acquisition: General constraints and specific detail. *Journal of Educational Psychology,82*, 315–318.

Singleton, D. (1999). *Exploring the second language mental lexicon*. Cambridge: Cambridge University Press.

Söderman, T. (1993). Word associations of foreign language learners and native speakers. The phenomenon of shift of response type and its relevance for lexical development. In H. Ringbom (Ed.), *Near-native proficiency in English* (pp. 98–182).Abo, Finland: Abo Akademi.

Stolz, W. S., & Tiffany, J. (1972). The production of "child-like" word associations by adults to unfamiliar adjectives. *Journal of Verbal Learning and Learning Behavior, 11*, 38–46.

Trembly, D. (1966). Laws of learning general and specialized vocabularies [Summary]. Proceedings of the 74th Annual Convention of American Psychological Association, 1, 229–230.

Wesche, M., & Paribakht, S. (1996). Assessing second language vocabulary knowledge: Depth versus breadth. *Canadian Modern Language Review, 53*, 13–40.

Wolter, B. (2001). Comparing the L1 and L2 mental lexicon. *Studies in Second Language Acquisition, 23*, 41–69.

Zareva, A. (2007). Structure of the L2 mental lexicon: How does it compare to native speakers' lexical organization? *Second Language Research, 23,* 123–153.

Zareva, A. (2008 – under review). Effects of lexical class and word frequency on L1 and L2 speakers' associative domains: Empirical findings and methodological concerns. *Applied Linguistics*.

Zareva, A., Schwanenflugel, P. , & Nikolova Y. (2005). Relationship between lexical competence and language proficiency: Variable sensitivity. *Studies in Second Language Acquisition, 27*, 567–595.

Zechmeister E. B., Chronis, A. M., Cull, W. L., D'Anna, C. A., & Healy, N. A. (1995). Growth of functionally important lexicon. *Journal of Reading Behavior, 27* (2), 201–212.

Zeno, S. M., Ivens, S. H., Millard, R. T., & Duvvuri, R. (1995). *The educator's word frequency guide. Touchstone Applied Science Association* (TASA).

Appendix A

List of the SWs used in the study grouped by LC and frequency of occurrence

Lexical category	*U-value*	*SWs*
Nouns	.0326	virtuosity
	.0424	bursar
	.0588	crampons
	.0653	abattoir
	.4153	parable
	.5556	inception
	.5651	edifice
	.5751	prefect
	.6393	solstice
	.6626	penance
	.6986	cassava
	.9539	promontory
	1.2348	custodian
	3.0000	bracelet
	9.8213	studio
Verbs	.1961	entrust
	.4292	unnerve
	.4498	forgo
	.5229	instill
	.9684	concede
Adjectives	.0326	amoral
	.0925	putative
	.2531	lackadaisical
	1.0000	naïve
	4.000	toxic