

CHAPTER TWO

LITERATURE REVIEW

When readers lack the adequate vocabulary, they will not be able to process reading materials successfully. Huang (1997: 323-324) points out that the total vocabulary words included in junior and senior high school English textbooks are about 1,050 and 2,800 respectively. This means that without extensive outside reading, a senior high school graduate may only know about 3,850 words. However, to prepare for the Scholastic Aptitude English Test (SAET) and the Department Required English Test (DRET), Cheng, Chang, Cheng, and Ku (2002: i) have published a word list of 6,480 words with which senior high school graduates should be familiar. The gap between what the students know and what is required for the tests often requires them to learn a large number of new words within a limited amount of time leading to these tests. Naturally, they would like a better and more efficient way to learn new words. This chapter is a literature review of the studies on English vocabulary teaching.

The first part looks into the importance of vocabulary teaching that includes reasons for the necessity of vocabulary teaching, levels of vocabulary processing, and guidelines for vocabulary instruction. The second part is the study of three intralexical factors (i.e., phonemic, semantic, and morphological features) in vocabulary teaching.

The third is an overview of four vocabulary teaching strategies that include traditional definition-based vocabulary teaching, vocabulary teaching from context, key-word vocabulary teaching, and morphologically-based vocabulary teaching. The last part is the study of the theoretical framework of vocabulary teaching: the iconic-morphological approach via commonly-used roots.

2.1 The Importance of Vocabulary Teaching

Due to its great impact on reading comprehension, vocabulary instruction plays a very important role in English learning. The following sections explore why it is absolutely necessary to teach vocabulary as well as introduce aspects of vocabulary processing and guidelines for vocabulary instruction.

2.1.1 Necessity of Vocabulary Teaching

Laufer (1986: 72) points out that, in the past, the acquisition of vocabulary was not considered extremely important because it was regarded as “inductive learning (the storage of linguistic information) and habit formation rather than as deductive learning (the discovery of linguistic information).” In addition, Hung (1991: 155) also indicates that the teaching of vocabulary has been greatly neglected in the past thirty years because of

- (1) the trend in linguistic theory, basically the audio-lingual methods,
- (2) the lack of understanding of how words are most efficiently learned,
- and (3) the impossibility of seeing immediate returns as compared to the teaching of grammar, pronunciation, etc.

Therefore, learning vocabulary is viewed as the greatest difficulty and the main obstacle in language learning today (Nation, 1982: 20; Meara, 1980: 221).

Inadequate vocabulary seriously hinders comprehension ability (Laufer, 1986; Nelson-Herber, 1986; Raimes, 1985; Martin, 1976). That is, vocabulary is a crucial

prerequisite to comprehension (Johnson & Johnson, 1986: 624) and vocabulary comprehension is highly correlated with reading acquisition (Cronin et al., 1986: 8-9). Today, vocabulary no longer serves merely to flesh out grammatical structures (Maiguashca, 1993: 84-85) but is, rather, even more important than grammar, phonology, and syntax (Gorman, 1979: 155-156).¹ For this reason, Jullian (2000: 37) notes that upper-intermediate and advanced learners usually do not make major grammatical mistakes but have limited lexicon. Moreover, Richards (1976: 78) claims that the native speakers, or advanced readers, continue to expand their vocabulary in adulthood, whereas there is comparatively little change in syntax.

2.1.2 Vocabulary Processing and Vocabulary Instruction

Henriksen (1999: 304-307) proposes three separate but related vocabulary dimensions of lexical competence for word processing:

(1) The Partial to Precise Knowledge Dimension: Beginning language learners first become vague about meanings and then precision will come later and the quantity of vocabulary (vocabulary size² or breadth) will be enriched. (2) The Depth of Knowledge Dimension: The concept of depth³ can be defined as “the quality of the learner’s vocabulary knowledge.” (3) The Receptive to Productive⁴ Dimension: Learners always acquire words from receptive use to productive use.

¹ For similar discussions, see Saville-Troike, 1984; Laufer, 1986.

² Many researchers have proposed different vocabulary sizes. Schmitt, Schmitt, and Clapham (2001: 56): 3000 words for reading authentic texts; Hirsh and Nation (1992: 695): 5000 words for pleasure reading; Grabe (1991: 380): 10,000-100,000 words for native speakers and 2,000-7,000 for ESL students; Huang (1997: 324): 50,000 words for native speaks; Johnson (1994: 65): 2,5000 for college students.

³ Schneider, Healy, and Bourne (2002: 419) also state that the nature of vocabulary knowledge is stressed as ‘complex’.

⁴ Similarly, Channell (1988: 84) assumes that receptive acquisition precedes productive acquisition. Consequently, much vocabulary still remains part of receptive competence without becoming part of productive capacity. Moreover, Crow and Quigley (1985: 499) state that one’s receptive lexicon (60,000 lexical units) outnumbers productive lexicon (2,000 lexical units), and the learning process of the former is much easier than that of the latter.

To facilitate vocabulary instruction, Stahl (1986: 664) defines three different levels of processing as follows:

(1) *Association processing* is when a student simply learns an association. Sometimes the association is between a word and a synonym, or sometimes between a word and a single text. (2) *Comprehension processing* is where more involvement with the previous association is utilized to better understand and manipulate the word. (3) *Generation processing* is when students take that comprehended association and then generate a novel product. Then, their own sentences can be created and better internalization of the words will be gradually built.

Besides, Carr and Wixson (1986: 588) set four guidelines for vocabulary instruction as follows:

(1) Instruction should help students relate new vocabulary to their background knowledge. (2) Instruction should help students develop elaborate word knowledge. (3) Instruction should provide for active student involvement in learning new vocabulary. (4) Instruction should develop students' strategies for acquiring new vocabulary independently.

Vocabulary instruction is important to learners. Stahl (1986: 663) ascertains that the students who are given vocabulary instruction outperform the students who are not given vocabulary instruction. Channell (1988: 94) encourages the instruction of lexical association when students learn new words. According to the statistical evidence provided by Nagy and Anderson (1984: 322), the reading vocabulary of students in grades 3 through 12 grows at a rate of 3,000-4,000 words per year or more, but only a small fraction is acquired through direct instruction.⁵ That is, most of students' vocabulary items are acquired outside of class. Therefore, vocabulary instruction plays an important role in vocabulary memorization. Undoubtedly, Carr and Wixson (1986: 588) assume that vocabulary instruction lessons can help students become spontaneous learners.

⁵ Haggard (1986: 639-640) provides statistical evidence that 40% of new words students have are course related, but 60% are noncourse related.

2.2 The Intralexical Factors in Vocabulary Teaching

Storkel (2001: 1333) states that when a new word is encountered, phonemic, semantic, and morphological representations are activated. They are intralexical factors of words and each factor has a certain influence on the teaching of lexical items.

2.2.1 Phonemic Features

Bloomfield (1933) states that phonemes are not sounds; but the features of sounds, which are signals that stand for certain sounds (Cited in Vachek, 1935: 245). Consequently, acquisition of spontaneous phonemic awareness is only true for very few children; the rest are trained through activities designed by teachers (Neuman, Cople, & Bredekamp, 2001: 11).

Phonemic ability is crucial to students. Wagner (1986: 624-627) presents three types of phonemic ability: “phonological awareness, verbal working memory, and rapid naming.”⁶ From Cornwall’s (1992: 537-538) and De Jong and Van der Leij’s (2002: 52) empirical evidence, phonological awareness and rapid naming are strongly related to word decoding. Besides, Hammill, Mather, Allen, and Rosberts (2002: 133) also state that rapid naming is more highly correlated with reading and is a better predictor of reading ability;⁷ however, the short-term verbal working memory indicates little or even no effects on word decoding.

⁶ *Phonological awareness* refers to the sensitivity for sounds in spoken words. Phonological coding in short term memory is denoted as *verbal working memory*. Finally, the phonological coding from long term memory can be denoted as *rapid naming* (Wagner, 1986).

⁷ There are more positive discussions between phonemic ability and reading ability, see Wagner, 1986; Goswami, 1991; Vitevitch and Luce, 1998; Zorzi, Houghton, and Butterworth, 1998; Sharon and Rolland, 2000; Dixon and Stuart, 2002.

Phonemic skills are crucial to beginning readers.⁸ These learners are found to rely more on phonemic processing skill than on other skills (Wagner, 1986: 623; Wood, 2002: 156). Conversely, older readers become less dependent on phonemic information (Oney, Peter, & Katz, 1997: 80). Moreover, with the delayed phonemic development, learners also display delayed acquisition of morphology, syntax, semantic, etc (Storkel, 2001: 1321; Hoffman & Norris, 2002: 230).

Reading words is not purely a visual process; it also involves intra-syllabic phonemic knowledge (Goswami, 1991: 1110). Therefore, the phonemic instruction of onset-rime is important especially for the readers who transfer from logographic (e.g., Chinese learners) to alphabetic reading and spelling (Layton, Deeny, Upton, & Tall, 1998: 38).

2.2.2 Semantic Features

A word is referred to as a symbol or a representation of an object rather than an object. The object is the *referent*. The reflection thought process of the referent occurring in our brain is called *reference*. The *symbol* is the name or word, which represents the product of thought association. The three are named the “semantic triangle”,⁹ the analysis of word meanings (Iacobelli, 1941: 16-17). According to Thomas, Campos, Shucard, Rammsay, and Shucard (1981: 802), the one from the end of the first year does not have to rely on the non-semantic linguistic factors (e.g.,

⁸ Sharon and Rolland (2000: 232) state that phonemic abilities are segmentation, blending, rhyming, first sound identification, and syllable deletion, which can be cultivated and applied to real reading.

⁹ The semantic triangle can be used to solve two major problems of language. The first one is the identification of words with things. The second one is recognition of abstract words and thoughts. Thus, with the aid of semantic, good communication will be generated as well as ambiguity and misunderstanding can be inhibited (Iacobelli, 1941: 17).

intonation) or nonlinguistic factors (e.g., object preference) but owns the signal detection to associate referents with symbols and the reference process.

Semantic instruction may set back vocabulary learning for beginning learners. According to the evidence provided by Tinkham (1997: 160-161),¹⁰ the instruction of semantic clustering impedes the learning of vocabulary owing to the confusion of semantic similarity.¹¹ Furthermore, Barcroft (2002: 354) states that the learning of semantic features analysis may facilitate the understanding of semantic (conceptual) properties, but inhibit the learning of the formal (phonemic and graphemic) properties of new words. Lipka (2002: 271) also claims that semantic information often plays a weaker role than syntactic information. Undoubtedly, most beginning learners fail to take advantage of semantic features (Lorsbach, 1982: 476).¹²

However, semantic development and conceptual development are closely intertwined (Gopnik & Meltzoff, 1986: 1052). Since familiar words usually develop stronger semantic representations, children with poor semantic system know fewer vocabulary words (Nation, Marshall, and Snowling, 2001: 255-256). Therefore, the method of combining semantic and phonemic processes can be recommended and applied to vocabulary teaching or learning (Nassaji, 2003: 261).

2.2.3 Morphological Features

Most English words are composed of two components: bases and affixes, which can be regarded as the “access codes”, through which one gains access to the

¹⁰ Lorsbach (1982: 479) also proposes similar findings.

¹¹ That is, the words paired with semantic related ones (e.g., apple, pear, peach....) are learnt with more difficulty than the words paired with unrelated ones (uncle, ice, market....) (Tinkham, 1997: 161).

¹² For similar detailed discussions, see Hammill et al., 2002: 121; Nassaji, 2003: 261.

information about the word during the process of recognizing new letter strings (Laudanna et al., 1994; Taft, 1994). Evidently, morphological relations of bases and affixes can facilitate the process of lexical acquisition. However, it is a great pity that English learners have paid little attention to morphology (Freyd & Baron, 1982: 293).¹³ Here are some reasons why.

First, semantically unrelated words with formal similarity (the so-called “deceptive transparency”¹⁴, e.g., *card/car*) easily confuse learners. The utilization of morphology may lead to false interpretation when words look as if they could be analyzed, but in fact they cannot (Laufer, 1990a: 151; Laufer, 1990b: 299). Besides, the similarity may slow down the recognition of the lexical units (Altarriba and Mathis, 1997: 563).¹⁵ Second, most morphological words are more than the sum of their parts. As a result, learners may not derive the correct meanings of the unknown words even if they know each morpheme (White, Power, & White, 1989: 302). Third, readers lack the knowledge of how to strip off the letter strings (Laufer, 1990a: 154; White et al., 1989: 287).¹⁶ Finally, the whole-word form procedures sometimes result in better facility with vocabulary memorization than the decomposed form procedures (Beauvillain, 1994; Zwitserlood, 1994; Taft, 1994).

¹³. For detailed discussions, see Sandra, 1994; Smith, 1995; Rueckl, Mikolinski, Raveh, Miner, and Mars, 1997; Mahony, Singson, and Mann, 2000.

¹⁴. For similar discussions about the deceptive transparency, see Zwitserlood, 1994; Rueckl et al., 1997,

¹⁵. Nagy, Anderson, Schommer, Scott, and Stallman (1989: 276) and Cole, Segui, and Taft (1997) also propose the same findings.

¹⁶. For example, *infallible* may be mistakenly interpreted as “something that cannot fall” (Laufer, 1990a: 154) or *unassuming* means “modest” rather than “not assuming” (White et al., 1989: 287). Moreover, for other similar discussions, see Taft, 1994; Cunningham, 1998.

2.2.3.1 Three Models of Morphologically-based Word Analysis

Nagy and Anderson (1984: 327) indicate that through morphological information, readers guess the appropriate meanings of the unfamiliar English words through context. Therefore, morphology can be the magic key that unlocks the meanings of hundreds, even thousands of words (White et al., 1989: 284). Chialant and Caramazza (1995: 58-65) conclude that to process morphological words successfully, English learners can use any of the three following models:¹⁷

- (1) Fully Non-decompositional Model: The whole-word representations are the only units of access in the lexicon (Hildreth, 1961: 110);
- (2) Morphologically Decomposed Model: A pre-access process of affix stripping precedes the access to the lexicon. All the morphemic units, such as roots, prefixes, and suffixes, are represented independently (Taft & Forster, 1975: 638);
- (3) Caramazza's Augmented Addressed Morphology Model (AAM Model): The mixed model combining the positive features of both the whole-word and the decomposed representation models provides a more reasonable theory of morphological process (Caramazza, Laudanna, & Romani, 1988: 300).

2.2.3.2 Frequency and Familiarity of Morphemes

Taft (1979: 263) assumes that the word “unhook” can be accessed by stripping off the prefix “un” and stem “hook”; nonetheless, “persuade” and “dissuade” can be accessed through the same lexical entry “suade” or separate lexical entries for “suade.” Consequently, some empirical researches¹⁸ conducted by Taft (1979) and Laudanna et al. (1994: 298) show that affixed words with a higher cumulative frequency (Word frequency < Morphemic frequency) seem to be represented in decomposed forms; on the contrary, affixed words with a higher surface frequency

¹⁷. Laudanna, Burani, and Cermele (1994: 296-297) and Cole et al. (1997: 312-313) also mention and illustrate the three models respectively.

¹⁸. Cole et al. (1997), Meunier and Segui (1999), and Hay (2001) also conduct similar studies and reach the same conclusion as Taft's.

(Word frequency > Morphemic frequency) seem to be accessed through their own representation. Namely, the stems of the whole-word frequency cause a null effect without leading to any facilitation or inhibition of memory (Meunier & Segui, 1999: 59). This phenomenon, thus, easily results in the rote memorization of the derived words with higher surface frequency (Cole et al., 1997: 321; Hay, 2001: 1062).

High frequency words prove more identifiable than low frequency ones.¹⁹ For instance, nurses respond faster to medical words than to engineering words, and engineers respond faster to engineering words than to medical words (Gardner, Rothkope, Lapan, & Lafferty, 1987: 26-28). Besides, Nelson-Herber (1986: 629) points out that technical vocabulary items always cause difficulty for learners not only because of its unfamiliarity and multi-syllable but also because of its collocation with other unfamiliar technical words. That is, familiarity and frequency affect the speed of access to the word (Jacoby, 1991: 530)²⁰ and determine the ease or difficulty of the lexical item (Sandra, 1994: 246). Therefore, only through either familiarity or unfamiliarity with the word and its meaning can the readers determine whether the word is easy or difficult (Hildreth, 1961: 109).

“Retention is enhanced by repetition, and repetition reduces false recognition and produces familiarity” (Seamon et al., 2002: 323). Owing to familiarity, morphology can facilitate recognition and production of a new word (Laufer, 1990b: 299) and lead to successful reading even without any strong contextual supports (Zutell, 1998: 233). Not surprisingly, Gardiner and Java (1991: 621) have found that

¹⁹. Meunier and Segui (1999: 58) also find the same empirical result.

²⁰. Hay (2001: 1065) also issues the similar statement.

with the aid of familiarity, recognition memory indicates little forgetting over the first 24 hours and usually lasts for more than 6 months.

“Few teachers know that those who possess the 2,000 most frequent words can perceive 95.38% of the text while the 4,500 least common words in the list of 10,000 most frequent words account for only 0.47% of the material” (Hildreth, 1961: 118).

This means that most English words do not appear frequently in text. Moreover, most of the least frequent words are analyzable and more numerous while most of the frequent words are simple and cannot be analyzed (Nagy & Anderson, 1984: 327);²¹ hence, morphological analysis is needed. Furthermore, Nagy et al. (1989: 264) even emphasize the importance of the frequency of the word and that of its morphological family.²²

2.2.3.3 Developmental Changes of Morphemes

English words with Latin or Greek affixes are more difficult for children to process because they store words according to the sounds, not the meanings; conversely, the learners who are older or who have higher English proficiency organize words with related meanings, not with the sounds (Hildreth, 1961: 129). That is, beginning learners are inclined to rely more on phonemic features than the older ones are (Hasher & Clifton, 1974).²³ Consequently, younger learners have a truly limited vocabulary which is expanded after they start to learn affixes and roots at the intermediate stage (Twaddell, 1973: 64-65).

²¹. White et al. (1989: 290) provide the statistics that affixed words outnumber the basic words at a ratio of almost 4:1.

²². Taft (1994: 277) and Cole et al. (1997: 313) also issue similar statements.

²³. For more similar discussions, see Naron, 1978; Fischler and Bloom, 1979; Freyd and Baron, 1982; Mahony et al., 2000.

It is, thus, suggested that the morphological instruction of affixes and roots begin in the third grade (Singson et al., 2000: 241-242; Carlisle, 2000: 169) or in the fourth grade (White et al., 1989: 303). It is at these grade levels that phonemic awareness gradually ends its contribution to vocabulary learning while morphemes begin to play a more important role in spelling development and vocabulary growth.²⁴ In addition, Carlisle (2000: 169) emphasizes that the contribution of morphological analysis is higher for fifth-grade students than for third-grade ones. Thus, Cunningham (1998: 194) indicates that a good speller should abandon a “memorize the letters” strategy (orthography) or a “write down the letters that stand for the sound” strategy (phonology) but develop a better strategy (morphology) -“What words may be related to this word?”

2.3 Vocabulary Teaching Strategies

Vocabulary teaching may be performed meaningfully or meaninglessly. Meaningful teaching means relating the new material or ideas to the cognitive structure already possessed by learners (Ausubel & Robinson, 1969);²⁵ thus, new information can be perceived quickly and more efficiently and can be retained longer (Hildreth, 1961: 112). Conversely, without any related cognitive structure, learners will find the learning material difficult and then meaningless learning may take place (Higa, 1965: 171).²⁶ Therefore, Dole, Sloan, and Trathen (1995: 452) emphasize

²⁴. Jacoby and Goolkasian (1973: 331) have found out that short-term memory is always sound-oriented based on the immediate recall tests and the delayed recall tests of the examinees.

²⁵. The cognitive structure already possessed by learners is also called “background knowledge.” For detailed discussions, see Laffey and Laffey, 1986; Thelen, 1986; Brown, 2000.

²⁶. Stahl (1986: 664) also states that only linking the L2 word form to its L1 equivalent without a greater connection between known and unknown, “deep” processing will not be constructed.

that the most effective vocabulary instruction is the kind that also improves comprehension, which, according to Nelson-Herber (1986: 629) or Johnson and Johnson (1986: 622), requires the meaningful integration of what is known and what is new.

For meaningful teaching and learning, Ausubel and Robinson (1969: 53)

propose the following three conditions:

- (1) The material to be learned must be capable of being related;
- (2) The learner must have relevant ideas to which this new material can be related;
- (3) The learner must intend to relate those new ideas to those known to him/her.

No wonder, Cornu (1979: 262-272) assumes that vocabulary teaching and learning should entail not only the correct presentation of meanings of words but also the ways to improve meaningful retention rather than mere assimilation of vocabulary lists.

As to vocabulary teaching and learning strategies and vocabulary size, they are highly correlated empirically (Gu & Johnson, 1996: 646). The more proficient students use various kinds of strategies more often than the less proficient ones (Fan, 2003: 233). Fan also synthesizes three types of strategies for vocabulary teaching and learning as follows:

- (1) Strategies which are perceived to be useful, often used, are used significantly more often by the most proficient students;
- (2) Strategies which are perceived to be useful and seldom used, but found to be related to high vocabulary proficiency;
- (3) Strategies perceived as not too useful and seldom used, but used significantly more often by more proficient students than by students with lower vocabulary proficiency.

Moreover, Foil and Alber (2002: 132) indicate that the strategies of vocabulary

instruction should incorporate the following four important principles:

- (1) employing a variety of methods for teaching vocabulary,
- (2) [designing] activity involving students in vocabulary instruction that endeavors to facilitate deeper levels of understanding,
- (3) providing instruction that enables students to see how target vocabulary words relate to other words, and
- (4) providing frequent opportunities to practice

reading and using vocabulary words in many contexts to gain deep and automatic comprehension of those words.

In summary, our educational goals should be the promotion of retention and transfer. The former is the memory ability to remember material at some later time and the latter is the ability to understand and use what was learned to solve new problems (Mayer, 2002: 226-228). To achieve these educational goals, teachers can help students facilitate their vocabulary memorization via the instruction of strategies. Subsequently, the four prevalent teaching strategies are presented as follows:

2.3.1 Traditional definition-based Vocabulary Teaching

The traditional definition-based vocabulary teaching method includes defining words, memorizing definitions, and assessing with matching and multiple-choice tests (Lisbeth, 2002: 314; Monroe & Orme, 2002: 140). Basically, the method is just memorizing key facts (Mayer, 2002: 227). However, without any vocabulary memorization techniques, rote learning emerges. Rote learning, like the process of acquiring isolated blocks, does not build relationships between the blocks (Brown, 2000: 83-84). Furthermore, Twaddell (1973: 76) has even argued that rote memory always leads to quick forgetting because it helps only short-term memory or some artificial purposes.²⁷

Chinese students are usually inclined to adopt the definition-based method of vocabulary learning. Cortazzi and Jin (1994: 15) state that Chinese students get more accustomed to the definition-based vocabulary learning, attributing this to “Confucian

²⁷. Prince (1996: 486-487) also statistically proves that rote memory always leads to quick forgetting.

values.” Woodard (1998: 7) also states that Asian students easily apply rote memory of facts because of the Confucian-style education of producing “test-aholics.” Therefore, without any vocabulary learning techniques, Chinese students are found to prefer oral repetition strategy, which is considered the strongest negative predictor of both vocabulary size and English proficiency (Gu & Johnson, 1996: 654). Empirically, Kang (1995: 52-53) indicates that the traditional Paper-and-Pencil group (i.e., a translation, definition-based approach to vocabulary manipulated by an instructor), scores the lowest in almost all the sessions for the three evaluation tasks (recall tests of vocabulary definitions, listening comprehension, and knowledge transfer) among all four groups (the Paper-and-Pencil, the Computer-based Word-for-Word, the Computer-based word-for-word plus Picture, and the Computer-based Context).

Nevertheless, rote learning still has its place in vocabulary instruction. Hung (1991: 160) says that rote learning is valuable at the beginning level. Moreover, Nation (1982: 160) claims that most teachers greatly underestimate learners’ capacity for vocabulary learning and memory; he also claims the following benefits of rote vocabulary learning: (1) Learners can memorize a large amount of vocabulary in a short period of time; (2) Learners can still have good retention and memory.

Rote learning needs constant repetition and multiple exposures to the words (James, 1952: 438; Karen, 1991: 5-6). Expectedly, repetition contributes to retention (Tussing & Greene, 1999: 523; Seamon et al., 2002: 323), but time is limited. Graves and Prenn (1986: 598-599) emphasize that the out-of-class time for teachers to prepare vocabulary instruction and the in-class time for students, with or without the aid of the teachers, are deficient and limited. In addition, Thelen (1986: 604) reiterates the impossibility for schools to teach all knowledge to students. No wonder, Foil and

Alber (2002: 138) state that merely requiring students to rehearse the definitions of words can become a tedious and time-wasting task. Obviously, the definition-based method appears ineffective and inefficient for vocabulary memorization.

2.3.2 Vocabulary Teaching in Context

Vocabulary knowledge and context comprehension are closely related (Nelson-Herber, 1986: 626). That is, the text-based vocabulary instruction can provide cumulative and varied exposures to target words that lead to more predictable and effective retention (Wesche & Paribakht, 2000: 208). By reading contexts, children can learn words in school settings in pretty much the same way they do at home (Miller and Gildea, 1987: 96). Thus, those who memorize words via the contextualized strategy always score high (Gu & Johnson: 1996: 668).²⁸ This is why students are encouraged to read newspapers, stories, magazines etc to expand their vocabulary size.

As to the selection of context, Harmon (1998: 519) suggests that students can read self-selected books at their own pace and for their own enjoyment. Foil and Alber (2002: 138) recommend that students be provided with personally meaningful context for using new words.

By combining contexts with the following strategies, learners can facilitate their vocabulary memorization. Knight (1994: 295) and Fan (2003: 234) state that reading, with or without a dictionary, is a useful way to increase vocabulary. Harmon (1998: 524) and Lisbeth (2002: 314) empirically prove that the integration of both contextual

²⁸Fan (2003: 233) also finds that only a small group of students score the highest in vocabulary learning just through extensive reading.

information and class-discussion can enhance the acquisition and retention of new words. Monroe and Orme (2002: 141) indicate that the strategy combining meaningful context approach with direct teaching method is found to be useful in teaching mathematic vocabulary.

However, guessing from context²⁹ is theoretically sound, but empirically opposed. Hopkins and Bean (1999: 279) state that students are struggling with inferring contextual word meanings.³⁰ Hulstijn, Hollander, and Greidanus (1996: 336) indicate that students often ignore unknown words or unsuccessfully infer their meanings of words in context-learning situations. Based on the following statement of Huckin and Haynes (Cited in Swaffar, 1995: 178-179), the use of contextual guessing by L2 learner is distinctly problematic.

(1) Problems in misrecognition lead readers to short-circuit the contextual guessing process; (2) Misuse of context is often due to inadequate grasp of the L2 vocabulary; (3) Inadequate grasp of the L2 passage context prevents readers from being able to fully define a word; (4) Inadequate grasp of the unfamiliar, nonlinguistic contexts often changes word meanings in passages; (5) Inadequate information is available to researchers about the degree to which guessing promote learning.

No wonder, Knight (1994: 295) provides the empirical evidence that subjects who guess from context not only learn less words but also achieve lower reading comprehension scores than those who use the dictionary only.

2.3.3 Key-Word Vocabulary Teaching

The key word technique, used to link the key term to another word that resembles the new one in form, meaning, or both, can facilitate memorization of

²⁹ Krashen (1989: 440) assumes that vocabulary is best acquired by guessing from context through the act of reading itself.

³⁰ For more similar discussions, see Raptis, 1997: 566; Monroe and Orme, 2002: 140; Prince, 1996: 488-489.

vocabulary items (Laufer, 1990a: 153). For instance, the mnemonically trained students empirically outperform the students with the instruction of definition-based or drill-and-practice method (Mastropieri, Scruggs, & Mushinski Fulk, 1990).³¹ Besides, the key word learners statistically benefit more from practice than do semantic-context learners (Wang and Thomas, 1995: 470). Even students with low receptive vocabulary³² who are assigned to the key word method can learn more words than those with high receptive vocabulary assigned to other methods (i.e., picture context, sentence-experience context, or learning independently) (Conduis, Marshall, and Miller, 1986: 612). Above all, the key word technique, in terms of association of form and meaning, is similar to the analysis of roots and affixes (Nation, 1982: 27). The familiar forms of the key word technique or of morphological chunks will enhance the recall and storage of the lexicon (Nattinger, 1988: 68).³³ The key word method especially benefits students with language disabilities (LD) (Mastropieri et al., 1990: 96). Here are some reasons why.

- (1) They do not need prior knowledge to optimize vocabulary learning;
- (2) The encoding of acoustic similarity is more effective with students with LD;
- (3) The key word method benefits students with LD who have more difficulty with abstract vocabulary;
- (4) The key word method provides students with LD with visual images.

Thelen (1986: 604), however, does not believe that the mnemonics of a key word is meaningful learning because no comprehension is involved. Furthermore, Hell and Mahn (1997: 534) state that a key word should sound or look like or be

³¹. For more detailed discussions, see Avila and Sadoski, 1996; Carney and Levin, 1998; Gruneberg, 1998; Uberti, Scruggs, and Mastropieri, 2003.

³². Crow and Quigley (1985: 510) also state that the teaching of key word does contribute to the learning of receptive vocabulary.

³³. Sharkin, Mohr, and Glover (1983: 251) also assume that the key word method can be used to significantly enhance the short-term memory and long-term memory.

semantically related to the target word, or the key word method may result in poor and delayed recall as well as rapid forgetting. This is why those who are assigned to the key word method score lower than those who receive the traditional definition-based method.³⁴ Moreover, the key word method is not suitable for the students under the age of 6-8 because of their lack of image creation (Avila & Sadoski, 1996: 392). No wonder, Lawson and Hogben (1996: 117) conclude that very few students (only 3 students) attempt to use the key word method for acquiring the meanings of new words. With such disadvantages, the mnemonic key word method cannot be applied to vocabulary teaching and learning.

2.3.4 Morphologically-based vocabulary Teaching

Denning and Leben (1995: 91) emphasize that “those who are skilled in both allomorphy (variation in form) and polysemy (multiplicity of meanings) will much better be able to recognize the forms and meanings of unfamiliar words.”³⁵ The application of morphologically-based orthographical similarity and polysemous core-meaning will be presented in the following sections.

2.3.4.1 The Application of Morphologically-based Orthographical Similarity

The organization of morphologically complex words is not viewed as meaningless separate letters (e.g., M+O+R+T+A+L), but meaningful morphemes (MORT+AL) (Beauvillain, 1994: 333). Evidently, morphologically related words can be identified and recognized by means of morpheme analysis. Thus, words similar in forms or meanings are easier to learn than those words without any similarity in forms

³⁴. Thomas and Wang (1996: 330) also conclude the same findings.

³⁵. Laudama et al. (1994: 298) also state similar statements about allomorphy and polysemy.

or meanings (Fischler & Bloom, 1979: 18; Nation, 1982: 19). Nagy and Anderson (1984: 332) indicate that there are more than 400,000 distinct words in printed school English (Grades 3-9).³⁶ However, these words can be semantically grouped into 88,533 distinct word families (e.g., *hunt/hunter; red/redness*). Nonetheless, if readers have knowledge of more complex morphology, there will only be 54,000 distinct word families (e.g., *ball/bald/ballot*). No wonder, Singson et al. (2000: 223) claim that morphological capacity (only 54,000 distinct words) in reading is more important than vocabulary knowledge (400,000 distinct words). Thus, morphologically complex words can be processed more rapidly than common words of the same length because of their orthographical similarity. Conversely, a study made by Altarriba and Mathis (1997: 557) has found that without the knowledge of morphology, both novice learners (or bilinguals) and proficient learners (or monolinguals) may experience confusion in words with orthographical similarity.

The shared morphemes, used as the prior knowledge (i.e., schema) to learn the new words with the same roots (Nelson-Herber, 1986: 627), facilitate vocabulary learning and memory (Freyd & Baron, 1982: 284).³⁷ Furthermore, Feldman and Prostko (2002: 14-23) indicate that the facilitation of morphological relatedness is much greater than that of semantic relatedness or orthographical relatedness; obviously, semantic and orthographical relatedness contribute little to long-term memory. Furthermore, Derwing et al. (1995: 20-21) state that orthographical

³⁶. These 400,000 words exclude proper nouns, numbers, and foreign words.

³⁷. Sandra (1994: 266) and Roelofs (1996: 863) also regard morphemes as schema to learn new words with the same roots.

similarity even produces inhibition rather than facilitation, without any morphological relatedness.³⁸

English is a morphologically related language because many English words are derived from Latin or Greek roots and affixes. Therefore, most words are closely related to Latin or Greek words in spelling, meaning, or pronunciation (Cunningham, 1998: 205). Thus, the examinees from Roman language backgrounds score higher than those from non-Roman backgrounds (Schmitt et al., 2001: 68-69); speakers of Western European languages appear more likely to use the similarity of cognates to infer the meanings of their unfamiliar words (Nation, 1982: 19). With the advantage of morphologically-based orthographical similarity, learners can memorize more related words and need less time to process difficult words (Rueckl et al., 1997: 83). This is why Derwing et al. (1995: 5) claim that the relationship between spelling and morphological knowledge is necessarily and definitely reciprocal.

2.3.4.2 The Application of Polysemous Core-meaning

Klein and Murphy (2001: 263) indicate that “the senses of the polysemous words are closely related because of the shared core-meaning.”³⁹ With the shared core-meaning, readers do not need much imagination to find the relationships among the meanings of the same morpheme (Denning & Leben, 1995: 93). Thus, polysemy can be predicted from every general cognitive principle and then words can be more easily kept in mind (Lehrer, 1990: 211). Empirically, Sandra (1994: 267) provides his

³⁸. Pastizzo and Feldman (2002: 39) also issue the similar statement.

³⁹. For example, *macro-* means ‘*long*’ and ‘*large*’; *path* means ‘*feel*’, ‘*suffer*’, and ‘*illness*’. Each morpheme has one shared core-meaning.

conclusive evidence that the semantically transparent derived words can be more easily learned through morphological structures.

The knowledge of a word entails the memorization of its meanings. However, a word usually has various meanings,⁴⁰ which may cause inhibition in vocabulary memorization if their meanings are unrelated. However, the facilitation of vocabulary memorization may emerge if their meanings are polysemously related. That is, not only can gaining meanings via polysemous relatedness lead to language comprehension⁴¹ (Laufer, 1990b: 302-303) but also it can reduce the amount of information to be stored (Sandra, 1994: 229-230).

2.4 Theoretical Framework of Vocabulary Teaching

The theoretical framework of vocabulary instruction combines the effects of both visual aids and morphological knowledge. The first part discusses the neurological bimodality learning and visual aids. The last part introduces the theoretical framework of vocabulary teaching: the iconic-morphological approach via commonly-used roots.

2.4.1 Neurological Bimodality Learning and Visual Aids

Karen (1991: 13) states that learners have different learning styles. Some learners are auditory, some tactile, others kinesthetic, or visual. However, visual learners are disadvantaged because most material is taught and learned verbally.

⁴⁰. The polysemous word has two or more related meanings, which can facilitate the memorization of meanings. However, the homonyms are different words happening to share the same form. The meanings of homonyms will be stored separately since they have no core-meanings and are unrelated to one another and unpredicted (Klein & Murphy, 2001: 260; Laufer, 1990b: 303).

⁴¹. Lehrer (1990), Denning and Leben (1995), Mahony et al. (2000), and Klein and Murphy (2001) also state that the polysemous related senses contribute much to comprehension and memory.

Verbal memory and visual memory are represented separately in the left-brain and the right-brain, respectively (Danesi, 1988: 17).⁴² The right hemisphere is associated with visual, tactile, and auditory images, while the left hemisphere is concerned with logical and analytical thought (Brown, 2000: 118). Additionally, the language functions of approximately 98% of right-handed individuals and almost 70% of left-handers are located in the left hemisphere. Obviously, people make use of left-brain processes more and almost ignore the functions of the right-brain. Therefore, the proposition and utilization of “neurological bimodality” can be effective and efficient in a complementary and cooperative function (Danesi, 1988: 13-15).⁴³

In addition, visual aids can benefit students because of the cooperation of the right and left hemispheres. Ausubel and Robinson (1969: 59) give a good example: initially, the word “dog” means nothing to the child without the association of the visual image of the dog. A study conducted by Levin and Kaplan (1972: 431) has found that images and pictures can help memory,⁴⁴ so words enriched by visual associations or images can be retained much longer in the memory. Undoubtedly, dictionaries with picture-word association are popular with young readers (Hildreth, 1961: 130) and the translation accompanied by a corresponding picture is superior to any of the alternatives (Nation, 1982: 25).

⁴². For similar discussions, see Karen, 1991: 13; Brown, 2000: 54.

⁴³. Karen (1991: 24) also emphasizes the cooperation of brain’s left and right hemispheres.

⁴⁴. Karen (1991: 22), Price and Finkelstein (1994: 34), Hung (1991: 160), and Plass et al. (1998: 30) also have similar findings about the effects of images or pictures in vocabulary memorization.

Empirically, Pressley (1977: 471) mentions that the second-grade subjects utilizing images remember 103% more lexical items than the key word control subjects and 189% more items than the no key word control subjects. The fifth-grade subjects utilizing images remember 92% more items than the key word control subjects and 83% more items than no key word control subjects. In addition, Pressley and Levin (1978: 367) indicate that key word users provided with pictures remember many more new vocabulary items (61.84%) than those provided with only pictures (28.95%) or with only words (13.60%).

2.4.2 An Iconic-Morphological Approach via Commonly-used Roots

Each Chinese character is an individual picture or icon since Mandarin Chinese uses a logographic writing system. Because of the writing system, the Chinese are more inclined to think and memorize by using the brain's right hemisphere. That is, Chinese students naturally regard an English word as an undivided icon,⁴⁵ just as they are accustomed to making "sound-to-symbol" correspondence of Chinese characters (Huang, 1997: 323; Everson, 1998: 201). Therefore, readers of logographic writing systems favor lexical access through direct recognition of word forms (Grabe, 1991: 387-388). Koda (1989: 214-215) also indicates that the Japanese subjects perform best in the graphically similar/phonologically dissimilar sets.⁴⁶ In summary, lacking the knowledge of morphological decomposition, Chinese students are more likely to regard an English word as a whole mass.

⁴⁵. English is an alphabetic writing system, different from the logographic writing system.

⁴⁶. Japanese and Chinese use the logographic writing system.

By means of morphological analysis, the meanings of a word can be inferred from its parts. Hildreth (1961: 114-115) mentions that readers should pay more attention to foreparts⁴⁷ and distinctive features, just as the listeners should not listen out for every syllable of words lest the main idea of the whole context be lost. For English vocabulary learning, the foreparts of words are the crucial clues and the consonants in the middle or ends can sometimes be ignored. Besides, Nagy and Anderson (1984: 315) estimate that at least 80% of the meanings of affixed words are what their parts suggest.⁴⁸ Therefore, readers are better able to memorize the derived forms (e.g., *solidify*, *solidity*, *consolidate*) from the underlying form (*solid*) (Richards, 1976: 80-81). The subjects, required to guess the meanings of words, intentionally neglect the abstract and difficult suffixes (even though they are quite frequent) and parse words into morphemes to find the roots (Freyd & Baron, 1982: 293).

Learners can facilitate their vocabulary memorization by tolerating the “orthographically or phonologically slight alteration and vagueness” (Twaddle, 1973: 77).⁴⁹ Thus, the application of allomorphs is proposed. For instance, in the pair of *sheep/shepherd*, *-ee-* and *-e-* alternate; *sheep* and *shep-* are allomorphs (Coates, 1999: 62-63). Similarly, some forms of the irregular past tense and their roots look alike (e.g., *made/make*; *swam/swim*) (Rueckl et al., 1997: 384). Even the meaning of “*jailbird*” can be inferred only from its part “*jail*” (Zwisterlood, 1994: 363).

⁴⁷. Laufer (1991: 328) also issues a similar statement that the initial elements of lexical items are salient and can be easily recognized and properly stored.

⁴⁸. The meanings of a word can be inferred from its parts (Zwisterlood, 1994: 366; Nagy et al., 1989: 264).

⁴⁹. For more similar discussions, see Wang, 1997: 124; Feldman and Prostko, 2002: 12.

Morphologically related words are usually related in forms (i.e., spelling and pronunciation) and meanings (e.g., *fire/fierce*; *space/spatial*; *message/messenger*), so spelling should be complementary to morphological knowledge (Derwing et al., 1995: 19; Rueckl et al., 1997: 399). To develop the habit of applying morphological information to vocabulary memorization, English teachers should always ask students the following questions:

“Do you know any words that look and sound like this word?”

(quoted from Cunningham, 1998: 204)

“Are any of these look-alike/sound-alike words related to each other?”

(quoted from Cunningham, 1998: 204)

“It looks something like another word I know.”

(quoted from Hildreth, 1961: 112)

“It has a word in it I know.”

(quoted from Hildreth, 1961: 112)

“Do you think that the word *teacher* ‘comes from’ the word *teach*?”

(quoted from Derwing et al., 1995: 7)

From the above questions, teachers can direct students to develop the morphological learning strategy in order to relate unknown words to known words. Consequently, students are capable of acquiring English lexical items on their own.

Mo (1993: 109; 1995: 204) refers to the application of sound-switching and the creation of mental images to memorize the lexical items with morphological relatedness (e.g., *crook/crookery/couch/crotch/crochet*). Moreover, Wang (1997: 563-564) and Chen (1999: 623) assume that learners can easily memorize a cluster of morphological related words (e.g., *bat/batter/baton/battle/combat/debate*) by using a combination of both morphology and phonetic symbolism. By making use of morphological references, learners can learn lots of derived words from the underlying forms while tolerating slight divergences in spelling and sound (Mahony et al, 2000: 194). No wonder, Derwing et al. (1995: 12-13) claim “those who can recognize the *cup* in *cupboard*, the *draw* in *drawer*, and the *lace* in *necklace* are

successful spellers.”

In this iconic-morphological approach via commonly-used roots, teachers can teach or students can learn English words with the aid of both morphological information and visual icons. However, there is an important limitation on the application of this approach to teaching or learning new words. Because some words are so simple and basic, they are hardly related to other commonly-used words with the same morpheme. In other words, students should memorize at least the basic 1,000-2,000 words. Then their capability of relating unknown words to commonly-used roots can be developed. After all, without any commonly-used roots, how can learners associate unknown words to known ones?