**2. Literature Review**

 To provide the theoretical basis of the study, this chapter begins by examining L2 vocabulary learning through reading. Next, the depth of processing theory which underpins the study is explored. Then, the framework, Involvement Load Hypothesis, is investigated as it is the first attempt to operationalize the depth of processing theory. A recently developed framework, technique feature analysis, is thoroughly discussed since the study observes its predictability in comparing vocabulary learning tasks. Finally, the impact of multimodal glossing and dual coding theory is discussed to provide background into the ideas that went into designing vocabulary tasks.

**2.1 L2 Vocabulary Learning through Reading**

 It is a widespread notion that L2 learners can learn vocabulary through reading (Waring & Nation, 2004). However, if reading text is too difficult, it is not possible for learners to actually read and guess from context to learn vocabulary. To determine how many words should be known for learners to be able to read, it is necessary to consider coverage rate of a text which refers to “the percentage of the vocabulary that is known by the reader” (Waring & Nation, 2004, p. 12). In early studies, 95% coverage was suggested (Laufer, 1989; Liu & Nation, 1985) but Hu and Nation (2000) suggested a higher rate of at least 98% for written discourse. In the light of the coverage rate, researchers have suggested adequate vocabulary size for reading. Nation (2006) suggested 8000-9000 word families to meet 98% coverage based on British National Corpus (BNC) to understand authentic texts such as a novel. A recent study proposed an optimal knowledge of 8,000-word families for 98% coverage and minimal knowledge of 4,000-5,000 word families for 95% coverage both including proper nouns (Laufer & Ravenhorst-Kalovski, 2010). However, Schmitt (2008) stated that even this figure of vocabulary size is underestimating the burden of vocabulary learning because it does not take into account of individual word forms including root form, inflections or derivations contained in each word family. To make matters worse, L2 learners generally have more possibility of encountering unfamiliar words in reading compared to first-language readers (Nagy, 1997). It is not surprising to hear that “learning such a large number of lexical items is one of the greatest hurdles facing learners in acquiring English” (Schmitt, 2008, p. 332). In other words, vocabulary learning can be quite a challenge for L2 learners. Without proper guidance, it may lead to unsuccessful foreign language learning. Therefore, to improve L2 vocabulary learning through reading, a more principled, systematic approach is crucial for researchers and language teachers.

**2.1.1 Incidental vocabulary learning**

 As this study concerns L2 vocabulary acquisition through reading, it is prerequisite to understand two approaches of vocabulary learning: incidental and intentional learning. Relevant to intentionality (Schmidt, 1994), researchers have long been trying to distinguish the term incidental and intentional learning in the area of vocabulary acquisition (Gass, 1994; Hulstijn, 2001). The terms can be elaborated in methodological and pedagogical meaning.

 In terms of study design, incidental learning refers to an experimental condition that requires participants to perform an instructional task without forewarning of an upcoming test while intentional learning allows learners to be aware of a test that asks for the recall of information they learned (Laufer & Hulstijn, 2001). Additionally, to find out which factors affect vocabulary learning, it is important for researchers to adopt an incidental learning design. In incidental vocabulary learning, researchers are capable of manipulating variables such as learning tasks or frequency of words to reveal the effectiveness of vocabulary learning. On the contrary, the intentional design does not allow them to control what learners actually do when they process words to put into memory. For example, no matter how carefully an experiment is designed to find out learners’ use of word part strategy in memorizing lexical items, it is impossible to ensure whether they actually use it or not. In other words, according to Laufer and Hultijn (2001), the nature of incidental learning enables researchers to have responsibility for empirical experiments, and eventually, contribute to vocabulary learning.

 In pedagogical view, incidental vocabulary learning is relatively a general term which incorporates “learning without an intent to learn” (Laufer & Hulstijn, 2001, p.10). It is directly connected explanations of vocabulary learning as the by-product of reading (Krashen, 1989). On the other hand, if any learning activity has a goal to memorize lexical items, it can be identified as intentional vocabulary learning (Hulstijn, 2001).

 Considering the two views, this study would focus on incidental vocabulary learning as the study design does not mention any vocabulary tests in advance. Additionally, reading along with learning tasks aims to comprehend a story without an explicit goal of memorizing target words.

**2.1.2 Empirical findings of incidental vocabulary learning**

 Researchers have generally agreed that in L1 contexts a large amount of vocabulary is acquired through reading (Krashen, 1985; Nagy, Herman, & Anderson, 1985; Nagy, 1997). Many studies of second language acquisition have also paid attention to seek advantage of incidental learning through reading for L2 learners. However, early studies showed quite discouraging results in incidental vocabulary acquisition through reading since it is a slow process with little gains of vocabulary knowledge (Hulstijn, 1992; Pitts, White, & Krashen, 1989). Researchers reviewed several studies with a conclusion of a low pick-up rate; roughly about one out of every twelve words tested (Horst, Cobb, & Meara, 1998). For example, reading a novel for an hour led to gain only two nadsat words out of 30 items (Pitts et al., 1989). Hulstijn (1992) found that learners learned one word out of twelve Dutch words after reading 907 words.

**2.1.3 Facilitating Vocabulary Learning**

Given the insufficient word gains from reading-only and the time limit in the classroom to memorize words explicitly, numerous studies have tried to discover what factors are involved in facilitating vocabulary learning. Some studies looked at the effect of gloss, specifically if it provides extra help for reading (Hulstijn, 1992; Hulstijn et al., 1996; Ko, 2012; Watanabe, 1997). For instance, Ko (2012) compared no-gloss, L1 gloss, and L2 gloss reading conditions on receptive vocabulary knowledge. The results illustrated that participants who read a glossed text outperformed those who read a text without gloss. It is assumed that providing explicit attention toward words by glossing can facilitate vocabulary learning (Schmitt, 2008).

 Several studies have investigated whether the use of a dictionary would make a difference in incidental vocabulary learning. In a study by Luppescu and Day (1993), learners who used a bilingual dictionary during reading appeared to perform significantly better than those who did not use the dictionary in the receptive vocabulary test. Knight (1994) also supported the use of dictionary showing better gains in vocabulary meaning test with Spanish students.

 There were studies examining the effects of various word learning tasks. Paribakht and Wesche (1997) compared a reading-only group who read several texts, and a reading plus group who did a variety of vocabulary exercises such as matching the definition exercise or multiple choice cloze exercise after reading. The results showed that the reading plus group learned more words than the reading-only group. Joe (1998) emphasized the importance of generative tasks that require learners to use words in original contexts. It was found that text-based tasks with a higher level of the generative process were more effective for incidental vocabulary learning than tasks with a lower level of the generative process. She argued that deeper engagement in tasks such as generation would be beneficial in vocabulary acquisition although there is no explicit vocabulary teaching. Production tasks were also beneficial for acquiring vocabulary (Atay & Kurt 2006; Newton, 2013; Webb, 2005). For example, it was shown that instruction with interactive post-reading tasks (e.g., retelling a story with target words) fostered vocabulary knowledge more effectively than instruction with discrete written tasks (e.g., matching target words with their definitions) (Atay & Kurt 2006).

 Taken together, many studies mentioned above have supported that L2 vocabulary acquisition through reading can be more effective when it is accompanied by various tasks rather than a reading-only condition. What is crucial now is to figure out which tasks are most influential for vocabulary learning (Hill & Laufer, 2003). According to Laufer and Hulstijn (2001), in terms of the effectiveness of tasks, researchers pointed out that the retention of vocabulary is conditional upon how deeply learners process lexical information based on the depth of processing hypothesis by Craik and Lockhart (1972).

**2.2 Depth of processing Theory**

In the field of learning and memory, depth of processing hypothesis was addressed by Craik and Lockhart (1972) to demonstrate factors in determining the retention of new information. They argued that the possibility of locating a new word in the long-term memory is decided not by the length of time being stored in short-term memory but by the degree of depth in processing new information. The processing was divided into levels in that phonological encoding was considered to be a shallow stage and semantic processing to be a deeper stage. In a nutshell, the more learners are involved in processing and manipulating lexical items; the more likely learners would be able to remember words (Lee & Hirsh, 2012).

 However, the hypothesis was criticized due to the ambiguity of the notion of depth (Baddeley, 1978; Craik & Tulving, 1975). Craik and Tulving (1975) pointed out that it failed to construct the independent measure of processing. Without the proper measurement, it was difficult to test its adaptability as a continuum of processing (Baddeley, 1978).

 Nevertheless, it provided a theoretical basis in the research of vocabulary learning (e.g., Brown & Perry, Jr, 1991; Joe, 1998; Stahl & Clark, 1987). In an attempt to provide an operational and measurable theoretical explanation, Laufer and Hulstijn (2001) proposed Involvement Load Hypothesis (ILH) based on the depth of processing theory.

**2.3 Involvement Load Hypothesis**

Although many studies had put an effort to explain various factors that would improve the effectiveness vocabulary learning (Joe, 1998; Knight, 1994; Paribakht & Wesche, 1997), it was necessary to establish a hypothesis that could explain the levels of lexical learning systematically and empirically. Therefore, Laufer and Hulstijn (2001) developed Involvement Load Hypothesis (ILH), a vocabulary learning framework, to operationalize the depth of processing theory.

 Three components constitute the construct of involvement loads: need, search, and evaluation. Each factor can be absent (0 points), moderate (1 point), or strong (2 points). The involvement loads will be quantified by the sum of the three factors. The researchers argued that the higher the involvement loads, the better the retention of vocabulary would occur.

 The first criteria need is a motivational construct which considers the “need to achieve” (Laufer & Hultijn, 2001, p. 14). The level is decided as moderate when the need for understanding a word is externally imposed by teachers. The need can be strong when the need for using the word is generated internally from learners. The cognitive construct search is the second component which can be discussed as moderate if the learner has an attempt to retrieve the meaning of a word. It is strong with an effort to look for the form of a word. The evaluation also reveals the cognitive dimension that includes assessing a certain word with other words. Moderate evaluation means to focus on comparing the meaning of a word to others while strong evaluation happens when choosing the best word that fits into a specific linguistic context. Overall, according to involvement load hypothesis, these three factors can be combined in diverse ways to generate involvement loads to predict successful retention of vocabulary.

 The hypothesis has been empirically examined by numerous studies mostly with adults (Folse, 2006; Kim, 2008; Hulstijn & Laufer, 2001). AN initial study by Hulstijn and Laufer (2001) supported the hypothesis with EFL young adults in Israel and Netherlands. The participants were provided three tasks with different involvement loads, the lowest in reading comprehension task, moderate in reading comprehension plus gap-filling activity, and the highest in composition writing task using target words. In lines with the hypothesis, the composition writing task with higher involvement loads resulted in the highest retention of vocabulary among three tasks. Kim (2008) also advocated the hypothesis by conducting two experiments on ESL adults with different proficiency levels. The first experiment lends support to the hypothesis as a task with a higher level of involvement (e.g., writing composition group) was superior to tasks with lower level (e.g., reading group and gap-filling group) on vocabulary learning. The second experiment showed similar results in different tasks with the same level of involvement loads in the immediate and delayed vocabulary post-tests.

 However, results were contradictory in some studies. Folse (2006) compared the effectiveness of three tasks consisting of various degree of involvement loads: one fill-in-the-blank activity, three fill-in-the-blanks activities, and one original sentence writing activity. Within controlled time-on-task, participants in the second task outperformed other tasks although sentence writing task was supposed to yield better results according to the original hypothesis. Keating (2008) pointed out that the task effects with different involvement loads diminished when time on task was considered. There was research that looked at the difference between input based tasks and output based tasks (Sarani, Negari, & Ghaviniat, 2013). Unlike the prediction, an output based task of sentence writing showed better results than an input based task of reading with multiple-choice gloss on vocabulary learning even though they represented a similar amount of involvement loads.

 With young learners, there are a few studies that have explored the theory. A recent study by Silva and Otwinowska (2017) examined the impact of the theory with three different combinations of tasks generating similar involvement loads to thirty-eight 10-year-old elementary EFL beginner level students in Poland. This study was meaningful in that it showed full support of the hypothesis with elementary level children. However, Alcaraz Mármol and Sánchez-Lafuente (2013) partially advocated Involvement Load Hypothesis with twenty-eight 10-year-old elementary EFL students in Spain. The sentence writing task with glosses outperformed the sentence writing task using a bilingual dictionary in both receptive and productive tests as opposed to the prediction of the hypothesis. The researchers attributed the results to participants’ low language proficiency or the complexity of tasks such as using a bilingual dictionary. Therefore, it seemed necessary to conduct such an experiment with careful consideration to young learners at a relatively low cognitive level.

With empirical findings, Involvement Load Hypothesis contributed to confirming the effectiveness of vocabulary learning tasks. Additionally, it can be concluded that not all activities guarantee equal quality of vocabulary learning within a different combination of involvement loads. Nonetheless, the hypothesis revealed limitations as previous studies showed contradictive evidence. Moreover, there was some contradicting evidence when time on task was considered. Therefore, it is necessary to investigate other existing frameworks of vocabulary learning. It was pointed out that Involvement Load Hypothesis with only three criteria lacked features to explain various aspects of vocabulary learning (Nation & Webb, 2011). To this end, Nation and Webb (2011) proposed Technique Feature Analysis (TFA) as a framework of vocabulary learning to help design effective vocabulary learning tasks.

**2.4 Technique Feature Analysis**

Technique feature analysis has been suggested by Nation and Webb (2011) in an attempt to operationalize the depth of processing theory as well as to compensate ILH by adding more criteria. Previously, Nation (2001) developed a vocabulary learning model with three components: noticing, retrieval, and generation. However, it was not quantified which made it difficult to assess diverse vocabulary learning tasks. To formulate more elaborative and quantitative design, motivation and retention were newly supplemented to TFA (Nation & Webb, 2011). Furthermore, in each component, specific criteria to quantify the elaboration are included. To sum up, there are five components and 18 criteria which lead to 18 scores in total to assess a vocabulary learning task. TFA is presented with its components and criteria in Table 1.

**Table 1: The checklist for Technique Feature Analysis**

**(adopted from Nation & Webb, 2011, p. 7)**

|  |  |
| --- | --- |
| **Criteria** | **Score** |
| **Motivation** |
| Is there a clear vocabulary learning goal?Does the activity motivate learning? Do the learners select the words? | 000 | 111 |
| **Noticing** |
| Does the activity focus attention on the target words? Does the activity raise awareness of new vocabulary learning? Does the activity involve negotiation? | 000 | 111 |
| **Retrieval** |
| Does the activity involve retrieval of the word? Is it productive retrieval? Is it recall? Are there multiple retrievals of each word? Is there spacing between retrievals? | 00000 | 11111 |
| **Generation** |
| Does the activity involve generative use? Is it productive? Is there a marked change that involves the use of other words? | 000 | 111 |
| **Retention** |
| Does the activity ensure successful linking of form and meaning? Does the activity involve instantiation? Does the activity involve imaging? Does the activity avoid interference? | 0000 | 1111 |
| **Total score** |  | 18 |

 Motivation has been recognized as a key factor in language learning (Dörnyei, 1994; Gardner, 2007). TFA describes motivation through three particular criteria; whether a task has a clear vocabulary learning goal, motivates learning, and provides learners opportunities to select target words. The second component is noticing which can be explained as support for learners to pay conscious attention to a certain word for successful vocabulary retention (Schmidt, 2012). A vocabulary task with attention to target words gets a score in TFA. Raising awareness of new vocabulary learning is another criterion for the component. Finally, if there is any possibility to negotiate with target words, a task gains a score. Retrieval refers to the process that strengthens the memory of a word through repetition (Nation, 2001). There is receptive and productive retrieval which may lead to the strong form-meaning link of a word (Baddeley, 1990). TFA suggests providing a point to retrieval factor and an extra point to productive retrieval. As Baddeley (1990) asserted, each opportunity to retrieve a word may be important for the form-meaning link, multiple retrieval gets another point. Generation is defined as a process to meet or use a word subsequently in a new context for reconceptualizing the meaning of a word (Nation, 2001). In TFA, an activity that includes generative use gets a point and productive generation is provided an extra point. Additionally, a point is given to an activity that involves the use of other words. Retention is described as four criteria in TFA including a successful form-meaning link, instantiation, imaging, and avoidance of interference. However, few studies investigate the effectiveness of vocabulary learning tasks from the perspective of Technique Feature Analysis.

In terms of comparing Involvement Load Hypothesis and Technique Feature Analysis, there are only two studies found. Nassaji and Hu (2016) conducted an experiment with 96 adult college-level students in Taiwan to compare four different vocabulary tasks based on the two frameworks to determine their predictive power of assessing vocabulary tasks. The results showed that TFA seems more conclusive in predicting the effectiveness of vocabulary tasks. There is another study that examined two frameworks in a different context with 90 Iranian high proficiency EFL students with various age from 15-25 (Gohar, Rahmanian, & Soleimani, 2018). In three vocabulary tasks, TFA was a better predictor than ILH in score change between pretest and posttest but not in the during-task performance.

There are hardly any studies that explore various vocabulary learning tasks based only on the TFA so far. Zou, Wang, Kwan, and Xie (2018) focused on the TFA framework with 125 undergraduate students to explore the effect of four vocabulary tasks combined with different types of annotation. The four tasks were reading comprehension with pictorial

annotations, cloze exercise with textual annotations, cloze exercise with pictorial annotations, and sentence writing with textual annotations. The researcher concluded that the results revealed TFA as a good determinant in probing various vocabulary tasks and pictorial annotation seemed effective for vocabulary learning. In sum, although there are not many studies that analyze vocabulary learning tasks with TFA, the results show consistent support for the framework.

Technique feature analysis has analyzed various factors that affect vocabulary learning and integrated them into a measurable index. Hence, it is necessary to investigate if it is a reliable indicator of effective vocabulary learning. Since the contexts examined so far are restricted from teenagers to adults with intermediate to advanced level of English, this study will explore vocabulary learning tasks through TFA with elementary school students. In addition, to specifically examine the effect of imaging in the criteria of retention in TFA, the researcher puts different types of glosses on the vocabulary learning tasks. The related studies of glossary based on dual coding theory will be reviewed in the next session.

**2.5 The effect of Multimodal Gloss for L2 Vocabulary Learning**

It seems like common sense to use glossary when materials need to provide extra information for unfamiliar words. A gloss refers to a brief explanation of a word in a text basically in the form of a definition or synonym (Nation, 2001). It has been considered to have benefits on language learning for several reasons. Glossing plays a role for language learners to deal with inefficient contextual cues while reading by representing lexical items with supplementary information (Ko, 2012). Additionally, it may enhance vocabulary learning through multiple exposures to lexical items when learners repeatedly pay attention to glosses to understand unknown words during reading (Watanabe, 1997).

 Given the positive impact of glossing on L2 vocabulary learning, researchers have examined whether various types of glossing have a differential effect on the retention of vocabulary. Some studies compared the effect of L1 and L2 glosses (Jacobs, Dufon, & Hong, 1994; Ko, 2012; Laufer & Hill, 2000; Yoshii, 2006). Additionally, some studies investigated the influence of multimodal glossing which can be defined as providing visual information such as pictures or illustrations with textual glossing (Acha, 2009; Boers, Warren, He, & Deconinck, 2017; Kost, Fost, & Lenzini, 1999). This line of research will be explained specifically as the study is related to multimodal glossing effects on L2 vocabulary acquisition.

 Many studies have talked about the benefit of adding pictures to textual glossing on vocabulary learning in computer-assisted settings (Akbulut, 2007; Chun and Plass, 1996; Plass, Chun, Mayer, and Leutner, 1998). However, Boers et al. (2017) pointed out that the computer-assisted contexts might be allowed to access pictorial and textual glossing separately by mouse-clicking, leading to more exposures than expected. Indeed, it is in doubt whether the positive impact is solely due to multimodality of annotation or multiple exposures to lexical items.

 However, the studies that combined textual and pictorial glossing in a single gloss will be discussed, since the present study follows the similar condition in a pen and paper setting. Kost, Fost, and Lenzini (1999) investigated university students learning German as a foreign language to find out whether L1 glossing, pictorial glossing, and a combination of L1 and pictorial glossing would benefit differently on incidental vocabulary learning. It was conducted in a pen and paper condition. The result of vocabulary learning was tested through three post-tests: providing L1 translation to L2 words, matching pictures to L2 words, and matching L1 translation to L2 words. It did not reveal significant differences among the three groups in the L1 translation test. The researchers attributed this finding to the lack of frequency of target words. However, the scores of a combination of textual and pictorial glossing were the highest in the third test (matching L1 translation with L2 words) suggesting a positive influence of multimodal glossing on receptive word knowledge. Moreover, there were studies conducted in computer-assisted environments that showed pictorial and textual glossing at the same time. As a replication study of Kost et al. (1999), Yoshii and Flaitz (2002) compared three conditions of glossing in the multimedia setting: text-only, picture-only, and text with a picture of a word. In accordance with the test format, the text with picture condition resulted in better scores than the text-only condition in a picture recognition test. The picture-only group gained relatively lower scores in a matching definition test and a providing meaning test. Additionally, the multimodal glossing group (text plus picture) outperformed those given single-mode glossing in a supplying definition test. Overall, the result showed that multimodal glossing might be conducive in acquiring receptive L2 vocabulary knowledge.

 In contrast, not all results showed support for the effectiveness of multimodal glossing. In Boers et al. (2017), university students from various countries were required to read an 800-word narrative text which contained six pseudo-words with marginal glosses. L2 learners were assigned to one of two glossing conditions: text-only glosses and multimodal glosses (textual information combined with pictures). Three trials were performed where trial 1 and 2 provided glossing with English (L2) translation while trial 3 with L1 translation as a textual gloss. To tackle both the receptive and productive vocabulary knowledge, a meaning recognition test and a form recall test was given. In all trials, adding pictures as support of textual gloss did not reveal statistical significance compared to text only gloss on the recall of meaning. Regarding form recall, the findings of all three trials showed the discouraging effect of adding visual information to textual gloss compared to text-only gloss. The researchers attributed the result to split-attention which may have occurred in multimodal glossing where learners were hindered to pay enough attention to the given words as the pictures needed attention as well.

Another study was conducted with ESL adult learners who were assigned to one of three conditions: text-only, picture-only, and text with picture glossing (Warren, Boers, Grimshaw, & Siyanova-Chanturia, 2018). It should be noted that the textual information was given in English (L2). Two types of posttest, a meaning recognition test, and a form recall test were analyzed for the acquisition of receptive and productive vocabulary knowledge. In the form recall test, although the picture-only group gained better results followed by a multimodal and text-only condition, the difference between the three groups did not appear to be statistically significant. Interestingly, in the meaning recognition test, the picture-only condition was superior to other conditions even though the statistical difference was revealed only between picture-only and text-only. Overall, providing a picture to L2 words without definition seemed beneficial in meaning retention of vocabulary while the effect of multimodality in gloss was not conclusive unlike the results of previous studies. The researchers explained that in multimodal condition, the difficulty of interpreting textual information might have hindered elucidating meaning from the picture whereas single-mode condition allowed full attention to the meaning given by one type of gloss. Another possibility was that processing visual information without text could be more engaging to learners, which may lead to better retention of vocabulary. In particular, through eye-tracking, it was revealed that the amount of attention given to words and pictures were not significantly affected by the type of gloss. When conducted with young L2 learners, it also showed a negative perspective on the effect of multimodal glossing (Acha, 2009). However, it should be noticed that this research was implemented in a multimedia setting. Students were divided into three groups where they received text-only glossing (L1 translation), picture-only glossing, and text-plus-picture glossing. A group of text-only glossing gained the best scores whereas the text-plus-picture glossing group gained the lowest scores in meaning recall tests of immediate and two-week delayed posttest.

 In sum, the research findings indicate that the effect of multimodal gloss seems to be unsettled yet requiring further examination in future studies. It is necessary to take a closer look at the ‘dual coding theory’ as it is theoretically related to processing visual and textual information altogether.

**2.6 Dual Coding Theory**

One of the well-established theories that considers the relationship between cognition and image is the Dual Coding Theory (Paivio, 1971). The initial theory was explained within unilingual contexts. It is a theory of human cognition which is mediated by dual code; verbal system and nonverbal system (Paivio, 1971). The verbal system can be described as the representation and processing of information related to language including speech and writing (Paivio, 1986). The nonverbal system, also called an image system, is conceptualized as processing information related to nonverbal objects and events, and creating mental images of the situations.

 According to the theory, the verbal system and nonverbal system work in both an independent and interconnected way when processing information in the brain (Paivio, 1971). The feature of independence refers to the possibility that one system can operate without the other during cognitive activities. For instance, we can perceive nonverbal objects only with an image by not thinking of the language (Paivio & Desrochers, 1980). The interconnectedness between two systems can be understood when verbal instruction activates a mental image of the events or vice versa.

Considering the EFL context of this study, it is necessary to understand the theory regarding bilingual memory. Paivio and Desrochers (1980) proposed an extended version of Dual Coding Theory that concerns bilingualism and multilingualism. The process of language input and output is shown in Figure 1. Within bilingualism, the assumption of Dual Coding Theory can be specialized as “the two verbal systems ($V\_{1}$ and $V\_{2}$) corresponding to the bilingual’s two languages ($L\_{1}$and $L\_{2}$) have referential connections to the image system that are partly shared and partly independent” (Paivio & Desrochers, 1980, p. 391). It means that the two verbal systems are independent as bilinguals understand the knowledge of the world in each language without depending on each other. The two verbal systems can also have partial interconnectedness when a bilingual switch language directly from one to another through translation. Additionally, the connection between three systems allows a bilingual to activate the image system as indirect access when translating one language to the other. For example, the English-Spanish bilingual may understand the $V\_{2}$ (Spanish) concrete word perro (dog) by accessing the image of a dog in the image system and at the same time, activating their $V\_{1}$ (English) word dog (Farley, Pahom, & Ramonda, 2014).

Mental image plays a crucial role in Dual Coding Theory especially to identify the difference between the process of understanding the meaning of concrete words and abstract words. For instance, to understand the abstract word true, it is necessary to think of associated words such as factual, not false, and so on (Sadoski, 2012). The concrete word tree may need verbal association likewise, but it can also be associative with mental images of tree which is more readily available than abstract words. In sum, Dual Coding Theory provides possibility that dual coded words may be better recalled than words that are not.

Dual Coding Theory has provided a long-standing theoretical background in the area of vocabulary learning. Many empirical studies have applied the theory to prove effective vocabulary learning by using pictures or mental imagery. Cohen and Johnson (2011) investigated whether imagery interventions make a difference for the recall of science vocabulary with 89 English speaking fifth-grade students in Long Island. Four different interventions were implemented: Word Only method, Picture Presentation + Word method, Image Creation + Word method, and Image Creation + Picture + Word method. The result showed that the students in groups using pictures or mental imagery outperformed the students with the word only method on the recall of vocabulary. It means that the results supported the principle of Dual Coding Theory as dual coded words were recalled better.

However, in the case of examining concrete words and abstract words separately, there is evidence that shows different performance. Shen (2010) set out to find the difference between verbal coding method and verbal plus imagery encoding method in acquiring isolated words with beginning level students learning Chinese as a foreign language. In verbal plus imagery encoding intervention, students were presented with Chinese words by adding pictures, pictographs, or acting out the meaning of words. The imagery plus verbal coding method was beneficial for learning the shape and meaning of abstract words whereas it was not superior for acquiring concrete words. The researcher assumed that insignificant results of imagery for concrete words were due to the association between concrete words and mental images already settled in the brain.

Farley, Ramonda, and Liu (2012), looked at whether adding visual imagery would contribute to acquiring the meaning of Spanish words with 87 American university students. The treatment involved 12 concrete words and 12 abstract words. Results indicated that participants in the picture group of abstract words outperformed the participants of the non-picture group for receptive knowledge of word meaning on both immediate post-test and delayed post-test. However, regarding concrete words, no significant effect was found. The researchers pointed out that this finding is attributed to the nature of concrete words which is already built in the learners’ mental lexicon. It can be connected to the predictions of Dual Coding Theory that abstract words, which do not possess direct nonverbal equivalents, would have more difficulty in creating a mental image than concrete words (Paivio & Desrochers, 1980). Therefore, abstract words would be more beneficial through vocabulary instruction using imagery than concrete words.