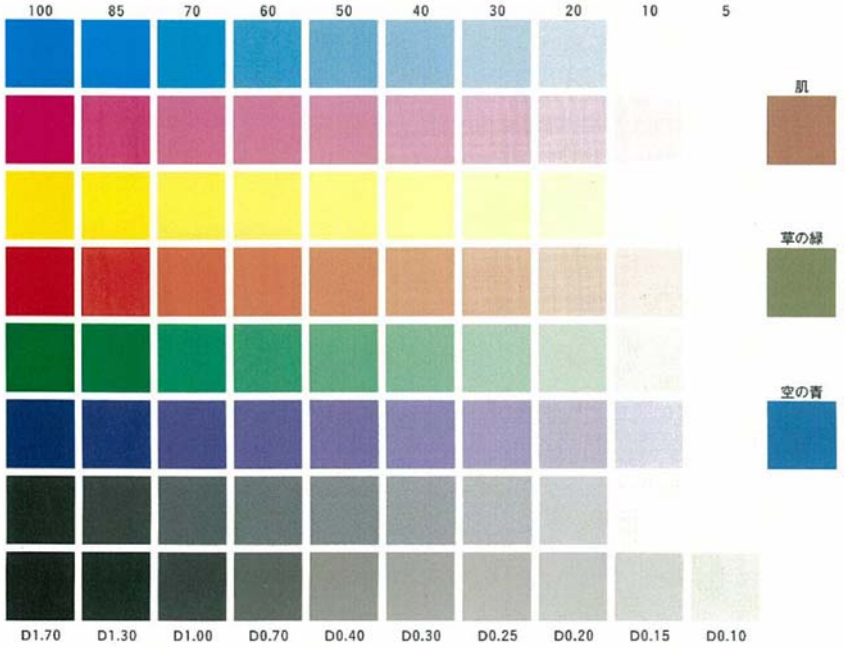


Studies on the Effect of Hand Gestures on the Recall of English Expressions

By

KAWAMURA, Yoshiharu

September 30, 2006

The Graduate School of Foreign Language Education and Research,

Kansai University

Studies on the Effect of Hand Gestures on the Recall of English Expressions

A Dissertation Submitted to

The Graduate School of Foreign Language Education and Research,

Kansai University

In Partial Fulfillment of the Requirements for the Degree

Doctor of Philosophy in Foreign Language Education and Research

By

KAWAMURA, Yoshiharu

September 30, 2006

This dissertation of KAWAMURA, Yoshiharu is approved.

Doctoral Committee:

Professor O. Takeuchi, Ph. D., Chair

Professor Y. Kitamura

Professor E. Yamamoto, Ph. D.

Professor S. Yamane, Ph. D.

Graduate School of Foreign Language Education and Research,

Kansai University, Osaka, Japan

2006

論文要旨（概要）

本論文は、英単語や英文の記憶再生を高めるために、ジェスチャーを用いて目標語の概念を視覚的かつ概念的に捉える手法の有効性を検証した五つの実証的研究の結果をまとめたものである。

第一章では、身体的な経験を基盤として人間の認知様式を考察する新たな認知論について言及し、語の概念を運動感覚的な手がかりを介して記銘する手法を提案する。

第二章では、言語記憶の観点から、イメージ、動作、認知の三者に関する先行研究をまとめて今後行うべき研究の方向性と課題を明らかにする。最初に、従来の記憶方略の検討として、記憶術的な方法であるキーワード法と身体活動を用いるTotal Physical Responseを取り上げて問題点を整理する。次に、1960年代後半以降のイメージと記憶をめぐる学説、イメージ論争、Paivioの二重符号化理論、情報の特定性および関係性と記憶との関連を取り上げる。三番目に、「認知の身体化」の問題に言及する。ここでは概念形成、認知と身体感覚の関連に関わる心理的現象、ジェスチャーの認知的側面を再考して人間の認知活動が身体活動と深く関連している事実を明らかにする。四番目に行為記憶の問題、つまり具体的な行為を示す表現は、聞いたり読んだりするよりその内容を動作で再現するとなぜ記憶が高まるかについて考察する。最後に認知意味論を取り上げ、抽象的概念は具象的概念をメタファー的に捉えて理解するという見解を言語（音声言語と手話言語）やジェスチャーの具体例を通して検討する。

第三章では、一番目の実証研究として、Paivioが二重符号化理論をバイリンガル（二言語併用者）に適用したバイリンガル二重符号化理論の有効性を検証するため、日本人大学生を被験者として追実験を試みた。バイリンガル二重符号化理論は、二つの言語システムと一つのイメージ・システムを想定して、それぞれのシステムの加算的効果が記憶向上に寄与すると仮定する。実験に使用する目標語は、まず語の具象性とイメージ性が高い英語の具象名詞の中から40個を選出した。統制群と実験群にそれぞれ20個の英単語と20個の日本語単語（英単語の翻訳語）を与え、日本語単語は英単語に翻訳し英単語はそのまま書き写すという作業を課して英単語の再生成績を調べた。記憶再生の手がかりは語とシンクロされたジェスチャーの映像で、ここでいうジェスチャーとはそれぞれの語の指示物を描写する写実的な動作である。統制群に文字の映像、実験群に文字と動作の映像を見せた。その結果、語彙の再生成績は、翻訳課題（バイリンガル条件）のほうが書き写し課題よりも優れていた。さらに、どちらに課題においても再生成績は映像提示あり条件のほうが提示なし条件よりも優れていた。また、課題の種類と映像提示の交互作用は有意であった。つまり、映像が提示されると書き写し課題の成績が翻訳課題の成績に近づいたからである。この結果から、映像が提示されるとイメージ・システムのみならず日本語システムまでも機能して加算的効果が働いたと推測できる。いずれの結果もバイリンガル二重符号化理論を支持し、言語システムとイメージ・システムの加算的効果が記憶再生を向上させるひとつの要因であることを示唆している。

第四章では、二番目の実証研究として、20個の具象名詞と20個の抽象名詞を目標語とし、動作映像の手がかりを与えて再生成績の違いを調べた。語のイメージ性をできるだけ均一にするために、具象語群と抽象語群はそれぞれ具象性とイメージ性の値に関して一定の範囲から選択された。統制群は40語の文字映像、実験群は40語の文字と動作の映像を見た。その後に行った再生テストの結果は、いずれの条件でも具象語の再生が抽象語の再生を上回り、映像提示あり条件のほうがなし条件よりも再生成績が優れていた。言語情報と非言語情報の加算的効果を示す結果であると考えられる。具象語のジェスチャーは指示物を描写する写実的な動作である一方、抽象語のジェスチャーはアメリカ手話のサインを基に考案し

た動作であった。例えば、ability は両手の拳を前に押し出す動作で示す。これは何かを成し遂げることができるという自信を表す。promote は両手を段階的に上方に移す動作である。この動作は、社会は上下構造からなり、重要な地位は上にあることを表す「高い地位は上、低い地位は下」という概念メタファーを反映している。このようにして、被験者が身体的経験や知識に基づく身体的イメージや空間的イメージを通して抽象概念を理解することが期待される。これらのイメージを共有できる経験と知識が日本人学習者にあれば、ジェスチャーは英単語の記憶再生の手がかりとして十分に機能する。

第五章では、三番目の実証研究として、文字情報を与える条件（言語課題）に加えて、実験者のジェスチャーを見る条件（実験者実演課題）と実際に被験者が動作を行う条件（被験者実演課題）での再生成績の違いを調べた。先行研究によれば（Engelkamp, 2001）、実験の条件により動作を実演するほうがいつも再生成績が良いとは言えないという。そこで、14 個の目標語を単純な動作を表す 7 個の動作動詞と 7 個の典型的な状態動詞で構成して実演効果を検証した。動詞の分類は行為性と自制性という意味属性に基づく。この実験では、言語課題グループはネイティブスピーカーが各語を発音する映像と字幕の文字を見る。実験者実演グループと被験者実演グループはネイティブスピーカーが各語を発音しながら動作を行う映像と字幕の文字を見る。二回目の映像が流れると、文字グループと実験者実演グループは文字を発音し、被験者実演グループは文字を発音しながら動作をまねる。その結果、再生成績は、言語課題グループ、実験者実演グループ、被験者実演グループの順で向上した。被験者実演グループが実験者実演グループを上回った理由は二つ推測された。一つは実際に体を動かすという運動情報の効果である。もう一つは実演する行為に関わる心理的側面である。体を動かすことは内面的にも課題と深く関与することを求めるからである。したがって、指示の仕方や教室の雰囲気次第ではジェスチャー行為がかえって心理的負担となって、記

憶上マイナスに作用する可能性もある。

第六章では、第四番目の研究として、記憶対象を語レベルから文レベルに移し、二つの実験を行って四つの異なるタイプの動詞を述語とする文の再生成績を調べた。各文の述語を動作化して、主語と目的語の再生数を調べた。述語動詞を動作する理由は、述語は文の中核として統語的にも意味的にも文全体に統制力を持つからである。したがって、述語を動作化することで語の概念内容を含む視覚情報を与えるならば、文全体の記憶再生の向上につながると推測できる。最初の実験では動作動詞と状態動詞を述語とする 14 個の文を用いた。次の実験では高頻度動詞と低頻度動詞を述語とする 14 個の文を用いた。動詞の選択は、ブリティシュ・ナショナル・コーパスにおいて 100 万語あたりの頻度が 250 回以上の動詞から選んだ語を高頻度語とし 25 回未満を低頻度語とした。統制群の映像はネイティブスピーカーが文を発音する様子と字幕が映り、実験群の映像はネイティブスピーカーが文を発音しながら動作を行う様子と字幕が映る。例えば、低頻出語である admire は拍手する動作で表し、criticize は×印を描く動作で表す。再生テストの結果、両方の実験で実演効果が確認され、動作を行う条件のほうがいずれのタイプの文でも再生成績は優れていた。動作動詞を述語とする文と状態動詞を述語とする文の再生には有意な差はなかったが、高頻度と低頻度を述語とする文の再生には有意差があった。一般に低頻度語は、高頻度語より具体性とイメージ性が低くいと判断される。その違いがそれらを述語とする文の再生の差になったと推測される。また、文再生における動作の実演効果に関しては、動作が高頻度動詞のみならず低頻度動詞にも記憶上の手がかりを与えたと考えられる。これらの結果は、述語動詞がどのような種類であろうと、述語を動作化することにより文の記憶保持が高まることを示唆する。

第七章では、五番目の実証的研究として、記憶対象をパラグラフにした。パラグラフを構成する四つの文を八つのアイデア・ユニットという統語的な単位に分

け、各ユニットから一語を選んで動作化した。統制群はネイティブスピーカーがパラグラフを発音する様子と字幕の映像、実験群はネイティブスピーカーがパラグラフを発音しながら動作を行う様子と字幕を見た。各ユニットから動作化した語と他の一語の再生を求めるテストの結果、ジェスチャーはパラグラフ全体の再生のみならず動作化した語の再生においても効果がなかった。その理由として、実験上の手続きの問題、ジェスチャーの質の問題、パラグラフの複雑な構造、記憶術における手がかりの特性が検討された。被験者が、ジェスチャー行為に加えて、統語的にも概念的にも複雑に構成されているパラグラフの理解に追われて注意を分散させてしまい、記憶効果があがらなかった可能性が考えられた。

最後に、第八章では、第三章から第七章までの五つの研究に共通した問題や限界点を指摘した後、各研究から得られた結果をまとめた。そして、結論では動作による英語学習の意義を述べた。

なお、本論の後に参考文献の一覧と研究に関わる資料を添付した。

Acknowledgements

This dissertation was made possible by the contributions of many people in a variety of ways. It is a great pleasure for the author to be able to acknowledge those individuals.

First, the author wishes to express his great gratitude to his two supervisors: Professor Yutaka Kitamura and Professor Osamu Takeuchi of Kansai University. Professor Kitamura, who was in charge of the author for the first two years, enlightened the way toward the theme of this dissertation. Professor Takeuchi, who had generously accepted the guidance of the author, supported him enduringly for the last two years by providing appropriate suggestions all the time.

Second, the author is very grateful to Lecturer Kazuto Sumita of Kanazawa Seiryo University for his technical assistance for statistical calculations, and Mr. Hiromitsu Imasaka of Seiryo High School for his technical assistance for editing video tapes. The author is also very grateful to his colleagues, Professors Colin Sloss and Richard Grassi for their warm friendships. Both of them took a part of an instructor pronouncing and enacting words in the experimental videos. Ms. Gabriella Gray also offered her co-operation as a model of an instructor in the video. Many Seiryo University staff members contributed to this dissertation in many ways. The author would like to express his deep appreciation to all of them. The author would like to thank students of Kanazawa University for their co-operation as subjects for experiments, and two students of Kanazawa Seiryo University, Miki Tanaka and Natsumi Fukuhara for their gesture performance.

Finally the author wishes to express special thanks to his family, his wife Akie and his three sons, Yu, Kai, and So for their warm encouragement. Without their support, the author could not have completed his dissertation.

Table of Contents

1. Introduction

The importance of vocabulary knowledge is evident to anyone who has ever read an English text. Vocabulary learning, then, tends to come first in the Japanese classroom. Vocabulary learning is repeatedly conducted and its progress is constantly measured. Some students try to memorize new words by simply writing spellings, some pronounce new words aloud, and some read a textbook again and again. In most cases, however, target English words are paired and memorized with Japanese equivalents, and reversely, Japanese equivalents are translated into target words. The method is conventional, but still seems to be dominant among Japanese students.

Mnemonic techniques have been developed for facilitating memorization. They assume that the recall of one idea promotes that of another, and the power of association is related to high imagery of a cue for recall. The keyword method, a mnemonic technique for learning vocabulary, constructs verbal and visual links between a target word and its translation equivalent. The technique is superior to rote repetition, but less popular and is rarely used in the classroom. Learners are required to find a keyword and create an associative image. The effort of producing those recall cues, which are the key to the mnemonic technique, is troublesome for beginning learners. The method may be effective in facilitating the recall of concrete words, but not abstract words. This is because a keyword is selected based on phonological similarity with a target word, and does not have an inherent connection with the concept of a target word. Total Physical Response, a method based on the coordination of speech and action, asks learners to follow commands (imperative sentences) by moving their bodies. The comprehension through physical behavior facilitates the recall of words and sentences. This method also finds difficulty with abstract words. Abstract concepts, it seems, are not easily conveyed through pictures and physical behavior. How to mediate abstract concepts is an unavoidable issue.

Traditionally, mind or reasoning has been accepted as being independent of bodily experiences such as perception and bodily movement, but the paradigm is changing now. A new perspective of cognition, *i.e.*, embodied cognition, assumes that cognitive processes are grounded in the body's interactions with the world. Language meaning, then, "includes patterns of embodied experience and perceptual structures of our sensibility (*i.e.*, our mode of perception, or orienting ourselves, and of interacting with other objects, events, or persons)" (Johnson, 1987: 14). The new paradigm regards mind as being inherently embodied, and abstract concepts as being largely based on metaphors of bodily and physical concepts (Lakoff & Johnson, 1999). Metaphors are not figures of speech, but patterns of cognition constituting the basis of abstract thought. If abstract concepts are mostly understood in terms of bodily experiences, conceptual metaphors can provide sensory-motor information related to conceptual content, and facilitate the recall of abstract words as well as concrete words.

This dissertation is an attempt to establish a new mnemonic technique to recall English expressions with hand gestures. The dissertation contains a literature review and five related empirical studies. In Chapter 2, the author provides a literature review of a wide range of studies on language and cognition. First, it deals with conventional mnemonic techniques and strategies for second language learning. Second, the chapter describes views related to imagery and memory by referring to the Imagery Debate of the 1970s, Paivio's dual coding theory, and distinctive-relational processing of information. Third, the chapter explains the interaction of body and mind by examining language acquisition, various psychological phenomena, and gesture. Fourth, the chapter deals with memory for action by asking why action phrases such as *opening the window* are better retained when people perform the phrases than when they just listen to or read them. Finally, the chapter reviews the background of embodied cognition by focusing on

metaphor and metaphorical features of language (signed and spoken). In Chapter 3, the author employs gesture images as recall cues and examines if the bilingual dual coding theory (the dual coding theory applied to bilinguals) is valid for Japanese university students learning English as a foreign language. After subjects conduct a task of translating Japanese words and copying English words with and without gesture images, their recall is tested. The results suggest that gesture images can function as effective recall cues. In Chapter 4, the author employs abstract words as well as concrete words as recall targets, and examines how word recall is influenced by word type and gesture-image presentation. The results suggest that gesture images can facilitate the recall of abstract words as well, though concrete words are better remembered either with or without gesture images. In Chapter 5, the author demonstrates the effect of gesture performance in recalling words. The control group watches a person pronouncing and enacting words, while the experimental group watches a person pronouncing and enacting words, and copies the actions. Gesture proves effective, and the reasons are discussed. In Chapter 6, the author extends recall targets form word level to sentence level. Four types of verbs are employed as the predicate of a sentence, and the predicates are enacted. The results suggest that gesture is effective in recalling sentences with a predicate of any type. In Chapter 7, the author further extends recall targets to paragraphs. A paragraph is divided into idea units, which are grammatical segmentations. One word per each idea unit is enacted. The results suggest that enactment of one word per each *idea unit* is ineffective in recalling the content of a paragraph. Many possible reasons are discussed. In Chapter 8, the author summarizes the findings of the five empirical studies elaborated in the preceding chapters, and states the significance of learning English with gesture as well as implications for further research.

2. Literature review

Mnemonic effects of imagery have been known since the Greek and Roman age. People have always looked for ways to help remember what they want to remember. In the late 1960s, cognitive research on memory and imagery began to bring forth scientific reports. Those reports indicated that people remembered linguistic items better when they learned the items through mental imagery. In the early 1980s, some linguists began to employ cognitive approaches for the investigation of conceptual organization of language. Those linguists, then, assumed that human mind is not separated from bodily capacities such as perception and motion. Language and perception are regarded as two overlapping cognitive domains based on bodily experiences. Investigation into mnemonic effects of mental representation now has to examine those findings and views.

This chapter thus aims to integrate findings from the previous research and reveal controversial points which deserve attention for future research. The issues to be discussed in this chapter are as follows: 1) mnemonics, 2) imagery and memory, 3) embodied cognition, 4) gesture, 5) action memory, and 6) cognitive semantics.

2.1 Mnemonic strategies

2.1.1 Mnemonics

Before the invention of printing, mnemonics or memory improvement techniques were used to retain vast stores of knowledge. Over the centuries, a great variety of memory-enhancing techniques were developed. In the 1960s, '*The Art of Memory*' by F. Yates (1966), which described a detailed history of mnemonics, was widely read. For example, the method of loci, a technique based on imagery, was introduced as one of the oldest mnemonic devices. The mnemonic method teaches people to mentally walk through a well-known place and associate things-to-remember with locations in the place one by

one. The publication of *The Mind of a Mnemonist* by A. R. Luria (1968), a Russian psychologist, was a great stimulus to English-speaking psychologists and promoted scientific research on memory and imagery. A man with a vast memory was characterized by his abnormally vivid visual imagination. Luria (1968: 30) says, "It was only natural, then, that the *visual quality of his recall* was fundamental to his capacity for remembering words. For when he heard or read a word it was at once converted into a visual image corresponding with the object the word signified for him" (emphasis in the original). Inspired by those stories, researchers and educators began to examine how mental imagery worked for improving memory performance, and devised new mnemonic techniques.

2.1.2 Memory strategies for foreign language vocabulary learning

Mnemonic devices aid people to remember information which is difficult to recall by integrating it into existing knowledge through some verbal or nonverbal cues. That is, verbal or nonverbal (usually visual) cues serve as cognitive mediators between what is known and what is to be remembered, and help people to retrieve target information. The role of those mnemonic devices is highly important for language learning because learning a foreign word is the process of associating an unknown word with an equivalent word and its concept in a language which learners know. Today, mnemonic devices are not just convenient tricks of remembering words and phrases; such devices can also be transformed into effective strategies for developing language knowledge and performance.

In *Language Learning Strategies* (1990), R. L. Oxford offers a comprehensive classification of language learning strategies. She divides language learning strategies into two main classes, *direct* and *indirect*. Direct strategies directly involve the target language. Direct strategies are further subdivided into three groups: *memory strategies*, *cognitive strategies*, and *compensation strategies*. *Memory strategies,* concerned with the present

paper, consist of four sets: *creating mental linkages*, *applying images and sounds*, *reviewing well*, and *employing actions* (Appendix A). The four sets further consist of sub-strategies.

The *Creating mental linkages* category is made up of *grouping*, *associating/elaborating*, and *placing new words into a context*. *Associating/elaborating* means "relating new language information to concepts already in memory" (p. 41) to create associations in memory. Oxford insists that associations must be meaningful to the learner. The association process is fundamental in image-using language learning. The *Applying images and sounds* category includes *using imagery, semantic mapping*, *using keywords*, and *representing sounds in memory*. *Using imagery* is a strategy for "relating new language information to concepts in memory by means of meaningful visual imagery, either in the mind or in an actual drawing" (p. 41). Oxford stated that the strategy can be used to remember abstract words by associating such words with a visual symbol or a picture of a concrete object. Is it really possible to associate abstract words with visual symbols and pictures? How to perceive and retrieve abstract words is a main concern of the present dissertation. *Using keywords* is a strategy for "remembering a new word by using auditory and visual links" (p. 41), and *representing sounds in memory* is for "remembering new language information according to its sound" (p. 42). In fact, auditory and visual cues are essential ideas of the *Keyword Method* (discussed later). The *Reviewing well* category contains one strategy called *structured reviewing*, which reviews new language information carefully at spaced intervals. The *Employing action* category contains two strategies: *using physical responses or sensations*, and *using mechanical techniques*. *Using physical responses* is defined as "physically acting out a new expression (*e.g.*, going to the door), or meaningfully relating a new expression to a physical feeling or sensation (*e.g.*, warmth)" (p. 43). The idea of *using physical responses* is the core of James

J. Asher's *Total Physical Response* (also discussed later).

2.1.3 Visual imagery approach: The keyword method

The most extensively studied mnemonic strategy for second language learning is the Keyword Method. The strategy was devised by Atkinson (1975) to learn foreign language vocabulary. The Keyword Method consists of two stages, a verbal association and a visual imagery association. First, a word (usually a concrete word) in the native language is chosen as a keyword, based on phonological similarity with a word to be learned in the target language. Next, a mental image (usually a visual image) is constructed between the keyword and the translation of the target word. Then, the sound of the target word triggers the activation of the sound of the keyword, which in turn activates the interactive image, resulting in the retrieval of the translation equivalent. By strengthening the associations, learners are expected to retrieve the meaning of a target word as soon as they hear the word pronounced.

Atkinson indicated that the Keyword Method enhances foreign language vocabulary learning by demonstrating that the method facilitated the learning of Russian vocabulary. He conducted an experiment of teaching 120 Russian words in three days. In the recall test carried out on the fifth day, the keyword group recalled 72% of the items compared to 46% by the control group, and in the delayed test six weeks later, the keyword group still recalled 43%, while the control group recalled 28%. The recall differences were significant.

At the same time, Atkinson raised the question of whether the experimenter should supply the keyword or learners should generate their own keyword because his experiments had indicated that the provided keyword worked better than the generated keyword. Paivio and Desrochers (1981) stated that the results conflict with the usual

superiority of subject-generated imagery or verbal mediators in paired-associate learning experiments, and suggested that "the difference may be attributable partly to the relative difficulty of discovering appropriate mediators" (pp. 781-782). Here, the problem of discovering appropriate mediators is not concerned with who originates the keyword, but with what associative image the keyword evokes. Without creating a mental image which firmly links a target word with its meaning, the keyword, experiment-provided or learner-generated, does not seem to facilitate the recall of foreign vocabulary.

Paivio and Desrochers raised another question about the type of vocabulary to which the Keyword Method can be applied. They speculated the possibility that the Keyword Method might be less effective for abstract words than concrete ones because imagery is an essential element of the method and the imagery-evoking quality of the target word would affect the procedure of the method. Pressley, Levin, and Miller (1981) conducted experiments that English speaking children learned and recalled 20 Spanish words (10 concrete items and 10 abstract items). The means of the correct responses in the control group was 26.8% for the concrete items and 17.9% for the abstract items, and those of the correct responses in the basic keyword technique was 58.4% for the concrete items and 42.1% for the abstract items. They concluded that the Keyword method's utility was not limited to concrete words. In their experiments, the keywords for the abstract words were highly concrete. Paivio and Desrochers (1981: 784) commented that concrete keywords might have stimulated the children to concretize the abstract words and made it easier to recall their meanings.

The third question about the Keyword Method is the quality of the keyword as a retrieval cue. It is a common idea to give prominence to the keyword. To enhance recall, a visual image between the keyword and the target word is usually required to be as unique as possible. Hulstijn (1997) described the tendency that the linking-image is constructed

"preferably in a salient, odd, or bizarre fashion" (p. 204). This is because the selection of the keyword is based upon phonological similarity with the target word; the keyword does not have an inherent connection with the target word. Brown and Perry (1991) argued that "semantic processing is not the main focus of the keyword method and should not produce better retention for the meaning of a new L2 word than a method that relies more heavily on semantic processing" (p. 658).

According to the levels-of-processing theory (Craik & Lockhart, 1972; Craik & Tulving, 1975), memory retention improves as the process of information moves from a shallow sensory level to a deeper semantic level, and elaboration at either level would further enhance retention. With this theory in mind, Brown and Perry (1991) compared three strategies: the keyword method, a semantic processing method, and a combined keyword-semantic condition. The keyword group, provided with a new word, its definition, and a keyword, practiced making interactive images. The semantic group, provided with a new word, its definition, and two example sentences, was required to answer a question. The keyword-semantic group, provided with a new word, its definition, a keyword, two example sentences, and a question, practiced making interactive images. The experiment showed that the keyword-semantic group had significantly better retention than the keyword group and slightly better retention than the semantic group. The results corresponded to the levels-of-processing theory, and suggested that the Keyword Method is rather weak at making semantic associations.

2.1.4 Motoric approach: Total physical response

James J. Asher, an American psychologist, originated Total Physical Response (TPR), a language teaching method modeled on the process of infants' language acquisition. Infants grow up in an interactive environment in which their caretakers often speak to them

in the form of commands and they respond physically. So, the method teaches language by coordinating speech and motor activity. Asher indicated that action enhances the comprehension of input items. In a TPR class, instructors direct a series of commands to learners, and learners listen and respond to the commands by physically performing the activities. For example, a typical TPR instruction contains commands such as *Stand up*, *Walk to the window*, and *Touch the window*. In a cooperative classroom environment, such activities can be both motivating and fun, and those linguistic expressions can be easily assimilated by learners.

In a TPR class, listening skills are emphasized far more than speaking skills. Asher believed that as listening comprehension develops, learners spontaneously begin to speak, because he hypothesizes that "the brain and nervous system are biologically programmed to acquire language, either the first or second, in a particular sequence and in a particular mode. The sequence is listening before speaking and the mode is to synchronize language with the individual's body" (Asher, 2000: part II, p. 4). He presumed that TPR takes advantage of the brain lateralization. It has been believed that the left hemisphere is used for dealing with verbal information while the right is for performing physical actions. So, learners in a TPR class are considered to activate both sides of the brain. Similar simplistic views on the brain lateralization and language acquisition can be found in articles from the 1980s (*e.g.*, Danesi, 1988; Joiner, 1984).

Today, TPR is very popular and often used, especially for teaching a second language to small children. But there are some problems with TPR. A most serious problem is how to teach abstract vocabulary. In response to the problem, Asher (2000) proposed two ways. One way is to use a flashcard with a foreign word on one side and a native equivalent on the other. The card is manipulated as a concrete item. For example, a Japanese teacher of English, when introducing a new word such as *government* to the

class, might give a student a command like "Tanaka, *government* wo Yamada ni nagenasai" (Tanaka, throw *government* to Yamada). In this approach, the act of throwing is not related to the meaning of a word at all. The other is to present an abstract word in context. For example, "Tokyo is the center of *government* for Japan." However, Asher confessed that he was not satisfied with these ways of teaching abstract words, and said, "this is an interesting problem for future researchers to discover" (Asher, 2000: part II, p. 21).

2.1.5 Strategies Japanese learners of English use

Schmitt (1997) was unique in reporting vocabulary learning strategies used by Japanese learners. He compiled many vocabulary learning strategies and completed a list of 58 strategies for vocabulary learning. He organized the list mainly according to the above mentioned system by Oxford and his own categories called *discovery* and *consolidation*. Discovery strategies enable learners to discover an unfamiliar word by guessing its meaning from various sources and contexts; on the other hand, consolidation strategies aid learners to retain words.

The survey project was conducted with Japanese junior and senior high school students, university students, and adult learners. Each group contained a total of 150 participants, bringing the total number to 600. All the subjects spoke Japanese as their first language and had learned or were learning English at school. The results showed that 40 strategies were used and evaluated by the subjects (Appendix B). According to the results, the most used strategies for consolidating meaning were *verbal repetition* (76%), *written repetition* (76%), *study the spelling of a word* (74%), and so forth, while least used ones were *use physical action* (13%), *use cognates in study* (10%), and *use semantic maps* (9%). The most helpful strategies for consolidating meaning, the participants believed, were *say*

new words aloud (91%), *written repetition* (91%), *connect word with synonyms/antonyms* (88%), and so forth, while least helpful ones were *image word's meaning* (38%), *use cognate in study* (34%), and the *Keyword Method* (31%). The results showed that the Japanese correspondents preferred techniques which focus on word form, and had little interest in mnemonic techniques using imagery and actions.

Schmitt (1997: 220) pointed out that the outstanding repetition of a word's verbal form or written form "can, at least, be attributable to the study style encouraged by the Japanese school system; students are required to memorize English grammar and vocabulary, usually through repetition." O'Malley and Chamot (1990) reported the effect of strategy training on learners with Hispanic and Asian backgrounds. The Hispanic experiment group who had learned the training exceeded The Hispanic control group in their vocabulary scores, while the Asian experimental group who learned the training achieved poorly compared to the Asian control group which used their traditional rote repetition strategy. Gu (2003: 18) concluded that, "strategies that work in one educational, cultural, and linguistic context might not work in another."

There is no doubt that role rehearsal such as repeating a word's verbal and written form is still widely employed in Japan. However, many scholars and teachers have tried different techniques. Kohno (1984), for example, reported that kinesthetic information such as facial expressions and gestures, as well as pictures, facilitated Japanese learners' English listening comprehension. Recently, teachers reported vocabulary teaching techniques through communicative activities for elementary school children (Naoyama, 2002), through illustrations (Teshima, 2002), and through illustrations and semantic maps for junior senior high school students (Ikemura, 2002). In Japan, it appears that vocabulary learning is changing from simply remembering words to learning words for use and concept-formation.

As mnemonic techniques like the Keyword Method and TPR have been introduced, simple rote repetition has become less popular than before. Repeating new words, however, is a simple and handy strategy for beginning learners encountering many new words, or even advanced learners trying to learn difficult words. On the other hand, mnemonic strategies also contain certain limitations. Gu (2003) explained the advantages and limitations of various vocabulary learning strategies in a comprehensive view of the subject. Some limitations related to the present dissertation are as follows: Mnemonic techniques tend to regard language learning as a memory problem and aim for retention of paired-associates. Mnemonic techniques also tend to emphasize the one-to-one relationship between form and meaning and neglect multiple meanings of a word in multiple contexts. Therefore, it appears that abstract words are not suitable for mnemonic techniques.

2.2 Imagery and memory

2.2.1 Imagery debate

Today no one denies that they experience seeing images in their mind and that the images are processed in the brain, but it is still not clear how mental images are represented. In the 1970s and 80s, the question was severely debated from two opposite perspectives. The argument was called *Imagery Debate* (Miyazaki, 1979). Some insisted that a mental image, in the format of a picture in the mind, depicted the outer world. Kosslyn (1981), for example, called mental images quasi-pictorial image representations. Others, for example Pylyshyn (1973), insisted that a mental image, in the format of a proposition or a sentence in the mind, described the world. He pointed out that the term *image* relied heavily on a picture metaphor, because people referred to the clarity and vividness of images and talked about the shapes and properties of the objects depicted in the images. He claimed that "when our recollections are vague, it is always in the sense that certain perceptual qualities or attributes

are absent or uncertain—not that there are geometrically definable pieces of a picture missing" (p. 10), and that "one's representation of a scene must contain already differentiated and interpreted perceptual aspects. In other words, the representation is far from being raw and, so to speak, in need of 'perceptual' (emphasis in the original) interpretation" (p. 10). Pylyshyn believed that sensory data are not stored as they are, but encoded into a finite number of concepts and relations, and are represented by logically independent descriptive propositions.

Anderson (1978), however, suggested that the imagery debate is empirically undecidable by claiming that "it will not be possible to establish whether an internal representation is pictorial or propositional" (p. 249). Image Scanning, for example, is used as evidence to support their own theory by both sides. Kosslyn, Ball, and Reiser (1978) asked their subjects to learn a map of locations, form an image of the map, and focus their attention at one place and then at another. They measured scanning time and found that the further apart two locations were, the longer it took subjects to report that they could see a second location in the image, though there was no time difference when they memorized a list of location names. In the scanning experiment, the subjects experienced distance between the imaged locations analogous to one between physical locations. Kosslyn et al. interpreted that scanning a mental image was similar to scanning an actual picture. On the other hand, Pylyshyn (1981) gave an alternative explanation of increased reaction time with increased map distance. He assumed that the subjects estimated the time the scanning would take because they already knew from their experience that it would take more time if they walked or moved their eyes around for a longer distance. Anderson (1978) already stated that "evidence for a particular representation was really evidence for a particular process and that there was no good reason to associate the process with a particular presentation" (p. 262). His conclusion is that "it is not possible to decide between imaginal and propositional

representations strictly on the basis of behavioral data" (p. 275).

Today, the debate about whether mental images are pictorial or propositional is no longer concerned with formats which represent imagery, but a matter of how images are perceived (Miyazaki, 1998). Reisberg and Chambers (1991) claimed that "images are meaningful depictions, that is, they carry certain specifications about how they are to be understood" (p. 350). They carried out two experiments. They showed 15 subjects an outline drawing of Texas, rotated by 90 degrees, and told them to imagine the figure. The subjects were then told to rotate the image at 90 degrees and identify the form. None of the subjects found Texas in the rotated image. Interestingly, however, 8 subjects were able to detect Texas when they drew a picture of the mental image. In the second experiment, another 15 subjects, instead of rotating the mental image at 90 degrees, were told to regard a particular side as the top of the form. Then, 7 subjects identified Texas. They were able to identify their mental images as Texas by deciding the position of the top. Some part of a mental image has a decisive meaning in the identification of the whole. The two experiments show that identifying mental images depends on how they are understood. The 8 subjects visualized their imagination and the 7 subjects used knowledge. Chambers (1993) stated that "both the depictive and descriptive aspects of images play a role in what is represented in images, subsequently, what can be discovered from the images" (p. 93). Mental representations are both imaginal and propositional, it is assumed.

2.2.2 Dual coding theory

In the last two decades, one of the most influential theories of mental imagery is surely Paivio's (1971, 1986) dual coding theory. Replying to the criticism that images can not represent general or abstract knowledge, Paivio (1977) claimed that it is common knowledge among cognitive psychologists that general ideas are represented in the form of

prototype images. The concept of a bird, for example, can be represented by the image of a typical bird, like a sparrow or blackbird, rather than a penguin (Rosch, 1973). He also claimed that an image need not carry the entire account for an abstract concept, because he assumed that human cognition is both verbal and nonverbal, and any representation theory should accommodate both functions. He then, proposed that verbal information and nonverbal information are coded and represented in two independent, but partially interconnected, symbolic systems: "one specialized for the representation and processing of information concerning nonverbal objects and events, the other specialized for dealing with language" (Paivio, 1986: 53). The nonverbal system is called imagery system because "its critical functions include the analysis of scenes and the generation of mental images (both functions encompassing other sensory modalities in addition to visual)" (p. 54). The independence of the two systems implies that cognitive activity is mediated in the nonverbal system or the verbal system alone, or simultaneously in both systems. The interconnection implies that activity in one system can activate the other system. It means that verbal information arouses imagery, and conversely, imagery can be expressed verbally.

If verbal information and nonverbal information are coded and represented in dual systems, how is our cognition or memory empirically explained? It is well known that verbal recall of items presented as pictures is generally higher than that of items presented as words. This is called *picture superiority effect*. The superiority of pictures in recall was well explained in the frame of the dual coding theory. Paivio and Csapo (1973) used pictures and words under four different conditions in their second experiment. Two groups of subjects were required to write the names of objects when pictures of them were flashed or to copy words when their spellings were flashed. Another two groups were required to make a rough sketch of flashed pictures or of objects which words referred to. Then, the

subjects were given an unexpected recall test. The task of sketching things drawn in the pictures and that of copying words were under a single coding condition, while the task of sketching things words refer to and that of writing the names of objects drawn in the pictures were under a dual coding condition. The mean numbers of recalls under the picture-word, the word-picture, and the picture-picture condition were more than double the mean number of recall under the word-word condition. As for the high performance under the dual condition, Paivio and Csapo believed that it could be attributed to the additive effects of the verbal and imagery systems of the dual coding theory. On the other hand, as for the high performance under the picture-picture condition, it was assumed that the subjects might have labeled the pictures of common objects in their mind spontaneously when they saw them. The low performance in the word-word condition was explained by the idea that words do not easily evoke mental images.

Different words, however, have different image evoking capacities. Paivio, Yuille, and Madigan (1968) asked subjects to rate 925 nouns along a 7-point scale in terms of imagery or a word's capacity to arise nonverbal images. Words (*e.g.*, *car*) which quickly and easily raised a metal image were given high imagery rating, while words (e.g., *criterion*) which did so only with difficulty or not at all were given low imagery rating. In addition to imagery, the difference between *car* and *criterion* were attributed to other factors. Paivio, Yuille, and Madigan also asked different subjects to rate the same items, this time, in terms of concreteness or directness toward sensory experiences. The word *car* was quite concrete, and *criterion* was very abstract in the sense that people can not directly or bodily experience it. The experiments found that imagery and concreteness were highly correlated to each other (e.g., *car* was 6.87 in imagery and 7.00 in concreteness, while *criterion* was 1.87 and 1.97).

It is well-known that words and phrases which were rated high in concreteness were

remembered better than word with low concreteness. This is called *concreteness effects*. Paivio and Csapo (1973)'s first experiment employed pictures, concrete words, and abstract words. They showed subjects slides of pictures, concrete words, and abstract words, and asked them to label the pictures and copy the words on a sheet of paper. Then, the subjects were given an unexpected recall test. Recall increased systematically from abstract words, to concrete words, to pictures. This fact can be interpreted as follows: pictures of common objects can readily evoke a verbal code as well as a nonverbal image, and can be verbalized easily, while words can stimulate a verbal memory code, but coding of words into imagery is assumed to be more difficult than coding of pictures into words, especially in the case of abstract words. Pavio (1986) said that "high imagery items readily evoke nonverbal imagery and that imaginal representations serve as a supplementary memory code for item retrieval" (p. 159).

In addition to imagery and concreteness, Paivio, Yuille, and Madigan (1968) rated the 925 nouns in terms of meaningfulness by requiring subjects to write words associated with each item in 30 seconds. They found that many words which are high in imagery and concreteness also tend to be high in meaningfulness (e.g., *car* is 6.63 in meaningfulness, while *criterion* is 3.00), but not the other way around. On the other hand, they found that some words are low in imagery and concreteness but relatively high in meaningfulness. Those which are high in imagery, concreteness, and meaningfulness are considered to be typical concrete words, and those which are low in imagery and concreteness but high in meaningfulness are considered to be typical abstract words. Pavio et al. assumed that "abstract items derive their meaning largely from intraverbal experience" (p. 8), and "highly concrete items, on the other hand, presumably are associated with both sensory and verbal experience, which are reflected in their high *I* and *m* values" (emphasis in the original: p. 8).

Paivio and Desrochers (1980) extended the dual coding theory to bilinguals and hypothesized that a bilingual owns two language systems corresponding to the bilingual's two languages and one imagery system. The bilingual dual coding theory assumes that the two verbal systems and the one imagery system are partly independent and partly interconnected. Bilingualism was speculatively categorized into two types: the compound and the coordinate (Weinreich, 1953). The compound type was presumed to control two languages sharing one conceptual system, while the coordinate was presumed to control two language systems and have two different conceptual systems, though Paivio regarded the compound-coordinate distinction as a matter of degree, changing with individuals and concepts. Paivio and Lambert (1981) tested the dual coding approach to bilingual memory. They provided French-English bilingual subjects with a mixed list of pictures, French words, and English words, and required them to write the English names of the pictures, translate the French words, and copy the English words. The subjects were unexpectedly asked to recall the generated English words. The word recall increased significantly from single verbal processing (copying English words) to verbal-verbal processing (translation) to verbal-nonverbal processing (labeling pictures). Paivio (1986) commented on the results by saying "verbal and nonverbal episodic trace components are independent and that the nonverbal (imaginal) component is mnemonically stronger than the verbal one" (p. 250).

2.2.3 Distinctiveness and relational processing

For memory retention in paired-associate learning, roles of both relational and item-specific or distinctive information have been emphasized (Hunt & Einstein, 1981). Relational information is common information to input elements or events, while distinctive information is unique to each element or event. In pair-associate learning, the relational information serves to outline a set of words from other pairs, and the distinctive

information serves to define the target word.

Marschark and Hunt (1989) tested concreteness effects (concrete words are remembered better than abstract words) in terms of the relational-distinctive processing rather than the dual coding theory. They assumed that concrete words are more distinctive than abstract words, but the effect of distinctiveness depends on the prior activation of relational information, and stated that "concreteness effects in memory depend on the activation of relational information in retrieval" (p. 717).

The role of relational and distinctive information or relational and distinctive processing of information was already mentioned in Paivio (1971). Based on results of recognition tests, he reported that recognition memory scores increased from abstract words, to concrete words, to pictures, and concluded that "imagery value is related to discriminability" (p. 288). He also focused on associative function of imagery in mnemonic techniques and stated that "the most effective 'conceptual pegs' (emphasis in the original) are those that evoke images that are well differentiated from each other and at the same time provide strong linkage to the appropriate response members of pairs" (p. 293). Differentiation and association are not mutually exclusive.

Paivio, Khan, and Begg (2000) tested the recall of concrete and abstract adjective-noun pairs under related and unrelated, and free and cued recall conditions. The results were that the recall advantage of concreteness remained highly substantial for unrelated as well as related pairs in both free and cued recall. In free recall, the ratio of concrete pairs and abstract pairs was 3 to 1 under unrelated conditions and 2 to 1 under related conditions. In cued recall, the ratio of concrete pairs and abstract pairs was 4 to 1 under unrelated conditions and slightly more than 2 to 1 under related condition. The concreteness advantage was seen to be significant with unrelated pairs as well as with related pairs in both tasks. The results were inconsistent with Marschark and Hunt's

assumption that concreteness effects in memory depend on the activation of relational information. In their assumption, the concreteness effects should have been greater with related pairs, especially in cued recall. Paivio reasoned from the dual coding theory that "imagery enhances both distinctiveness (e.g., item-specific recognition) and relational processing" (p. 156).

There is, however, another explanation for concreteness effects from the view point of relational processing: a context availability model. This model assumed that abstract materials have a disadvantage in memory and comprehension because "persons find it more difficult to retrieve the associated contextual information from knowledge base needed to augment comprehension for abstract concepts than concrete concepts" (Schwanenflugel, 1991: 243) and "in meaningful, supportive contexts, abstract sentences and words are comprehended as quickly as concrete ones" (p. 239).

Sadoski, Goetz, and Avila (1995) compared recall testing of concrete and abstract paragraphs about two historical figures (Michelangelo and James Madison). The paragraphs varied in familiarity of information. According to the context availability model, meaningful and familiar paragraphs, either concrete or abstract, would be equally recalled. On the other hand, the concrete paragraphs would be recalled better than the abstract paragraphs according to the dual coding theory, though the familiarity effect would work independently. The recall of the Michelangelo concrete paragraph exceeded its equally familiar abstract counterpart, and the recall of the Madison concrete paragraph was equivalent to that of its much more familiar abstract counterpart. The results were more consistent with the dual coding theory that the verbal system and the imagery system, which are independent and additive in processing information, can contribute to both the distinctiveness and the relational processing of information.

2.2.4 Neuropsychological evidence for concreteness

Kiehl, Liddle, Smith, Mendrek, Forster, and Hare (1999) investigated the neural pathways involved in processing concrete and abstract words, using functional magnetic resonance imaging (fMRI). Subjects, presented with concrete words and pseudo-words, or abstract words and pseudo-words, were asked to decide if an item appearing is a real English word. Analysis of fMRI indicated significant activation in the right anterior temporal cortex in the processing of abstract words. Concrete words, on the other hand, were recognized more rapidly and accurately than abstract words. They reasoned that "execution of a lexical decision is associated with performance of the semantic processing required to distinguish between concrete and abstract words" (p. 231) and "more extensive semantic processing is required for the recognition of abstract words" (p. 231). Their hypothesis, in which imagery was not taken into account, was consistent with the context availability model in that additional contextual information facilitates a reduction of the speed difference in the processing of concrete and abstract words.

Measuring the brain's electrical activity, Holcomb, Kounios, Anderson, and West (1999) examined event-related potentials (ERPs) during the processing of concrete and abstract words which were semantically congruent, anomalous, or neutral with respect to sentence context. The neutral sentences were congruous but did not provide any supportive context for their final word (concrete or abstract). In the experiments, Holcomb et al. focused on the so-called N400 wave, a component of ERPs peaking at approximately 400 msec. after the onset of a stimulus. They assumed that "the N400 reflects the process of integrating semantic information into a relatively high level discourse" (p. 723). They found that when the sentences were anomalous, concrete final words induced larger N400 than abstract final words. A similar effect was observed even with the neutral sentences. The results were interpreted as "a more effortful or involved integration process" (p. 723).

On the other hand, the congruous sentences did not induce any significant ERP differences between concrete and abstract final words. It implied that the presence of a supportive context can contribute to the semantic integration of abstract words. Concrete effects, however, cannot be reduced to the context availability because they found that "in the anomalous sentences, the differences between concrete and abstract words varied systematically across the scalp" (p. 730). It suggests that both word concreteness or imagery and linguistic context can contribute to sentence processing in different ways. The results were consistent with the dual coding theory rather than a single coding theory like the context-availability model.

Jessen, Heun, Erb, Granath, Klose, Papassotiropoulos, and Grodd (2000) also excluded a single coding theory and suggested a combination of both the dual coding theory and the context availability model in order to explain concreteness effects. They used even-related fMRI to detect brain areas which serve for concreteness effects. They found greater activation in the lower right and left parietal lobes, and in some other areas, during encoding of concrete nouns compared to abstract ones. Jessen et al. considered that concreteness effects may result from greater verbal contextual information reflected by the activation of the left parietal lobe and additional nonverbal information reflected by the activation of the right parietal lobe. The right parietal lobe is known to be related to spatial processing. They suggested that spatial imagery can be a cause of concreteness effects because concrete objects are easily associated with space.

2.3 Embodied cognition

2.3.1 Color concepts

Color perception is a good example to show the involvement of embodiment in our cognition. Color is a physical phenomenon of light. Its electromagnetic spectrum is filtered

out through a complex neuropsychological process and visually perceived as colors. But color experience is not only perceptual but also cognitive. Linguistic categories for color were different in cultures because different languages seem to divide the visible spectrum arbitrarily. Berlin and Kay (1969), however, found that there are eleven basic color terms across cultures and stated that "color categorization is not random and the foci of basic color terms are similar in all languages" (p. 10). Rosch (1973) satated that Dani, a New Guniea language, has only two basic color terms, *muli* (a category for dark and cool colors) and *mola* (a category for light and warm colors). In a series of experiments, she found that when Dani people were asked for best examples for their two color categories, they chose the focal colors, for example, black, green, or blue for *mulil*, and white, red, and yellow for *mola*. She also found that those basic colors could be learned more rapidly and remembered more easily than peripheral colors by Dani people who did not have names for the basic colors.

The above research shows that boundaries between color ranges in the color spectrum are different from language to language, but the best examples or the focal colors which speakers in different languages recognize for basic color terms are very similar. This is because colors do not objectively exist in the world without being related to human beings. These facts implicate two considerable hypotheses: one is that linguistic concepts are embodied in the sense that their acquisition is conditioned by human biological factors; the other is that human beings acquire linguistic concepts through daily experiences in the culture in which they live. In other words, language and concept could not be discussed without consideration of the embodiment of cognition.

2.3.2 Early evidence on embodied cognition

Early researchers have already found out some interesting evidence that human bodily or sensorimotor experiences constitute a part of their cognition. Today, the phenomenon is called *embodied cognition*. Kaden, Kapner, and Werner (1955), for example, showed that both objects and words which have upward directional orientation are seen relatively higher than those which have downward directional orientation. They first measured subjects' eye line which a subject experienced as his or her own eye level. They then placed pictures of hands pointing upward or downward on the wall at each subject's eye level and instructed them to adjust the location of the pointing hands to their eye level if they felt that the pictures were not at their eye levels. When they were asked to adjust, the subjects shifted the upward-pointing hand lower than their eye level and the downward-pointing hand higher. This is because the subjects felt that they had perceived the upward-pointing hand above their eye line and the downward-pointing hand below. Kaden et al. did the same experiment with verbal items, such as *climbing* and *rising* and *falling* and *plunging*. The results were strikingly similar to those found in the picture experiment.

Another interesting piece of evidence which suggests embodied cognition is *semantic satiation*. It is no doubt that the form and meaning of a word is inseparable: word form denotes word meaning and word meaning is transmitted through word form. People, however, often experience the loss of the meaning of a word when the word is repeatedly perceived or produced. This is called semantic satiation. Cognitive stability dissolves as a result of verbal repetition. Miller (1963) demonstrated that physical activity related to word meaning delays semantic satiation effect. He employed two simple action verbs, *push* and *lift*, and instructed subjects to repeat pronouncing each word until it lost its meaning under the following four conditions. In the first condition, subjects repeated each word

performing the denoted action. In the second, they repeated pronouncing each word performing action opposite to the denoted meaning (*e.g.*, pulling action with the repetition of *push*). The third condition was to repeat the pronunciation of each word performing an irrelevant action (*e.g.*, lifting action with the repetition of *push*). The fourth condition was no action accompaniment. He found that physical activity semantically consistent with the repeated words caused the greatest delay in semantic satiation, activity opposite to word meaning was the second, activity irrelevant to word meaning was the third, and semantic satiation occurred in the shortest time with no action.

2.3.3 Concept formation in infants

Developmental psychologists have pointed out the importance of body-environment interaction for concept formation. Werner and Kaplan (1963), for example, reported the interactive process with the following examples. A 25-month-old infant referred to milk by two distinct names: she referred to 'milk as drunk out of a bottle' as *mimi*, and 'milk as drunk out of a cup' as *te*. Both words described total situations in which milk was drunk by the infant. Another infant used the word *pin*, by referring not only to the pin itself but also to a piece of breadcrumb, to a fly, to a caterpillar, to all in the context of being picked up from the floor. That is, the concept of the word *pin* for her contained something small to be picked up with the fingers. Werner and Kaplan (1963) concluded that "the early names do not depict stable, circumscribed things but rather refer to global events in which things and the agent's (the speaker's) action upon things are intimately fused" (p. 118). Their comment indicates that actions infants perform with objects constitute an important part of their concepts of the objects.

Kobayashi (1997) emphasized the importance of action in forming concepts of objects by demonstrating that young children used action information to make inferences

about the identification of solid objects. In one experiment, she showed three groups of children an egg-shaped glass object and called it *muta* (a nonsense name). She then demonstrated the action of rolling the object (shape-related action) to the first group and the action of looking through it (material-related action) to the second group. To the third group, she did not show any action, but just introduced the object as *muta*. In the test, she handed all the children both an egg-shaped styrofoam object (an object of the same shape, but of a different material) and a pyramid-shaped glass object (an object of a different shape, but of the same material), and asked "which was *muta*?" The children's inference was influenced by the action information. The first group who had been shown the rolling action tended to relate the shape of the object to the name *muta*, while the second group who had been shown the looking-through action tended to relate the material of the object to the name. The introduction of the name did not influence the third group's inference about their choice significantly.

Masataka (2000), using the same experimental format, investigated whether an adult's gesture (pantomime of action) related to the shape and material properties of objects would influence children's identification of objects. In his experiment, an experimenter pretended to perform actions (for example, rolling an object in case of a ball-like object). The results showed that the children significantly utilized the adult's gesture for object categorization much like children in Kabayashi (1997) did. Focusing on the role of action in cognition, Masataka stated that "developmentally, what young children are required to do to learn a word is to bundle various modality-specific information stemming from recurring experience with the hearing of the word, and to somehow combine the information to represent a coherent and amodal notion" (p. 47).

The above research suggests that young infants tend to categorize objects and events in terms of actions they perform on objects. Werner and Kaplan (1963) already explained

children's concept formation process as follows: "the (pre-symbolic) world of the very young infant is primarily one of things-of-actions, articulated in terms of affective-sensory-motor patterns" (p. 18). According to Werner and Kaplan, sensory-motor and affective response patterns which an infant experience, at the next stage, "become the means by which he comes to know objects, to reflect on them, to present them to himself." In other words, the cognitive stage called "the internalization of sensory-motor patterns" has the function that "objects are given form, structure, and meaning through inner-dynamic schematizing activity which shapes and intertwines the sensory, postural, affective, and imaginal components of the organismic state" (p. 18). The inner sensory-motor patterns which Werner and Kaplan observed correspond to what is now called *image schemas* (discussed later) in the context of cognitive semantics.

2.4 Gesture

2.4.1 Early gesture

The emergence of gesture can be regarded as a process of symbol formation. Werner and Kaplan (1963) described how gesture changes from responsive, expressive activity through an imitative, depictive stage to context-independent, imaginative activity under the heading of *distancing*. Distancing is a process in which a symbol form becomes less identical and contiguous to the referent it stands for. Werner and Kaplan referred to Piaget's observations. Piaget's two children at the age of nine to eleven months imitated the opening and closing of the experimenter's eyes by alternately opening and closing their hands, mouth, or eyes. Another nine-month girl imitated Piaget's action of sticking out his tongue by consistently raising her forefinger. If a symbol is a unit of form (the signifier) and concept (the signified), those gestures seem to have some distance or indirectness

between the form and the content, because they were performed based on the sensory-motor patterns the infants had acquired.

When one of Piaget's daughters became two years and two months old, she let some gravel trickle through her fingers and said, "It's raining." The girl, without being restricted to the real situation, represented the imaginative event. Werner and Kaplan (1963) stated that "at the beginning stages of symbolic play, where play sphere and reality sphere are little differentiated, we may assume that from the child's viewpoint he *is* what he represents; there is little awareness of the body as being a *medium* (emphasis in original) of representation" (p.95). And more and more *distanced* gestures appear in make-believe situations as infants grow older. For example, "one child represented a house by standing erect with the arms slightly raised sidewise to indicate the roof; similarly, by bending down he tried to represent a mountain" (p. 97). Those representative gestures were imaginatively performed independently of the real situation. Little children who have learned to imitate actions and events can later represent actions and events with gestures without the support of concrete objects and real context.

2.4.2 Classification of gesture

The word *gesture* generally means natural body movements (mainly hands and arms) during speech. People gesture when they speak, and gesture and speech occur almost simultaneously, and gestures often have similar meaning. Gesture has been classified into different types by different researchers (*e.g.*, Ekman & Friesen, 1969; McNeill, 1992). The classification adopted here is a modification of the system employed by Kita (2000a) .

Gestures are generally divided into conventionalized gestures and spontaneous gestures. Conventionalized gestures are sign-like gestures whose hand shapes and meanings are firmly conventionalized, and are often called *emblems* (Ekman & Friesen,

1969). The *OK* sign is a good example of this type. Spontaneous gestures acquire their timing or content from accompanied speech. Rhythmic gestures, which are called *beats* (McNeill, 1992), are related to the rhythm of speech. Representational gestures, on the other hand, are related to the meaning of speech. Deictic gestures, called *deictics* (McNeill, 1992), refer to things, persons, or locations by pointing. Depictive gestures represent forms and meanings of the referents in the context. Iconic gestures, called *illustrations* (Ekman & Friesen, 1969) or *iconics* (McNeill, 1992), depict concrete objects and events. On the other hand, metaphoric gestures, called *metaphorics* (McNeill, 1992), visualize abstract meaning in terms of metaphor.

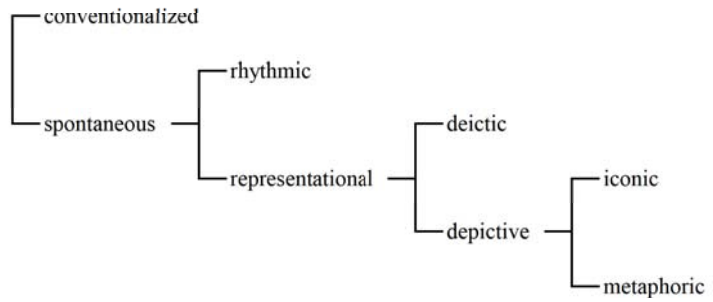


Figure 2-1 Gesture classification (kita, 2000a)

2.4.3 Features of gesture

McNeill and Duncan (2000: 142) observed a speaker raise his hand upward to depict a character in a story climbing up, saying "and he climbed up the pipe." The upward gesture coincided with 'up the pipe'. They inferred that "the speaker was thinking in terms of a *combination* (emphasis in the original) of imagery and linguistic categorical content; 'thought' (emphasis in the original) was in the form of a holistic image of the character

rising upward coordinated with the analytic, linguistically categorized meanings of 'up' and 'the pipe' (emphasis in the original)" (p. 142). Then, they stated that "the gesture and its synchronized co-expressive speech express the same underlying idea unit but do not necessarily express identical aspect of it" (pp. 142-143). As for gesture and language, De Ruiter (2000) insisted that "most gestures cannot be associated with single lexical items" (p.291) and concluded that "gestures do not have lexical affiliates but rather 'conceptual affiliates' (emphasis in the original)" (p. 291). It is assumed that the human brain can process holistic images and analytic linguistic components simultaneously, but natural gesture is characterized as holistic and may not be decomposed into a lexical item or unit.

Another feature of natural, depictive gesture lies in its idiosyncrasy. McNeill and Duncan (2000) commented that "they (gestures, added by the author) are not held to standards of good form; instead they are created locally by speakers while they are speaking" (p. 143) and "idiosyncratic gestures differ from arbitrary sound-meaning pairings in which meaning plays a role in determining signifier shape" (p. 143).

It is true that natural depictive gestures accompanying speech have no standards of form because a particular individual makes a particular speech in a particular situation, but different individuals often perform similar gestures. The concept of bigness, for example, is usually demonstrated by moving one's hands apart to show that the size of an object is big. This is because the notion of bigness is derived from one's daily experiences. In short, natural gesture is idiosyncratic because it is not based on any code, however, it can also be similar because people experience common physical and social activities.

2.4.4. Cognitive functions of gesture

A great deal is already known about what kind of gestures speakers perform, but little is known about why people gesture when they speak. *Emblems* like the *OK* sign and

deictic gestures such as referring to an object by pointing are communicatively intended and effective in conversation. Depictive gestures are intentionally performed to convey information to listeners in a particular situation, for example, in a noisy factory.

It is widely accepted that gesture is so closely integrated with speech that it should be investigated as an integral part of the utterance. Kendon (1988) stated that the function of gesture may be "looked at from a communicational point of view, whereby it appears as one of the resources a speaker has for conveying meanings" (p.133). Then, an answer to why people gesture when they talk is that gesture may play a role in the process of speech production. One account is that gesture facilitates access to lexical memory. Rauscher, Krauss, and Chen (1996) stated that gesture restriction affects speech fluency. They found that people gestured more when the content of their speech was spatial than when was not, and that people under the condition of preventing gesture spoke slowly when the content was spatial than when was not. Another account is that gesture facilitates speech fluencey. Rauscher et al. also found that restricting gesture led to increase in speech dyfluency (*e.g.* long and short pauses, filled pauses, uncompleted and repeated words, and restarted sentences).

Gesture surely plays a communicative role in speech, but there is some evidence to suggest that communication is not the only function gesture serves. For example, why do people often gesture on the phone? The most striking evidence is that blind people spontaneously gesture when they talk (Iverson & Goldin-Meadow, 1998). They found that congenitally blind children and adults gestured when they talked to sighted listeners. This fact denies the possibility that congenitally blind speakers gesture simply because they have seen others' gesture and have learned how to move their hands when they talk. They also found that congenitally blind speakers gestured even when they talked to blind listeners. This fact denies the possibility that congenitally blind speakers gesture because

they understand that gesture can convey useful information to the listener. Gesture, therefore, requires no model. Iverson and Goldin-Meadow (1998) referred to the possibility that "gestures that accompany speech may reflect, or even facilitate the thinking that underlies speaking" (p. 228).

Alibali, Kita, and Young (2000) presented the hypothesis that gesture is assumed to play a role in the process of conceptualization. Referring to McNeill (1992)'s ideas that gesture, together with language, helps constitute thought, and that gesture reflects imagistic mental representations which are activated at the moment of speaking, they hypothesized that "gesture is involved in the *conceptual planning* (emphasis in the original) of the message to be verbalized, in that gesture helps speakers to *package* (emphasis in the original) spatial information into units appropriate for verbalization" (pp. 594-595). To prove the hypothesis, Alibali et al. executed two tasks with five-year-old children. In the explanation task, children were required to explain whether some materials have the same quantity in different conditions (*e.g.*, children first see two glasses of sand, then they observe one glass of sand poured into a bowl). In the description task, the same material was used in the same procedure, but children were simply asked to describe how two items (*e.g.*, a glass of sand and a bowl of sand) look different. The explanation task demanded more constrained and complex thinking than the description task. The researchers, therefore, supposed that children's gesture production would differ in their replies if gesture was involved in their conceptual planning. In fact, children in the explanation task produced more gestures which described perceptual dimensions (*e.g.*, height, width, shape) of the items, and more gestures which differed in content from the accompanying speech. The results suggest that gesture plays a role in the conceptual planning of speech and contributes to speech production.

McNeill (1992) regarded an utterance not as a static sentence in speech or in writing, but as a process that has an internal development and has a sentence as its final stage. He called the peak of the internal development or the underlying starting point of the sentence *growth point* (p. 219). He hypothesized that the growth point is "a combination of image and word, of image and linguistic meaning category" (p. 220), from which speech and gesture are produced. In other words, speech and gesture are a single system or two aspects of one process. Kita (2000b) conceptualized speech and gesture from the viewpoint of separate streams of cognition. He hypothesized two modes of thinking: *analytic thinking* and *spatio-motoric thinking*. Analytic thinking organizes information by structuring abstract concepts like a 'script' (Schank & Abelson, 1977) or equivalents into linguistic items. On the other hand, spatio-motoric thinking is "the type of thinking normally employed when people interact with the physical environment, using the body (*e.g.*, the interaction with an object, locomotion, and imitating somebody else's action)" (Kita, 2000b, p. 164) and "can also be applied to the virtual environment that is internally created as imagery" (pp. 164-165). In short, analytic thinking deals with propositional knowledge, while spatio-motoric thinking deals with imagery based on physical experiences. Whether speech and gesture are explained in terms of a single conceptual process (as postulated by McNeill) or dual (as hypothesized by Kita), the content of a word or a sentence consists of propositional information and imaginal information. Speakers, therefore, can organize a wider range of information for speaking by performing representational gestures and actions in the virtual environment.

2.5 Action memory

In earlier laboratory research, memory retention was often investigated in terms of verbal stimuli like word lists and visual stimuli like pictures, but performing an action was

hardly considered to be relevant. Research on action memory began in the 1980s. In typical action memory research, experimental groups are required to read action phrases such as *clap your hand*, and *stretch your legs* and perform the denoted actions (subject-performed tasks or SPTs), or listen to an experimenter read the phrases and watch the person perform the actions (experimenter-performed tasks or EPTs), while control groups are asked to read the phrases or listen to an experimenter read the phrases (verbal tasks or VTs). It is acknowledged that recall in SPTs and EPTs is better than that in VTs. The memory enhancement by action performance is called *enactment effect*. The results of action memory experiments seem to be clear to everyone, but explanations of enactment effect are controversial.

2.5.1 Controversial views on enactment effect

Saltz and Donnenwerth-Nolan (1981) assumed that the effect of enactment on action memory was due to storage of motor images, not motor images stimulating visual images, and tested their assumption by means of selective interference. Selective interference effect means that when experimental subjects simultaneously do two tasks which use the same coding system (*e.g.*, two visual tasks or two motor tasks), they perform more poorly than when doing two tasks which use different coding systems (*e.g.*, a visual task and a motor task). This is because storage capacity for motoric memory, for example, may be overloaded by another motor task, not by a visual task, and vice versa. Saltz and Donnenwerth-Nolan found that the recall was better in doing tasks using different coding systems than using the same coding system. They insisted that visual processing and motor or enactment processing call for a different coding system.

Engelkamp (2001) argued enactment effect from the viewpoint that motor activities create different sensory experiences compared to verbal and visual experiences. He then

doubted Paivio and Csapo (1973)'s second experiment (mentioned in **2.2.2.**) employing pictures and words. In their dual coding theory, the task of sketching flashed pictures and that of sketching the referents of flashed words were regarded as visual encoding, but Engelkamp pointed out the possibility that the behavior of sketching had a strong effect on the enhanced recall performance by stating that "motor performance improves conceptual and/or motor information or even other modality-specific information" (p. 60). One way to test whether enhanced memory after enactment is at least partially due to motor information is to compare recalls in SPTs and EPTs. According to Englkamp (2001: 62), however, recall after SPTs are sometimes better and sometimes not, depending on the experiments' design and stimuli.

Kormi-Nouri, Nyberg, and Nilsson (1994) tested the effect of retrieval enactment on recall of SPTs and VTs. They designed different encoding/retrieval conditions: dual enactment (motor encoding and motor retrieval), single enactment (motor encoding and verbal retrieval, verbal encoding and motor retrieval), and no enactment (verbal encoding and verbal retrieval). The results showed an effect of encoding enactment but no additional effect of dual enactment. The effect of retrieval enactment was restricted to action phrases including word referring to body parts. Kormi-Nouri and Nilsson (2001) criticized the motor encoding view by Saltz and Donnenwerth-Nolan (1981) and Englkamp (2001), and stated that "the reason for the superior memory performance after motor encoding, compared to only verbal encoding, is due to better self-involvement in the former than in the latter case" (p. 98). They then assumed that "motor retrieval cues are not effective because the stored information is in a verbal rather than in a motor code" because "if the encoded information had been in motor code, motor cues would have been more effective and an encoding specificity advantage would have been found" (p. 100).

Instead, Kormi-Nouri and Nilsson considered enactment effect from two aspects. First, enactment, as an action event, facilitates interaction between the subject and the environment. The subjects are more involved in learning in SPTs than in VTs. "A better self-involvement in action events helps a rememberer to be more aware of his action, self-knowing, thereby leading to a better episodic memory" (p.105). Second, enactment facilitates linking a particular action and a particular object into a memory unit with the help of semantic knowledge (*e.g.*, verb-object relation). In other words, encoding enactment improves relational processing by integrating item-specific information. Believing in a single coding processing, Kormi-Nouri and Nilsson insisted that action performance facilitates the integration of action information with episodic memory.

2.5.2 Neuropsychological evidence for motor processing

Human cognition and action performance seem to be far more complicatedly interrelated than people once believed. Brain imaging technology can bring new data on brain functions for the processing of words and phrases associated with action.

Using Positron Emission Tomography (PET), Grafton, Fadiga, Arbib, and Rizzolatti (1997) investigated how differently 1) observation of tools (*e.g.*, razor), 2) silent naming of the tools, and 3) silent naming of their use (*e.g.*, to shave) would activate the brain. Tool observation activated the left dorsal premotor cortex, and silent tool-naming activated Broca's area (mainly related to speech production) without any additional activation in the dorsal premotor cortex. On the other hand, silent tool-use naming activated Broca's area, the dorsal premoter cortex, and the ventral motor cortex. The data shows that premotor cortex can process action information from action–related objects. Grafton et al., then, stated that "to categorize an object, it is not enough to have a description of its visual characteristics" (p. 235) and that "the premotor activations found in the present study may

subserve the motoric aspects of object semantics" (p.235). Neural activity difference between silent tool-naming and silent tool-use is very suggestive for the comprehension and retrieval of object concepts.

Using Transcranial Magnetic Stimulation, Oliveri, Finocchiaro, Shapiro, Gangitano, Caramazza, and Pascual-Leone (2004) investigated the following hypothesis: activation in motor cortex increases during the retrieval of action verbs like *throw* and action-related nouns like *key*, compared to that of non-action verbs like *belong* and non-action-related nouns like *cloud*. They found that the retrieval of action-related words, either verbs or nouns, induces greater activation of the motor cortex than that of non-action words. Activation in the motor cortex is related to motor properties, not grammatical categories of words. Oliveri et al., then, proposed two possibilities for the neural activation at the time of retrieving action words. One is that "motor schemata associated with a given word are embedded within its cortical representation, and are activated automatically whenever that word is retrieved" (p.377). The other is that "word production sometimes leads to the corollary generation of mental images related to the concept being retrieved. It might then be the case that the increased amplitude of the motor response is related to these (epiphenomenal) motor images, rather than to the semantics of the word as such" (p. 377).

Russ, Mack, Grama, Lanfermann, and Knoph (2002) classified views on enactment effect into two types: motor information reactivation view and action representation view. The former view hypothesizes that motor information, in addition to verbal and visual information, facilitates the encoding and retrieval of action phrases, while the latter view focuses on motivational and semantic components (self-awareness, intention to act, action schema, object knowledge and so on) involved in performing actions. Russ et al., then, investigated the brain activity of subjects recognizing action phrases which they had enacted, using fMRI. Neural activity was not limited to the motor cortex and neighboring

areas. That suggested that complex multi-modal associative processing was involved in enactment effect. Russ et al., then, assumed that "the beneficial memory effect of encoding by performance is mediated by a highly complex neuronal network, integrating representations of actions in external space with intentions and object knowledge" (p. 503). The cause of enactment effect may not be reduced to one element.

2.6 Cognitive semantics

Color concept mentioned in **2.3.1** indicates the involvement of bodily experiences in cognition. Embodiment in cognition or embodied cognition is explained from the perspective of cognitive linguistics (*e.g.*, Johnson, 1987; Lakoff, 1987). They assumed that concepts are highly motivated by bodily experiences in the world. Human body structure, for example, gives an experiential basis for understanding the world. Generally, language (*e.g.*, English, Japanese) has a group of lexical items which refers to asymmetrical axes of spatial orientations, such as *up-down*, *front-back*, and *right-left*. Their referential asymmetry is related to the asymmetry of human body structure (Hill, 1982). According to those embodied spatial concepts, people can identify the orientation of objects and locations. In other words, the world is conceptualized in the way human beings are embodied.

2.6.1 Image schemas

Based on the study of spatial concepts, cognitive linguists such as Lakoff and Johnson proposed that a small number of cognitive patterns provide a foundation for thought, reasoning and imagination. They assumed that a substantial part of human knowledge is not static and propositional, but is grounded in bodily experiences and structured by various patterns of sensory-motor interactions with the environment. These patterns are called *image*

schemas. Johnson (1987) defined an image schema as "a recurring, dynamic pattern of our perceptual interactions and motor programs that gives coherence and structure to our experience" (p. xiv). Among such image schemas are schematic structures such as UP/DOWN, FRONT/ BACK, CONTAINER, PATH, CIRCLE, SCALE, LINK, BALANCE, and FORCE. These image schemas cover a wide range of everyday experiences. For example, the CONTAINER schema, which represents in-out orientations, is pervasive in everyday experiences. Johnson (1987) described our daily activities as follows: "you wake *out* of a deep sleep and peer *out* from beneath the covers *into* your room. You gradually emerge *out* of your stupor, pull yourself *out* from under the covers, climb *into* your robe, stretch *out* your limbs, and walk *in* a daze *out* of the bedroom and *into* the bathroom " (emphasis in the original) (p. 30-31).

Image schemas are cognitive, not static or propositional, in the sense that they are embodied. Johnson (1987) assumed that "(1) meaning in natural language begins in figurative, multivalent patterns that cannot typically be reduced to a set of literal concepts and propositions; (2) the patterns and their connections are embodied" (p. 5). Johnson, at the same time, emphasized close relations between image schemas and propositional contents of utterances by stating "*this propositional content is possible only by virtue of a complex web of nonpropositional schematic structures that emerge from our bodily experience*" (emphasis in the original) (p. 5). On the other hand, image schemas are not concrete mental images, either. Concrete images are images of some particular things, but image schemas are supposed to be more abstract and general imagery. They contain structural features common to many different bodily movements, objects, and events. In short, image schemas "operate at a level of mental organization that falls between abstract propositional structures, on the one side, and particular concrete images, on the other" (p. 29).

2.6.2 Basic level

A coin in the pocket may be called *money*, a *coin*, or a *dime*. A dog on the chain may be called a *mammal*, a *dog* or a *bulldog*. What determines a name given to a child? This is the question Brown (1958) discussed. The brevity and frequency of the name may matter, but he pointed out that "the best generalization seems to be that each thing is first given its most common name. This name seems to categorize on the level of usual utility" (p. 19). The level of usual utility is *basic level categories* in Rosch, Mervis, Gray, Johnson, and Boyes-Braem (1976).

Rosch et al. defined a category, for example, *dog* and *seat* as basic level categories. They characterized basic level objects or categories, in contrast to more general categories and more specific categories, in terms of attributes common to the category, motor movements common to the interaction with the object, and the identification of the average shape of the object. They found that basic level objects are first seen and recognized as members of their basic category, and that basic object names are daily used by adults and acquired by young children. Rosch et al., then, concluded that "basic objects are the most general classes at which attributes are predictable, objects of the class are used in the same way, objects can be readily identified by shape, and at which classes can be imaged. Basic objects should generally be the most useful level of classification. Universally, basic object categories should be the basic classifications made during perception, the first learned and first named by children, and the most codable, most coded, and most necessary in the language of any people" (p. 435).

Lakoff (1987) summarized basic level categories in four respects: 1) as members of basic level categories are prominent in perception, they can be represented by a single mental image, 2) when people interact with basic level category members or objects, they make many common bodily and muscle movements, 3) as labels for basic level categories

are usually short and commonly used, the terms are contextually independent or neutral and first learned by children, 4) the basic level is the most inclusive level of categorization at which numbers of a category share many attributes. His view can be reduced to the following comment that "basic-level categorization depends upon experiential aspects of human psychology: gestalt perception, mental imagery, motor activities, social function, and memory" (p. 37).

According to Pansky and Koriat (2004), many studies of memory showed that it loses specific details and becomes more abstract over time, while some studies showed that abstract or general memory can be retrieved in more specific terms. They, then, hypothesized that retained information may converge on a level between the most general and the most specific in taxonomy, the basic level, saying "If the BL (basic level, added by the author) is cognitively optimal for perception, categorization, communication, and knowledge organization, is it also optimal for retaining episodic information over time?" (p. 53). They presented subjects with stories with target items of the three abstraction levels embedded; for example, *clothes* at the superordinate level, *pants* at the basic level, and *jeans* at the subordinate level, and then tested the recall level immediately and again, one week later. The results demonstrated that retained information tends to converge on the basic level whether the information is at the superordinate level or the subordinate level at the time of encoding.

The above research and comments indicate that the basic level is a desirable level of abstraction for perceiving and retaining information. Probably the basic level categories (words and concepts) will also be the most appropriate items for gesture or expressive meaningful bodily movements because they can be represented by a single mental image and have a strong association with motor movement.

2.6.3 Metaphor

Metaphor is traditionally viewed from the following three aspects: 1) conveying ideas which are difficult to convey in literal expressions, 2) giving a compact sketch of complex information, 3) giving a detailed and vivid image of experiences. Metaphor is viewed as a way or strategy of conveying what is difficult, complex, and abstract by means of what is plain, simple, and concrete (Gibbs, 1994: 124-125).

Lakoff and Johnson (1980), from the viewpoint that metaphor is a conceptual phenomenon, stated that "the essence of metaphor is understanding and experiencing one kind of thing in terms of another" (p. 5). In their view, metaphor is not just a matter of language, but a matter of thought and experience, too. They classified conceptual metaphors into three types: *structural metaphor*, *orientational metaphor*, and *ontological metaphor*.

Structural metaphors are conventional metaphors. Many expressions related to time such as "I lost a lot of time when I got sick" are constructed through a structural metaphor TIME IS MONEY. The conceptual metaphor conceives an abstract concept of time in terms of a concrete concept of money. Consequently, "we understand and experience time as the kind of thing that can be spent, wasted, budgeted, invested wisely or poorly, saved, or squandered" (p. 8).

Orientational metaphors and ontological metaphors, however, are different from structural metaphors in their existential basis. Orientational metaphors are based on basic experiences of human spatial orientation such as up-down, front-back, and in-out. One example is HAPPY IS UP. Happy feeling is associated with up-orientation such as *I'm in high spirits*, while sad feelings are linked with down-orientation such as *I'm feeling down*. This is an example of SAD IS DOWN metaphor. A person in sadness and depression is typically in a drooping posture, but goes into an erect posture when in a positive emotional state. Lakoff and Johnson (1980) stated that "since there are *systematic correlates* (emphasis in the

original) between our emotions (like happiness) and our sensory-motor experiences (like erect posture), these form the basis of orientational metaphorical concepts (such as HAPPY IS UP)" (p. 58). On the other hand, ontological metaphors are metaphors in which abstract concepts are understood and experienced in terms of concrete concepts such as objects, containers, and substance. For example, VISUAL FIELDS ARE CONTAINERS metaphor, which regards "visible field" as a kind of bounded area, constructs expressions such as *The ship is coming into view*. This metaphor is also based on an existential correlation between what a person sees and a bounded physical space.

Lakoff (1987: 8) explained the structure of conceptual metaphors by saying that "each metaphor has a source domain, a target domain, and a source-to-target mapping" and that "metaphor is *natural* in that it is *motivated by the structure of our experience*" (emphasis in the original). In the above conceptual metaphors, mood and visual field are understood in terms of image schemas such as UP, DOWN, and CONTAINER. Mood and verticality, and visual field and containment have a correlation in human experiences, and *up* and *down* (represented by verticality) and *in* and *out* (represented by containment) are the basic level categories. Therefore, metaphor, which has been regarded as abstract reasoning, is constructed on the basis of common mental and bodily experiences. Lakoff and Johnson (1999: 58) re-defined conceptual metaphors as "cross-domain mapping, from a *source domain* (the sensorimotor domain) to a *target domain* (the domain of subjective experience)" (emphasis in the original,). If "many, not all, of our abstract concepts are defined in significant part by conceptual metaphor" (p. 128), they can be readily understood according to people's daily experiences.

Conceptual metaphors, motivated by physical experiences, are a surely useful mental model for explaining human concept structure, but the mental model will have both its advantages and limitations. The present dissertation will consider conceptual metaphors to

be an effective conceptual strategy for interpreting a wide range of human concepts.

2.6.4 Metaphor in gesture

In gesture, too, abstract concepts are metaphorically represented by concrete concepts. Metaphoric gestures are, as already mentioned in **2.4.2**, one type of spontaneous gestures. Metaphoric gestures are originally iconic, but their pictorial content metaphorically represents abstract ideas. McNeill (1992) referred to a speaker who was talking about a cartoon with his hands rising up as if offering an object to a listener, and stated that the speaker's gesture represented the genre of the cartoon by displaying it as a physical object. This is an instance of the *conduit* metaphor (Reddy, 1979). The conduit metaphor is a principal metaphor for communication which regards ideas, meanings, and knowledge as objects, linguistic expressions, stories, and genres as containers, and communication as the act of sending the containers. So the conduit metaphor is a family of related conceptual metaphors such as IDEAS ARE OBJECTS and COMMUNICATION IS SENDING. People, on the other hand, often say in daily conversation, "Your words are hollow" or "His words carry little meaning." That is, the concept of communication is conceptualized by the same conceptual metaphors in gesture and in speech. Spontaneous gesture accompanying speech can show evidence of the psychological reality of conceptual metaphors.

Cienki (1998) found, based on a series of interviews with American college students, that they used gestures depicting a straight line when referring to honesty and truthfulness in speech. The combination of the concept of straightness and the concept of honesty or truthfulness is identical to linguistic expressions such as *He is straight as an arrow* and *He gave a straight talk*. Phrases such as *bend the truth* and *twist words* also suggest the combination. Cienki (1998) also analyzed a series of videotaped academic lectures and discovered that gestures and speech were motivated by common conceptual metaphors.

She observed that a speaker held out one or both open hands with their palm upwards toward a listener when the speaker put forward arguments. The gesture, she assumed, iconically represented the act of holding out a physical object to the listener and metaphorically represented a comment on the ongoing discourse. Another observation was that a speaker extended one or both open hands with their palm outwards towards a listener. She explained that the gesture iconically represented a barrier to an offered object and metaphorically represented a request for the listener not to speak. It is clear that those gestures were motivated by conceptual metaphors, IDEAS ARE OBJECTS and COMMUNICATION IS SENDING, respectively.

In short, gestures often bear a close cognitive structure to linguistic expressions they accompany. In other words, gestures agree with a cognitive linguistic hypothesis that human conceptualization is grounded in physical experience and that abstract concepts are metaphorically understood in terms of concrete concepts.

2.6.5 Metaphor in American Sign Language (ASL)

Sign languages are not a set of natural gestures or pantomimes, but a language based on a visual mode, and they have their own syntax and semantics. Signs are composed of several physical and spatial elements: the shape, location, orientation, and motion of articulating hands. As a visual language, many signs are iconic and close to natural gestures and mimes in form. The ASL sign for a tree, for example, is articulated by placing the elbow of the upright right hand on the palm of the left hand stretched horizontally. The flat left hand represents the ground level, the upright right hand represents the trunk, and the five fingers represent branches and leaves. The sign's form directly resembles that of a tree on the ground.

Iconic properties or iconicity of signs can be described as a transparent connection between form and meaning. Form represents meaning, but an iconic form is not an objective representation of an object. Taub (2001), for example, hypothesized that an iconic sign can be created through complex processes, image selection, schematization, and encoding. The lexicalization process of the ASL sign for a tree is explained as follows: the concept of a tree is associated with many different sensory images (*e.g.*, visual, motor, auditory). So, first, a prototypical image of a tree is selected, for example, an image of a trunk with branches and leaves growing from the ground. Next, the detailed image of a tree is schematized into a flat level surface, a tall vertical shaft, and a complex branching structure. The schematized image, then, is mapped to linguistic knowledge and coded into a sign form.

ASL also has an abundance of metaphoric signs for abstract meaning (Taub, 2001; Wilcox, 2000). For instance, the articulation of the ASL sign corresponding to English verb *inform* is that the fingertips of both hands are positioned on the forehead, and then both hands are swung down towards the viewer and opened as if taking knowledge from the head and passing it to the viewer. In this sign, knowledge is regarded as an object in the head, and communication is regarded as sending the object to the viewer. This conforms to the conduit metaphor, which consists of the conceptual metaphors, IDEAS ARE OBJECTS and COMMUNICATION IS SENDING. Such conceptual metaphors are seen not only in English, but also in natural gestures in North America, and ASL.

Conceptual mapping like that is not random. For example, the power and consistency of conceptual metaphors can be well observed in the conceptualization of time. Taub (2001: 115) stated that "it is extremely common in the languages of the world for *time* to be metaphorically understood in terms of *space*" (emphasis in the original). Time is conceptualized as a moving object passing by a person by means of TIME IS A MOVING

OBJECT metaphor, or conceptualized as a location through which a person moves by means of TIME IS A LANDSCSAPE WE ARE MOVING THROUGH metaphor. In both conceptualizations, the location of the person represents present time, future time is ahead of the person, and past time is behind. English-speaking people say, "We are looking forward to the arrival of Christmas" and "We are coming up on Christmas," and Japanese speakers say, "haru ga kita (spring has come)" and "haru ni chikazuku (We are approaching spring)." If time is conceptualized as a moving object (human beings as a location), it is oriented with the direction of its advance, and if time is conceptualized as a location (human beings as a moving object), it is oriented with the front and back of human beings. Both time and human beings, either a moving object or a location, encounter and pass by each other. Two related conceptual metaphors, FUTURE IS AHEAD and PAST IS BEHIND are also true of ASL. The ASL sign for future is articulated by moving the right hand ahead of a signer, and the sign for past is articulated by holding the right hand above the right shoulder and throwing it back over the shoulder.

It is evident that many linguistic items, both words and signs, and natural gestures can share the same mental images and conceptually correspond to each other. That fact, therefore, suggests that Japanese learners of English can use the conceptual images to understand English words if the images are common to their daily experience.

2.6.6 Bodily experience

Language is concrete and abstract. Abstract words, as mentioned in **2.2.2**, are typically rather high in meaningfulness but low in imagery and concreteness. Generally speaking, abstract concepts are considerably harder to remember than concrete ones because they are considered to be indirectly tied to sensory-motor experiences, resulting in evoking poor imagery. Many abstract concepts, however, are assumed to be structured by mapping from

sensory-motor experiences to non-sensory-motor or subjective experiences. If those specific sensory-motor experiences get awaken and refreshed in the mind of learners, related abstract concepts will be better recognized and retained.

One principle of conceptual metaphor theory is that the metaphorical mapping is motivated by aspects of bodily experiences. Lakoff and Johnson (1980) explained the structure of metaphors from the view point of experiential correlation. MORE IS UP/LESS IS DOWN metaphors, for example, conceptualize the concept of quantity in terms of verticality, and motivate expressions such as *My income rose last year* and *He is under age*. The metaphors are grounded in human experiences. If more of some objects are piled, the level goes up, and the level goes down if some are taken out form the pile. The ASL sign corresponding to English verb *increase* is, roughly speaking, performed by a motion of adding to a pile, and the sign for *decrease* is a motion of taking some of a pile.

Johnson (1997) introduced the idea of *conflation* into the question of how metaphor is motivated. He stated that young children encounter scenes conflating metaphorical meanings of words with their literal meanings. For example, the use of *see* in "Oh, I see what you wanted", an utterance a parent may use in response to a child's request to go and get a toy (p. 164). The utterance can be interpreted in multiple ways, as the speaker's visual experience, cognitive experience, or conflation of both. Grady and Johnson (1997) hypothesized that young children can acquire metaphorical meanings in a conflated situation like that above. Conflation can motivate many abstract concepts. The meaning of affection, for example, is motivated by experiences of being held affectionately. The sensory-motor experience (body temperature or warmth) is correlated with the subjective experience (affection). The mapping between both experiential domains underlies expressions such as *They greeted me warmly* (Grady, 1997: 293). The ASL sign

corresponding to English noun *affection* is performed by a motion of pulling something in as if giving it a hug.

Sweetser (1990) examined semantic changes of perception verbs in English and other Indo-European languages, and assumed that "metaphor is a major structuring force in semantic change" (p. 19). Verbs of vision, for example, are often derived from verbs of manipulation: *behold, perceive* (<Lat, *-cipio* "seize"), *examine* (<Lat, *ex-agnen-* "pull out from a row"), *discern* (<Lat, *dis-cerno* "separate"). Viewing is conceptualized in terms of manipulating objects. This conforms to the generalization that "change proceeds from concrete to abstract" (p. 29).

In speech activity, which is metaphorically structured in terms of sending objects, participants exchange objects (ideas) with each other. Sweetser (1987) pointed out that the metaphorical structure is also seen in etymologies like *suppose* (<Lat, *sub+ponere* "put under") and *propose* (<Lat, *pro+ponere* "put forward"). In the former, an object (idea) is manipulated in the head of one participant, and the object (idea) is offered to another in the latter. What should be considered here is that many expressions, words and sentences, are metaphorically related to each other by the conceptual metaphor, COMMUNICATION IS SENDING AN OBJECT. That is, speech-related words such as *suppose*, *propose*, *inform, accept*, and *refuse* are all conceptually associated with each other, visually perceived, and in addition, can be represented by gesture. Through metaphorical thinking, abstract words in the same conceptual domain could be encoded and retrieved systematically.

In the following chapter (Study 1), the author first of all tests whether gesture images or motor images can be effective recall cues in the framework of the dual coding theory. In the previous studies on the dual coding theory, static images like pictures were recall cues. In the Chapter 4 (Study 2), the author tests whether gesture images can work as effective recall cues for abstract words. Abstract words are metaphorically represented by iconic

gesture images. In Chapter 5 (Study 3), the author tests whether gesture performance is different from or the same as gesture-image presentation in terms of recall cues. Enactment effect is expected. In Chapter 6 (Study 4), recall targets are extended from word level to sentence level, and the predicate of a sentence is enacted. In previous studies on action memory, well-known action verbs were enacted. In Study 4, then, state verbs as well as action verbs, and low frequency verbs as well as high frequency verbs are enacted by gesture. In Chapter 7 (Study 5), recall targets are extended up to paragraphs. A paragraph is usually composed of complex rhetorical relations such as cause-effect and problem-solution. Mnemonic techniques, on the other hand, usually emphasize one cue to one target relation. A paragraph is divided into idea units, and one word per each idea unit is enacted for facilitating the recall of the whole paragraph. Gestures are generally motoric, concrete, and specific. The theme of the present dissertation, therefore, is to investigate if such gestures can be effective in recalling English expressions in various conditions.

3. Study 1

This study tested the validity of gesture images or motor images as recall cues in the bilingual dual coding theory. As previously mentioned in **2.2.2**, Paivio and Desrochers (1980) extended the dual coding theory to bilinguals, and Paivio and Lambert (1981) tested the approach to bilingual memory. They required French-English bilinguals to learn a mixed list of pictures, French words, and English words by writing the English names of the pictures, translating the French words, and copying the English words. The test indicated that word recall increased from monolingual processing to bilingual processing to verbal-nonverbal processing. The results can be interpreted differently from other theories. Assuming that two verbal systems and one imagery system are partly independent and party interconnected, the bilingual dual coding theory explained that some modal-specific component of the nonverbal stimuli (visual imagery in this case) strengthened the link between the imagery and the two verbal systems, resulting in the superiority of the dual coding to the verbal coding.[1]

This study applied the bilingual dual coding theory to Japanese college students who learn English as a foreign language, and investigated how imagery contributed to word recall both in monolingual processing and bilingual processing, by examining the recall of generated English words in a translation task and in a copy task, with and without gesture images, in an immediate test and a delayed test.

3.1 Purpose

The study has three purposes. The first was to confirm whether bilingual processing would be more effective than monolingual processing. The second is to confirm the effects of imagery stimuli in monolingual processing and bilingual processing. The third is to judge whether gesture images would work as stimuli for facilitating memory retention.

3.2 Subjects

Subjects were 60 undergraduate students who belonged to the faculty of medicine at a national university in the Hokuriku district. All the subjects, 38 males and 22 females, began to study English at junior high school. Two groups of 30 subjects were chosen from two different English classes. The level of students' English in both classes was considered to be similar by the teacher in charge.

3.3 Design

A 2×2 factorial design was used for this study. The first factor was either presenting gesture images or not, which was the between-subjects variable. The second factor was translating Japanese words into English or copying English words. That was the within-subjects variable.

3.4 Material

A total of 40 simple English nouns were selected as target words on the basis of their imagery and concreteness ratings (Paivio, Yuille, & Madigan, 1968). Paivio et al. measured 925 nouns along a 7-point scale by defining imagery in terms of a word's capacity to arouse nonverbal images and concreteness in terms of directness of reference to sense experience. All the selected 40 words had an imagery rating which was more than 6.14 (1 = low, 7 = high imagery value), and a concreteness rating which was more than 6.00 (1 = low, 7 = high concrete value). The 40 target words belonged to 20 categories, and every category contained two words. Either of the two words of each category was translated into Japanese and was given to the control group or the experimental group. The control group was presented with 40 words, 20 English words and 20 Japanese words. The 20 Japanese words were Japanese translations of 20 English words for the experimental group. The

experimental group was also given 40 words, 20 English words and 20 Japanese words, with gesture images. The 20 Japanese words were Japanese translations of the 20 English words given to the control group (Appendix C). This was done to balance language load in translation[2]. The 40 words were digitally processed on a computer and recorded on a videotape for the control group. Gestures for the words were digitally videotaped, and synchronized with the 40 words on a computer. Then, the visual materials were recorded on a videotape for the experimental group. Gestures for the words were pictorial gestures. Gestures for *dawn*, *fork*, *grass*, *oven*, *pipe*, *shoes*, *sugar*, and *sunset* were based on the ASL signs corresponding to those words, but those signs are iconic signs which depict the words' referents or human behaviors associated with the referents.


Figure 3-1. *book*


Figure 3-2. *sky*


Figure 3-3. *grass*


Figure 3-4. *sunrise*

3.5 Procedure

The subjects were informed that they would be shown an educational video of 20 English words and 20 Japanese words (with gesture images in the case of the experimental group). The subjects, then, were required to do two tasks: to translate a Japanese word into English and write its English equivalent on a sheet of paper, and copy an English word on the paper. Stimulus images remained on the screen for five seconds. Prior to the learning task, two sample words (one Japanese word and one English word), with gesture images in the case of the experimental group, were shown to the subjects to provide a chance to become familiar with the task. The target English and Japanese words were presented

alternately. No sound was included.

Five minutes after the tasks, the subjects were unexpectedly given a five-minute recall test. They were asked to recall as many of the generated English words as possible in any order. One week later, 50 of the 60 subjects took the delayed test. They were unexpectedly given a five-minute recall test again.

3.6 Results

Table 3-1 shows all mean numbers and standard deviations for the recalled English words under all the conditions in both the immediate test and the delayed test. Table 3-2 is an analysis of variance for the results of the immediate test, and Table 3-3 is for the results of the delayed test.

Table 3-1. Means and standard deviations for the immediate test and delayed test

stimuli	M		SD	
	immediate	delayed	immediate	delayed
translation (T)	9.73	5.88	2.16	1.90
copy (C)	5.55	3.48	2.58	1.81
T and image	13.30	10.84	2.10	2.29
C and image	12.23	9.64	1.92	2.94

Table 3-2. Analysis of variance for the immediate test

	SS	*df*	MS	F	*p*
gesture image (A)	790.53	1	790.53	101.97	0.00
task type (B)	208.03	1	208.03	104.62	0.00
interaction (A × B)	73.63	1	73.63	37.03	0.00
error S(A)	449.67	58	7.75		
error S(A) × B	115.33	58	1.98851		

In the immediate test, an analysis of variance revealed that there was a large and

significant main effect for gesture-image presentation, F (1, 58) = 101.97, $p < 0.05$. The main effect for task difference (translation or copy) was also significant in the predicted direction, F (1, 58) = 104.62, $p < 0.05$. The effect of the interaction of the image presentation and the task difference was significant, F (1, 58) = 37.03, $p < 0.05$. As for the delayed test, an analysis of variance revealed that there was still a large and significant main effect for gesture-image presentation, F (1, 48) = 91.60, $p < 0.05$. The main effect for task difference was significant, F (1, 48) = 41.85, $p < 0.05$. The interaction of both variables had a minor effect, F (1, 48) = 4.65, $p < 0.05$.

Table 3-3. Analysis of variance for the delayed test

	SS	*df*	MS	F	*p*
gesture image (A)	772.84	1	772.84	91.60	0.00
task type (B)	81.00	1	81.00	41.81	0.00
interaction (A × B)	9.00	1	9.00	4.65	0.04
error S(A)	405.00	48	8.44		
error S(A) × B	93.00	48	1.94		

3.7 Discussion

The results of the present study supported the bilingual dual coding theory that two verbal systems corresponding to a bilingual's two languages and one nonverbal imagery system are independent functionally but are interconnected referentially. Both the main effect for gesture-image presentation and the main effect for task difference between a bilingual task and a monolingual one were statistically significant. The results suggest that gesture images can play an effective role in word memory retention.

The role of imagery in learning and memory is well known. Paivio's series of studies on the dual coding theory have demonstrated the mnemonic significance of imagery. The imagery stimuli employed, however, were mostly pictures. Action memory research,

which began in the 1980s, has provided evidence that enactment of action phrases like *close the door* is effective in recalling the phrases. The memory enhancement is called enactment effect. In this study, gesture images were employed as imagery stimuli for English nouns. Gesture has been considered to be related to language in some ways[3], but only a few researchers paid attention to gesture for educational purposes.[4]

The dual coding view could explain how the additive effect of the gesture images and the verbal information contributed to the better recall by the gesture images evoking mental images and facilitating the association between the mental images and the verbal information. English words employed for this study were typical concrete words. Concrete nouns such as *house* and *flower* intrinsically contain high concreteness and imagery. The gesture performance was presumed to activate the high imagery.

The multi-modal coding view (*e.g.*, Engelkamp in **2.5.1**) could explain that the motor information by gesture performance, in addition to the visual information and the verbal information, contributed to the enhanced recall. The modal-specific information, moreover, might have activated the motor component involved in the concept of words such *hammer*, *spray*, *car* and *ship* because sensory-motor information is an important part of the concept of motion-related words, as previously mentioned in **2.3.2** and **2.3.3**. The motor information might also have stimulated other sensational components of the concept, sense of taste in case of *apple*, *lemon*, *pepper*, and *sugar*.

The single conceptual coding view (*e.g.*, Kormi-Nouri & Nilsson in **2.5.1**), might explain the motivational aspect of the gesture performance. Gesture performance demands strong self-involvement, such as self-awareness, intention to perform, and even a sense of success or failure. As many subjects took part in the experiment together, emotional factors might have complicated the retention of the words.

The significant effect of the interaction of the gesture images and the task difference is

supposed to be related to the stronger effect of gesture images on the copy task. In the immediate test, for example, the means of recalled words in the translation task increased from 9.73 to 13.30, while that of recalled words in the copy task jumped form 5.53 to 12.23. The gesture images might have strengthened the associative link between the imagery system and the verbal systems, and the associative link between the verbal systems. The results suggest that gesture can facilitate Japanese students in learning English words, at least concrete nouns, without translations.

3.8 Conclusion

This study was conducted to test the validity of gesture images as recall cues in the frame of Paivio's dual coding theory. Gesture images were employed as nonverbal cues for word recall. The findings of this study can be summarized as follows: first, bilingual processing (Japanese and English) was more effective for the recall of English concrete nouns than monolingual processing (English). Second, gesture-image presentation was effective for the recall of English concrete nouns, either in monolingual condition or in bilingual condition. Third, gesture images can be effective cues for recalling concrete English nouns. English concrete nouns can be learned without Japanese equivalents if they are provided with gesture images.

Notes

1 Amedt and Gentile (1986) tested the dual coding theory with French-English bilinguals using pictures as visual cues, and their results supported the bilingual dual coding theory. Matsumi (1994) tested the dual coding theory with Japanese subjects by instructing them to make up mental images.

2 Norio Matsumi (Hiroshima University) suggested to the author that language load between an experimental group and a control group should be balanced.

3 McNeill (1992) assumed that gesture and speech are closely linked in meaning and function and demonstrated that gestures can depict abstract ideas as well as concrete ideas.

4 Riseborough (1981) showed the role of gesture in listening comprehension, and suggested that gesture information made listening tasks more vivid by evoking mental images stored in the mind of listeners. Kohno (1984) also showed that kinesic information such as gesture and body motion helped listening comprehension. Kellerman (1992) and Antes (1996) discussed the role of gestures in second language education.

4. Study 2

Vocabulary learning techniques usually work well on concrete words. However, they have difficulty in facilitating recall of abstract words. Gu (2003) summarizes that "not all words are equally suitable for mnemonic mediation" (p. 13). In one experiment, Paivio and Csapo (1973) used pictures, concrete words, and abstract words as dual coding materials. Recall increased systematically from abstract words to concrete words to pictures. The results suggest that items with high imagery value are better remembered than ones with less imagery.

Thus, this study examines the recall of abstract nouns with gesture-image presentation. The effect of gesture images on the recall of abstract nouns depends on how abstract concepts are visualized. Abstract words are low in concreteness and imagery, but are considered to be relatively high in meaningfulness (**2.2.2**). If gesture, therefore, can illustrate abstract concept through motor images linked with prior experience and knowledge, it will be a good recall cue. It is surely difficult to come up with concrete and familiar images for abstract words. Frick-Horbury (2002) showed that self-generated gestures were effective for the recall of abstract words as well as concrete words.[1] In this study, however, gesture images used for the abstract words are based mostly on American Sign Language (ASL) signs corresponding to the words. Metaphorical gesture images borrowed from ASL signs are expected to be good visual and conceptual mediators for Japanese subjects.

4.1 Purpose

The purpose of the study was to test the effects of hand gesture images on recall of English abstract words. The validity of gesture images would be discussed.

4.2 Subjects

The subjects were 40 undergraduate students learning English as a foreign language. All the subjects, 24 males and 16 females, began to study English at junior high school. They belonged to the faculty of medicine at a national university in the Hokuriku district. Two groups of 20 subjects were chosen from two different English classes. The level of students' English in both classes was considered to be similar by the teacher in charge.

4.3 Design

A 2×2 factorial design was used for the experiment. The first factor was either presenting hand gesture images or not, which was the between-subjects variable. The second factor was copying concrete words and abstract words, which was the within-subjects variable.

4.4 Material

A total of 40 target words, 20 concrete words and 20 abstract words, were selected on the basis of their imagery and concreteness ratings (Paivio, Yuille, & Madigan, 1968). The 20 concrete words had an imagery rating which was more than 6.40 (1 = low imagery value, 7 = high imagery value), and had a concreteness rating which was more than 6.90 (1 = low concrete value, 7 = high concrete value). 17 of the 20 concrete words were items used in Study 1. On the other hand, the 20 abstract words had an imagery rating which was less than 5.00 and had a concreteness rating which was less than 4.00 (Appendix D) . Gestures for the concrete words were pictorial gestures which describe the objects and actions which the words represent. The 40 words were digitally processed on a computer and recorded on a videotape for the control group. Gestures for the abstract words were based on the ASL signs corresponding to the words. Gesture performance was digitally

videotaped, and synchronized with the 40 words on a computer. Then, the visual images were recorded on a videotape for the experimental group. No sound was included.


Figure 4-1. *tower*


Figure 4-2. *hammer*


Figure 4-3. *power*


Figure 4-4. *promotion*

4.5 Procedure

All the subjects were informed that they would be shown an educational video of 40 English words (with gesture images in the case of the experimental group). All the subjects were asked to copy each word on a sheet of paper after viewing the video. Each word remained on the screen for five seconds. Prior to the learning task, two sample words (one concrete word and one abstract word), with gesture images in case of the experimental group, were shown to the groups in order to become familiar with the task. Five minutes after the task, the subjects were unexpectedly given a five-minute recall test. They were required to recall as many words as possible in any order. After the recall test, a translation test was given to the subjects. The subjects were asked to translate the 20 abstract words into Japanese in another five minutes.

4.6 Results

Table 4-1 shows means and standard deviations for the results in all the conditions. As table 4-2 shows, an analysis of variance revealed that there was a significant main effect for gesture-image presentation, $F(1, 38) = 78.47, p < 0.05$. The main effect for word type was significant, $F(1, 38) = 10.31, p < 0.05$. The effect of the interaction of gesture-image

presentation and word type was not significant, $F(1, 38) = 0.35, p > 0.05$. Table 4-3 shows that there was no significant difference between the mean numbers of correct translations of the 20 abstract words.

Table 4-1. Means and standard deviations

Stimuli	M	SD
concrete (C)	7.85	1.67
abstract (A)	6.75	1.71
C + image	11.35	2.01
A + image	9.75	1.67

Table 4-2. Analysis of variance

	SS	*df*	MS	F	*p*
gesture image (A)	211.25	1	211.25	78.47	0.00
word type (B)	36.45	1	36.45	10.31	0.00
interaction (A × B)	1.25	1	1.25	0.35	0.56
error S(A)	102.3	38	2.69		
error S(A) × B	134.3	38	3.53		

Table 4-3. Means of correct translations

	M	SD
control	18.85	0.99
experimental	18.90	0.97

4.7 Discussion

The study examined the effects of gesture images on the recall of English concrete and abstract words. As predicted, the subjects recalled more concrete words than abstract words, and the gesture images facilitated the recall of both concrete and abstract words.

Gestures for the concrete nouns were considered to be natural and describe the referents of the words. Gestures for the abstract words were mostly modeled on ASL signs

corresponding to those words. *Ability* was enacted by pushing the fists forward to show one's ability. The motion is based on the ASL sign corresponding to English *able* and *can*. *Agreement* was enacted by pointing at the forehead with an index finger and then putting both index fingers side by side. *Amount* was enacted by outlining a pile of something. *Anger* was enacted by moving bent fingers upward to the face as if strong emotion were welling up. *Cleanness* was enacted by a hand wiping something off the other hand. *Co-operation* was enacted by moving joined hands in a circle. *Development* was enacted by moving a hand upward as if something were appearing and growing. *Discovery* was enacted by a motion of picking up something. *Dream* was enacted by an index finger touching the forehead and then pulling away in a wiggling way. The wiggling finger stands for thought springing out. *Effort* was enacted by a motion of pushing something forward as if removing an obstacle. *Expression* was enacted by a motion of placing both hands on the chest and then opening them forward as if opening one's heart. The gesture was invented by the author. *Grief* was enacted by a motion of palms facing the eyes and then moving down as if the eyes were drooping. *Hatred* was enacted by a motion of pushing something aside as if removing something unwanted. *Health* was enacted by a motion of placing hands against the chest and then pulling them forward as if strength were springing out from the body. *Justice* was enacted by a motion of hands with the palm upward moving up and down as if the two dishes of a scale were moving up and down. *Pleasure* was enacted by hands circling up alternately in front of the chest as if happy feelings were welling up. *Power* was enacted by drawing an arc over the muscle of the arm as if representing a strong muscle. *Pressure* was enacted by a motion of pressing down the fist of the other hand. *Promotion* was enacted by hands moving upward stage by stage. *Thought* was enacted by pointing at the forehead.

Some gestures are iconic, and some are metaphorical. Gestures for *amount*, *cleanness*,

discovery, *grief*, *hatred*, and *power* are very iconic, and their forms represent their meanings. The gesture for *hatred*, for example, is a natural gesture since people push away something disgusting. Gestures for *agreement* and *promotion* are metaphorical. The gesture for agreement consists of two parts. One part is pointing at the forehead. This is the HEAD IS MIND metaphor regarding the head as a container of thought. Gestures for *dream* and *thought* are also based on the same metaphor. The other is placing index fingers side by side. Generally, things in similar shapes are grouped together. The index fingers placed together represent the concept of sameness. The repeated upward movements for *promotion* represent levels of promotion. The gesture is based on the idea of the HIGH STATUS IS UP metaphor (Lakoff & Johnson, 1980) regarding the society as a vertical organization. The metaphors underlie conventional thinking and language, and they are usually not noticed. All the gestures are more or less iconic and metaphorical, and iconicity and metaphorical property are a matter of degree. *Ability* was enacted by a motion of pushing fists forward. It is conventional knowledge that fists represent will power and hands represent some skills because people use their hands to do things.

Each gesture visualized some components of each concept. Some gestures visualized abstract concepts in terms of metaphor and metonymy. In this sense, *the concrete stands for the abstract*, and *the part stands for the whole*. What should be noted is that the cues represented by the gestures are related to bodily experiences in daily life, whether the gestures are iconic or metaphorical. As an image of pushing an obstacle forward represents *effort*, abstract concepts are represented by small fragments of bodily experiences. Common bodily experience, therefore, is considered to be a reason for the better recall of the abstract nouns.

Most gestures for abstract nouns were modeled on ASL signs. Some people, then, may say that as the transparency of the concept depended on the ASL signs, the signs

played an effective role in recalling the words, not the gestures. As previously mentioned in **2.6.4**, many ASL signs bear a striking resemblance to the things they represent. The resemblance or iconicity of signs suggests that the concepts of linguistic items (signs or words) are motivated by human experiences. *Effort,* for example, etymologically means *out + strong* (*Oxford Concise Dictionary of English Etymology*, 1986). It is probable then that the ASL sign and the English word *effort* originated form the same experience. The gesture for *effort* was not a copy of the ASL sign, but based on the concept visualized by the sign.

4.8 Conclusion

This study was conducted to test the effectiveness of gestures for recall of both concrete and abstract nouns. Gesture images employed for abstract nouns were based on the ASL signs corresponding to those words. The findings of this study can be summarized as follows: first, concrete nouns were remembered better than abstract nouns, as predicted. Second, gesture-image presentation was effective for the recall of both English concrete nouns and abstract nouns. Third, gestures based on the ASL signs were effective cues for the recall of abstract nouns.

As an effective cue for word recall, gesture is associated with a target word by two modes of thinking: the part for the whole; and the concrete for the abstract. Gestures as a recall cue visualize a component of the concept of a target word. A specific image of a strong muscle represents the concept of *power*. Gestures for abstract nouns visualize a component of an abstract concept in terms of metaphor. The metaphor is not a figure of speech, but a systematic way of thinking. The gesture of repeated upward movements for *promotion,* for example, is based on the conceptual metaphor HIGH STATUS IS UP. The metaphorical idea is a part of common knowledge. In other words, the metaphorical

gestures for the abstract words are supposed to be based on daily experiences common to the Japanese subjects, as the gestures for the concrete words.

Studies 1 and 2 showed that gesture images work well as recall cues for abstract nouns as well as concrete nouns. The gestures provided for abstract nouns are visual and meaningful. The dual coding theory assumes that verbal information and nonverbal information were coded in different systems and additive in effect. In the following chapter (Study 3), gesture performance is employed as a recall cue. In addition to verbal and visual information, bodily information is expected to contribute to better recall.

Note

1 Frick-Horbury (2002) examined the effect of hand gestures as self-generated cues for recall of both concrete and abstract words. Subjects were first asked to describe target words, and their description was videotaped for producing gesture cues. In the learning task, the experimental group was cued with their own gestures, and other groups were cued with someone's gestures or not cued at all. The results found that the experimental group excelled the other two groups in recalling both concrete and abstract words. The results also found that meaningful gestures (iconic and metaphoric) facilitated the recall of abstract words. Self-generated gestures were effective for the recall of abstract words as well as concrete words because Frick-Horbury assumed that "they are subjective and distinctive cues for personal memory and do not necessarily rely on visual imagery as the only aspect of cueing" (p. 4). The study indicates that the meaningfulness is an important component for better recall of abstract words. Japanese learners of English, however, cannot be expected to produce their own gesture cues especially for abstract English words. What they need will be systematic gesture models which are both imaginal and meaningful.

5. Study 3

Study 3 employs gesture performance as a cue for word recall. Recall targets are both action verbs and state verbs. Previous studies on memory of action assume that action memory is better after listening to (or reading) action phrases and performing the denoted actions than just listening to (or reading) action phrases. However, explanation for the enactment effect is controversial (**2.5.1**). Some insist that the effect comes from motor encoding, some assume that deep involvement in action performance facilitates the integration of action information to episodic memory. Yet others assume that multi-modal processing is responsible for enactment effect.

Another controversial issue about enactment effect is whether SPTs (subject-performed tasks) are more effective than EPTs (experimenter-performed tasks). Although Engelkamp (2001) acknowledged that SPTs are not always better, depending on experiment design, and material, he insisted that "SPT learning cannot be equated with visual-imaginal encoding" (p. 62). The superiority of SPTs in recall may be true of Japanese learners, too. In the Japanese classroom environment, students are usually required to sit still and focus on what their teacher is saying. They sometimes watch educational programs, but self-performance is very rare in the class. Taking these conditions into account, in the present study which involved Japanese subjects, recall in SPT learning was expected to be better than recall in EPT learning.

5.1 Purpose

This study has two purposes. One is to investigate the difference between recall in SPT learning and recall in EPT learning in a Japanese classroom, compared to VT (verbal tasks) learning. The other is to compare recall of action verbs with that of state verbs in the three learning conditions.

5.2 Subjects

A total of 78 university students belonging to the faculty of engineering at a national university in the Hokuriku district participated in the experiments. All of them were learning English as a second language. All the subjects began to study English at junior high school. 24 subjects (15 males, 9 females) had a VT experiment, 29 (17 males, 12 females) had an EPT experiment, and 25 subjects (14 males and 11 females) had a SPT experiment. Each group was chosen from a different English class. The level of students' English in three classes was considered to be similar by a teacher in charge.

5.3 Design

A *t*-test analysis was used for this study to investigate recall difference in the following three conditions: either to watch a native speaker pronouncing words subtitled (VT learning), or to watch a native speaker pronouncing and enacting words subtitled (EPT learning), or to watch a native speaker pronouncing and enacting words subtitled and copy the enactment (SPT learning).[1] Recall difference between action verbs and state verbs in the three conditions were also investigated.

5.4 Material

Enactment effect in action memory experiments is usually measured by recall of action phrases including typical action verbs, but this study focused on word recall, using action verbs and state verbs as recall materials. A total of 14 target words (Table 5-1), seven action verbs and seven state verbs, were selected based on Suzuki and Yasui (1994)'s verb classification (p. 60). They classified verbs into four categories in terms of two semantic features: *active* or *non-active*, and *self-controllable* or *not* (Table 5-2). Verbs in a-zone (active and self-controllable) are considered to be typical action verbs. Verbs in d-zone are

considered to be typical state verbs (non-active and not controllable). The seven action verbs were selected from the a-zone, while the seven state verbs were selected from the d-zone. Two video tapes were produced. One for the VT group was a tape of a native speaker (aged 23, female, British) pronouncing the 14 target words. The other for the EPT group and the SPT group was a tape of the same native speaker pronouncing and enacting the actions and states denoted by the 14 target words. Each scene on both tapes was synchronized with the subtitle of each target word at the lower part of the frame. Each scene was repeated twice in succession. All the processes were edited on a computer.

Table 5-1. A list of action verbs and state verbs

action verbs	catch, cut, drive, paint, pull, push, throw
state verbs	believe, fear, fit, hear, own, understand, want

Table 5-2. Verb classification

	active	non-active
s-c	a. slice, listen to, ask, and so on (unmarked)	b. remain, stand, stay, huddle, keep, squat, sit and so on (marked)
nsc	c. fall, encounter, and so on (marked)	d. know, hear, think that, believe that and so on (unmarked)

s-c: self-controllable, nsc: not self-controllable
* Suzuki and Yasui (1994: 58)'s table was revised as Table 5-2

5.5 Procedure

Three groups were informed that they would be shown an educational video of a native speaker pronouncing 14 English words. The VT group was requested to pronounce each word aloud during its second appearance. The EPT group was also requested to

pronounce each word aloud during its second appearance. On the other hand, the SPT group was requested to pronounce each word aloud and copy the native speaker's gesture at its second appearance. Prior to the task learning, two sample scenes, with gesture images in the case of the EPT and the SPT group, were shown to each group in order to become familiar with their task. Five minutes after the task, the subjects were unexpectedly given a five-minute recall test. They were required to recall as many words as possible in any order.

5.6 Results

Table 5-3. Means and SDs for recall results in three learning conditions

stimuli	M	SD
VT learning	7.06	2.16
EPT learning	9.93	1.75
SPT learning	11.16	1.52

Table 5-4. Analysis of *t* test (A-B) for recall difference

A/B	VT learning	EPT learning	SPT learning
VT learning	- - - -	-5.28 0.00	-7.70 0.00
EPT learning	- - - -	- - - -	-2.76 0.01
SPT leaning	- - - -	- - - -	- - - -

* upper is t-value, lower is p-value

Table 5-3 shows the means and standard deviations for recall results in the tree learning conditions. Table 5-4 shows that there was a statistical difference of mean scores between EPT learning and SPT learning, $t(42) = -2.76$, $p < 0.01$ (the significance level is adjusted by the Bonferroni method). A significant difference was confirmed between VT learning and EPT learning, and between VT learning and SPT learning.

Table 5-5. Means and SDs for recall results of action verbs and state verbs

	VT learning		EPT learning		SPT learning	
	action	state	action	state	action	state
M	3.42	3.63	4.79	5.17	5.52	5.56
SD	1.21	1.24	1.01	1.10	1.00	0.91

Table 5-5 shows means and standard deviations for recall results of action verbs and state verbs in the three learning conditions. There was no significant difference in VT learning, EPT learning, and SPT learning, $t(45) = -0.58$, $p > 0.05$, $t(56) = -0.58$, $p > 1.36$, and $t(48) = -0.45$, $p > 0.05$, respectively.[2]

5.7 Discussion

The present study focused on the recall difference between EPT learning and SPT learning. The recall in SPT learning was better than that in EST learning. As previously mentioned, there are two possible explanations for the better recall. One is the contribution of motor encoding because the SPT group actually performed actions, while the EPT group just watched the experimenter's performance. As tool-use gestures activated the premotor area, the supplementary motor area, and so forth in the brain (Choi, Na, Kang, Lee, Lee, & Na, 2001), the gesture performance might have been encoded into the brain. The other possibility for the better recall in SPT learning is the degree of involvement in the task. Kormi-Nouri and Nilsson (2001) referred to the possibility of better self-involvement in SPT learning as a possible reason for enactment effect compared to in VT learning. The similar reason might be applicable to the recall difference between SPT learning and EPT learning in this study because SPT learning demands a stronger

awareness and intention to perform in Japanese classroom environment. As a matter of fact, the attitude of the SPT group seemed to be quite relaxed during their participation in the task. The motivational aspect cannot be underestimated.

The present study also showed that there was no significant recall difference between action verbs and state verbs in all learning conditions. Semantic features of activeness and self-controllability did not affect the recall of verbs. The results may seem a bit surprising, because the retrieval of action verbs like *throw* and action-related nouns like *key* induced greater activation of the motor cortex than that of non-action verbs like *belong* and non-action-related nouns like *cloud* (Oliveri, Finocchiaro, Shapiro, Gangitano, Caramazza,& Pascual-Leone, 2004). The author suspects that enhanced activation in motor areas, though degree of activation cannot be ignored, may not directly relate to enhanced recall of action-related words.


Figure 5-1. ***throw***


Figure 5-2. ***drive***


Figure 5-3. ***believe***


Figure 5-4. ***fear***

As for gestures of the target words, the action verbs were enacted by pictorial gestures which represent the denoted actions. On the other hand, gestures for the state verbs were performed as follows: *believe* was enacted by a natural praying gesture. *Fear* was enacted by a natural gesture of protecting the body with both hands. *Fit* was enacted by a motion of both hands with fingers spread apart coming together. *Hear* was enacted by bringing an index finger to the ear. *Understand* was enacted by pointing at the forehead. *Want* was enacted by a motion of grasping something and pulling it toward the body. The gestures for

fear, *fit*, *hear*, and *want* were modeled on the ASL signs. State verbs as well as action verbs were recalled better with the help of gestures. Whether based on the signs or not, all the gestures for the state verbs were quite natural and iconic.

5.8 Conclusion

This study was conducted to investigate recall difference between SPT learning and EPT learning in a Japanese classroom, compared to VT learning. Action verbs and state verbs were used as recall targets. Recall of action verbs was compared with that of state verbs in the three learning conditions. The findings of this study can be summarized as follows: 1) word recall in SPT learning (motor encoding) was better than in EPT learning, as predicted, 2) gesture-image presentation was effective for the recall of both action verb and state verbs. There was no significant difference between the recall of both verb types in the three learning conditions, 3) gesture was an effective cue for the recall of state verbs.

There were two possible factors for the better recall in SPT in this study: addition of motor encoding, and psychological influence. Taking the Japanese educational environment into consideration, motivational factors cannot be underestimated. Since physical performance is very rare in Japanese classrooms, gesturing demanded much stronger self-involvement.

Notes

1 In Studies 1 and 2, the subjects in VT learning were provided with verbal information only. The study assumed that recall of verbal items was better in EPT learning than in VT learning. In this study, the subjects in VT learning watched a video of a person pronouncing words entitled, while the subjects in both EPT and SPT learning watched a video of a person pronouncing and enacting words entitled.

2 After the recall test, a translation test was conducted. As the translation of one word was scored as one point, the total amount of marks for the test was 14 points. The average score for VT group was 13.79, the average score for EPT group was 13.93, and the average score for SPT group was 13.96.

6. Study 4

In the previous three studies, lexical items were used as recall materials. The targets were concrete nouns in Study 1, concrete and abstract nouns in Study 2, and action and state verbs in Study 3.This study, then, employed sentence constructions, tested the recall of sentences using different types of verbs as the predicate, and it also tested the effect of gesture on the recall of those sentences.

In Experiment 1, the action and state verbs used for Study 3 are employed as the predicate verb in each sentence. As the action and state verbs are recalled to a similar degree, sentences with an action verb as the predicate and sentences with a state verb as the predicate are also expected to be recalled similarly. As gesture is effective for the recall of both action and state verbs, it is expected to be effective for the recall of both types of sentences.

In Experiment 2, high frequency verbs and low frequency verbs are used for the predicate. High frequency verbs are considered to be familiar and concrete, while low frequency verbs are considered to be rather difficult and abstract. The recall of sentences with a familiar word as the predicate, therefore, would be better than that of sentences with a low frequency verb as the predicate. Gesture is expected to be effective for both types of sentences because the recall of both concrete and abstract nouns were facilitated by gesture in Study 2.

6.1 Experiment 1

6.1.1 Purpose

Experiment 1 has two purposes. One is to investigate recall difference between action verb sentences and state verb sentences. The other is to investigate the effect of gesture on the recall of both types of sentences.

6.1.2 Subjects

A total of 56 first-year students, 42 males and 14 females, served as subjects. They belonged to the faculty of science at a national university in the Hokuriku district. All of them began to study English at junior high school. The subjects were chosen from two different English classes. The level of students' English in both classes was considered to be similar by a teacher in charge.

6.1.3 Design

A 2×2 factorial design was used for the experiment. The first factor was either performing gestures or not, which was the between-subjects variable. The second factor was either recalling action verb sentences or state verb sentences, which was the within-subjects variable.

6.1.4 Material

A total of 14 sentences with subject-verb-object construction were produced and checked by a native speaker of English (Appendix E). The action and state verbs used for Study 3 were employed as the predicate of each sentence. Two video tapes were then produced. One for the control group was a tape of a native speaker (aged 46, male, British) pronouncing the 14 target sentences. The other for the experimental group was a tape of the same native speaker pronouncing the target sentences and enacting the actions and states denoted by the 14 verbs. Each scene on both tapes was five seconds long with the subtitle for each sentence at the lower part of the frame. Each scene was repeated twice in succession. Action verb sentences and state verb sentences alternately appeared. All the processes were edited on a computer.

6.1.5 Procedure

Both the control group and the experimental group were told in advance that they would be shown an educational video of the native speaker pronouncing 14 English sentences (with gestures in the case of the experimental group). They were also told that they would be given a recall test of the sentences. The control group was required to read each sentence during its second appearance, and then write it down on a sheet of paper. The experimental group was required to read each sentence and copy the speaker's action during its second appearance, and then write it down on a sheet of paper. Prior to the task learning, two samples appeared. Soon after each experiment, both groups were asked to recall as many sentences as possible, and write them down on a new sheet of paper within five minutes. Predicate verbs were printed as cues on the paper.

6.1.6 Results

The recall of the subject and the object of each sentence was scored as one point, respectively. The total amount of marks for each task (seven action sentences or seven state sentences) were then 14 points. A simple spelling error was judged as a correct recall.

Table 6-1. Means and standard deviations of the results of recall test

stimuli	M	SD
action verb (AV)	9.46	2.59
state verb (SV)	10.21	2.48
AV + gesture	12.54	1.43
SV + gesture	12.86	1.35

Table 6-1 shows all the mean numbers and standard deviations for the recalled English words under all the conditions of the recall test. Table 6-2 shows the analysis of variance for the results of the recall test. It revealed that there was a large and significant

main effect for gesture performance, $F(1, 54) = 36.26$, $p < 0.05$.The main effect for word type was not significant, $F(1, 54) = 3.75$, $p > 0.05$. The main effect for the interaction of gesture performance and word type was not significant, $F(1, 54) = 0.60$, $p > 0.05$.

Table 6-2. Analysis of variance for the results of Experiment 1

	SS	*df*	MS	F	*p*
gesture performance (A)	228.57	1	228.57	36.26	0.00
word type (B)	8.04	1	8.04	3.73	0.06
intraction (A × B)	1.29	1	1.29	0.06	0.44
error S(A)	340.39	54	6.3		
error S(A) × B	115.68	54	2.14		

After the experiment, students in two other classes, who were also freshmen in the same faculty, had a translation test of the target sentences in the same experimental condition in order to suggest that the control group and the experimental group had a similar level of word knowledge. The correct translation of the subject, verb, and object of each sentence were scored as one point, respectively. The total amount of marks for the 7 sentences of each predicate type was 21 points. There was no significant difference between the results.[1] The results suggest that the subjects of Experiment 1 seemed to have a similar level of word knowledge, and the difference in experimental condition did not affect their comprehension.

6.2 Experiment 2

6.2.1 Purpose

In the previous studies on action memory, simple and familiar action verbs were employed as target words. Experiment 1, then, tested action and state verbs. In Experiment 2, high frequency verbs and low frequency verbs are to be tested. Experiment 2 has two

purposes. One is to investigate recall difference between sentences with high frequency verbs as their predicate and sentences with low frequency verbs as their predicate. The other is to test the role of gesture on the recall of both types of sentences.

6.2.2 Subjects

A total of 44 first-year students, 34 males and 10 females, served as subjects. They belonged to the faculty of science at a national university in the Hokuriku district. All of them began to study English at junior high school. The subjects were chosen from two different English classes. The level of students' English in three classes was considered to be similar by a teacher in charge.

6.2.3 Design

A 2×2 factorial design was used for the experiment. The first factor was either performing gestures or not, which was the between-subjects variable. The second factor was either recalling sentences with a high frequency verb as the predicate verb or recalling sentences with a low frequency verb as the predicate verb, which was the within-subjects variable.

6.2.4 Material and procedure

Referring to Leech, Rayson, and Wilson (2001), which was based on the British National Corpus, a sample of some 100 million words, seven high frequency words and seven low frequency words were chosen for the experiment as Table 14. The number after each word showed the frequency score of occurrences per million words. The high frequency words were selected from words with more than 250 occurrences per million words, while the low frequency words were selected form words with less than 25

occurrences per million words. The high frequency verbs had short syllables and represent simple actions and feelings, while the low frequency verbs had relatively long syllables and represent complex behaviors and mental attitudes. Interestingly, the high frequency word group seems to correspond to vocabulary for junior high school students, and the low frequency word group seems to correspond to vocabulary for high school students.[2]

Table 6-3. High frequency verbs and low frequency verbs

high frequency verbs	call (535), follow (460), like (424), read (284), take (1797), tell (775), want (945)
low frequency verbs	admire (22), assemble (17), classify (16), criticize (17), exhibit (18), guard (14), tolerate (12)

A total of 14 sentences were produced for the experiment and were checked by a native speaker of English (Appendix F). The high and low frequency verbs chosen were used as the predicate of each sentence. 13 sentences had subject-verb-object construction, but one sentence had subject-verb-adverbial phrase construction.[3] The process of making videotapes for experiment 2 was the same as for experiment 1. A native speaker of English (aged 47, male, American) pronounced target sentences, and enacted the predicates. The procedure of experiment 2 was also the same as that of experiment 1.

6.2.5 Results

The rating system for sentence retention was the same as for Experiment 1. The retention of each subject and object (the object of the preposition in No. 6 target sentence) was scored as one point, respectively. Table 6-4 shows means and standard deviations for the results of the recall tests. Table 6-5 shows the analysis of variance for the results of

experiment 2. It revealed that there was a significant main effect for gesture performance, F (1, 42) = 13.58, $p < 0.05$. The main effect for word type was also significant, F (1, 42) = 66.16, $p < 0.05$. The main effect for the interaction of gesture performance and word type was not significant, F (1, 42) = 0.76, $p > 0.05$.

Table 6-4. Means and standard deviations for the results of the recall tests

stimuli	M	SD
high frquency verb (HV)	9.64	1.87
low frequency verb (LV)	5.41	2.81
HV + gesture	11.32	2.19
LV + gesture	7.91	2.78

Table 6-5. Analysis of variance for Experiment 2

	SS	*df*	MS	F	*p*
gesture performance (A)	96.18	1	96.18	13.58	0.00
word type (B)	320.73	1	320.73	66.16	0.00
intraction (A × B)	3.68	1	3.68	0.76	0.39
error S(A)	297.41	42	7.08		
error S(A) ×B	203.59	42	4.85		

After the experiment, students in two other classes, who were also freshmen in the same faculty, had a translation test of the target sentences in the same experimental condition. There was no significant difference between the results.[4] The results suggest that the subjects of experiment 2 seemed to have a similar level of word knowledge, and the difference in experimental condition did not affect their comprehension.

6.3 Discussion

The purpose of Study 4 was to test the role of gestures in recalling sentences by performing denoted actions and states represented by predicate verbs. In the paradigm of

action memory, predicate verbs employed in the subject-verb-object construction were typical action verbs such as *pinch*, *kick*, and *rip* (*e.g.*, Saltz & Donnenwerth-Nolan, 1985). In this study, Experiment 1 employed action and state verbs, and Experiment 2 employed high and low frequency verbs.

The results of Experiment 1 were consistent with the prediction that action verb sentences and state verb sentences would be recalled at a similar rate. There was no significant main effect for predicate type. The results of Experiment 1 were also consistent with the prediction that the enactment of both action and state predicates would be effective for sentence recall. The main effect for gesture performance was significant. The action verbs and state verbs employed were different in their semantic features of activeness and self-controllability, but the difference did not affect sentence recall. The results suggest that gesturing a predicate, either an action verb or a state verb, facilitates sentence recall at a similar rate.

The results of Experiment 2 corresponded to the prediction that the recall of high frequency sentences would be better than that of low frequency sentences. There was a significant main effect for predicate type. The results of Experiment 2 also corresponded to the prediction that the enactment of both high and low frequency predicates would be effective for sentence recall. The main effect for gesture performance was significant. If high frequency words are regarded as familiar words in terms of syntax and semantics, their usage and meanings are better known than those of low frequency words. According to Paivio (1965), auditory familiarity of words was correlated with imagery and meaningfulness (verbal association), and concrete nouns were higher in familiarity than abstract nouns. If word frequency is defined in terms of concreteness, high frequency words are familiar words with high imagery and meaningfulness, and low frequency words are less familiar words with relatively low imagery and meaningfulness. The results of

Experiment 2, then, suggest that imagery and meaningfulness of the predicate facilitate sentence recall.

The predicate verb governs a sentence syntactically and semantically. A transitive verb has two core arguments: subject and object. The subject-verb-object order is syntactically determined, and the content of the subject and object is semantically restricted. *Tolerate,* the lowest frequency verb for Experiment 2, demands the grammatical subject and object. A person who has the ability to bear something unnecessary or undesirable generally goes to the position of the subject, and its object is typically an unnecessary or undesirable thing. In other words, the predicate verb, as the core of a sentence, contains relational information influencing the structure of a sentence.

Begg and Paivio (1969) referred to an integrative role of imagery in concrete sentences. They stated that a concrete sentence such as *The fat boy kicked a girl*, "can be imaginally represented as an action picture in which the meaning of the entire sentence is summarized as one organized unit, or complex image" (p. 821). From the viewpoint of the dual coding theory, the sentence was coded by both the verbal and the imagery systems. On the other hand, abstract sentences were assumed to remain linked more with the verbal system. Begg and Paivio, then, hypothesized that "the most effectively coded, stored, and retrieved aspects of a concrete sentence will be those related to the sentence as a whole unit" (p. 821). As the predicate, *kick* demands two grammatical elements: the subject and the object. Semantically, *kick* demands actor-action-object relation. If the sentence *The fat boy kicked a girl* is readily comprehended as one organized unit or complex image, this may be because *the fat boy* smoothly fits the role of the actor and *a girl* fits the role of a kicked person. As a result, the high imagery may have facilitated organizing the sentence into a unit.

Gesture performance was effective in recalling sentences with a low frequency

predicate. *Classify* in *The librarian classified the books* was enacted by both hands with the palms facing each other moving from side to side in a series of small arcs. The gesture was invented, based on the two ASL signs corresponding to English *group* and *arrange*. The motion visualizes the idea of thing being grouped and placed in order. The gesture seems clear because people classify things by dividing and grouping them. *Criticize* in *The journalist criticized the music* was enacted by outlining an X with an index finger. The gesture, based on the ASL sign corresponding to *criticize*, visualizes the idea of crossing out something unnecessary or wrong. Form represents meaning. *Assemble* in *The students assembled in the hall* was enacted by wiggling both hands' fingers and bringing them toward one another. The gesture, based on the ASL sign corresponding to *assemble*, visualizes the idea of many people gathering in a place (Sternberbg, 1994). *Exhibit* in *The artist exhibited the picture* was enacted by moving both hands forward with an index finger pointing at the palm of the other. The gesture, based on the ASL sign corresponding to *exhibit*, visualizes the idea of focusing the attention on something, and bringing it forward (Sternberbg, 1994). The gesture is commonly seen when people show or demonstrate something. *Guard* in *The men guarded the star* was enacted by holding both arms with fists in front and then pushing them forward. The gesture, based on the ASL sign corresponding to *guard*, visualizes the concept of defense. *Admire* in *People admired the President* was enacted by clapping hands. The gesture was invented because people express their admiration with applause on many occasions. *Tolerate* in *The patient tolerated the pain* was performed by wiggling fists in front. The gesture visualizes the concept of endurance.

All the gestures seemed specific and relational because they reflected people's daily experiences, and were directly associated with feelings and objects in the scenes. The gesture performance, by providing the predicate with a concrete image and meaningful

information, might have facilitated organizing the sentence into a unit or complex image.


Figure 6-1. *call*


Figure 6-2. *want*


Figure 6-3. *admire*


Figure 6-4. *assemble*

6.4 Conclusion

This study was conducted to test the role of gestures in recalling sentences by performing denoted actions and states represented by predicate verbs. Four different types of verbs, action verbs, state verbs, high frequency verbs, and low frequency verbs, were employed as the sentence predicate. The findings of this study can be summarized as follows: 1) enactment was effective for the recall of sentences with any type of predicate; 2) there was no significant recall difference between sentences with an action verb as the predicate and sentences with a state verb as the predicate; 3) sentences with high frequency verbs were remembered better than those with a low frequency verb as the predicate.

In study 3, there was no significant difference in recall results of the action verbs and state verbs in VT learning, EPT learning, and SPT learning. The results were parallel to the recall results of sentences with the same action verbs and state verbs as the predicate in Experiment 1. Enhancement in word recall by enacting the denoted actions and states was also parallel to that in sentence recall by enacting the actions and states denoted by the predicates.

The predicate verb governs a sentence syntactically and semantically. The predicate verb, as the core of a sentence, contains relational information influencing the structure of a sentence. It was assumed then that the predicate with high imagery facilitates organizing a

sentence into a unit or complex image. The results of Experiment 2 indicated s a significant deference in recall between sentences with high frequency verbs as the predicate and sentences with low frequency verbs as the predicate. It was assumed that familiarity of words was correlated with imagery, and meaningfulness and familiarity was higher with concrete words than abstract words (Paivio, 1965). High frequency words, then, are regarded as familiar words with high imagery and meaningfulness, and low frequency words are less familiar words with relatively low imagery and meaningfulness. The difference in imagery and meaningfulness (imagery and meaningfulness are correlated) between the high and low frequency verbs is considered to be the cause of the difference in recall between the sentences with high frequency verbs as the predicate and the sentences with low frequency verbs as the predicate.

Enacting the actions and states denoted by high and low frequency verbs enhanced the recall of the sentences with both types of verbs as the predicate. The results suggested the possibility of enhanced recall of both high and low frequency verbs by gesture. The analysis of gestures for the low frequency predicates suggested that all the gestures for the seven predicates were based on common bodily experiences. It is assumed that predicate verbs, either high frequency verbs or low frequency verbs, can facilitate the recall of sentences if they are enacted by gestures based on people's daily experiences.

Notes

1 The average score of action verb sentences for the first class (19 students) was 20.79, and the average score for the second class (22 students) was 20.86; the average score of state verb sentences for the first class was 20.84, and the average score for the second class was 20.91.

2 JACET 8000, a new word list for Japanese students, classified English words from level one (1-1000) to level eight (7000-8000). Words in level one and level two (1000-2000) are considered to be words for junior high school students and the first year students of senior high school, and words in level three are for the second and third year students of senior high school. The words classified as the high frequency verbs in Study 4 all belong to the level one, and the words classified as the low frequency verbs belong to level three (2000-3000) and four (3000-4000) except *guard*.

3 The target sentence, *The students assembled in the hall*, has the subject-verb-adverbial (prepositional) phrase structure, not the subject-verb-object structure. The definition of *assemble* is that a group of people come together in a place. As for *assemble*, a phrase expressing a location is not always necessary in terms of syntax, but the concept of location is a part of the concept of *assemble*. In this sense, the object of the preposition, *hall*, was considered to be a significant component of the sentence.

4 The correct translation of the subject, verb, and object of each sentence was scored as one point, respectively. The average score of high frequency sentences (HFSs) for the first class (17 students) was 19.94, and the average score of HFSs for the second class (20 students) was 20.05. The average score of low frequency sentences (LFSs) for the first class was 15.59, and the average score of LFSs for the second class was 16.10.

7. Study 5

Study 4 extended a recall target from word to sentence, and tested the effect of gesture by enacting the predicate verb of a sentence. Study 5, in turn, employs a paragraph as a recall target. The paragraph is analyzed into small parts called *idea units* (Carrell, 1985).[1] The idea unit is a syntactic unit which expresses a thought or idea. Dividing a paragraph into idea units, Study 5 tests the effect of gesture in recalling a paragraph. The enactment of one word per each idea unit is expected to enhance the recall of the whole paragraph.

7.1 Purpose

The purpose of the study is to test the role of gesture in recalling an English paragraph.

7.2 Subjects

A total of 60 first-year college students served as subjects. They belonged to the faculty of engineering at a national university in the Hokuriku district. They began to study English at junior high school. 29 subjects, of which 22 were males and seven were female, constituted the control group, and 31, of which 25 were males and six were females, constituted the experimental group. The subjects were chosen from two different English classes. The level of students' English in three classes was considered to be similar by a teacher in charge.

7.3 Design

A *t*-test analysis was used for this study to investigate the difference between the recall of a paragraph in VT learning and that of a paragraph in SPT learning

7.4 Material

An English paragraph with four sentences about business innovation (Appendix G) was constructed, and checked by a native speaker of English (aged 23, female, British).[2] The four sentences were analyzed into eight idea units. One lexical item was selected from each idea unit for enactment: seven verbs and one adverb. Two videotapes were then produced. One for the control group was a tape of the native speaker pronouncing the paragraph. The other for the experimental group was a tape of the same native speaker pronouncing the paragraph and enacting the eight words. The paragraph was segmented into eight scenes of eight idea units, and each scene was subtitled. Before the experiments, the eight enacted words appeared on the screen with their Japanese translations, and the native speaker pronounced the words (with gestures for the experimental group).

7.5 Procedure

All the subjects were informed that they would be shown an educational video of a native speaker of English pronouncing a paragraph (with gestures in the case of the experimental group). The control group was required to read the paragraph aloud at the second appearance of the paragraph, while the experimental group was required to read the paragraph aloud and copy the person's gestures. Prior to the task learning, the 8 enacted words appeared with Japanese translations (with gestures in the case of the experimental group). Five minutes after the task, both groups were given an unexpected recall test. The test was a fill-in-blank format. Two words (including each enacted word) per idea unit had been deleted. The subjects were required to fill in the 16 blanks in five minutes.

7.6 Results

The author measured recall by the number of the words supplied.

Table 7-1. Means and SDs for the results of the recall test

stimuli	M	SD
pronouncing (P)	10.59	3.58
P + gesturing	11.74	3.18

Table 7-2. Means and SDs for the results of the recalled enacted words

stimuli	M	SD
pronouncing (P)	5.52	1.72
P + gesturing	6.16	1.70

Table 7-1 shows the means and standard deviations for the results of the recall test. There was no statistical difference in mean scores between VT learning and SPT learning, t (58) = - 1.32, $p > 0.05$. Table 7-2 shows the means and standard deviations for the results of the recalled enacted words. There was no statistical difference in mean scores between VT learning and SPT learning, t (58) = - 1.49, $p > 0.05$.

7.7 Discussion

The previous studies indicated that enactment was effective in recalling words and sentences. The purpose of study 5, then, was to test the role of gestures in recalling a paragraph. The target paragraph was divided into eight idea units, and the recall of the paragraph was examined by counting the recall of two words per idea unit, totally sixteen words. The results showed that gesture performance was not significantly effective. Even recall of the eight words enacted was not significantly better than that of the same words not enacted. What were the possible reasons for the ineffectiveness of enactment, even for the recall of enacted words?

There were four possible reasons for the ineffectiveness of gesture in recalling. The first is three technical problems in the experiment. Prior to the task learning, the eight

words, which were enacted for the experimental group, were shown to the subjects of both groups with their Japanese translations. The word presentation might have strongly influenced at least the recall of enacted words. Another possible technical problem was that the subjects were not firmly required to memorize the paragraph. Unless mnemonic cues should be intentionally created and used, they will not be effective. The rest possible technical reason was that the native speaker tended to pronounce the enacted words slowly then she spoke with gestures. The speed difference might have interfered with the subjects' smooth understanding.

The second possible reason was related to the quality of the gestures. All the gestures provided for the enacted words were modeled on the ASL signs corresponding to those words. *Decline* was performed by moving a hand downward in a series of short movements to represent the condition of something getting worse. Lakoff and Johnson (1980) referred to conceptual metaphor GOOD IS UP/BAD IS DOWN in English. "Things are looking *up* and things are at an all-time *low*" (emphasis in the original, p. 16). The downward motion seemed understandable because goodness-badness is represented in terms of verticality in Japan, too. *Realize* was enacted by pointing at the head with an index finger. It is widely accepted that the head is the location of intelligence. *Do* was enacted by pushing the fists forward to show one's ability. *Attract* was enacted by a motion of bringing something together to one point. *Develop* was enacted by moving one hand upward to represent something appearing and growing. *Later* was enacted by moving an open hand forward to represent something ahead in the future. The motion is based on the ASL sign corresponding to the English word *later*. This is an example of the conceptual metaphor FUTURE IS AHEAD. *Increase* was enacted by moving an open hand upward. The gesture is an example of MORE IS UP. *Anticipate* was enacted by index fingers pointing at the eyes, and then moving forward as if the eyes looking into the future. This is another example of

FUTURE IS AHEAD. Most gestures were metaphorically associated with word concepts. The gestures seemed understandable to the subjects because similar modes of thought were found in Japanese language.


Figure 7-1. *later*


Figure 7-2. *realize*


Figure 7-3. *anticipate*


Figure 7-4. *develop*

The third possible reason was related to the syntactically and conceptually complicated structure of a paragraph. The target paragraph, composed of four sentences was analyzed into eight idea units. The units consisted of various grammatical structures: one single clause, three main clauses, two subordinate clauses (one noun clause and one adverb clause), one to-infinitive phrase, and one prepositional phrase. The target words for recall were also composed of various grammatical components. The eight enacted words were seven verbs and one adverb. The rest eight recall words, seven nouns and one adjective, were four grammatical subjects, two grammatical objects, one item in a prepositional phrase, and one item in an adverb clause. The complicated syntactic structure of the paragraph might have disturbed the subjects' learning.

Generally, the comprehension of a paragraph requires word knowledge, syntactic knowledge, knowledge of text structure, and general knowledge of topic. For example, Carrell (1985) stated that "the more tightly organized patterns of *comparison*, *causation*, and *problem/solution* generally facilitate the recall of specific ideas from a text more than the more loosely organized *collection of descriptions* pattern" (emphasis in the original, p. 730). In the target paragraph, the down sales and the development of a new product should

be linked in terms of problem-solution, and the new product and the up sales should be linked in terms of cause-effect. All the events need to be linked to each other in a temporal relation. It might have been not appropriate to divide up such a grammatically and semantically complex body into idea units and enact one word per idea unit.

The forth possible reason was related to the limitation of mnemonic techniques. Mnemonic techniques in general emphasize a one-to-one relationship. In the present study, gesture mediates between the form and meaning of a word. Although gesture can convey both distinctive and relational information, it may not be able to cope, in a fundamental way, with a complex structure of information like a paragraph. Performing one gesture per idea unit might have conflicted with coherent comprehension of content and thus confused the subjects. Sadoski and Paivio (2001) stated that "the only necessary and sufficient condition for textual coherence appears to be that the sentences in a text conform to an image of a possible world in the experience or imagination of the reader" (p. 126). The eight gestures, then, might have provided the subjects with collections of segmented images, not a unified one.

7.8 Conclusion

In previous studies on memory, linguistic targets were mostly words and phrases. Study 5 then extended a recall target to a paragraph. The study was conducted to test the role of gestures in recalling the content of a paragraph. The paragraph was divided into eight idea units. One word per each idea unit was enacted, and two words (including each enacted word) per each idea unit were recall targets. The results showed that gesture performance was not significantly effective, even for the recall of the enacted words.

Several possible reasons for the ineffectiveness of gesture were considered. The first was three technical problems in the experiment. Prior to the task learning, the subjects were

presented with the eight enacted words and their Japanese translations. The presentation might have impressed the subjects strongly. The other possible technical problem was that the subjects were not strongly required to memorize the content of the target paragraph. Mnemonic cues should be intentionally employed. The rest technical problem was that the model native speaker tended to pronounce the enacted words slowly when she spoke with gestures. The irregular speed might have interfered with the subjects' smooth understanding.

The second possible reason for the ineffectiveness of gesture was the quality of the gestures. All the gestures provided for the enacted words were modeled on the ASL signs corresponding to those words. *Decline,* for example, was performed by moving a hand downward in a series of short movements to represent the condition of something getting worse. Lakoff and Johnson (1980) referred to conceptual metaphor GOOD IS UP/BAD IS DOWN in English. The downward motion seemed understandable because goodness-badness is represented in terms of verticality in Japanese language, too. Most gestures were metaphorically associated with word concepts. The understanding of the gestures depended on whether the subjects accepted the gestures according to their daily experiences.

The third possible reason was related to the syntactically and conceptually complicated structure of a paragraph. The target paragraph, composed of four sentences was divided into eight idea units. The sixteen target words for recall were composed of various grammatical components. The eight enacted words were seven verbs and one adverb. The rest eight recall words, seven nouns and one adjective, were four grammatical subjects, two grammatical objects, one item in a prepositional phrase, and one item in an adverb clause. The complicated syntactic structure might have disturbed the subjects' learning.

A paragraph is also conceptually complicated. The comprehension of a paragraph requires word knowledge, syntactic knowledge, knowledge of text structure, and general knowledge of topic. The target paragraph was a business story. The down sales and the development of a new product should be linked in terms of problem-solution, and the new product and the up sales should be linked in terms of cause-effect. All the events need to be linked in a temporal relation. It might have been not appropriate to divide up such a grammatically and semantically complex body into idea units and enact one word per idea unit.

The forth possible reason was related to the nature of mnemonic techniques. Mnemonic techniques generally emphasize a one-to-one relationship. They associate one target item with one cue. In the present study too, one gesture was provided for one word. Gesture may not be able to cope with a complex structure of information like a paragraph. Performing one gesture in one idea unit might have conflicted with coherent comprehension of content and have confused the subjects.

Note

1 Carrell (1985) stated, "each idea unit consisted of a single clause (main or subordinate, including adverbial and relative clauses). Each infinitival construction, gerundive, nominalized verb phrase, and conjunction was also identified as a separate idea unit. In addition, optional and /or heavy prepositional phrases were also designated as separate idea units" (p: 757).

2 The paragraph was constructed, based on a paragraph in Wattenmaker and Shoben (1987, p 142).

8 Conclusion and implications

This final chapter refers to a limitation common to the five studies and summarizes their findings. It also mentions some implications for further research, and finally proposes a model of teaching English with gesture.

8.1 Limitations of the studies

In the experiments, the English ability of the control groups and the experimental groups was not formally tested before conducting the experiments. Translation tests were conducted subsequently to the experiments in Studies 2 and 3, and students who were considered to be equivalent to the subjects had translation tests in Study 4 in order to test whether the control group and the experimental group had a similar level of word knowledge. As for Studies 1 and 5, no measure was taken except for the teachers' judgment. The equality of subjects' English ability was only based on experience, because simple nouns were target words in Study 1, and main words were shown with Japanese translations before the experiment in Study 5. This might have affected the results of the studies reported above.

Visual information for all studies was first digitally videotaped, and then was edited on a computer. Noise, poor lighting, and spatial constraints were involved in the filmmaking, and some parts were not technically well-edited. The model native speakers tended to speak a bit slowly when they spoke with gestures. Some gestures were poorly performed, and some gestures were not well-organized. Enacted English words (nouns and verbs) were not randomly selected from English lexicon. A word whose ASL equivalent was iconic or metaphorical was a good candidate for enactment.

With these limitations in mind, the findings of the five studies are briefly reviewed in the following section.

8.2 Summary of the findings

Study 1 (Chapter 3) examined Paivio's (1986) bilingual dual coding theory that two verbal systems and one nonverbal system are independent and additive in effect, by supplying Japanese EFL students with gesture images as nonverbal cues for word recall. The study confirmed that the bilingual processing (translation of Japanese words into English equivalents) was more effective for the recall of English words (concrete nouns) than the monolingual processing (copying English words). The study revealed that gesture images were effective in recalling English words in both conditions, and English words were similarly well remembered even in the monolingual processing if gesture images provided. The results suggest that gesture images are useful cues for word recall, and English concrete nouns can be learned without Japanese equivalents if gesture images are provided.

Concrete words which represent simple objects and events have sometimes been presented with pictures and gestures, but abstract words are only verbally explained and comprehended. How to visualize abstract concepts is a challenging issue. In Study 2 (Chapter 4), the effectiveness of gesture images for the recall of abstract nouns as well as concrete nouns were tested. The study revealed that gesture images were effective even for recalling abstract nouns. Abstract words were represented by gestures modeled on the ASL signs corresponding to each word. Some ASL signs iconically represent the referents, and some signs metaphorically represent abstract concepts. Their iconic and metaphorical motions were applied to the enactment of abstract words. Previous studies on gesture and language (both signed and spoken) show that gestures and linguistic items are mostly conceptualized through human bodily experiences (Chapter 2). Common bodily experiences, then, are assumed to provide Japanese EFL students with cues for comprehending and retrieving English abstract words.

Study 3 (Chapter 5), using action verbs and state verbs as recall targets, compared the recall in SPT (subject-performed task) condition with that in EPT (experimenter-performed task) condition. The action and state verbs were similarly recalled. Semantic features (activeness and self-controllability) did not affect their recall. Performing gesture (motor encoding) was more effective than watching somebody else's gesture (visual encoding) in recalling words. In addition to the motor component, psychological aspects involved in gesture performance were discussed as a possible reason for better recall. Gesturing demands much stronger self-involvement than expected because physical performance is very rare in Japanese classrooms. Students' emotions, however, are often influenced by class atmosphere. There is a possibility for worse recall. Some Japanese students who are not used to physical performance may try to avoid gesturing before their peer students. The effectiveness of gesture performance may not always be predictable in Japanese class rooms.

In Study 4 (Chapter 6), recall targets were extended from word level to sentence level. The predicate verb was enacted because the predicate of a sentence governs a sentence syntactically and semantically. In other words, the predicate has relational information influencing a whole sentence. A transitive verb, for example, has two core arguments. subject and object. The subject-verb-object order is syntactically determined, and the content of the subject and object is semantically restricted. It is, therefore, predicted that better comprehension and recall of the predicate verb should facilitate better recall of the whole sentence. The enactment of the predicate was effective in recalling sentences with four types of predicates: action verbs, state verbs, high frequency verbs, and low frequency verbs. Low frequency verbs were regarded as less familiar words with low imagery and concreteness. The concept of each low frequency verb was enacted and visualized by a gesture modeled on the ASL sign corresponding to the verb. Based on common bodily

experiences and knowledge, the gestures were supposed to familiarize subjects with abstract concepts.

In Study 5 (Chapter 7), the recall target was a paragraph. The target paragraph was divided into eight idea units, and one word per idea unit was enacted. This time, however, the enactment of the words did not facilitate the recall of a paragraph significantly. There were four possible reasons: technical problems in the process of the experiment, the quality of the gestures as a recall cue, the complex structure of the paragraph, and the general characteristics of mnemonic techniques. Generally a paragraph is grammatically and conceptually complicated compared to words, phrases, and sentences. The target paragraph was composed of various grammatical elements and complex rhetorical relations such as cause-effect and problem-solution. On the other hand, the gestures emphasized a one-to-one relation by providing one word in an idea unit with one recall cue. The results may suggest that the effective role of gesture as a mnemonic cue is limited to a word, phrase or sentence level at the longest.

8.3 Implications for further research

It has been shown that gesture is an effective cue for recalling abstract words as well as concrete words, and for recalling simple sentences if predicate verbs were enacted. Concrete concepts are pictorially described, and abstract concepts are metaphorically represented by some common gestures based on daily experiences. Further research, then, should analyze as many English words as possible in terms of visualization and conceptual metaphor, and test if words in the same conceptual domain can be enacted systematically. Etymology may be useful for conceptual analysis of abstract words, as both *suppose* and *propose* are analyzed and connected by COMMUNICATION IS SENDING metaphor.

Further research also should investigate how teachers present and explain words with

gesture. Using the gesture method suggested by the present dissertation, teachers can show their students the role of gesture in leaning and recalling English vocabulary. Many Japanese teachers of English, sorry to say, are not fully aware of gesture as a useful aid for language teaching and learning. They tend to teach English only through a verbal channel, mainly through Japanese. They should know that they can explain word meaning both verbally and nonverbally, and that their gesture, including their facial expressions, can encourage students to gain access to unknown words. Students may also have stereotype images of gesture; gesture is only physical, childish, and useless for language learning. Gesture performance by students demands strong self-awareness, especially at the presence of peer students. Teachers, then, are expected to offer the gesture method, as one of a variety of methods, and demonstrate the effect of gesture on comprehension and retrieval of English words by visualizing concepts and connecting them with students' sensory-motor experiences. Their daily demonstrations in classrooms will then lighten students' psychological burden and change their concept of gesture.

8.4 Conclusion: learning English with gesture

Mnemonic techniques, classical and modern, all share the assumption that the recall of one idea promotes that of another, and the power of association is related to the high imagery of a recall cue. In the present study, gesture is proposed as a both imaginal and conceptual cue which strengthens associative ties between a target word, its Japanese equivalent, and a learner's prior experience and knowledge. Gesture picks up a sensory-motor aspect of a concept, and visualizes it through a specific image. *Cry*, for example, is visualized by index fingers wiggling down the cheeks. Producing tears is a typical image of crying and easily associated with the concept of *cry*, this is because the gesture is based on a bodily experience that tears roll down the cheeks. If *cry* is

synchronized with the gesture and learned, the word will be coded both verbally and nonverbally, linked with the bodily experience, and readily remembered.

Gesture is effective for the recall of abstract words as well. The visualization of abstract concepts has been an unsolved issue, but the present thesis proposes to represent abstract English words by gestures modeled on corresponding ASL signs. It is assumed that natural gestures and language (signed or spoken) basically stem from bodily experiences, and even abstract concepts, in many cases, are metaphorically represented on the basis of the bodily experiences. Time, for example, is conceived as a linear object in terms of space with a person in the center. This is FUTURE IS AHEAD/PAST IS BEHIND metaphor. Based on the conceptual metaphor, *forecast (*as well as *anticipate)* is visualized by index fingers pointing at the eyes, and then moving forward as if the eyes are looking into the future. *Postpone* is represented by moving a hand forward as if placing something ahead into the future. In contrast, *recollect* is represented by fingers pointing at the eyes and then moving over the shoulder as if the eyes were looking back at the past. *Forecast*, *postpones*, *recollect*, and phrasal verbs such as *look forward to* and *look back* are all related and comprehended in terms of the same concept metaphor, FUTURE IS AHEAD/PAST IS BEHIN. Abstract words are low in concreteness and imagery, but relatively high in meaningfulness. Gesture, therefore, visualizes abstract concept in terms of metaphor linked with prior experience and knowledge.

Language education is widely thought of as being verbally conducted. In the classroom, the four skills, speaking, listening, reading, and writing are repeatedly drilled primarily through the verbal code. Language acquisition, however, is widely accepted as being multi-modal. Infants acquire language through bodily interacting with the physical world. Sensory-motor activities are a fundamental aspect of language acquisition and concept formation. In this sense, language is grounded in bodily experiences. If

sensory-motor information is indispensable for language acquisition, isn't it also true of second language learning? Learners may have little or no knowledge of a second language, but they have acquired many bodily and mental experiences. Their experiences can give them opportunities to conceive the second language from both sensory-motor and conceptual aspects. Although gesture has no standard form or code, based on common experiences, gesture can be a visual and conceptual cue for second language learning.

As has been explained, the role of gesture in English classes in Japan has received little attention. Many teachers consider gesture to be one mode of non-verbal behavior and not directly related to language education. Mental imagery, however, is depictive and descriptive. Language and gesture can represent the same image and concept, and the production and perception of both activities are grounded in the same mental processes. Reducing language learning to the verbal channel is not desirable in terms of both cognition and communication. Japanese teachers of English, then, should notice that their teaching with gesture can facilitate the comprehension and retrieval of English language. Language and gesture are mostly based on common bodily experiences, and students' prior experiences can readily lead them to the concepts of unknown English words and expressions. Gestures based on daily experiences, therefore, can be a promising cue for word recall.

References

Aizawa, K., Ishikawa, S., & Murata, M. (2005). *JACET 8000 eitango*. Tokyo: Kirihara.

Alibali, M., Kita, S., & Young, A. J. (2000). Gesture and the process of speech production: We think, therefore we gesture. *Language and Cognitive Processes*, *15*(*6*), 593-613.

Anderson, J. R. (1978). Arguments concerning representations for mental imagery. *Psychological Review*, *85*, 249-277.

Antes, T. A. (1996). Kinesics: the value of gesture in language and in the language classroom. *Foreign Language Annals*, *29* (*3*), 439-448.

Arnedt, C.S & Gentile, R. (1986). A test of dual coding theory for bilingual memory. *Canadian Journal of Psychology*, *40*, 290-299.

Asher, J.J. (2000). *Learning another language through action* (6th ed.). Los Gatos, California: Sky Oaks Publications.

Atkinson, R.C. (1975). Mnemotechnics in second-language learning. *American Psychologist*, *30*, 821-828.

Begg, I. & Paivio, A. (1969). Concreteness and imagery in sentence meaning. *Journal of Verbal Learning and Verbal Behavior*, *8*, 821-827.

Berlin, B. & Kay, P. (1969). *Basic color terms: Their universality and evolution*. Berkeley: University of California Press.

Brown, R. (1958). How shall a thing be called? *Psychological Review*, *65* (*1*), 14-21.

Brown, T. S. & Perry, F. L. (1991). A comparison of three learning strategies for ESL vocabulary acquisition. *TESOL Quarterly*, *25*(*4*), 655-670.

Carrell, P. (1985). Facilitating ESL reading by teaching text structure. *TESOL Quarterly*, *19* (*4*), 727-745.

Chambers, D. (1993). Images are both depictive and descriptive. In B. Roskos-Ewoldsen, M. J. Intons-Peterson, & R. E. Anderson (Eds.), *Imagery, creativity, and discovery: A cognitive perspective* (pp. 99-121). The Netherlands: Elsevier Science Publishers.

Choi, S.H., Na, D.L., Kang, E., Lee, K.M., Lee, S.W., & Na, D.G. (2001). Functional magnetic resonance imaging during pantomiming tool-use gestures. *Experimental Brain Research*, *1* (*39*), 311-317.

Cienki, A. (1998). Metaphoric gestures and some of their relations to verbal metaphoric expressions. In Jean-Pierre Koening (Ed.), *Discourse and cognition: Bridging the gap* (pp. 189-204). Stanford: CSLI Publications.

Craik, F.I.M. & Lockhart, R. (1972). Levels of processing: A framework for memory research. *Journal of Verbal Learning & Verbal Behavior*, *11*, 671-684.

Craik, F.I.M. & Tulving, E. (1975). Depth of processing and the retention of words in episodic memory. *Journal of Experimental Psychology: General, 104,* 268-294.

Danesi, M. (1988). Neurological bimodality and theories of language teaching. *Studies in Second Language Acquisition*, *10*, 13–31.

De Ruiter, J.P. (2000). The production of gesture and speech. In D. McNeill (Ed.), *Language and gesture* (pp. 282-3119). New York: Cambridge University Press.

Ekman, P. & Friesen, W.V. (1969). The Repertoire of nonverbal behavior: Categories, origins, usage, and coding. *Semiotica*, *1*, 49-98.

Engelkamp, J. (2001). Action memory: system-oriented approach. In H. D. Zimmer, R. L. Cohen, M. J. Guynn, J. Engelkamp, R. Kormi-Nouri, & M. A. Foley. *Memory for action: A distinct form of episodic memory?* (pp. 49-96). New York: Oxford University Press.

Frick-Horbury, D. (2002). The use of hand gestures as self-generated cues for recall of verbally associated targets. *American Journal of Psychology*, *115* (*1*), 1-22.

Gibbs, R. (1994). *The Poetics of mind: Figurative thought, language, and understanding.* Cambridge: Cambridge University Press.

Grady, J.E. (1997). *Foundations of meaning: Primary metaphors and primary scenes*. Ph.D. dissertation, University of California, Berkeley.

Grady, J.E. & Johnson, C. (2003). Converging evidence for the notions of *subscene* and *primary scene*. In J. Monxley & M. Juge (Eds.), *Proceedings of the twenty-third annual meeting of the Berkeley Linguistic Society* (pp. 123-136). Berkeley: Berkeley Linguistic Society.

Grafton, S.T., Fadiga L., Arbib, M.A., & Rizzolatti, G. (1997). Premotor cortex activation during observation and naming of familiar tools. *Neuroimage, 6*, 231-236.

Gu, P.Y. (2003). Vocabulary learning in a second language: Person, task, context and strategies. *Teaching English as a Second or Foreign Language*, *7* (*2*), 1-25, 03/06/2006, http://www-writing.berkeley.edu/TESl-EJ/index.html.

Hill, C. (1982). Up/down, front/back, left/right: A contrastive study of Hausa and English. In J. Weissenborn & W. Klein (Eds.), *Here and there: Cross-linguistic studies on deixis and demonstration* (pp. 13–42). Amsterdam: John Benjamins.

Holcomb, P. J., Kounios, J., Anderson, J. E., & West, W. C. (1999). Dual-coding, context-availability, and concreteness effects in sentence comprehension: An electrophysiological investigation. *Journal of Experimental Psychology: Learning, Memory and Cognition, 25,* 721-742.

Hulstijn, J. (1997). Mnemonic methods in foreign language vocabulary learning: Theoretical considerations and pedagogical implications. In J. Coady & T. Huckin (Eds.), *Second language vocabulary acquisition* (pp. 203-224). Cambridge: Cambridge University Press.

Hunt, R. R. & Einstein, G. O. (1981). Relational and item-specific information in memory. *Journal of Verbal Learning and Verbal Behavior*, *20*, 497-514.

Ikemura, D. (2002). Koko no Jyugyou de. *The English Teachers' Magazine*, *50* (*12*), 13-15.

Iverson, J.M. & Goldin-Meadow, S. (1998). Why people gesture when they speak. *Nature*, 396, 228.

Jessen, F., Heun, R., Erb, M., Granath, D. O., Klose, U., Papassotiropoulos, A., & Grodd, W. (2000). The concreteness effect: Evidence for dual coding and context availability. *Brain and Language*, *74*, 103-112.

Johnson, C. (1997). Metaphor vs. conflation in the acquisition of polysemy: The case of *see.* In M. K. Hiraga, C. Sinha, & S. Wilcox (Eds.), *Cultural, typological, and psychological perspectives in cognitive Linguistics* (pp. 155-169). Amsterdam: John Benjamins.

Johnson, M. (1987). *The body in the mind: The bodily basis of meaning, imagination, and reason.* Chicago: Chicago University Press.

Joiner, E. G. (1984). Listening from the inside out. *Foreign Language Annals*, *17* (*4*), 335-338.

Kaden, S., Kapner, S., & Werner, H. (1955). Studies in physiognomic perception: II, effect of directional dynamics of pictured objects and of words on the position of the apparent horizon. *The Journal of Psychology*, *39*, 61-70.

Kellerman, S. (1992). I see what you mean: The role of kinesic behaviour in listening and implications for foreign and second language learning. *Applied Linguistics*, *13* (*3*), 239-258.

Kendon, A. (1988). How gestures can become like words. In F. Poyatos (Ed.), *Cross-cultural perspectives in nonverbal communication* (pp. 131-141). New York: C.J. Hogrefe.

Kiehl, K. A., Liddle, P. F., Smith, A. M., Mendrek, A., Forster, B. B., & Hare, R. D. (1999). Neural pathways involved in the processing of concrete and abstract words. *Human Brain Mapping*, *7*, 225-233.

Kita, S. (2000a). *Gesture*. Tokyo: Kaneko shobo.

Kita, S. (2000b). How representational gestures help speaking. In D. McNeill (Ed.), *Language and gesture* (pp. 162-185). New York: Cambridge University Press.

Kobayashi, H. (1997). The role of actions in making inferences about the shape and material of solid objects among Japanese 2 year-old children. *Cognition*, *63*, 251-269.

Kohno, M. (1984). Is Kinesic information perceived as an extralanguage or as a language itself ? *Study of Sounds*, *20*, 134-146.

Kormi-Nouri, R., Nyberg, L., & Nilsson, L.-G. (1994). The effect of retrieval enactment on recall of subject-performed tasks and verbal tasks. *Memory & Cognition, 22,* 723-728.

Kormi-Nouri, R & Nisson, L.-G. (2001). The motor component is not crucial! In H. D. Zimmer, R. L. Cohen, M. J. Guynn, J. Engelkamp, R. Kormi-Nouri, & M. A. Foley (Eds.), *Memory for action: A distinct form of episodic memory?* (pp. 97-111). New York: Oxford University Press.

Kosslyn, S.M. (1981). The medium and the message in mental imagery: A theory. *Psychological Review*, *88*, 46-66.

Kosslyn, S. M., Ball, T. M., & Reiser, B. J. (1978). Visual images preserve metric spatial information: Evidence from studies of image scanning. *Journal of Experimental Psychology: Human Perception and Performance, 4* (*1*), 47-60.

Lakoff, G. (1987). *Women, fire, and dangerous things*. Chicago: The University of Chicago Press.

Lakoff, G. & Johnson, M. (1980). *Metaphors we live by*. Chicago: The University of Chicago Press.

Lakoff, G. & Johnson, M. (1999). *Philosophy in the flesh: The embodied mind and its challenge to Western though*. New York: Basic Books.

Luria, A.R. (1968). *The mind of a mnemonist: A little book about a vast memory*. New York: Basic Books.

Marschark, M. & Hunt, R. R. (1989). A reexamination of the role of imagery in learning and memory. *Journal of Experimental Psychology: Learning, Memory and Cognition, 15*(*4*), 710-720.

Masataka, N. (2000). Information from speech and gesture is integrated when meanings of new words are categorized in normal young children, but not in children with Williams syndrome. *Cognitive Studies*, *7*, 37-51.

Matsumi, N. (1994). Processes of word memory in second language acquisition: A test of bilingual dual coding theory. *The Japanese Journal of Psychology*, *64* (*6*), 460-468.

McNeill, D. (1992). *Hand and mind*. Chicago: The University of Chicago Press.

McNeill, D. & Duncan, S.D. (2000). Growth points in thinking-for-speaking. In D. McNeill (Ed.), *Language and gesture* (pp. 141-1619). New York: Cambridge University Press.

Miller, A. (1963). Verbal satiation and the role of concurrent activity. *Journal of Abnormal and Social psychology*, *3*, 206-212.

Miyazaki, H. (1979). Mental image wa e ka meidai ka. *Annual Report of Educational Psychology in Japan*, *19*, 112-124.

Miyazaki, H. (1998). Image wa sozo wo sasaeuruka? In S. Maruno （ed.）, *Shinrigaku no naka no ronsou* (pp. 121-148). Kyoto: Nakanishiya.

Naoyama, Y. (2002). Souki eigo kyouiku deno sidou. *The English Teacher's Magazine*, *50* (*12*), 15-16.

O'Malley, J.M. & Chamot, A.U. (1990). *Learning strategies in second language acquisition*. New York: Cambridge University Press.

Oliveri, M., Finocchiaro, C., Shapiro, K., Gangitano, M., Caramazza, A., & Pascual-Leone, A. (2004). All talk and no action: A transcranial magnetic stimulation study of motor cortex activation during action word production. *Journal of Cognitive Neuroscience, 16*, 374-381.

Oxford, R.L. (1990). *Language learning strategies: What every teacher should know*. Boston: Heinle & Heinle.

Paivio, A. (1965). Abstractness, imagery, and meaningfulness in paired-associate learning. J*ournal of Verbal Learning and Verbal Behavior*, *4*, 32-38.

Paivio, A. (1971). *Imagery and verbal processes*. New York: Holt, Rinehart, & Winston.

Paivio, A. (1986). *Mental representations*. New York: Oxford University Press.

Paivio, A. (1977). Images, propositions, and knowledge. In J. M. Nicholas (Ed.), *Images, perception, and knowledge* (pp. 47-71). Dordrecht/Boston: Reidel.

Paivio, A. (1986). *Mental representations*. New York: Oxford Science Publications.

Paivio, A., Yuille, J.C., & Madigan, S.A. (1968). Concreteness, imagery, and meaningfulness values for 925 nouns. *Journal of Experimental Psychology*, *76* (*1*), Monograph Supplement Part 2, 1-25.

Paivio, A. & Csapo, K. (1973). Picture superiority in free recall: Imagery or dual coding? *Cognitive Psychology, 5*, 176-206.

Paivio, A. & Desrochers, A. (1980). A dual-coding approach to bilingual memory. *Canadian Journal of Psychology*, *34*, 390-401.

Paivio, A. & Desrochers, A. (1981). Mnemonic techniques in second-language learning. *Journal of Educational psychology*, *73* (*6*), 780-795.

Palvio, A. & Lambert, W. (1981). Dual coding and bilingual memory. *Journal of Verbal Learning and Verbal Behavior*, *20*, 532-539.

Paivio, A. Khan, M. & Begg, I. (2000). Concreteness and relational effects on recall of adjective-noun pairs. *Canadian Journal of Experimental Psychology*, *54* (*3*), 149-159.

Pansky, A. & Koriat, A. (2004). The basic-level convergence effect in memory distortions. *Psychological Science*, *15* (*1*), 52-59.

Pressley, M., Levin, J. R., & Miller, G. A. (1981). The Keyword method and children's learning of foreign vocabulary with abstract meanings. *Canadian Journal of Psychology*, *34*, 283-87.

Pylyshyn, Z. W. (1973). What the mind's eye tells the mind's brain. *Psychological Bulletin*, *80* (*1*), 1-24.

Pylyshyn, Z.W. (1981). The imagery debate: Analogue media versus tacit knowledge. *Psychological Review, 88*, 16–45.

Rauscher, F. H., Krauss, R. M., & Chen, Y. 1996. Gesture, speech, and lexical access. *Psychological Science*, *7* (*4*), 226-231.

Reddy, M. (1979). The conduit metaphor. In A. Ortony (Ed.), *Metaphor and thought* (pp. 240-251). Cambridge: Cambridge University Press.

Reisberg, D. & Chambers, D. (1991). Neither pictures nor propositions: What can we learn from a mental image? *Canadian Journal of Psychology*, *45* (*3*), 336-352.

Riseborough, M. G. (1981). Physiographic gestures as decoding facilitators: Three experiments exploring a neglected facet of communication. *Journal of Nonverbal Behavior, 5*, 172-183.

Rosch, E. H. (1973). On the internal structure of perceptual and semantic categories. In T. E. Moore (Ed.), *Cognitive development and the acquisition of language* (111-144). New York: Academic Press.

Rosch, E., Mervis, C.B., Gray, W.D., Johnson, D.M., & Boyes-Braem, P. (1976). Basic objects in natural categories. *Cognitive Psychology*, *8*, 382-439.

Russ, M.O., Mack, W., Grama, C.R., Lanfermann, H., & Knopf, M. (2003). Enactment effect in memory: Evidence concerning the function of the supramarginal gyrus. *Experimental Brain Research*, *149* (*4*), 497–504.

Sadoski, M., Goetz, E.T., & Avila E. (1995). Concreteness effects in text recall: Dual coding or context availability? *Reading Research Quarterly*, *30* (*2*), 278-288.

Sadoski, M. & Paivio, A. (2001). *Imagery and text*. New Jersey: Lawrence Erlabum Associates.

Saltz, E. & Donnenwerth-Nolan, S. (1981). Does motoric imagery facilitate memory for sentences? A selective interference test. *Journal of Verbal Learning and Verbal Behavior*, *20*, 322-332.

Schmitt, N. (1997). Vocabulary learning strategies. In N. Schmitt & M. McCarthy (Eds.), *Vocabulary: Description, Acquisition, and Pedagogy* (pp. 199-227). Cambridge: Cambridge University Press.

Schwanenflugel, P. J. (1991). Why Are Abstract Concepts Hard to Understand? In P. J. Schwanenflugel (Ed.), *The psychology of word meanings* (pp. 223-250). New Jersey: Lawrence Erlbaum Associates.

Schank, R. C. & Abelson, R.P. (1977). *Scripts, plans, goals, and understanding: An inquiry into human knowledge structures.* Hillsdale, NJ: Lawrence Erlbaum.

Sternberg, M.L.A. (1994). *American sign language concise dictionary*. New York: HarperCollins Publishers, Inc.

Sweetster, E. E. (1990). *From etymology to pragmatics*. New York: Cambridge University Press.

Sweetser, E. E. (1987). Metaphorical models of thought and speech: A comparison of historical directions and metaphorical mappings in the two domains. In J. Aske, N. Beery, L. Michaelis, & H. Filip (Eds.), *Proceedings of the thirteenth annual meeting of the Berkeley Linguistics Society* (pp. 446-459). Berkeley: Berkeley Linguistic Society.

Taub, S.F. (2001). *Language from the body*. Cambridge: Cambridge University Press.

Teshima, M. (2002), Tyugakuko no Jyugyou de. *The English Teachers' Magazine*, *50* (*12*), 13-15.

Wattenmaker, W. D. & Shoben, E. J. (1987). Context and the recallability of concrete and abstract sentences. *Journal of Experimental Psychology: Learning, Memory and Cognition, 13*(*1*), 140-150.

Weinreich, U. (1953). *Languages in contact: Findings and problems*. New York: Linguistic Circle of New York.

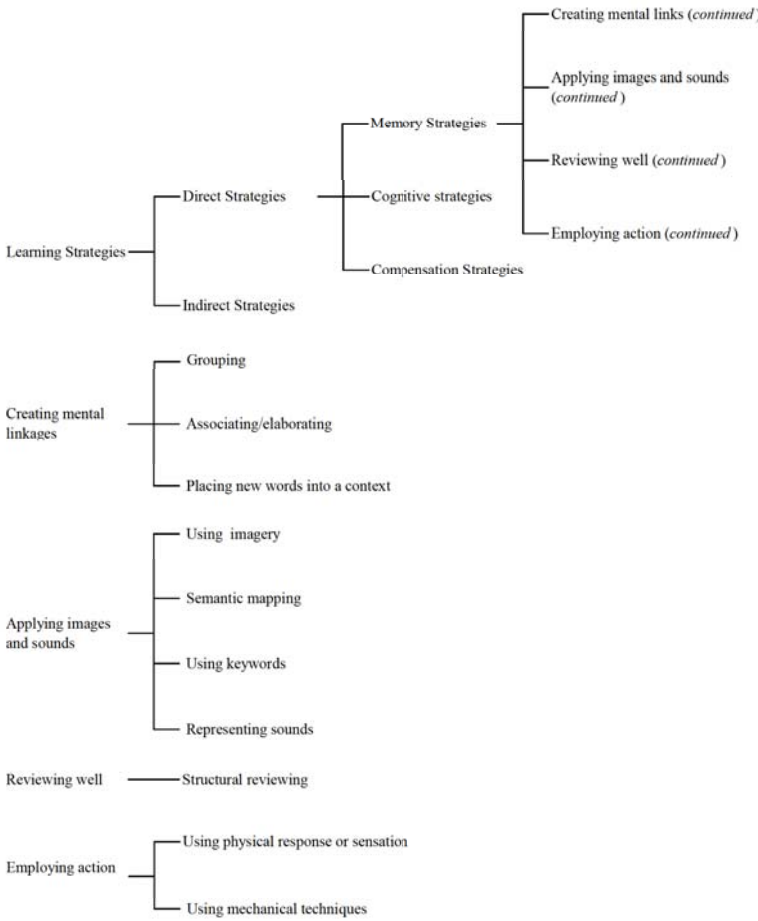
Werner, H. & Kaplan, B. (1963). *Symbol formation*. New York: Wiley.

Wilcox, P.P. (2000). *Metaphor in American sign language*. Washington D.C.: Gallaudet University Press.

Yates, F. (1966). *The art of memory*. Chicago: The University of Chicago Press.

Appendices

Appendix A Memory strategy diagrams revised from Oxford (1990: pp. 16-18)



Appendix B　A list of vocabulary learning strategies used by Japanese EFL learners revised from Schmitt (1997: pp. 207-208)

strategies	use (%)	helpful (%)
Strategies for the discovery of a new word's meaning		
Analyse part of speech	32	75
Analyse affixes and roots	15	68
Check for L1 cognate	11	40
Analyse any available pictures or gestures	47	84
Guess from textual context	74	73
Bilingual dictionary	85	95
Monolingual dictionary	35	77
Ask teacher for an L1 translation	45	61
Ask teacher for paraphrase or synonym of anew word	42	86
Ask teacher for a sentence including the new word	24	78
Ask classmates for meaning	73	65
Discover new meaning through group activity	35	65
Strategies for consolidating a word once it has been encountered		
Study and practice meaning in a group	30	51
Teacher checks students' flash cards or word lists for accuracy	3	39
Image word's meaning	50	38
Connect word to a personal experience	37	62
Associate the word with its coordinates	13	54
Connect the word with its synonyms and antonyms	41	88
Use semantic maps	9	47
Use 'scales' for gradable adjectives	16	62
Use new word in sentences	18	82
Study the spelling of a word	74	87
Study the sound of a word	60	81
Say new word aloud when studying	69	91
Image word's form	32	22
Use Keyword Method	13	31
Affixes and roots (remembering)	14	61
Parts of speech (remembering)	30	73
Paraphrase word's meaning	40	77
Use cognates in study	10	34
Learn the words of an idiom together	48	77
Use physical action when learning a word	13	49
Verbal repetition	76	84
Written repetition	76	91
Word lists	54	67
Flash cards	25	65
Take notes in class	64	84
Use the vocabulary section in your textbook	48	76
Skip or pass new word	41	16
Continue to study word over time	45	87

Appendix C　A list of English words and Japanese words for Study 1

category	control group		experimental group	
	English	日本語	English	日本語
1 building	house	塔	tower	家
2 geography	sea	川	river	海
3 body	arm	唇	lip	腕
4 figure	circle	四角	square	円
5 mature	fire	水	water	火
6 goods	fork	オーブン	oven	フォーク
7 plant	flower	草	grass	花
8 stationary	book	鉛筆	pencil	本
9 animal	bird	猫	cat	鳥
10 vehicle	car	船	ship	車
11 house	door	窓	window	戸
12 human	boy	少女	girl	少年
13 tool	hammer	スプレー	spray	ハンマー
14 cooking	pepper	砂糖	sugar	胡椒
15 children	baby	子供	child	赤ん坊
16 sun	dawn	日の入り	sunset	夜明け
17 fruit	lemon	リンゴ	apple	レモン
18 space	sky	星	star	空
19 furniture	chair	テーブル	table	チェア
20 clothing	dress	クツ	shoes	ドレス

Appendix D　A wordlist of concrete words and abstract words for Study 3, revised from Paivio et al. (1968pp. 10-25)

	concrete word	I	C	abstract word	I	C
1	bird	6.67	6.96	ability	2.67	2.63
2	book	6.43	6.96	agreement	3.33	2.93
3	chair	6.63	7.00	amount	2.73	3.62
4	clock	6.50	6.94	anger	4.87	1.70
5	door	6.60	7.00	cleanness	4.77	3.63
6	dress	6.53	6.93	co-operation	3.40	1.92
7	flower	6.57	6.96	development	3.07	2.82
8	folk	6.57	6.94	discovery	4.13	3.06
9	grass	6.63	6.96	dream	4.6	3.03
10	hammer	6.73	6.96	effort	3.33	2.22
11	house	6.67	6.93	expression	3.57	2.48
12	lemon	6.83	6.96	grief	4.70	1.86
13	ocean	6.70	6.90	hatred	3.97	1.52
14	oven	6.40	6.96	health	4.10	3.54
15	pipe	6.43	6.90	justice	3.60	2.18
16	ship	6.67	6.93	pleasure	4.80	2.10
17	sugar	6.57	6.96	power	4.47	2.73
18	table	6.50	7.00	pressure	4.10	3.63
19	tower	6.53	6.96	promotion	3.33	3.56
20	water	6.60	6.96	thought	2.77	1.28

I: imagery, C: concreteness

＊Imagery and concreteness were rated on 7-point scales. *Water*, for example, was rated as 6.60 and 6.96. Both imagery and concreteness were very high.

Appendix E　A list of sentences for Experiment 1 in Study 4

練習文

1 The cook washed the pot.（行為動詞）
2 The fruit contained the vitamins.（状態動詞）

課題文

1 The fan caught the ball.（行為）
2 The singer heard her song.（状態）
3 The mother cut the cake.（行為）
4 The children believed the story.（状態）
5 The racer drove the car.（行為）
6 The students feared the teacher.（状態）
7 The pilot pulled the lever.（行為）
8 The boy wanted the soccer ball.（状態）
9 The salesman pushed the button.（行為）
10 The coat fitted the gentleman.（状態）
11 The boy threw the stone.（行為）
12 The baby understood the gesture.（状態）
13 The father painted the fence.（行為）
14 The farmer owned the land.（状態）

Appendix F　A list of sentences for Experiment 2 in Study 4

練習文

1 Santa Claus came to the house.（高頻度動詞）
2 The army surrendered to the enemy.（低頻度動詞）

課題文

1 1 The teacher called the name.（高）
2 The librarian classified the books.（低）
3 Fans followed the player.（高）
4 The journalist criticized the music.（低）
5 The girl liked the cat.（高）
6 The students assembled in the hall.（低）
7 The doctor told the truth.（高）
8 The artist exhibited the picture.（低）
9 The boy wanted the bike.（高）
10 The men guarded the star.（低）
11 The robber took the bag.（高）
12 People admired the President.（低）
13 The mother read the story.（高）
14 The patient tolerated the pain.（低）

Appendix G A task paragraph of eight idea units for Study 5

The company had been declining. The president realized that they had to do something. They developed a new product to attract more customers. Two years later, the sales increased more than they had anticipated.

1 The company had been **declining**.
2 The president **realized**.
3 that they had to **do** something.
4 They **developed** a new product
5 to **attract** customers.
6 Two years **later**,
7 the sales **increased**
8 more than they had **anticipated**.

* The bold words are enacted and recalled.
* Underlined words are also recall targets.