Fall 11-15-2017

THINKING-FOR-SPEAKING AND THE EFL MIND: FACE-TO-FACE DIALOGUE TO TALK ABOUT VERTICAL SPACE

Kunisawa
The University of New Mexico

Tae Kunisawa
Candidate

Language, Literacy, & Sociocultural Studies
Department

This dissertation is approved, and it is acceptable in quality and form for publication:

Approved by the Dissertation Committee:

Dr. Jill Morford, Primary Co-Chairperson

Dr. Holbrook Mahn, Co-Chairperson

Dr. Sherman Wilcox

Dr. Shoko Hamano

THINKING-FOR-SPEAKING AND THE EFL MIND: FACE-TO-FACE DIALOGUE TO TALK ABOUT VERTICAL SPACE

By

Tae Kunisawa

B.A., British and American Literature, University of Kochi, Japan, 1974
M.A. in Education, East Carolina University, USA, 1995

DISSERTATION

Submitted in Partial Fulfillment of the
Requirements for the Degree of

Doctor of Philosophy
Educational Linguistics

The University of New Mexico
Albuquerque, New Mexico

December, 2017

DEDICATION

For Harunori, Sayoko, and Eiji Kunisawa

ACKNOWLEDGMENTS

I wish to acknowledge the contributions of the following individuals to various aspects of this dissertation project. I could not have completed this project without their generous contributions and expertise. First of all, I hold Jill Morford in my highest regards. She has offered me years of unconditional mentorship, including guiding me through a proper research direction, statistical methods, and in creating research questions and hypotheses in my dissertation. Her insightful feedback on this dissertation was invaluable, and thus, I am deeply indebted to her. She has properly led me to the door where one of the goals in my life had stayed, while I was walking an extremely long and unbelievably winding road in completing this dissertation.

I wish to thank Holbrook Mahn who looked at my descriptions in more detail regarding the relationship between gesture and linguistic relativity and gave me proper suggestions about it. I would also like to thank Sherman Wilcox for his willingness to be part of this project, despite the fact that he has various different kinds of obligations in his professional and family life. I am also grateful for his suggestions. I am extremely grateful for Shoko Hamano who has helped me not only with linguistic analysis in Japanese, but has made a suggestion about the relationship between gesture and linguistic relativity.

Additionally, I would like to offer my sincerest gratitude to the researchers who gave me insightful input at the University of New Mexico and at various conferences: James Degnan, Adele Goldberg, Vera John-Steiner, Vsevolod Kapatsinski, Sotaro Kita, James Lantolf, Brian MacWhinney, Kyoko Masuda, Sharon Newbill, Lourdes Ortega, Richard Plass, and Eve Sweetser. Furthermore, I also would like to express my sincere thanks to the following classmates, roommates, and friends who discussed issues in my research with me and kindly offered their editing help: Jack Bogusch, Hafiz Fazalehaq, Margaret Johnson, Susan Metheny, Patrick Osahor, Angelica Plass, and Ryan Smith. Lastly, and the most prominently, I am grateful for unconditional support offered by my parents, Harunori and Sayoko Kunisawa, and my son, Eiji Kunisawa, although my parents were already deceased before I completed this project.

THINKING-FOR-SPEAKING AND THE EFL MIND: FACE-TO-FACE DIALOGUE TO TALK ABOUT VERTICAL SPACE

By

TAE KUNISAWA

B.A., British and American Literature, University of Kochi, Japan, 1974
M.A. in Education, East Carolina University, USA, 1995
Ph.D., Educational Linguistics, University of New Mexico, USA, 2017

ABSTRACT

Studies of thinking-for-speaking (Slobin, 1987) and of linguistic relativity (Gumperz & Levinson, 1996) in multilinguals have been attracting more attention (Ortega, 2015). I propose the incorporation of sociocultural theory and linguistic relativity as a novel research approach in second language acquisition (SLA). Japanese learners of English go through a process in which word meaning develops from a single to binary semantic categorization as they learn to express vertical spatial operations in their second language (L2). Japanese has a nonobligatory distinction between contact and noncontact relationships when expressing vertical space (single semantic categorization), whereas English has an obligatory contrast (binary semantic categorization) (Munnich et al., 2001). The expression of vertical spatial relationships in Japanese and English is further influenced by language typology. Japanese, an SOV language, uses postpositions while English, an SVO language, uses prepositions.

Vygotsky (1987) argues that verbal thinking (the internalization of speech) is tied with word meaning, and thus, as Japanese EFL high school students learn to express the obligatory contact-noncontact feature of vertical spatial configurations in English, moving

from a single to a binary semantic categorization, verbal thinking will also develop. Vygotsky (1987) further claims that verbal thinking has sociocultural origins. In this dissertation, I investigate whether gesture can be instrumental in overcoming the constraints imposed by linguistic relativity. Vygotsky (1998) states, “Speech becomes the means for thinking mainly because it reflects an objectively occurring intellectual operation. This is a moment of major importance in the development of speech and thinking, which discloses the secret of the development of verbal thinking as a whole” (p. 114). I predict that a distinct worldview and the “development of cognitive processes” (Matyushkin, 1997b, p. 272) arise together when Japanese EFL students learn vertical spatial structure with the Gesture Listening Higher Concept Approach, which leads to “a qualitatively new mental formation that develops according to completely special laws and is subject to completely different patterns” (Vygotsky 1998, p. 34).

The purpose of this study: (1) To pursue the new research path regarding incorporating linguistic relativity into SLA in sociocultural theory; (2) to explore whether the concurrent use of iconic co-speech co-thought gesture (ICSCTG) and listening practice can help Japanese high school students learn to express vertical spatial relationships in English more than they would learn from either treatment alone; (3) to investigate whether teaching ICSCTG and listening practice together will help Japanese EFL learners preserve knowledge of how to express vertical spatial relationships in English for a month after the intervention. I employed quantitative methods to accomplish the goals noted above. Results in this study suggest that the Gesture Listening Higher Concept Approach is an effective, evidence-based theoretical and pedagogical framework, which can facilitate L2 learning and conceptual change at the high school level. The effect of the Gesture Listening Higher Concept

Approach on long-term foreign language learning would be a valuable avenue for future research.

Slobin, D. I. (1987). Thinking for Speaking. Proceedings of the Thirteenth Annual Meeting of the Berkeley Linguistics Society, 13, pp. 435-445.

Table of Contents

List of Figures

List of Tables

Chapter 1

Introduction

The language used to talk about space differs markedly between cultures, despite the fact that a physical space is the same regardless of culture (Bowerman, 1996a, 1996b; Levinson, 2003; Pederson, Danziger, Wilkins, Levinson, Kita, & Senft, 1998; Levinson & Wilkinson, 2006a, 2006b). I personally experienced the difference when I taught Japanese to English-speaking high school students. The use of the Gesture Listening Higher Concept Approach[1] had a significant effect on students' success in learning Japanese. This approach refers to the simultaneous use of concurrent iconic co-speech co-thought gesture (ICSCTG) and listening practice with an accumulation of knowledge in a level of scientific and categorical thinking (a higher concept) (Vygotsky, 1987). I saw a similar effect of this approach on students' ability to learn about vertical spatial structures in an English as a foreign language (EFL) Japanese high school class. (Refer to Appendices A and B for a list of acronyms in this study).

This experimental study investigates whether the use of the Gesture Listening Higher Concept Approach has a significant effect on teaching English as a foreign language (EFL) to high school students in Japan. The following includes a structure of this dissertation that consists of five chapters. Chapter 1 provides purposes of the current study, a statement of the problems, and research questions and hypotheses. Subsequent sections demonstrate reasons why these research questions and hypotheses are crucial in experimental gesture studies in

[1] Vygotsky (1987) argues that a "higher concept presupposes both a hierarchical system and concepts subordinate and systematically related to the given concept" (p. 192). For example: Some EFL learners have difficulties with putting knowledge together in formulating speech, even though they have grammatical and semantic knowledge. This approach could help those students.

EFL, and how these questions and hypotheses are motivated by prior empirical, hypothetical, and theoretical work. Additionally, this chapter includes a brief explanation of the component of the Gesture Listening Higher Concept Approach. Chapter 2 reviews prior literature that experimentally and theoretically strengthens the hypotheses in this study. Based on studies of sociocultural theory, gesture research in first language (L1) and FL, and contemporary linguistics research, this chapter includes an experimental, theoretical, and hypothetical discussions of this approach. Chapter 3 includes an investigation of effective methods to achieve the purposes in this study, which built upon studies of contemporary linguistics, of gestures in L1 and second language (L2), of linguistic relativity (language shapes thought), of second language acquisition (SLA), and of sociocultural theory. Based on collected data in an EFL high school class in Japan, this chapter also includes a discussion of statistical methods to obtain accurate results. Chapter 4 provides the results of this experimental study, which both support my hypotheses and answer my research questions. This chapter also describes an effect of a positive correlation that links to explanations of how a higher concept emerges. Finally, Chapter 5 revisits the research questions and hypotheses investigated in this dissertation. This chapter also includes the discussions and conclusions of this study based on findings, results, and novel discoveries. On the basis of the experimental, hypothetical, and theoretical investigations, I conclude that the Gesture Listening Higher Concept Approach is one of the most effective and evidence-based teaching methods in Japanese EFL high school education that led to accomplishing three goals to be described in section 1.1. Furthermore, this chapter proposes pedagogical suggestions, as well as implications of future research in experimental gesture studies in FL.

1.1 Purpose of This Study

This experimental study has three goals: First, this study will respond to the novel research trend regarding incorporating a moderate version of linguistic relativity that is Thinking-For-Speaking (TFS) into SLA in sociocultural theory. The studies of TFS in L2 and of a weak version of linguistic relativity in multi-linguals continue to attract increased attention (Ortega, 2015), and are a novel research direction.[2]

Unfortunately, scholars indicated that applied linguistics failed to incorporate insights from the heyday of empirical linguistics research, including the relationships between language and culture investigated by Whorf (1956 as cited in Kramsch, 2006). In the 1950s and early 1960s, some structuralists shared specific approaches to studying and analyzing language processes with behaviorism, including the audio-lingual approach, based on structuralist linguistics and behaviorist psychology. Whorf (1956) was a structuralist, but not interested in behaviorism. Chomsky (1959) attacked behaviorism and B. F. Skinners'(1938) proposal that language acquisition was subject to social conditioning and habitual processes. This period was not a time when SLA research included examination of linguistic relativity. Instead, contemporary studies on language and thought introduced novel fields into the established disciplines of the social sciences.

Vygotsky (1987) suggests that verbal thinking has sociocultural origins, which indicates that the dissimilarity of social beings triggers different verbal thinking patterns. This is an equivalent claim to that of the linguistic relativity principle that shares a similar idea with TFS, which states that "users of markedly different grammars…must arrive at

[2] Refer to Kunisawa (under review) for the differentiations among TFS, the weak and strong versions of linguistic relativity.

somewhat different views of the world" (Whorf, 1956, p. 221). Therefore, the incorporation of sociocultural theory into TFS can respond to a novel research pathway in SLA.

The second goal of the current study is to includes an exploration of whether the concomitant use of iconic co-speech co-thought gestures (ICSCTGs) and listening practice can help Japanese high school students learn about a single word meaning (Vygotsky, 1987) [3] more than they would learn from either treatment alone. This issue remained unclear to date in experimental gesture studies in foreign language (FL) acquisition. To explore this issue, an experimental study was completed in an EFL class for Japanese high school students.

The third goal is to investigate whether learning is retained a month later using these methods. Literature review indicated that the concurrent iconic gestures and listening practice had an effect on learning a single word meaning, but for no more than 5 days, even though researchers did not clarify this limitation of their study results (Macedonia, Muller, & Friederici, 2011).[4] To explore the goals noted above, the current study created the Gesture Listening Higher Concept Approach that refers to the simultaneous use of ICSCTGs and listening practice with an accumulation of knowledge in a level of scientific and categorical thinking (i.e., a higher concept). This approach built upon the researcher's teaching and research experiences in the United States, Canada, and Japan, which led to the incorporation of these experiences into studies of TFS, of SLA in sociocultural theory, of gestures in L1 and FL, and of contemporary linguistic analyses. The study includes a pilot, which lasted approximately 5 months in an EFL class for Japanese high school students, to investigate the three goals by utilizing this approach.

---

[3] See Appendix C for the definition of the term *single word meaning*.
[4] Refer to Chapter 5 for more detail.

The domains in this study are research in:

- Experimental gesture studies in FL acquisition (Kelly, McDevitt, & Esch, 2009; Macedonia, Muller, & Friederici, 2011; Quinn-Allen, 1995; Tellier, 2008),
- Gesture research in L1 (Chu & Kita, 2012, 2016; Goldin-Meadow, 2001, 2003, 2009; Kita, 2000; Kita, Alibali, & Chu, 2017; Kita & Ozyurek, 2003; McNeill, 1992, 2000, 2008, 2005, 2012),
- Linguistic relativity (TFS) (Boas, 1911; Lucy & Wertsch, 1987; Slobin, 1987, 1996; Whorf, 1956),
- SLA (Cummins 2000, 2001, 2005; Vygotsky, 1987, 1993, 1997a, 1997b, 1998, 1999),
- Contemporary linguistics (Comrie, 1989a; Langacker, 1987, 1993, 2008, 2009; Huddleston & Pullum, 2002; Iwasaki, 2013; Kunihiro, 1967, 1986; Kuno, 1973; Martin, 1975; Shibatani, 1990; Tanaka, 1997; Tyler & Evans, 2003).

Helpful hypotheses are:

- The gesture-for-conceptualization hypothesis (Kita, Alibali, & Chu, 2017),
- The Information Packaging Hypothesis (IPH) (Kita, 2000; Kita et al., 2017),
- The Interface Hypothesis (IH) (Kita, 2014;Kita & Ozyurek, 2003),
- The forgetting curve hypothesis (Ebbinghaus, 1985/ 1913; Heller, Mack, & Seitz, 1991; Murre & Dros, 2015).

Levinson (1996b) suggested that spatial conceptualization was fundamental to social cognition. Additionally, Chilton (2015) argues, “Research into spatial preposition in general indicates that geometry is essential for their description” (p. 16). Many spatial operations take place in three dimensions, including vertical, lateral, and sagittal orientations. This study includes a discussion of spatial conceptualizations by focusing on vertical axis operations, but not lateral and sagittal positioning, although the current study includes discussions about both lateral and sagittal orientations briefly to explain how Japanese speakers conceptualize vertical space.

Although both Japanese and English speakers are shown the same pictures of vertical spatial relationships, the languages that they employ are substantially different, which explains the dissimilar vertical axis conceptualization (See Appendix H). Furthermore, acquiring spatial language semantics is a special challenge (Baek, 2015; Celce-Murcia, Larsen-Freeman & Williams, 1999; Coventry et al., 2012; Ijaz, 1986; Jarvis & Odlin, 2000; Munnich & Landau, 2010). Thus, a majority of EFL students will face challenges when learning about vertical relationships, specifically if their L1 typologically differs from English in syntax and has non-obligatory contrast between contact and noncontact distinctions as in Japanese. The following section 1. 2 uncovers five problems in FL research.

1.2 Statement of Problem

1.2.1 Effect of ICSCTG Accompanied by Listening Practice

Researchers have found that co-verbal gestures have an effect on learning not only in speakers' L1, but also their FL (Kelly et al., 2009). Furthermore, an effect of the concurrent iconic gestures and listening practice on learning artificial words and maintaining knowledge of them included exploration at a college level (Macedonia et al., 2011). However, whether the simultaneous ICSCTGs and listening practice can facilitate learning about vertical spatial structure more than either treatment alone remains unclear to date. To extent to which the concurrent use of these two techniques can have an effect on learning about vertical axis operations is uncertain to date, as well (M. Gullberg, personal communication, July 9-10, 2014).[5] Chapter 5 includes a discussion of these two issues.

1.2.2 Spatial language

Researchers in second language (L2) learning of spatial language published only a few studies, despite growing interest in the challenges faced by L2 students in learning spatial language (Celce-Murcia et al., 1999; Coventry et al., 2012; Ijaz, 1986; Jarvis & Odlin, 2000; Munnich & Landau, 2010; Pavlenko, 2014). Four potential problems which Japanese EFL students encounter in talking about vertical axis structures in English are: (a) omitting prepositions (Schumann, 1986, p. 277); (b) not making a contact-noncontact distinction that lacks proper developed meaning of words in EFL[6] because of the confusion of an appropriate use of a prepositional phrase that consists of a single word meaning and word order; (c) creating an erroneous relationships regarding a trajector being higher than the corresponding to landmark (TRHLM) or a trajector being lower than the corresponding to landmark (TRLLM) relationships, although students express the contact-non contact distinctions; (d) improper choice of a preposition.

1.2.3 Development of Meaning of Words in FL

Despite that the concept of development of meaning of words is central to the theory of thinking and speech (Vygotsky, 1987), a structure of developed meaning of words in FL remains uncertain to date when EFL learners express vertical axis relationships in SLA. The

---

[5] See Chapter 5 for a discussion of these two issues.
[6] See Appendix C for the definition of the term *developed meaning of words.*

concept of the single and binary semantic categorization[7] can allow one to study the structure when they learn about vertical axis relationships in a Japanese EFL class.[8]

1.2.4 Vertical Spatial Conceptualization in Japanese

Several published research studies include convincing examination of whether Japanese speakers conceptualize vertical space with a non-obligatory contrast in Japanese linguistics,[9] although studies of horizontal orientations exist in Japanese linguistics (Kabata, 2000, 2014; Kabata & Rice, 1997; Kumashiro,1994a, 1994b; Kunihiro, 1967a, 1967b, 1982, 1986; Martin, 1975; Masuda 2002; Tanaka, 1997; Yamada, 1981; Yamanashi, 1994). On the basis of corpora analyses in Japanese, studies presented by Kunihiro (1986), Munnich et al. (2001), Munnich and Landau (2003),[10] Tanaka (1997), and Vygotsky (1987), Japanese speakers have non-obligatory contrast between contact and noncontact in expressing vertical spatial operations (see Kunisawa, under review).

1.2.5 EFL Instruction in Japan

The Educational Testing Service (2013) indicated that the Test of English as a Foreign Language (TOEFL) score obtained by Japanese speakers who took the test in 2013 were ranked fifth to last among Asian countries. Moreover, according to the U.S. Foreign Service Institute (FSI) (2014), structural differences between Japanese and English are significant. Students at FSI are adult native speakers of English.[11] The following table 1 includes an

---

[7] See Appendix C for the definition of the term *single and binary semantic categorization*.
[8] Refer to Kunisawa (under review) for a discussion of developed meaning of words and single and binary semantic categorization.
[9] Psychological experimental studies presented by Munnich et al. (2001) discussed the non-obligatory contrast in vertical spatial operations for Japanese speakers and they (2001) identified that Japanese speakers had non-obligatory contrast between contact and noncontact by “an observation by Japanese…linguists and informants”, but not based on published papers in linguistics regarding the non-obligatory contrast when Japanese speakers are talking about vertical spatial relationships (E. Munnich, personal communication, November, 5, 2016).
[10] Munnich and Landau (2003) summarized a study of Munnich et al. (2001).
[11] FSI teaches over 70 different foreign languages (e.g., FSI, 2014, p. 176).

explanation of a degree of difficulty in learning FLs for English speakers based on categories which specified amount of time in acquiring a certain level of proficiency in each foreign language.

Table 1:

Amount of Time in Acquiring a Certain Level of Proficiency

Category I: 24 weeks	
Danish Dutch French (30 weeks) Italian	Norwegian Portuguese Romanian Spanish Swedish
Category II: 36 weeks	
German, Indonesian	Malay, Swahili
Category III: 44 weeks	
Most non-Romance/ Germanic	except Arabic, Chinese (Cantonese and Mandarin,) Japanese and Korean
Category IV: 88 weeks	
Arabic, Cantonese, Mandarin	Japanese, Korean

(FSI, 2014, p. 176)

FSI (2014) categorized Japanese as one of the most difficult languages for English speakers, and thus, differences between Japanese and English are significant regarding speaking, reading, and listening. Japanese EFL learners face additional challenges of learning about vertical spatial operations. I argue that the Gesture Listening Higher Concept Approach can be an effective and evidence-based teaching method for learning about vertical spatial relationships in Japanese EFL education.

1.3 Research Questions and Hypotheses

Based on the goals in the current study, the problems explained above, prior studies, and the concept of the Gesture Listening Higher Concept Approach, the research questions and hypotheses for this study are presented below:

Research Question 1: The major research question in this study is: Does teaching students ICSCTG facilitate learning how to express vertical spatial relationships in English for Japanese EFL high school students?[12]

Research Question 2: A secondary research question in this study is: Does listening practice facilitate learning how to express vertical spatial relationships in English for Japanese EFL high school students with or without ICSCTG?

Hypothesis 1: Teaching ICSCTG will facilitate learning how to express vertical spatial relationships in English for Japanese EFL high school students.

Hypothesis 2: Teaching ICSCTG and listening practice combined will facilitate learning how to express vertical spatial relationships in English for Japanese EFL high school students more than either treatment alone.

Hypothesis 3: Teaching ICSCTG and listening practice combined will help Japanese EFL learners maintain knowledge of how to express vertical spatial relationships in English.

The following sections explain reasons why these research questions and hypotheses are crucial in experimental gesture studies in FL, and how these questions and hypotheses are motivated by prior empirical, hypothetical, and theoretical work.

---

[12] It has to be noted that in oral speech conditions, there are different answers in Research Question 1 and two different situations in supporting Hypothesis 1, including post-test situations, and between post- and delayed setting (effect of time). Refer to later Chapter 4.

Despite the fact that Moskowitz (1976) suggested that FL teachers obtained higher evaluation from their students when employing nonverbal communication skills, a very few experimental gesture studies in FL investigating the use of co-speech gestures in teaching L2 have been conclusive (Kelly et al., 2009; Sueyoshi & Hardison, 2005; Tellier, 2008,). Additionally, as addressed earlier, the simultaneous use of ICSCTGs and listening practice has been little discussed to date, even though gesture researchers in FL collected and analyzed data with the concomitant use (Macedonia et al. 2011). Additionally, this study draws on my approximately 30 years of teaching experience to scrutinize the effects of the simultaneous ICSCTGs, listening practice, and higher concepts. One problem my previous American high school students in a Japanese as a Foreign Language (JFL) class faced was the considerable typological differences between the two languages in syntax and semantics. Shifting linguistic knowledge in English to that of JFL alleviated the problem by re-conceptualizing students' habitual thought in English to that of JFL with the Gesture Listening Higher Concept Approach.

In spite of the insufficient experimental work in gesture studies in FL explained, my experiential background described above, and significant evidence of hypothetical, theoretical, and empirical work in L1 explained below led to investigating the research questions and hypotheses noted above. This study incorporates my nearly 30 years' experience as an EFL teacher in Japan, as a JFL instructor in the United States, and as a researcher in the United States and Canada into EFL high school education in Japan to explore the effect of this approach. The following describes how the research questions and hypotheses noted above are motivated by prior empirical evidence and hypotheses in gesture study in L1, FL, and bilingualism, and studies in sociocultural theory.

1.4 Prior Empirical Evidence, Hypotheses, and Theory

This section 1.4 includes discussions of existing empirical studies and hypotheses in gesture studies in L1, FL, and bilingualism. Subsequently, based on sociocultural theory, this section explains a theoretical factor of an effect of listening practice. In the process of examining these research questions and hypotheses, a crucial factor of a higher concept presented by sociocultural theory has been discovered. Accordingly, the Gesture Listening Higher Concept Approach built upon the prior studies in gestures research in L1, FL, sociocultural theory, and my nearly 40 years' experience in academia.

1.4.1 Gesture Studies in L1

Kita et al. (2017) summarize existing gesture studies in L1 and propose the gesture-for-conceptualization hypothesis, which is analogous to the Gesture Listening Higher Concept Approach concerning the effects of regenerating ICSCTGs at a high school level. For instance, this hypothesis suggests that representational gestures (i.e., iconic and metaphoric)[13] are involved, not only in speaking, but also learning and problem solving. Furthermore, this hypothesis proposes four basic functions of gesture: maintaining activation; manipulation; exploration; and packaging, which share the schematic nature of gestural representation. "In all cases, information reduction plays a key role" in schematization (Kita et al., 2017, p. 257). The following sections explain empirical and hypothetical gesture studies in L1 in the gesture-for-conceptualization hypothesis that partially motivated this study to generate the research questions and hypotheses.

[13] McNeill, 1992, p. 377.

1.4.1.1 Maintaining Activation of Spatio-Motoric Representations: Co-Speech Gestures

The gesture-for-conceptualization hypothesis claims, “gestures can help maintain the activation of spatio-motoric representations that are already active, so that these representations do not decay during speaking or thinking” (Kita et al., p. 247). Willem and Hagoort (2007) contend that some co-verbal gestures has a clear semantic connection to speech when co-expressing gestures and speech, which is activation (lexical retrieval[14]) in the gesture-for-conceptualization hypothesis. Additionally, this hypothesis argues that there are two types of effects regarding activation. One of them is “gesture maintains pre-existing spatio-motoric representations (i.e., help them resist decay)” (Kita et al., 2017, p. 247), which this study discuses.

Additionally, this hypothesis is consistent with a study presented by (Goldin-Meadow et al., 2001), which suggests that generating co-speech gestures accelerate remembering vocabulary mathematical class in L1, lightening the cognitive load of speaking (Goldin-Meadow et al., 2001). Moreover, generating co-speech gestures can have an effect on language development both in L1 and bilingual education, suggesting that co-speech gestures can help students with conceptualization in certain situations (Goldin-Meadow, 2003, 2009; Mayberry & Nicoladis, 2000; Nicoladis, Pika, & Marentette, 2009). Similarly, experimental gesture studies in FL argue that teaching (co-speech) gestures helps students learn foreign vocabulary (Kelly et al., 2009; Tellier, 2008),[15] and thus, co-speech gestures can stimulate

---

[14] The concept of lexical retrieval in the gesture-for-conceptualization hypothesis differs from that of the lexical retrieval hypothesis proposed by Krauss et al. (2000) that claims that gesture and speech develop in parallel with no collaboration (cf. Kita, 2000; Kita & Ozyureck, 2003).

[15] Tellier (2008) taught illustrative gesture, but not iconic gestures. However, there is no clear difference between them. Refer to Chapter 2 in this regard. Additionally, Tellier (2008) and Macedonia et al. (2011) did

learning (Broaders et al., 2007; Golden Meadow et al., 2009; Gullberg, 2006, 2008; Kita et al., 2017). Importantly, Ozcaliskan, Lcero, and Goldin-Meadow (2016) claim that there is "no effect of language on gesture when it was produced on its own – silent gesture" (p. 10), and thus, this study does not include silent gestures in the research questions and the hypotheses, although some experimental gesture studies in FL use silent gestures in assessing participants' knowledge (Tellier, 2008).

1.4.1.2 Manipulation: Co-Thought Gesture Studies in L1

Hegarty and Waller (2006) claim, "The strongest factor, spatial visualization … was first defined by Guilford and Lacey (1947) and is typically described as an ability to mentally manipulate, rotate, twist, or invert objects without reference to one's self" (p. 127). On the basis of this argument, the gesture-for-conceptualization hypothesis suggests that gestures helps with object manipulation, by generating co-thought gestures. For instance, in Shepard and Metzler's (1971) mental rotation tasks, participants matched the stimulus object with one of the two mirrored three-dimensional objects in a non-communicative mental task by generating co-thought gestures. Chu and Kita (2012) conclude that co-thought gestures in silent problem solving enhanced participants' performance in spatial problem solving, even if participants did not verbalize how they rotated the object. Moreover, they produced co-thought gestures when a silent problem-solving task was harder than when less demanding (Chu & Kita, 2008, 2011, 2012; Kita et al., 2017). Needless to say, learning about vertical spatial structures is a highly demanding task for Japanese EFL students. Additionally,

---

not use *co-speech gestures* in assessing their participants' knowledge to express a single word, and thus, a parenthesized term (*co-speech) gestures* will be used when this study discusses their studies.

Vygotsky (1998) and contemporary researchers claim form-meaning pairings.[16] Therefore, I argue that recreating ICSCTGs alone facilitates learning about vertical axis operations with head-initial SVO word order (grammatical and semantic structures) in English by solving problems to a certain extent, even though students did not verbalize the grammatical and semantic structures.

1.4.1.3 Exploration of Spatio-Motoric Information

The gesture-for-conceptualization hypothesis claims, "four lines of evidence converge to build a case for this function of gesture" (Kita et al., 2017, p. 250), which is to explore spatio-motoric information. This study focuses on exploring spatio-motoric representation, via trial-and-errors by seeking information useful for speaking and thinking. I hypothesized that Japanese students would learn how to express vertical spatial operations through trial–and errors in formulating speech and thinking while recreating ICSCTGs.

1.4.1.4 Packaging: Co-Speech Co-Thought Gestures

Kita et al. (2017) contend, "the information packaging hypothesis holds that gesture helps speakers package spatio-motoric information into units appropriate for verbal encoding" (p. 246). Alibali et al. (2017) suggest that speakers who cannot produce gesture should have difficulty planning syntactic units that correspond to units for speech formulation (see Bock & Cutting, 1992). Therefore, gestures facilitate chunking information into units that can be coded in a clause and/ or a phrase (e.g. *on the table*). Furthermore, based on experimental studies, Kita et al. (2017) conclude that difficulty in information packaging triggers gesture. The gesture-for-conceptualization hypothesis argues that:

> gesture packages spatio-motoric representations for thinking and speaking. This idea builds on the information packaging hypothesis that has been put forward for co-

[16] Refer to section 2.3.2.6.

> speech gesture, and that has received extensive empirical support…The current proposal extends this idea to thinking more generally, encompassing findings from both co-speech and co-thought gestures (p. 259).

I predicted that Japanese EFL high school students would integrate the four functions in the gesture-for-conceptualization hypothesis by schematizing given images in expressing vertical spatial operations, which led to creating the research questions and hypotheses in this study. Thus, this hypothesis is harmonious with the Gesture Listening Higher Concept Approach with regard to reproducing gestures in Japanese EFL high school education.

1.4.2 Viewing Gestures

Of note, an effect of viewing gestures on comprehending messages "has been the topic of many research studies over the previous 35 years, and there has been little consensus" (Hostetter, 2011, p. 297). This view is shared by experts in the field of gesture research (S. Kita, personal communication, November, 20, 2016).

Furthermore, there are different cortical activations between children and adults in the processing of co-speech gestures (Dick, Goldin-Meadow, Solodkin, & Small, 2012). An age difference exists regarding an effect of viewing gestures. For instance, the non-gesture group in Tellier's (2008) study may have improved because of participants' age (M age = 5.5) without viewing and regenerating gestures. Conversely, observing gestures had an effect on learning French as a Foreign language (FFL)[17] at a college level (Quinn-Allen, 1995). Quinn-Allen (1995) used gesture alone as a classroom intervention, but not listening practice. The

[17] It seems that English speakers use both terms such as *a FL* and *an FL* in everyday life, and for this reason, both terms should be accepted. The term starts with a vowel sound needs an article *an*, but not *a* including *an FM radio station* and *an hour,* and thus, a term *an FL* is correct, although there are exceptions such as a year. I discussed this issue with Jackelyn Van Buren on October 11, 2015. There is a term *an FFL* that refers to a federal firearms license. However, I have not found a term *an FFL* that refers to French as a Foreign Language.

following section 1.4.3 includes a theoretical aspect of listening practice in Research Question 2 and Hypotheses 2 and 3 in sociocultural theory.

1.4.3 Listening Practice

The processes of listening comprehension and auditory perception in cognitive science are highly complex (Suvorov, 2007; Lotto & Holt, 2011). Thus, it is not surprising that no commonly accepted hypothesis exists regarding not only listening comprehension, but also auditory perception. Additionally, MacWhinney stated that there were very few published research studies on listening in SLA (B. MacWhinney, personal communication, October 7, 2015).[18]

Based on sociocultural theory and studies presented by contemporary researchers, this study includes a discussion of an effect of listening practice that consists of inner and overt speech. Vygotsky (1987) suggests, “the phenomenon of inner speech is fundamental to … logical forms of thinking” (p. 75). Importantly, Vygotsky (1987) suggests, “verbalizing and displaying his inner speech to see the profound similarity between the adult’s overt verbal thinking” (p. 72), and thus, inner speech relates to verbal thinking.

When learners reproduce overt speech what they heard, they copy what they did with inner speech. Overt speech often plays a role in the primacy effect on remembering (Baddeley, 2007; Vygotsky, 1997b). Additionally, Rauschecker (2012) claims that solid psychological evidence shows that listeners accurately tap along to auditory sequences.

---

[18] MacWhinney stated, “The biggest focus on this literature is on tasks such as the serial reaction task (SRT) in which people learn a sequence of tapping or rhythms implicitly. People like Darlene Howard have dozens of papers using this and similar tasks in which the items are tones instead of pictures. You can Google for “auditory serial reaction task”. Of course, these are meaningless sequences. We also learn meaningful ones as in words…there is the big literature on “statistical learning” by people like Saffran, Newport, Aslin, Thiessen etc. This is for both infants and adults…Luca Onnis has a bit of work on statistical learning for SLA, but you are right that there is not much there yet” (B. MacWhinney, personal communication, October 7, 2015).

Consequently, listening practice can help students with logical verbal thinking and remembering the sequence of word order to a certain extent in this study because of the concept of form meaning-pairing.

In short, the creation of research questions and hypotheses were motivated by gesture studies in L1, experimental gestures studies in FL and bilingualism, and a study of inner speech and overt verbal thinking in sociocultural theory. The following section 1.4.4 includes a brief explanation of how positive correlations of two different types of test scores measured by two alternative assessments influence an accumulation of knowledge of a single word meaning in a level of scientific thinking, leading to a higher concept. The higher concept is one of the central tenets of the Gesture Listening Higher Concept Approach.

1.4.4 Higher Concept

Section 1.4.4 presents the argument of an effect of a positive correlation that can lead to accumulation of knowledge that increases in a level of scientific thinking. Furthermore, I argue that scientific and categorical thinking is synonymous, which refers to a higher concept in this study.

1.4.4.1 Effect of Positive Correlation Between Two Alternative Types of Test Scores

In addition to an effect of the concurrent ICSCTGs and listening practice, of ICSCTGs alone, and of listening practice only, there should be one more effect on learning about a vertical axis structure (a single word meaning) in this study.[19] The following section includes an explanation of how an accumulation of knowledge can occur.

[19] Refer to Chapters 4 and 5 for more detail.

1.4.4.2 Accumulation of Knowledge

Memory is essential to growth of our knowledge. Vygotsky (1999) explains how memory shifted and argues, "memory is transformed from passive recording into a function[20] of active selection and active and intellectual remembering" (p. 25). He (1998) further indicates, "memory is no longer a store for single image, but is an archive of knowledge" (p. 88), stressing the crucial factor of linking memory to knowledge. He (1987) also stated that the "accumulation of knowledge leads directly to an increase in the level of scientific thinking" (p. 168).

Thus, based on sociocultural theory, one can conclude that the transformation from passive recording into the functions of active selection and intellectual remembering links to an accumulation of knowledge. This accumulation of knowledge increases in a level of scientific and categorical thinking, suggesting that memory ties with knowledge and thinking.

1.4.4.3 Scientific and Categorical Thinking

Terms such as *scientific thinking* and *categorical thinking* are synonymous in this study, and refer to a higher concept. Vygotsky (1999) claims that an aphasic patient who lost categorical thinking of color shade behaved as if they were an ape, and thus, categorical thinking is considered to be higher thinking.[21] He further suggests, "The strength of the scientific concept lies in the higher characteristics of concepts, in conscious awareness and volition" (Vygotsky, 1987, p. 220). It appears that both scientific and categorical thinking are classified as higher concepts that are higher thinking. Thinking occurs in concepts (Vygotsky, 1993).

---

[20] That is, "operating."
[21] See Chapter 2 for detail.

Accordingly, I argue that scientific thinking, categorical thinking, and a higher concept are synonymous in sociocultural theory.

In short, the Gesture Listening Higher Concept Approach, which is harmonious with the gesture-for-conceptualization hypothesis and sociocultural theory, is an experimentally and theoretically effective framework in helping Japanese EFL students learn about vertical axis relationships. Additionally, there could be one more effect on learning about vertical spatial structures that these research questions and hypotheses do not describe.

1.5 Conclusions

This chapter has clarified reasons why the research questions and hypotheses in this study are particularly essential in experimental gesture studies in FL, and how these questions and hypotheses are motivated by previous empirical, hypothetical, and theoretical work in L1, bilingualism, and FL. I address current issues in experimental gesture studies in FL and SLA. I propose that the use of the Gesture Listening Higher Concept Approach can be a potential solution of these issues in Japanese EFL education. This approach also can have a significant impact on responding to a new research path in SLA that incorporates studies in the thinking-for-speaking hypothesis into those of sociocultural theory. In the following chapters, I argue in support of the gesture-for-conceptualization hypothesis and the Gesture Listening Higher Concept Approach that can be a promising empirical and theoretical framework for helping students effectively learn about vertical spatial operation in FL classes at the high school level. Additionally, I argue that seeking the commonality between sociocultural theory and linguistic relativity can have a crucial impact on answering a new research trend in SLA.

Chapter 2

Review of the Literature

2.1 Introduction

A goal of this literature review is to help to understand the backgrounds of experiential, hypothetical, and experimental gesture studies in L1, FL (EFL, bilingualism), and studies of sociocultural theory in discussing the Gesture Listening Higher Concept Approach. This literature review uncovers a novel effect of a higher concept in experimental gesture studies in FL. Additionally, this literature review reinforces Hypothesis 1which states that the use of ICSCTG alone will facilitate learning about vertical axis structures in a Japanese EFL high school class. Studies presented by gesture researchers in L1 and FL experimentally support Hypothesis 1 and evidence presented by gesture researchers in L1 hypothetically support Hypothesis 1 as well. Furthermore, sociocultural theory and gesture studies in L1 and FL together respond to the new research direction in SLA by supporting Hypotheses 2 and 3 in this study. Responding the new research path leads to explaining three different components in the Gesture Listening Higher Concept Approach.

2.2 Experiential Background

The present study draws on my teaching experience to investigate the effects of concurrent ICSCTGs and listening practice, and a higher concept. As explained earlier, one problem my former American high school students in a JFL class encountered was the substantial typological differences between the two languages in syntax and semantics. Restructuring linguistic knowledge in English to that of JFL lessened the problem by re-conceptualizing students' habitual thought in English to that of JFL with the Gesture Listening Higher Concept Approach. I will incorporate my experiences as an EFL teacher in Japan, as a JFL instructor in the United States, and as a researcher in the United States and Canada into Japanese EFL high school education to investigate the effect of this approach.

2.3 Theoretical and Experimental Backgrounds

The following discussion allows one to understand the commonality between sociocultural theory and linguistic relativity regarding relationships between language and thought, and of the Gesture Listening Higher Concept Approach. In this way, one of the goals stated in Chapter 1 can be accomplished theoretically and experimentally.

2.3.1 Linguistic Relativity, Verbal Thinking, Development of Meaning of Words in FL

In the following, I present arguments of how a study of linguistic relativity can be incorporated into SLA in sociocultural theory, and then follow with an explanation of how the meaning of words develops in an FL class. In subsequent sections, I explain theoretical, hypothetical, and experimental backgrounds of the Gesture Listening Higher Concept Approach.

Verbal thinking that ties in meaning of words is sociocultural in origin (Vygotsky, 1987), indicating that the dissimilarity of social beings initiates distinct verbal thinking patterns. Acquiring developed meaning of words in FL tied with developed verbal thinking in FL education is central in attaining a different worldview from that of L1 to help students successfully learn about vertical spatial operations in a certain type of FL class. Sociocultural theory argues, “verbal thinking is not a natural but a socio-historical form of behavior” (Vygotsky, 1987, p. 120), which is an equivalent claim to that of linguistic relativity (TFS).

Linguistic relativity argues that language shapes conceptual categories or patterns of thought, which involves a direct influence of language on thought and action (Boas 1911; Slobin 1996). Similarly, Whorf (1956) suggests that:

> Human beings do not live in the objective world alone, nor alone in the world of social activity as ordinarily understood, but are very much at the mercy of the particular language which has become the medium of expression for their society (p. 134).

In other words, language is a social existence specific to each speech community by influencing the thought of individuals in the community. Learning about vertical axis operations in English can influence Japanese EFL students' verbal thinking that needs to shift the meaning of words from the single to binary semantic categorization to fit in English speech communities. For this reason, sociocultural theory and linguistic relativity (TFS) agree that the language that one speaks can have an effect on the way in which one thinks.

Based on studies by Kunihiro (1986), Munnich et al. (2001), Munnich and Landau (2003), Vygotsky (1987), and Tanaka (1997), I propose the concept of the single and binary semantic categorization, which leads to the understanding of developed meaning of words in EFL. For instance, English speakers structure vertical axis operations with an obligatory contrast (the binary semantic categorization) as in Munnich and Landau (2003, p. 29), whereas Japanese conceptualize vertical spatial structures with a non-obligatory distinction between contact and noncontact (the single semantic categorization). Students whose L1 has the non-obligatory contrast must shift their thinking from the single to binary semantic categorization when they learn vertical axis operations in an EFL class, leading to development of the meaning of words in FL (cf. Kunisawa, under review).

This study incorporates linguistic relativity into SLA in sociocultural theory in a methodological design, too.[22] For instance, Berman and Slobin (1994) used a picture book, namely *Frog, Where Are You?* (Mayer, 1969) to examine how children in different countries

---

[22] See Chapter 3.

explained stories about the same sequence of images. Berman and Slobin (1994) compared descriptions of several scenes expressed by those children whose L1 are English, German, Spanish, and Hebrew. Their focus was on expressions of temporal and spatial relationships. Participants in their experiment were preschool (three-five years), school age (nine years) children, and adults (Berman & Slobin, 1994). Their finding was that even preschoolers demonstrated evidence of language-specific patterns of TFS in each language (Berman & Slobin, 1994).

The current study employs a similar experimental research design to that of TFS, including examining how Japanese EFL students restructure their worldview from Japanese to English in expressing vertical spatial structures at a lexicon-grammar level (a prepositional phrase). However, this study did not employ narratives in German, Spanish, and Hebrew, but dialogues in English on the basis of sociocultural theory.[23]

Concisely, verbal thinking has sociocultural origins as in sociocultural theory, which leads to a dissimilar conceptualization of vertical spatial operations between Japanese and English. In other words, English shapes speakers' thought with the obligatory contrast, whereas Japanese structures speakers' thought with non-obligatory contract. Accordingly, one can argue that languages influence thought, and thus, sociocultural theory and TFS agree that language can shape thought, which is one of the three goals in this study. For this reason, the experimental method in the present study is the incorporation of sociocultural theory and TFS, leading to theoretically and experimentally accomplishing this goal.

[23] See Chapter 3.

2.3.2 Gesture Listening Higher Concept Approach

In the following, I make arguments for the central tenants of the Gesture Listening Higher Concept Approach, which are (a) a higher concept built upon sociocultural views of memory, knowledge, and scientific and categorical thinking; (b) listening practice based on sociocultural theory; (c) gesture that sociocultural theory has commonalities with hypotheses and theories presented by gesture researchers in L1; and (d) the integration of regenerating ICSCTGs and listening practice. The following includes a discussion of a higher concept first, and then of listening practice. Subsequent sections reveal an effect of ICSCTGs and address the issue of the integrating the regeneration of ICSCTGs with listening practice.

2.3.2.1 Higher Concept

As introduced earlier, a higher concept is defined as higher thinking that arises as a result of accumulation of knowledge in a level of scientific and categorical thinking, and remembering. In the following, I present the experimental and theoretical analysis of relationships among memory, knowledge, and scientific and categorical thinking to understand how a higher concept emerges. As a result, a higher concept helps students learn about vertical spatial operations. The following sections include an explanation of the experimental analysis and then follow with discussion about theoretical analysis of the relationships, based on sociocultural theory.

2.3.2.1.1 Experimental Evidence

2.3.2.1.1.1 Tellier’s (2008) Study

The non-gesture group in Tellier (2008) did not view gestures during instruction of English words in class. Nonetheless, a graphical analysis suggests that participants’ knowledge of how to express those words improved between Assessment 1 and 3b (a passive recording) (cf. Figure 1).[24] Their test scores in an active selection[25] also increased between Assessment 2 (3rd week) and 3a (4th week).[26] Graphical data analyses in Figure 1 indicated that participants' test scores of the words in both gesture and non-gesture groups are positively correlated. Unwin (2016) suggests, "You can't prove anything, but you can disprove some things and the same goes for graphics … Use graphics to discover information that is difficult to investigate statistically" (p. 276-7). This positive correlation led to active and intellectual remembering of English words by transforming a passive recording (a silent task of pairing a word with nonverbal information) into an active selection (verbalizing a word matching to nonverbal information that was gesture). Vygotsky (1999) contends, "memory is transformed from passive recording into a function of active selection and active and intellectual remembering" (p. 25).

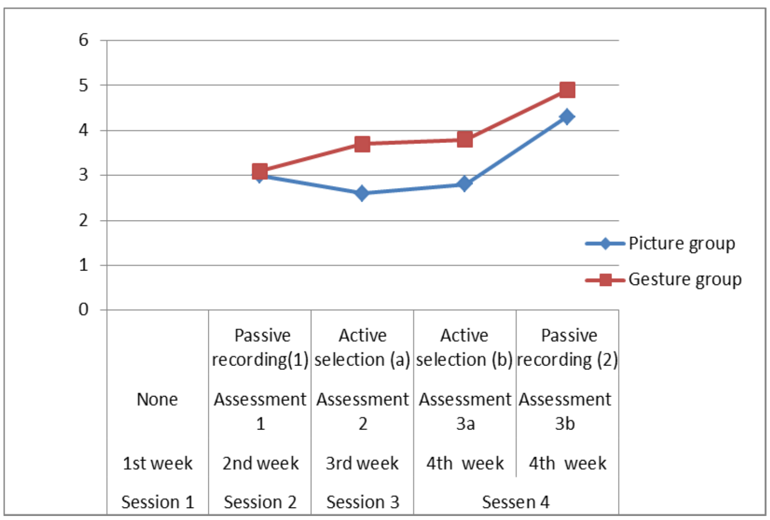


Figure 1. Positive Correlation between Two Different Types of Test Scores Which were Assessed by Two Alternative Assessments. Based on data presented by Tellier (2008, p.228), the author generated Figure 1. Telllier (2008) assessed participants' knowledge by using both

---

[24] When an investigator verbalized an English word, participants showed a picture card or perform a gesture, which corresponds to the English word.
[25] When the investigator showed a picture card or made a gesture, which corresponded to an English word, they verbalized the English word.
[26] Their test score in active selection of English word decreased between Assessments 1 (2nd week) and 2 (3rd week) (cf. Figure 1).

passive recording (passive knowledge in Tellier's term) and active selection (passive knowledge in Tellier's term).

- **The red line in figure 1 demonstrates the improvement of the gesture group:** The passive recording was done by asking participants to create gestures when they heard an English word produced by a video on a computer. Participants performed the active selection when they were asked to verbalize an English word corresponding to a gesture on a computer screen.
- **The blue line in figure 1 shows the improvement of the picture group:** The passive recording was done by asking participants to show a picture that corresponds to an English word when they heard the English word, whereas participants performed the active selection when they were asked to verbalize an English word that corresponds to a picture shown by the investigator.

As explained earlier, "memory is an archive of knowledge and not a storehouse of images" (Vygotsky, 1998, p. 98), and thus, knowledge and remembering of English words co-evolved in Tellier (2008). However, why was the gesture group superior to the non-gesture group? Co-speech gestures which Tellier (2008) used during class may have created a higher concept in gesture groups' mind. Vygotsky (1987) argues, "the motor processes associated with speech play an important role in facilitating the thinking process in particular, in improving the subject's understanding of difficult verbal material" (p. 44). Additionally, Vygotsky (1987) concludes, "the accumulation of knowledge leads directly to an increase in the level of scientific thinking and that this, in turn, influences the development of spontaneous thinking" (p. 168). Co-speech gestures may have generated a higher concept than non-gesture did by assisting the gesture group with thinking processes, leading to understanding difficult verbal materials in a higher level of scientific thinking than that of non-gestures. Accordingly, I argue that the accumulation of knowledge of English words in Tellier (2008), created by the positive correlation, led to an increase in students' level of scientific thinking (a higher concept) as Vygotsky suggests.

2.3.2.1.1.2 Macedonia et al.'s (2011) Study

The arguments proposed by Vygotsky (1987) explains above can also be applicable to findings in Macedonia et al. (2011). For instance, Figure 2 suggested that participants' remembering of artificial words learned by iconic gestures transformed from a passive recording into a function of an active selection and active and intellectual remembering when students assessed participants' knowledge of artificial words for 5 days.

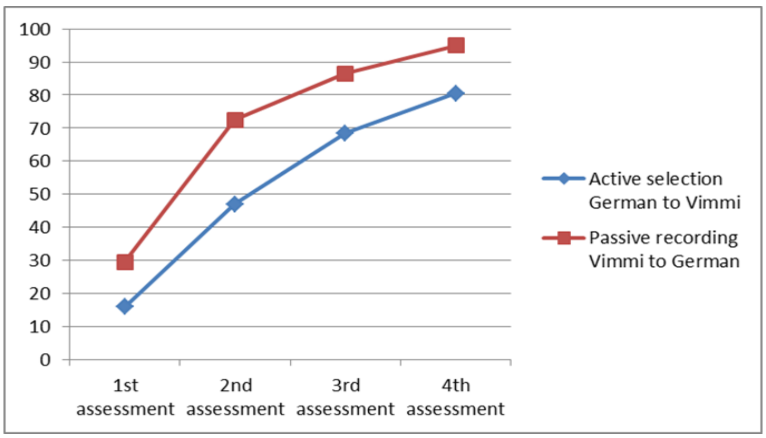


Figure 2. Results of 4 Day Training in Artificial Words Learned by Iconic Gestures. The blue line in Figure 2 refers to the training results of the written translation test from German into artificial words (Vimmi), whereas the red line is for the training results of written translation test from artificial words (Vimmi) into German.

In Macedonia et al. (2011), participants' remembered artificial words learned by iconic gestures when they translated Vimmi (artificial words) into their L1 (German) and vice versa (an active selection). Thus, Macedonia et al.'s (2011) findings suggest that accumulation of knowledge of artificial words obtained by the positive correlation led to an increase in their participants' level of scientific thinking (a higher concept) as Vygotsky claims.

2.3.2.1.1.3 No Positive Correlation

Conversely, Quinn-Allen (1995) employed one type of the assessment tool of translating 50 French expressions into English. Ebbinghaus (1885/1913) proposed the forgetting curve hypothesis and then Heller et al. (1991), and Murre and Dros (2015) replicated Ebbinghaus' classic forgetting curve and found their results similar to those of Ebbinghaus' original data.[27] Ebbinghaus (1885/1913), Heller et al. (1991), and Murre and Dros (2015) used the saving measure of the eight rows of 13 one-syllable nonsense words in their experiments.[28] The use of one type of an evaluation of participants' performance led to a lack of correlation of two different types of test scores in assessing their knowledge in the two studies noted above. The purpose of the forgetting curve hypothesis, however, was to measure participants' forgetting (not accumulation of knowledge), but Quinn-Allen did not intend to assess participants' forgetting.

Accordingly, significant memory decay resulted. For instance, of the students' knowledge in the 50 French expressions (60.5% to 9.1% in 77 days on average) in Quinn-Allen (1995) and of the eight rows of 13 one-syllable words (51% to 17.78% in 31 days on average) resulted.[29] These results are the consequences of the lack of two different assessment methods, leading to the absence of interaction of passive recording and active

[27] Ebbinghaus, who is an originator of the forgetting curve hypothesis, tested his own memory of nonsense syllables and then plotted the results (Baddeley et al., 2007). "This study [a study presented by Heller et al. (1991)] has been published only in German, without an English abstract, and is not easily accessible; at the time of writing, it is not available in electronic format (it is not available online) and it has never been cited in international journals in English. It is, however, a thorough study and an excellent replication attempt. Where the Ebbinghaus … texts are unclear about certain detail, we have mostly followed Heller et al. … as a guideline so that we can also compare our results with theirs " (Murre & Dros, 2015, p. 3).

[28] "Savings is defined as the relative amount of time saved on the second learning trial as a result of having had the first. Suppose, one has to repeat a list for 25 times in order to reach twice perfect recollection and that after one day, one needs 20 repetitions to relearn it. This is 5 less than the original 25; we can say that on relearning we saved 20% with respect to the original 25 rehearsals (5/25 = 0.2 or 20%)" Murre and Dros (2015, p. 3).

[29] Refer to Table 18 and Figure 7 for Quinn-Allen (1995) and Table 17 and Figure 9 for the forgetting curve hypothesis.

recording in remembering expressions and one-syllable words. Consequently, there was a limited accumulation of knowledge of those expressions and words.[30] The inevitable consequences were that participants in their studies may not have had a limited ability in generating a higher concept in learning about those expressions and words. As a result, the use of two different types of assessment tools is essential in the growth of knowledge, which increases in a level of scientific thinking.

2.3.2.1.2 Theoretical Framework

A positive correlation between memory and accumulation of knowledge are two components in the theoretical framework of a higher concept in this study. I argue that a higher concept emerges when students learn a developed meaning of words in EFL by using two different types of assessment tools at minimum to increase the accumulation of knowledge. Moreover, accumulation of knowledge leads directly to an increase in the level of scientific and categorical thinking (semantic and grammatical structures). As a result, effective learning can take place. The following sections include discussion of these two components to understand how a higher concept emerges in this study.

2.3.2.1.2.1 Positive Correlation with Memory

As stated earlier, Vygotsky (1999) argues that memory transforms from passive recording into a function of active selection and active and intellectual remembering. Thus, the positive correlations between two different types of test scores in studies presented by Macedonia et al. (2011) and Tellier (2008) explained in this literature review can be understood as the transformation of participants' memory of a single word meaning (cf. Figures 1 and 2). As

[30] Other factors can be involved in memory decay. However, one of the causes of memory decay could have been the use of a single assessment tool, although more research is required on possible contributing factors such as untranslatable idiomatic expressions in French causing attrition of the expressions (Tellier, 2008).

stated earlier, Vygotsky (1998) argues that memory links an archive of knowledge, but not a storehouse of images, suggesting that the transformation of participants' memory can lead to accumulation of knowledge.

As argued earlier, Vygotsky (1987) contends that the accumulation of knowledge leads to a growth in the level of scientific thinking and that this, in turn, effects the development of thinking. While the transformation takes place, the creation of higher-level scientific thinking is a key for growth of knowledge.

2.3.2.1.2.2 Accumulation of Knowledge

Vygotsky (1987) also states that:

> The development of scientific concepts begins with the verbal definition. As part of an organized system, this verbal definition descends to the concrete; it descends to the phenomena which the concept represents (p. 168).

Thinking occurs in concepts (Vygotsky, 1993), which leads to defining the development of scientific thinking. Verbal definition is crucial, which allows one to understand the representation of a concept. In this study, the verbal definition of scientific thinking refers to the concepts of the form-meaning couplings (The SVO word order with meaning), of the single and binary semantic categorization, and of TRHLM-TRLLM distinctions as in categorical thinking. Moreover, Vygotsky (1998) claims that categorical thinking refers to higher thinking stating that:

> [an aphasic patient] was able to select precisely the same shade of red, but could not select red of a different shade and behaved much more concretely than a normal person…Thus, the genetic key to understanding the behavior of the aphasic is the assertion that his categorical thinking is disrupted and yields to a more primitive, genetically earlier stage of concrete-visual relation to reality (pp. 129-132).

Categorical thinking that contains meaning entails a higher concept as the example noted above. Vygotsky (1999) also claims that:

> the methods of carrying out a task by the aphasic and the human-like ape are similar on the whole and coincide in the most essential points. This fact thus confirms our idea that speech plays an important role in the organization of higher mental functions (p. 17)

Therefore, categorical thinking can take place on the basis of a higher concept accompanied by speech and meaning. Vygotsky (1987) further suggests, “The strength of the scientific concept lies in the higher characteristics of concepts, in conscious awareness and volition” (pp. 220),[31] and thus, both scientific and categorical thinking include classification as a higher concept accompanied by word meaning.

Similarly, the increase of knowledge of word meaning can be central in a level of scientific and categorical thinking (Vygotsky, 1987), leading to the synonymous use of scientific and categorical thinking. In other words, participants in this study can acquire the ability to express vertical spatial operations according to their abilities of scientific and categorical thinking. For example, in this study participants who understand the concept of the TRHLM-TRLLM distinction, of SVO word order, and of the transformation of vertical axis semantic categorization from the single to binary and who utilize these concepts can gain a higher concept. Additionally, the unique form of cooperation between an instructor and students can lead to the growth of knowledge in a level of scientific and categorical

---

[31] Vygotsky (1987) criticizes the view of scientific thinking in children proposed by Piaget and stated that:

> The problem of causality is definitive for the logic of Piaget’s scientific thinking…this new social research led Piaget to the conclusion that causality, in the true sense of the word, is not present in the child’s representation of the world … ”(p. 80).

Vygotsky (1987) contends that:

> The development of the scientific social science concept … it (The maturation of the child’s higher mental functions) occurs through the adult’s assistance and participation. In the domain of interest to us, this is expressed in the growth of the relativeness of causal thinking and in the development of a certain degree of voluntary control in scientific thinking. This element of voluntary control is a product of the instructional process itself … This is also why the level of development of scientific concepts forms a zone of proximal possibilities for the development of everyday concepts (p. 169).

Therefore, scientific thinking that children will learn can assist them with solving problems in a proximal zone by obtaining help from more capable adults.

thinking because in this process, an instructor's knowledge is transferred to the students (Vygotsky, 1987). Consequently, a higher concept helps students with growth of knowledge when they learn EFL.

2.3.2.2 Listening

A structure of listening is one of the central tenets in the Gesture Listening Higher Concept Approach, however, the procedures of listening comprehension and auditory perception in cognitive science are highly intricate (Lotto & Holt, 2011; Suvorov, 2007). For this reason, it is not surprising that no commonly accepted hypothesis exists regarding not only listening comprehension, but also auditory perception. As explained in Chapter 1, MacWhinney (2015) suggested that very few published research studies exist on listening in SLA regarding how students learn word order, although there are studies in L1 regarding the serial reaction task (SRT) (B. MacWhinney, personal communication, October 7, 2015). Thus, on the basis of studies of sociocultural theory, this study includes theoretical discussion of how listening practice can help students learn word order in English.

The salience of listening practice lies in both inner and external speech, which can occur during listening practice in this study. Ushakova (1994) stated that inner speech inherently developed from listening (p. 140, de Guerrero, 2006; Ellis, 2001, Sokolov, 1972). Vygotsky (1987) states, “Inner speech is speech for oneself. External speech is speech for others. This is a fundamental functional difference in the two types of speech that will have inevitable structural consequences” (p. 257).

Inner speech takes place not only in sensory and motor processes,[32]but also in reading. On the basis of studies of dysarthric patients and anarthria[33]with silent reading tasks, Baddeley (1997) claims that inner speech plays a role in the development of cognition. In the current study, I discuss inner speech in listening practice only, but not inner speech in sensory motor processes and silent reading tasks.

2.3.2.2.1 Inner Speech and Verbal Thinking

Vygotsky (1987) argues, “inner speech is an internal plane of verbal thinking which mediates the dynamic relationship between thought and word” (p. 279). Additionally, Vygotsky (1987) indicates that as inner speech develops, verbal thinking also develops with the development of speech and intellect. John-Steiner (2007) contends that:

> Vygotsky’s strong emphasis on the centrality of inner speech in thought is paralleled by some influential linguists such as Benjamin Whorf and Edward Sapir. They have written extensively about the reciprocal relationship between language and thought (p. 143).

That is, the development of inner speech can lead to the formation of novel concepts of vertical spatial operations, such as the binary semantic categorization in English, which is distinct worldview. Importantly, Vygotsky (1987) suggests that:

> … the thesis that inner speech is an entirely unique, independent, and distinctive speech function, that it is completely different from external speech. This justifies the view that inner speech is an internal plane of verbal thinking which mediates the dynamic relationship between thought and word (p. 279).

More importantly, Vygotsky (1987) states that:

> … a thesis of great methodological significance for the correct statement of the problem of thinking and speech. This thesis stems from our comparison of the development of inner speech and verbal thinking in man with the development of

[32]Sokolov (1972) stated that listening and reading texts associated with inner speech. Vygotsky (1987) argues, “sensory and motor processes do indeed have a subordinate significance in inner speech” (p. 256).
[33]Patients with anarthria have characteristics of loss of speech.

> speech and intellect … The very type of development changes. It changes from a biological form of development to a socio-historical form of development (p. 120).

That is, as inner speech develops, verbal thinking can also develop with the development of speech and intellect, which can lead to helping students with obtaining a different worldview from that of their L1.

Vygotsky (1987) also notes that inner speech is fundamental to logical thinking, remembering, and verbal thinking. It is important for listeners to speak for themselves in generating inner speech by repeating expressions after they listen to them. This develops meaning of words in FL in the auditory materials. In this way, verbal thinking of developed meaning of words in EFL could further develop in their mind, while remembering develops to some degree. As a result, students could maintain knowledge and further enhance their ability to speak in a level of scientific and categorical thinking in a Japanese EFL high school class to a certain extent.

Furthermore, vertical spatial relationships should be logically taught on the basis studies of linguistics, gesture, TFS, and sociocultural theory, including a single and binary semantic categorization with SVO word order prior to listening practice. Likewise, Vygotsky (1987) states, “It was also found that inner speech plays a role as a facilitating factor in the transition from thought to overt speech”[34] (p. 44). Vygotsky also (1987) suggests that:

> … the more closely inner speech is connected with external speech in behavior, the more similar they become. For example, inner speech may take a form identical with external speech[35] when it occurs in the preparation of an upcoming speech (p. 115).

[34]Vygotsky (1987) argues, “To state the problem more directly, there is no evidence that the transition from external overt speech to inner covert speech moves through the whisper” (p. 112). That is, terms *overt* and *external speech are* interchangeably used (Vygotsky, 2001).
[35] Refer to footnote 20 noted above.

That is, after students speak for themselves with inner speech while listening to auditory materials, they copy their actions by reproducing overt speech while internalizing inner speech (verbal thinking). For this reason, qualities of inner speech are crucial in listening practice.

2.3.2.2.2 Repetition with External Speech

Listening practice in this study involve verbally repeating expressions in audio material accompanied by verbal thinking (external speech). Accordingly, “Learning requires frequent repetition” (Vygotsky, 1987, p. 210). Vygotsky (1997b) suggests, “one must have certain experience in remembering. It has been observed in experiments how this remembering occurs, and it is understandable that it grows stronger with repetition” (p. 116).

Baddeley (2007) suggests that overt rehearsal[36] (overt articulation or overt vocalization) often plays an important role in remembering, and claims, “most subjects in free recall are to rely on semantic rather than phonological coding” (p. 113). For this reason, this study focuses on spatial semantics, but not on phonological coding. Additionally, Rauschecker (2012) concludes that robust psychological evidence demonstrates that listeners precisely tap along to auditory sequences (p. 31).

Consequently, when listeners were finger tapping along to auditory sequences, participants in this study were sequencing the SVO word order in learning about vertical spatial relationships in a Japanese EFL class by re-conceptualizing vertical spatial operations. As stated earlier, word order is crucial to express proper vertical axis structures in English.

[36] In six volumes of the Collected Works of L. S. Vygotsky published between 1987 and 1999, Vygotsky does not use a term *rehearsal* that has been used in studies of the Psychology of Memory.

Properly expressed word order leads to learning about the developed meaning of the words in English that express a proper vertical axis relationship between two referents.

2.3.2.3 Gestures

Kita (2015) states, "in the last several years, research on the role of gesture in language development has been flourishing."[37] Moreover, Paggio (2012) points out that there is "a growing interest for gesture studies" (p. 281). To explore the hypotheses and the research questions in this study, I present the literature review in experimental gesture research in FL and bilingualism and then follow with an explanation of the gesture-for-conceptualization hypothesis proposed by gesture researchers on the basis of summary of gesture studies in L1.

2.3.2.3.1 Experimental Gesture Studies in FL and Bilingual Education

On the basis of prior studies, I examine the following three different issues in this section: (a) to what degree ICSCTGs alone can facilitate learning a single word meaning; (b) whether an experiment presented by Tellier (2008) shows a positive correlation in her data, and (c) whether viewing gestures can help students learn languages.

The first issue is that despite Moskowitz (1976) reporting that FL teachers received higher evaluation from their students when utilizing nonverbal communication skills, a surprising scarcity of experimental gesture studies exist in FL that investigate the use of co-speech gestures in teaching L2 (Kelly et al., 2009; Sueyoshi & Hardison, 2005; Tellier, 2008).[38] The following addresses another issue and examine two areas of studies and in experimental gesture studies in FL based on studies presented by Tellier (2008) and Macedonia et al. (2011).

---

[37] An email sent to the members of the International Society of Gesture Studies on March 4, 2015.
[38] Refer to Block (2014), Gullberg (2010), McCafferty and Stam (2008), and Salvato (2015) for summaries of gesture studies in SLA.

2.3.2.3.1.1 Illustrative Gestures[39] Accompanied by Two Alternative Assessments without Listening Training

Tellier (2008) indicates that teaching gestures has an effect on learning a single word meaning in a French EFL pre-school class. Tellier's (2008) study leads to[40] partially answers Research Question 1 and somewhat supports Hypothesis 1 in the post-oral speech setting at the pre-school level because of the lack of co-speech gestures in Tellier's assessments and participants' age difference between the two studies. Nonetheless, Research Question 1 will be fully responded and Hypothesis 1 will be completely supported by incorporating findings in Tellier's study (2008) into outcomes offered by Mayberry and Nicoladis (2000) with regard to co-speech gestures.[41]

The second issue addressed in this section is that Tellier (2008) did not ask participants to reproduce (co-speech) gestures in assessing their active recording and passive selection, which led to the lack of the concept of the instruction-assessment dualism proposed by sociocultural theorists. That is, the zone of proximal development (scaffolding) is the key concept in learning that includes an assessment inasmuch as students can improve while assessments take place (cf. later in this chapter to be discussed shortly). According to Ozcaliskan, Lucero, and Goldin-Meadow (2016) there is "no effect of language on gesture when it was produced on its own – silent gesture" (p. 10). Thus, Tellier's (2008) study lacks

[39]Refer to a later chapter for the illustrative gesture.

[40] Tellier did not ask participants to use co-speech gestures in assessing participants' active knowledge. Additionally, Tellier did not use a term *co-speech gestures* in the study even though participants used co-speech gestures during class, but not in assessments and therefore, this study includes the apprentices in describing the term.

[41] Refer to Chapter 5 for my arguments on the age difference in experimental gesture study in FL.

the concept of the instruction-assessment dualism proposed by sociocultural theorists because she uses a silent gesture that gesture researchers in L1 have criticized.

It appears that Mayberry and Nicoladis (2000) employed the concept of the instruction-assessment dualism by using co-speech gestures in assessing participants' knowledge, even though they do not use this term. However, one cannot observe a positive correlation in the data presented by Mayberry and Nicoladis (2000), which leads to the argument that both studies complement each other. The results in Tellier's (2008) study are comparable with those of this study. For instance, Tellier's (2008) data suggested that no statistically significant difference in group means between the gesture and the non-gesture groups existed regarding passive knowledge. However, as discussed earlier, a positive correlation between two different types of test scores measured by two alternative assessments (passive and active knowledge) was observed at a preschool level as well (cf. Tellier, 2008, Figure 1).

2.3.2.3.1.1.1 Brief Overview of Procedures in Tellier's (2008) Study

Tellier (2008) examines the effect of teaching gestures on learning eight English words with French pre-school children (Mean age 5.5). Half of the group (N = 10) learned the words with pictures (the non-gesture group) and the other half (N=10) included teaching strategies by using regenerating illustrative gestures (the gesture group). All participants repeated the words when they listened to them briefly. Tellier's finding was that reproducing gestures significantly influenced the memorization of meaning of words in a French EFL pre-school class.

(a) Structure of Data Collection:

One session took place for 1 hour once a week during 4 weeks. In Session 1, Tellier (2008) taught students the words (‘house’, ‘swim’, ‘cry’, ‘snake’, ‘book’, ‘rabbit’, ‘scissors’ and ‘finger’) without articles.[42] Gestures and pictures were pre-recorded in a computer. The picture group viewed eight different pictures corresponding to the words; the gesture group observed eight alternative gestures matching to each of the words on a computer screen, respectively. A video clip demonstrated a gestures, pronunciation of a word, and a picture equivalent to each word, according to their group identities, which took place in silence.

During the first three sessions, all students heard and repeated each word exactly five times to ensure that they received the same input and then the following classroom activities took place.

(i) The gesture group reproduced gestures and verbally repeated the corresponding to words while viewing the gestures (co-speech gestures).

(ii) The picture group repeated the matching to words while watching the pictures of the words.

(b) Assessments in Sessions 2, 3, and 4

In Session 2, the first assessment:[43] All participants heard the words in a different order from that of Session1 and had to show an associated picture or gesture depending on their group identities (assessing their passive knowledge in the gesture group).[44]

---

[42] Tellier (2008) taught only a small number of words based on studies of Dempsey (1981) and Gaonac’h (1995) in a preschool class.

[43]Tellier (2008) used the term *gesture*, but not a term *co-speech* gestures for her assessments and thus I employ the same term *gestures* as that of Tellier (2008) in explaining her assessments, but not a term *co-speech gestures.*

[44] Tellier (2008) did not offer any discussion about what is *active and passive knowledge*. However, to explain data presented by her, I utilize the terms *active and passive knowledge* to avoid any confusion in this chapter,

In Session 3, the second assessment: All participants saw pictures or gestures for the words according to their group identity and had to verbalize the corresponding to English words (evaluating their active knowledge). She did not ask the gesture group to regenerate (co-speech) gestures.

In Session 4, the third assessments: Tellier measured their passive and active knowledge as in the first and the second assessments without the rehearsal or warming up. Tellier did not ask the gesture group to regenerate (co-speech) gestures.

2.3.2.3.1.1.2 Similarities between the Two Studies

(a) Effect of Re-Generating (Co-Speech) Gestures:

Tellier's (2008) study includes the gesture and the non-gesture groups. Additionally, these groups in the two studies receive regular repetition practices that usually take place in FL class, but Tellier did not give any assignment of listening practice to the participants in the entire data collection processes.

In Tellier's study (2008), the gesture group (M = 4.7) outperformed the non-gesture group (M = 3.6) in learning English words with statistical significance [t (18) = -2.433, *p* <.0256] in Assessment 3a.[45] However, Tellier (2008) did not use co-speech gestures in the assessments. Thus, Tellier's study partially answered Research Question 1 and somewhat supported Hypothesis 1.

although based on sociocultural theory, Chapter 5 will use different terms from those of Tellier's (2008) in this chapter.

[45] Gesture researchers emphasized the importance of teaching gestures in L1, bilingual, FL education, and SLA (Goldin-Meadow, 2001, 2003, 2009; Gullberg & De Bot, 2010; Gullberg, 2014; Kelly et al., 1999, 2009; Macedonia, Muller, & Friederici (2011); Macedonia, 2014; Mayberry & Nicoladis, 2000; McCafferty & Stam, 2008; Salvato, 2015).

(b) No Mean Difference:

Children's test scores in measuring their passive knowledge of Tellier's (2008) study had no statistically significant difference in group means between the gesture and the non-gesture groups regarding passive knowledge (See Table 2).

Table 2:

Summary of Assessments in Tellier (2008)

	Session 1	Session 2	Session 3	Session 4	
	1st week	2nd week	3rd week	4th week	4th week
	None	Assessment 1	Assessment 2	Assessment 3a	Assessment 3b
		(passive)	(active)	(active)	(passive)
Picture group		3.0	2.6	2.8	4.3
Gesture group		3.1	3.7	3.8	4.9
Difference		x	t(18) = −2.108	t(18) = −2.433	t(18) = −1,579
(t-tests)		x	p < .0493*	p < .0256*	p < .1318

Note. The author modified Table 2 based on data presented by Tellier (2008, p. 228).
Source: Tellier, M. (2008). The effect of gestures on second language memorisation by young children. *GESTURE 8*(2): 219-235.

(c) Effect of Time:

Figure 1suggests that the picture/the non-gesture (control) group improved their memory of the words between the second and the fourth week in passive knowledge, despite no-observation of gestures.[46]

(d) Gesture Type

Tellier (2008) used illustrative gestures in her study while the current study employs iconic gestures. The definition of illustrative gesture is unclear even though Vygotsky (1987) explains, "illustrative gesture is required to further define the object relatedness of a

[46]Tellier (2008) did not assess participants' knowledge with co-speech gestures in Assessments 2 and 3a. Additionally, children's' knowledge did not follow the forgetting curve proposed by the forgetting curve hypothesis (1885/1931, 1991, 2015), even though outcomes in Quinn-Allen's (1995) study followed the forgetting curve to a certain degree.

particular meaning” (p. 155). Tellier (2008), however, does not define illustrative gesture.

Feyereisen suggests that:

> Terminology used for gesture classification is not fixed and varies depending on the authors. I think that "illustrative" gesture has a more extensive meaning than "iconic". It encompasses iconic and metaphoric gestures (which differ depending on the concrete/abstract nature of the referent). It is synonymous to "representational" or "ideational" or "lexical" gesture, as opposed to meaningless beat gestures (P. Feyereisen, personal communication, December, 8, 2016).

There was no significant difference between illustrative and iconic gestures, according to Feyereisen (2016), and therefore, the difference can be ignored. Refer to Feyereisen and de Lannoy (1991, p. 12).

2.3.2.3.1.1.3 Differences Between Two studies

(a) Instruction-Assessment Dualism

Tellier (2008) did not ask children to reproduce co-speech gestures in measuring their active knowledge. However, this study employs co-speech gestures in assessing students’ knowledge in oral speech conditions. If participants in Tellier’s study would have regenerated co-speech gestures in assessments, the gesture group in Tellier’s study (2008) could have obtained higher scores than the current ones (Chu & Kita, 2012; Mayberry & Nicoladis, 2000). As stated earlier, the concept of the instruction-assessment dualism is crucial inasmuch as students can improve during assessment processes (cf., Daniels, 2007; Hamers & Resing, 1992; Poehner, 2008). That is, the zone of proximal development (scaffolding) is the key concept in learning that includes an assessment. Vygotsky (1987) states that:

> We have seen that instruction and development do not coincide…Instruction is only useful when it moves ahead of development. When it does, it impels or wakens a whole series of functions that are in a stage of maturation lying in the zone of

> proximal development. This is the major role of instruction in development. This is what distinguishes the instruction of the child from the training of animals (p. 212).

An assessment is considered to be one of the instructional forms, which can be ahead of students' development. Ratner (1998) suggests that:

> Vygotsky recommended that psychological assessment of higher functions be conducted in a social interaction between the subject and tester. In such an occasion, the tester can prompt the subject, demonstrate examples of what he is to do, and ask questions about the reasons for his action (xiii).

It appears that instruction and assessments should take place in the zone of proximal development, and thus if co-speech gestures are taught, co-speech gestures should be also employed in assessing participants' knowledge. In this way, participants' ability can also improve during assessments.

(b) Age Difference:

Participants' age differs between the two studies. For instance, the mean age is 5.5 in Tellier (2008) and 17.15 in the present study. Consequently, an age difference between the two studies can influence the outcomes of two experimental studies. Vygotsky (1987) argues that:

> The practice of cultivating several languages in the child at an early age has shown that the mastery of two or three languages does not slow language learning…I mention this only to illustrate how sharp the young child's memory is. It cannot even be compared with the memory of the adolescent or adult (pp.306-307).

A students' neuroanatomical pattern for co-speech gestures in the present study falls somewhere between children and adults. Their neuroanatomical pattern for co-speech gestures in Tellier's study (2008) can be categorized as a child, although their age in Tellier's study (2008) is younger than that of in Dick et al. (2012).

Dick et al. (2012) concludes that the age difference between participants (seven females, range = 8-11 years, M age = 9.5 years /12 females, range = 18-38; M age = 23.0

years) resulted in dissimilar processing of co-speech gestures and self-adaptors in cortex.[47] They gave a task of viewing and listening to a story accompanied by hand movements to participants and concluded that the developing children's brains used different cortical regions from those of adults when processing co-speech gestures. Therefore, the age difference has significant effect on language learning with gestures.

(c) Viewing (Co-Speech) Gestures

Unlike the picture group in Tellier's (2008) experiment, all students observed gestures in the present study. As introduced earlier, in gesture studies in L1 and FL, little consensus subsists whether gestures help with comprehending a message in either children or adults (Hostetter, 2011). However, the non-gesture group in Tellier's (2008) study improved knowledge of a single word meaning in an absence of viewing gestures with statistical significance. The non-gesture group(s) in this study which observed gestures also learned a single word meaning with statistical significance.[48]

(d) Sample Size:

A sample size differed between this and Tellier's (2008) study, as Tellier (2008) had 20 and the current study 126. Consequently, the two studies have five main commonalities, including the positive correlation of two different assessments as discussed earlier and four major differences.

2.3.2.3.1.1.4 Limitation of Tellier's Study:

Two limitations in Tellier's study exist. First, Tellier's descriptions of passive and active knowledge are insufficient. According to Tellier (2008):

[47] Self-adaptor gestures refer to gestures which encompass pointing to oneself, putting one's hands to oneself, including scratching a head (Dick et al., 2012; Streek, 2008).
[48] See Chapter 5 for this issue.

(a) Passive knowledge: Participants have the ability to pair a picture or a gesture with the corresponding English word, when they hear the English word.

(b) Active knowledge: Students have the ability to verbalize an English word matching to a picture or a gesture when the English word has the corresponding nonverbal signs.

Discussions about active and passive knowledge made by Tellier (2008) need theoretical backgrounds, and thus[49] in Chapter 5 I will discuss this issue based on sociocultural theory. Second, the lack of the concept of the instruction-assessment dualism is a crucial factor in Tellier's (2008) study. In short, Tellier's (2008) study led to partially answering Research Question 1 and to relatively supporting Hypothesis 1 at the pre-school level.

2.3.2.3.1.2 Bilingual Children's Language Development with Co-Speech Gesture[50]

A previous study suggested that that oral speech and gestures were independent of each other (Pettito, 1992, p. 55). However, Mayberry and Nicoladis (2000) show evidence that gestures help bilingual children with simultaneous language and oral speech development in two languages.

Mayberry and Nicoladis (2000) tested whether gestures linked to oral speech with five bilingual boys in French and English to examine the argument made by Pettito (1992). In their longitudinal study (one year and a half) with the boys, the measuring device of their

[49] For instance, according to Hulstijn (2010), the classification of knowledge in FL is more complicated than that of Tellier's (2008) study. For instance, the cloze test is used to measure a combination of productive vocabulary knowledge, receptive knowledge of grammar, productive knowledge of grammar (to the extent that the forms of the words to be supplied express grammatical information), orthographic knowledge, knowledge of semantic, pragmatic and discourse (text comprehension) (cf. pp. 191-192).

[50] Their research design was similar to that of Ronget (Vygotsky, 1997b).

language development was the average length of an utterance. A finding of Mayberry and Nicoladis (2000) was that:

> The children began to produce iconic and beat gestures while speaking only after their utterances became longer than two words. As a result, utterances accompanied with iconic or beat gestures were more linguistically complex than utterances accompanied with points or no gestures at all (p. 194).[51]

At 2 years of age, 81% of the gestures made by participants were generated while speaking, whereas by age 3½, this percentage increased to 90%. Importantly, participants did not abandon co-speech gestures as they learned to talk. For the reasons explained above, Mayberry and Nicoladis (2000) proved that the hypothesis proposed by Pettito (1992) that gestures were not linked to oral speech was incorrect. Besides, on the basis of studies presented by sociocultural theory, I argue that participants' thinking arose when they generated iconic co-speech gestures in communicating with their interlocutors. Two reasons account for this:

First, Vygotsky (1987) states, "Word meaning is a phenomenon of verbal thought or of the meaningful word. It is a unity of word and thought "(p. 244). Vygotsky (1987) also suggests, "Is word meaning speech or is it thought? It is both at once and the same time; it is a unit of verbal thinking" (p. 47). For the reasons explained above, word meaning encoded by iconic and/ or beat gestures in the study presented by Mayberry and Nicoladis (2000) unified words and the participant's thought, leading to the formation of syntax with word meaning in their speech.

Second, Willem and Hagoort (2007) claim that some co-speech gestures have a clear semantic relationship to speech when gestures and speech were co-expressed. Contemporary

[51] This co-speech gesture may have contained co-thought gesture.

researchers (Bybee et al., 1994; Goldberg, 1995; Lackoff, 1987; Langacker, 2009) and sociocultural theorists (Masuda & Arnett, 2015a; Vygotsky, 1998) agree upon the fact that meaning co-evolved with grammatical construction (lexicon- grammar), leading to the concept of form-meaning couplings. Furthermore, IPH proposed that co-speech gestures might help speakers package information when constructing their utterance with multiple words at clause-level combinations (Bock & Cutting, 1992; Kita, 2000; Mol & Kita, 2012; Nicoladis et al., 2009). Thus, producing co-speech gestures helps participants learn multiple words with proper syntax to a certain extent, by assisting them with organizing their thinking to formulate proper speech (syntactic constructions) (Mol & Kita, 2012).

Consequently, a study presented by Tellier (2008) and Mayberry and Nicoladis (2000) complements each other, which led to strengthening Hypothesis 1and properly responding to Research Question 1, although participants' age differences exist among a study presented by Mayberry and Nicoladis (2000), Tellier (2008), and the current study. However, an effect of viewing gesture is still inconclusive. Three issues in experimental gesture studies in FL have been addressed.

2.3.2.3.2 Gesture-for-Conceptualization Hypothesis: Studies of Gesture in L1

As explained earlier, Kita et al. (2017) proposed the gesture-for-conceptualization hypothesis. This hypothesis is built upon gesture studies in L1 (Alibali et al., 2000; Chu & Kita, 2011, 2012; Goldin-Meadow, 2003, 2015; Kendon, 1994, 2004; Kita & Ozyurek, 2003; McNeill, 1992, 2005).

Kita et al. (2017) argue that representational gestures schematize information in the four functions; gestures manipulate, activate, package, and explore. The following sections include an explanation of commonalities between the gesture-for-conceptualization hypothesis and the Gesture Listening Higher Concept Approach regarding an effect of regenerating ICSCTGs. Section 2.3.2.3.2 includes a discussion of how manipulation in producing co-thought gestures can help Japanese EFL students with learning word order in English to a certain extent first. Subsequent sections include descriptions of co-speech co-thought gestures (exploration) and then follow with an explanation of how gestures help students activate lexical access and package information to formulate proper oral speech to some extent. I further outline how gesture is instrumental in overcoming the constraint imposed by the Whorfian version of linguistic relativity.

2.3.2.3.2.1 Co-Thought Gestures: Manipulation of TRHLM-TRLLM Distinction and Word Order

This section includes a discussion of manipulation of co-thought gestures. Using Shepard and Metzler type mental rotation tasks, Chu and Kita (2012) asked participants to match the stimulus object to one of the two mirrored three-dimensional objects in a non-communicative mental task. An example of the manipulation and exploration nature of generating gestures:

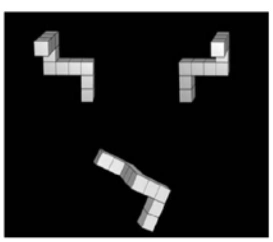


Figure 3: Shepard and Metzler Type Mental Rotation Tasks in Chu and Kita (2008). (Chu & Kita, 2008, p. 717).

While Chu and Kita[52] asked participants in a gesture encouraged group to solve the problems alone in a room, the investigators videotaped participants' spontaneous co-thought

[52] Other groups are a gesture-prohibited group and a gesture-allowed group in this experiment.

gestures by using a hidden camera. Participants used their hands that represented a stimulus object and then rotated their hands to represent the object rotation in the task. Chu and Kita (2012) concluded that co-thought gestures in silent problem solving enhanced their performance in spatial problem solving, suggesting that even if participants did not verbalize how they rotated the object, co-thought gestures helped them solve the problem by manipulating the information explained above.

Moreover, participants created co-thought gestures when a silent problem-solving task was harder than when it was less demanding (Chu & Kita, 2012). Needless to say, learning about vertical spatial operations is a challenge for EFL students. Furthermore, Kita et al. (2017) claim that abandoned gestures were a consequence of abandoned speech and that participants produced abandoned gestures more often in more difficult problems. This was considered to be unsuccessful exploration.

Additionally, Kita and Ozyurek (2003) proposed that the interface hypothesis examined cognitive processes in cross-cultural differences in creating spontaneous gestures with English, Turkish, and Japanese speakers. For instance, Kita and Ozyurek showed participants in their study a short video of a cat swinging on a rope from one building to another in an attempt to catch a bird. The investigators asked participants to tell a story about what they had seen and then the investigators videotaped their gestures and narratives for analysis.

Kita and Ozyurek (2003) claimed that none of participants verbalized left and right in expressing the lateral direction of the swing, although the participants systematically produced gestures by showing the left-right horizontal orientation of the observed motion. That is, overt speech does not always link to a specific gesture. Additionally, gestures suggest the thoughts of the speaker (de Ruiter 2007). Similarly, Vygotsky (1987) states, “speech is not necessarily linked to a specific material carrier [gesture]” (p. 106).

Based on gesture studies in L1 and sociocultural theory, thinking can link to wordless motor processes by advancing a problem solving ability of Japanese EFL high school students in expressing vertical spatial relationships. That is, although students did not verbalize the contact-noncontact and the TRHLM-TRLLM distinctions with the head-initial SVO word order in English, co-thought gestures helped them with learning about the obligatory contrasts with manipulating the contact-noncontact and TRHLM-TRLLM distinctions, and the head-initial SVO word order to some extent. This is considered to be successful manipulation and exploration. However, if students would be unsuccessful in the manipulation of co-thought gestures in expressing vertical axis operation, students abandon co-thought gestures, which can influence formulating their oral speech.

2.3.2.3.2.2 Co-Speech Co-Thought Gestures

This section includes a discussion of how co-speech and co-thought gestures co-evolve by packaging information to formulate proper oral speech, which is an analogous claim to that of the Gesture Listening Higher Concept Approach. As discussed earlier, co-speech gestures play an essential role in helping speakers with thinking processes, which leads to helping participants understand difficult verbal materials. Vygotsky (1987) contends, “the motor processes associated with speech play an important role in facilitating the thinking process in particular, in improving the subject’s understanding of difficult verbal material” (p. 44).

Additionally, the action generation hypothesis (AGH) presented by Chu and Kita (2016) claims a positive link between the co-thought and co-speech gestures created from the representational utilization of the action generation processes. Mittelberg and Waugh (2009) note that thought in co-speech gestures (co-thought and co-speech gestures) demonstrates learners' conceptualization processes.

Moreover, based on congenitally blind speakers, Goldin-Meadow (2003) argues that co-speech gesture is not only an integral part of communication, but may also be an integral part of thinking. Thus, I argue that participants in Mayberry and Nicoladis (2000) generated ICSCTGs, even though they did not discuss co-thought gestures in their study. Mayberry and Nicoladis demonstrated that producing ICSCTGs assisted participants with oral speech and language development accompanied with linguistically complex expressions by bundling information when the participants were structuring their oral speech with multiple words. Mayberry and Nicoladis (2000) indicate that regenerating ICSCTGs would help Japanese EFL students with the formation of semantics and syntax by packaging information to formulate oral speech to a certain degree. Mayberry and Nicoladis' study (2000) is an analogous claim to that of the gesture-for-conceptualization hypothesis (Kita et al., 2017) regarding packaging information.

2.3.2.3.2.3 Lexical Access: Activation and Package

The gesture-for- conceptualization hypothesis suggests that representational gesture makes a speaker's semantic image of a word more highly activated and consequently more accessible (Kita et al., 2017). Vygotsky (1997a) suggests, "the sign gives birth to meaning; meaning sprouts in consciousness" (p. 137). As stated earlier, Willem and Hagoort (2007) claim that some co-verbalized gestures have a clear semantic relationship to speech when gestures and speech are co-expressed. Additionally, Kita et al. (2017) argue that a schematized representation (iconic gesture) is more suitable for packaging into units or chunks (lexical item) for speaking or thinking (p. 258), and thus, recreating ICSCTGs helps students with lexical access when they talk about vertical spatial structures.

2.3.2.3.2.4 Exploration

As introduced earlier, the gesture-for-conceptualization hypothesis suggests the four lines of evidence that converge in building a case for this function of gesture, indicating that gestures explore spatio-motoric information. This study focuses on exploring spatio-motoric representation, via trial-and-error, by pursuing information useful for speaking and thinking. It is most likely that Japanese students would learn how to express vertical axis operations via trial–and-error in formulating speech and thinking while regenerating ICSCTGs.

Moreover, this hypothesis claims that representational gestures (iconic and metaphoric) are involved, not only in speaking, but also learning and problem solving, which is analogous to the Gesture Listening Higher Concept Approach regarding an effect of ICSCTGs. As explained earlier, word meaning unifies thinking and speech (Vygotsky, 1987). Regenerating ICSCTGs can lead to internalizing developed meaning of words in EFL to a certain extent by helping participants activate lexical access, and package, manipulate, and explore schematized information in learning how to express vertical spatial operations by unifying thinking and speech.

Co-speech and co-thought gestures can co-evolve when participants engage with difficult verbal materials. For this reason, the more students recreate ICSCTGs, the higher achievement they can attain in learning about a prepositional phrase in vertical spatial operations. Thus, I argue that reproducing ICSCTGs helps them learn grammatical and semantic structures in expressing vertical axis structures to a certain degree. In this way, the answer to Research Question 1 can be strengthened, supporting Hypothesis 1.

Consequently, the Gesture Listening Higher Concept Approach shares the similar thought to that of the gesture-for-conceptualization hypothesis with regard to the effects of regenerating ICSCTGs. The following explains how gestures help Japanese EFL student reduce the level of constraint of linguistic relativity.

2.3.2.4 Gesture and Linguistic Relativity

The Whorfian version of the linguistic relativity hypothesis and the thinking-for-speaking hypothesis predict that second language learners may be constrained in their ability to adapt their thinking to a new language (McNeill & Duncan, 2000; Whorf, 1957). For example, Japanese students' conceptualization of vertical spatial operations is based on their L1, which does not encode the contact-noncontact distinction, whereas English does. This linguistic difference creates difficulties with conceptualizing vertical axis structures.

Regenerating ICSCTGs can lead to internalizing developed meaning of words in EFL by helping participants package, manipulate, and explore schematized information to express vertical spatial operations by unifying thinking and speech. Moreover, recreating ICSCTG facilitates lexical access (on, over, above, and under) by visualizing contact and non-contact distinctions and TRHLM-TRLLM differentiations, which leads to activating lexical representations. Furthermore these four functions of gestures link to maintaining information by lightening cognitive load (Kita et al., 2017). This study focuses on the challenges that EFL learners have in adapting their thinking to the novel language, exploring whether the four functions of gestures identified in studies of L1 will also apply to L2. Additionally, it appears that gesture reflects one's thinking patterns among different cultures (Kita & Ozyurek, 2003). Therefore, clearly gesture can be the instrument that leads to shifting students' thinking in L1 to that of L2 (EFL) when Japanese EFL students learn about vertical spatial relationships to a certain extent. The following section 2.3.2.5 includes an explanation of the integration of reproducing ICSCTGs with listening practice.

2.3.2.5 Integration of Regenerating ICSCTG with Listening Practice

Very few papers exist in regard to the theoretical framework of the incorporation of recreating ICSCTGs into listening practice, although some researchers have published papers with regard to multimodality (Hosteller, 2011; Muller, Cienki, Fricke, Ladewig, McNeill & TeBendorf, 2013). Some FL professionals use this technique to help students acquire a certain level of proficiency. Nonetheless, researchers have not established the clear theoretical framework in this regard to date.

2.3.2.6 Form and Meaning Pairing

As explained earlier, contemporary linguists, neuro scientists, applied linguists, and psycholinguists reduce the separation between lexicon and grammar. For this reason, the researchers noted above agree with the concept of form-meaning pairing or couplings.

Vertical axis operations should be taught by incorporating the binary semantic categorization with proper TRHLM-TRLLM distinctions into The SVO word order in a Japanese EFL class. Otherwise, erroneous expressions can result when Japanese EFL students express vertical spatial operations (See Appendix F).

Langacker (2009) suggests, “Grammar and lexicon form a continuum fully describable as assemblies of symbolic structures (form-meaning pairings)” (p.38). Goldberg (1995) indicates, “the lexicon is not neatly differentiated from the rest of grammar” (p.4). Goldberg also states that the term *lexicon-grammar* is unacceptable (A. Goldberg, personal communication, May 8, 2015). Lakoff (1987) suggests, “grammatical constructions come with meaning” (p.256). Bybee et al. (1994) states, “the development of grammatical materials is characterized by the dynamic coevolution of meaning and form" (p.20). Davidse and Lamiroy (2002) propose a coupling of form and meaning (p. 1).

Furthermore, researchers suggest the existence of "grammatical categories in the lexicon" (MacWhinney, 2013, p.216). Ellis and Robinson (2008) recognize, "form and meaning come together" (p.5). Hagoort and Poeppel (2013) claim that:

> there is no single computational entity called 'syntax' and no unstructured operation called 'semantics,' just as there is no single brain area for words or sounds. Because these are structured domains with considerable internal complexity, unifications, or linking operations *as outlined* in the MUC [memory, unification, and control model][53] perspective above, is necessary (p.254).

The concept of form-meaning pairing "applies to both language production and language comprehension" (Hagoort, 2005, p.416). Similarly, Vygotsky (1998) states, "In adult speech, in words-concepts, it has its material in grammar and syntax, its normal rules of formulation" (p.76).[54] Thus, the concept of form-meaning couplings or pairing is essential in teaching languages. Applied linguists have conducted research with the concept of couplings of form and meaning (Cardierno, 2008; Ellis, Masuda & Annet, 2015). Furthermore, some have integrated cognitive linguistics and sociocultural theory in their studies (Masuda & Annet, 2015). Consequently, the concept of form-meaning couplings is crucial in language learning, which sociocultural theory supports.

2.4 Summary

A goal of the literature review in the present study is to guide one to understand backgrounds of experiential, hypothetical, and experimental gesture studies in L1, FL, EFL, bilingualism,

---

[53] The author inserted.

[54] It is clear that Vygotsky (1998) supports the concept of form-meaning couplings, even though Vygotsky (1987) states, "the semantic and the verbal syntax arise neither simultaneously nor together" (p. 253). Taking into consideration when he was alive, he might have thought that morpho-syntax was not syntax. Thus, on the basis of this claim (1998), he considered that there was a linkage between semantic and syntax.

and studies of sociocultural theory in discussing the Gesture Listening Higher Concept Approach. This literature review reveals a novel effect of a higher concept in experimental gesture studies in FL. Furthermore, this literature review strengthens the support of Hypothesis 1 that corresponds to Research Question 1, which experimentally, hypothetically, and theoretically directs to answering to the new research path in SLA. However, Hypotheses 2 and 3 remain uncertain. Therefore, Chapter 4 will test Hypotheses 2 and 3, leading to discussing these hypotheses in Chapters 4 and 5 based on the results in this study and findings in other literature reviews, which this chapter does not demonstrate.

Chapter 3

Methods

3.1 Introduction

The methods outlined in this chapter include an explanation of a comprehensive framework within which to test the hypotheses and to investigate the research questions addressed in Chapter 1. As will be shown, effects of the Gesture Listening Higher Concept Approach were found at a high school level in Japanese EFL education with inferential analyses on theoretical and experimental bases, fulfilling the goals in this study.

On the basis of research in TFS, sociocultural theory, and gesture study, section 3.2 includes an explanation of study design that uses quantitative methods. Subsequently, section 3.3 regards data collection procedures. Section 3.4 includes an explanation of the outline of the process that is used in categorizing participants' responses by coding them. The coding is done based on studies presented by Bowerman (1996a), Herskovits (1985), Huddleston and Pullum (2002), Lakoff (1987), Martin (1975), Tyler and Evans (2003), Corpus of contemporary American English (COCA), studies of language transfer (positive and negative transfer in SLA). It also incorporates some constraints which perspectives in sociocultural theory and SLA clarify. Evaluators whose L1 is English double-decoded ten percent of the data from the oral speech and cloze tests to increase reliability (Loewen & Plonsky, 2015). And finally, section 3.5 includes a description of statistical methods in this study, increasing the likelihood of a finding a significant difference between group means (ANCOVA models) (McMillan & Schumacher 2009).

3.2 Study Design

This section includes an explanation of structures of theoretical and hypothetical framework in conducting this experimental study. Section 3.2.1 justifies employing a quantitative method and then section 3.2.2 comprises of a research design in this study.

3.2.1 Justification of Study Design

McMillan and Schumacher (2009) write that only recently the mixed method (the use of both quantitative and qualitative approaches) has been given serious consideration by scholars.[55] The disadvantage of mixed methods is that one method may be used superficially, thus causing difficulty in writing reports and forming conclusions. Accordingly, the use of quantitative approaches employing both experimental and non-experimental research designs can be helpful in strengthening the objectivity of test results. As a result, I argue that the quantitative methods will lead to attaining the goals in the current study.

I employed a quasi-experimental (no random assignment of participants) design. Quasi-experimental designs are similar to randomized experimental designs in that the purpose of both is to determine cause and effect. For an experiment, I controlled classroom interventions and experimental treatments using several classes within a single school to determine the effectiveness of teaching methods (McMillan & Schumacher, 2009). However, the classes had already been established according to grade, students' proficiency and subject matter by school officials. Accordingly, a random selection of participants was not possible. Three classrooms of students were assigned to experimental groups. A fourth classroom was the control group for examining the results of treatments.

---

[55] *Journal of Mixed Methods* started in 2007.

3.2.2 Research Design of This Study

On the basis of theoretical, hypothetical framework, and literature review, this study structures a research design with a quantitative method, including a non-experimental research design (a survey) that is frequently used in educational research (McMillan & Schumacher, 2009). A survey examines relationships between distinct phenomena in the absence of any direct manipulation (cf. Appendix L). This section includes an explanation of the target population of the current study first, and then follows with explanation of sample size, of instrumentations with theoretical and hypothetical framework and gesture study in L1, of internal and external validity, and of quasi-experimental procedures.

3.2.2.1 Population of Present Study

The target population in this study was 11^{th} and 12^{th} graders at a public high school in one prefecture in Japan.[56] The enrollment of 11^{th} and 12^{th} graders at public high schools in the prefecture was approximately 9,400 in 2010 and 2011. These students were enrolled in general and specialized courses (nursing, math and science, music, EFL, agriculture, and engineering) in the prefecture.

The investigator was born in the prefecture, worked there for approximately 20 years, taught English as a foreign language (EFL) at eight different public schools in the prefecture, and was an administrator of an educational organization for approximately six years in the prefecture, and thus, the investigator is familiar with the characteristics of the population. Furthermore, familiarity with the population group made it easier to draw student samples for the study.

---

[56] The prefecture is a geographical unit and administrative subdivision roughly corresponding to a state of the U.S.A.

3.2.2.1.1 Selection of High School and Groups of Participants

Based on standardized test scores[57] in EFL administered by the prefecture board of education, a high school and a sample population were identified for this study. When selecting the samples for the study, the standardized test scores, the degree of difficulty learning about vertical axis relationships at a lexicon-grammatical level (developed meaning of words in EFL), and the use of established classes were considered to maximize validity and reliability.[58] Permission to use the standardized test scores was obtained from the participants, their parents, the teachers, and the principal at the chosen high school.

I selected high proficiency (HiP) students at the selected high school to examine the effects on iconic co-speech co-thought gestures (ICSCTG) and listening practice because their EFL standardized test scores suggested their EFL proficiency were neither too high nor too low to test the effects. For instance, the EFL standardized test scores of 11th and 12th graders at the selected high school were lower (mean = 40.2, N≈409) than that of the overall average score in the prefecture (mean=45.75, N≈9,400). The mean of HiP group was 42.54 (N ≈ 126), which was higher than the entire 11th and 12th graders at the selected high school, although a little lower than that of the overall average score in the prefecture. Thus, the sample in this study allowed for generalizations to be drawn regarding the population of the prefecture.

[57] When standardized tests are administered, “careful attention has been paid to the nature of the norms, reliability, and validity. This results in instruments that are ‘objective’ and relatively uninfluenced or distorted by the person who administers them and the standardized test does not focus directly on the variable of interest in the study ... The alternative is to develop your own instrument” (McMillian & Schumacher, 2009, p. 189). Thus, I developed revised versions of existing instruments.

[58]“Validity is a characteristic based on the reasonableness of inferences drawn from a set of scores. Evidence is gathered to support an argument that the inferences are reasonable…Consistency is the essential element of reliability. It can be determined by using several sources of evidence. Stability and internal consistency are the most common types of reliability emphasis, though internal consistency can result in too many items.” (McMillan & Schumacher, 2009, p.185).

3.2.2.1.2 Selection Process for Individuals: High Proficiency Students Only

Two different EFL student groups, HiP and LoP (a low proficiency group), had already been determined by the teachers at the high school where the current study took place. These groups had been determined on the basis of term examinations consisting of reading comprehension, writing, and grammar tests, but not oral speech. These examinations take place six times a year at the high school. The enrollment of 11th and 12th graders at the school was roughly 400 when this study took place.

Given markedly distinct linguistic differences between English and Japanese and lexicon-grammar complexity of spatial terms indicated by researchers, the mere 150 minutes of instruction (three lessons) allotted to the investigator by the school for the study would have been insufficient for LoP students to learn vertical spatial operations. Accordingly, I selected the HiP group in the 11th and 12th grades to test the effects of teaching ICSCTG and listening practice in learning about vertical axis relationships.

(1) Random Group Sampling

Four of these HiP EFL classrooms, consisting of approximately 30 students each, were randomly selected from the six HiP EFL classes. However, individual random sampling was not administered because of the pre-determined set up of classrooms, as noted above.

(c) Sample Size

One-way ANOVA, one-way ANCOVA including the repeated-measures ANOVA, and a paired sample t-test were used to conduct statistical analyses. The number of participants determined by G*Power analysis[59] necessary to one-way ANOVA was 112 (ES[60] = .4; α err prob = .05; Power (1-β err prob) = .95; critical F=2.69, p<.05, participants=69 females and 57 males, range=16-18; *M* age=17.15).

The number of students defined by G*Power analysis required to run one-way ANCOVA was 124 (ES = .38; α err prob = .05; Power (1-β err prob) = .95; critical F=2.68, p <.05).

Likewise, the number of participants defined by G*Power analysis crucial to run the paired sample t-test was 122 (ES = .33; α err prob = .05; Power (1-β err prob) = .95; critical t= 1.97976, p <.05). Accordingly, recruiting 126 students was sufficient for data analyses with the ANOVA, the ANCOVA, and the paired sample t-test approaches.

3.2.2.2 Instrumentation[61]

Five major but differing types of instruments were used in the present study to assess the effects of instructing ICSCTG accompanied by listening practice on developed meaning of words in EFL when students were taught vertical axis structures in a Japanese EFL high school class. One of them is a professionally generated assessment tool (participants' standardized tests consisting of three different scores and following ACTFL guidelines). The others were instruments which I generated based on my theoretical framework and teaching and research experience in Japan, the U.S., and Canada. These included (a) a 20-picture book

[59] This is a tool that was designed as a general stand-alone power analysis program for statistical tests commonly used in social and behavioral research by researchers at the Heinrich Heine University Düsseldorf in Germany (Faul et al., 2007).
[60] ES refers to an effect size. American Psychological Association (APA) suggested including effect sizes and confidence intervals (CIs) and/ or statistical significance levels (2010). Population of effect size is: f=.5 is large, f=.3 is medium, and f=.1 is small. Refer to Cohen (1969/1988).
[61] Instrumentation is done based on Creswell (1994).

with audio material for a quasi-face-to-face dialogue oral speech test, (b) a written cloze test, and (c) a background survey.

3.2.2.2.1 Professionally Generated Instrument

(a) Standardized Test Scores

The standardized test scores for 11^{th} and 12^{th} graders at the selected high school were used as independent variables, which allowed one to assess the effects of dependent variables on developed meaning of words in EFL. The standardized test scores consisted of students' general EFL abilities (reading, writing, and listening comprehension skills) and their general language abilities in L1. It should be noted that all public high school students in the prefecture are required to take standardized tests consisting of Japanese as L1, EFL, and mathematics.

The test of Japanese as L1 examined their lexicon-grammar knowledge, vocabulary in context with Chinese characters, and reading comprehension for both classic and modern Japanese. The EFL test investigates their lexicon-grammar knowledge, reading and listening comprehension. They usually take the tests seven months following their acceptance by the high school.

- **Participants' Abilities in Japanese as a First Language and EFL**

Cummins (2001) suggests, "The developmental interdependence hypothesis proposes that the level of L2 competence which a bilingual child attains is partially a function of the type of competence the child has developed in L1 at the time intensive exposure to L2" (p. 75). I statistically examined how students' L1 and EFL scores in this standardized tests can have an effect on learning about developed meaning of words in EFL in this study.

(b) Oral Proficiency

The American Council on The Teaching of Foreign Language (ACTFL) guidelines published in 2012 were employed to assess the participants' EFL oral proficiency prior to any of the interventions[62] described below. These were unofficial assessments by teachers who had taught them at high school for one to two and half years. ACTFL guidelines help one determine students' oral proficiency levels, including Novice high, Novice mid, or Novice low (ACTFL guidelines 2012).

3.2.2.2.2 Instruments Created by Investigator

This section includes a description of three instruments: two major instruments to generate dependent variables (an oral speech test and cloze tests), and a survey to elicit demographic variables.

(a) Oral Speech Tests

I presented a quasi-face-to-face dialogue with a 20-page picture book and audio material as an instrument to test oral speech proficiency. The instrument was generated on the basis of the theoretical and hypothetical framework and existing instruments, including a moderate version of linguistic relativity (TFS), L1 gesture studies, and research of language and thought, motor processes, memory and knowledge, and dialogue in sociocultural theory.

The picture book was a modified version of the *frog story* implemented by Berman and Slobin (1994) for a TFS project and an adjusted version of Oregon Japanese Oral Performance Assessment (OJOPA).[63] The theoretical background for generating the

[62] Some researchers define *intervention* as a manipulation over time. There is no agreement among researchers regarding the definition of *intervention* and *treatment* (cf. Appendix C).
[63] It assesses oral proficiency for American high school students in Japanese as a Foreign Language (JFL) classes at a secondary school level. The investigator had completed training to assess secondary school students' oral proficiency of JFL, and thus, I have modified OJOPA to fit teaching EFL at a high school level.

instruments in the present study were gesture studies in L1 (Chu & Kita, 2012, 2016; Golden Meadow, 2003; Kita, 2000; Kita & Ozyurek, 2003), TFS hypothesis proposed by Slobin (1987, 1991, 1996, 1997, 2000, 2001, 2003, 2004, 2006) and sociocultural theory. The instruments were used to obtain dependent and independent variables. The picture book was designed to assess participants' abilities to formulate developed meaning of words in EFL in context when students talk about vertical spatial relationships in English.

OJOPA is designed to measure the oral proficiency of American high school students in a Japanese class in face-to-face dialogues in Japanese about their everyday life. However, OJOPA is not designed to employ images and the principle of linguistic relativity as an assessment tool.

- **Thinking-for-Speaking Project**

Conversely, TFS is closely associated with the principle of linguistic relativity presented by Whorf (1956). The TFS hypothesis proposes that language learners need to acquire strategies for constructing and interpreting extended discourse which are deeply influenced by specific modes of thinking and speech (Slobin, 1987). The TFS hypothesis can be viewed as a contemporary moderate version of linguistic relativity (Slobin, 1996). In the TFS project, Berman and Slobin (1994) used plausible stories that explain images illustrated on a wordless picture book, which enables one to understand how different languages (English, German, Spanish, Turkish, and Hebrew) express differently a scene illustrated in a wordless picture book.

Age range of participants in their studies was from three years old to adult. Berman and Slobin (1994) employed narratives, and reported developmental trends in forms and functions both within and across languages. Their study "has burgeoned considerably—as demonstrated by the collection edited by Stromqvist and Verhoeven (2004)" (Berman 2008, p. 121), which leads to significant influence on research about linguistic relativity. Their narrative method has been recognized as an outstanding way to elicit participants' utterances to examine cognitive consequences of linguistic relativity in expressing the cross-linguistic patterning of relative clauses.

However, I employed a quasi-face-to-face dialogue, but not narratives. Vygotsky (1987) write, "dialogue is the most natural form of oral speech… that language reflects its true nature only in dialogue" (p. 272). For this reason, a face-to-face dialogue was employed so that students could specify their answer, unlike in narratives. Besides, it was considered to be a more appropriate method for assessing the naturalness with which they restructured their speech and thinking with developed meaning of words in EFL, rather than using narratives, in learning about vertical relationships in an EFL class.

The investigator's L1 is Japanese rather than English, and thus, a face-to-face dialogue with Japanese students did not seem natural or authentic. Therefore, a quasi-face-to-face dialogue was used where the investigator did not speak to participants directly, but instead used audio material that was recorded by a native English speaker. The context for the recorded dialogue was a conversation between a Japanese exchange student in the US and an American student about how to organize furniture in his room at his host family's house.

According to Alibali et al. (2001), the rate of gesture generation differed between a quasi-face-to-face and a direct-face-to-face condition. Even though ICSCTG' reproduction rate in the current study would be lower than that of a direct-face-to-face condition, using an audio instrument was sufficient for assessing an effect of ICSCTG inasmuch as all four groups were tested with the quasi-face-to-face condition. Thus, the quasi-face-to-face recorded dialogue was used in conjunction with the 20-page picture book I had devised. This instrument assessed their learning of lexicon-grammar to talk about vertical axis relationships in English. Slobin (1997) writes, “Most work in linguistic typology focuses on lexicon and grammar, and does not concern itself with large units” (p. 23). Accordingly, the picture book allowed one to elicit not only lexicon-grammar, but also large units of developed meaning of words in EFL with grammatical structures.

Concisely, the instrument in this study, a 20-page picture book, named *Potluck*[64] with the audio material was an effective tool to assess learning processes of developed meaning of words in EFL with TRHLM-TRLLM distinctions and The SVO word order in talking about vertical axis relationships (cf. Appendix J).

(b) Cloze Test

For a long time, language researchers considered that cloze tests were a suitable instrument to examine learners’ general language proficiency. For instance, Koda (1990a/ 2005) notes that grammar knowledge associates more highly with cloze performance than multiple-choice test scores.[65] This study examined whether the cloze test (i.e., the polar opposite of oral speech) help participants learn about vertical spatial operations at a lexicon and grammar level. Lantolf and Thorn (2006) indicate that the use of a cloze procedure is effective in assessing a proper use of articles in writing essays “in the precise lexical contexts in which they had exhibited problems prior to the tutorial sessions” (p. 290).

The theory behind the cloze test is the principle of reduced redundancy (Babaii & Shahri, 2010; Grotjahn, 2013; Oller, 1972, Tayler, 1953) (cf. Appendix I for examples of

---

[64] Berman and Slobin (1994) used a 24-page-wordless picture book for their TFS project.
[65] Refer to Lantolf and Appel (1994) as well and Grotjahn (2013).

cloze test in this study). There have been several variations on the cloze procedure. I installed a Multiple-Choice Cloze test whereby participants filled in each blank in a sentence on the basis of vertical axis operation analysis at a lexicon-grammar level. This cloze test was used to examine participants' learning processes about the developed meaning of words in EFL in vertical axis operations at a lexicon-grammar level.[66]

Koda (2005) suggests that cloze tests have been used to estimate reading ability and that the purpose of cloze test is to measure local text understanding, and does not test global text-coherence building. The cloze procedure has been widely used because of its relative ease in test construction, administration, and scoring. However, the procedure is not problem-free. For instance, if a text segment is not understood, a blank obviously cannot be filled. Additionally, if the deleted word is not in participants' productive vocabulary, the slot can be left blank, even if they understand a text segment, and thus, this study avoided these problems noted above in generating cloze tests.

Tayler (1953) generated the cloze test and coined the term "cloze", linking it explicitly to the term "closure." Classic cloze tests are considered integrative or holistic tests which assess proficiency in L1 and L2. Oller (1972) theoretically criticizes the discrete-point tests[67] and states that the cloze test (i.e., cloze procedure) helps improve proficiency in reading comprehension tests.

---

[66] Refer to Lantolf and Appel (1994) as well.

[67] Discrete-point tests are related to multiple choice and true/ false layouts, which have been criticized for testing only recognition knowledge and assisting guessing. However, if they are used for a suitable purpose and if the test questions are well structured, discrete-point tests can be used for successful teaching and learning. However, if only discrete-point tests are administered for a long period of time, it could be problematic in language learning.

Furthermore, as discussed earlier, sociocultural theorists propose the instruction-assessment dualism. Even though cloze tests have been used assessments, cloze tests can be an effective pedagogical strategy in helping students learn about developed meaning of words in EFL, if instructors properly use the tests considering the zone of proximal development.

(c) Survey[68]

The purpose for using the survey was: (a) to identify relevant background variables that might interact with the experimental manipulation; and (b) to collect the data of the demographics and characteristics of the sample population, which could affect variables (McMillan & Schumacher, 2009). The survey questions consisted of close-ended, dichotomous questions (Yes/ No) and four-point Likert scales from 1=strongly agree to 4=strongly disagree. The survey was given to all the participants following the completion of the delayed post-tests.

In short, two major but divergent types of instruments were utilized to assess the effects on teaching ICSCTG accompanied by listening practice in talking about vertical axis relationships for Japanese EFL high school education.

3.2.2.3 Internal and External Validity

In general, there are different types of potential threats to internal validity (McMillan & Schumacher, 2009) and dissimilar categories of threats to external validity in an experimental design (Gravetter & Forzano, 2011). I minimized the threats to both internal and external validity, as discussed below.

---

[68] Refer to Appendix L.

3.2.2.3.1 Internal Validity

To warrant a causal conclusion in the present study, minimizing systematic errors or bias improved the level of internal validity during data collection with participants. I identified different experimenter and subject effects which threaten the internal validity and address them as discussed below.

(a) Experimenter effects: The same instructor taught them in four different classes (groups), and for this reason influencing students' responses during pre-, post-, and delayed post-tests was avoided. Additionally, the time when they took the pre-, post, and delayed-post tests were taken into account when giving random assignments. In this way equal opportunity to take the test at a different time such as after school or during class was ensured, and thus, experimenter effects were minimized.

(b) Subject (Hawthorne) effect: A subject effect means that participants change their behavior because they know that they are participating in an experiment. Mackey and Gass (2005) suggested that it may be difficult for researchers to separate subject (Hawthorne) effect from experimental variables.

3.2.2.3.2 External Validity[69]

External validity refers to the extent to which the results of this study can be generalized. McMillan and Schumacher (2009) argued that in both social sciences and education, a majority of generalizations were limited to specific times and locations. Additionally, the social world changed more quickly than the physical world, social generalizations typically had shorter life spans than generalizations in the physical world.

[69] As oppose to external validity, there is internal validity. McMillan and Schumacher (2009) suggested, "internal validity is hard to complete control of unrelated variables in educational research is difficult" (p. 106), and thus, I will not describe internal validity in this study in detail.

Additionally, two or three experimental studies similar to this study are required for generalization of outcomes created by this study. However, very few previous studies regarding the concurrent ICSCTG and listening practice at a high school level have motivated this study. Hence, generalization of results in this experimental study will not be intended (cf. Yano & Long, 1994).

3.2.2.4 Quasi-Experimental Design

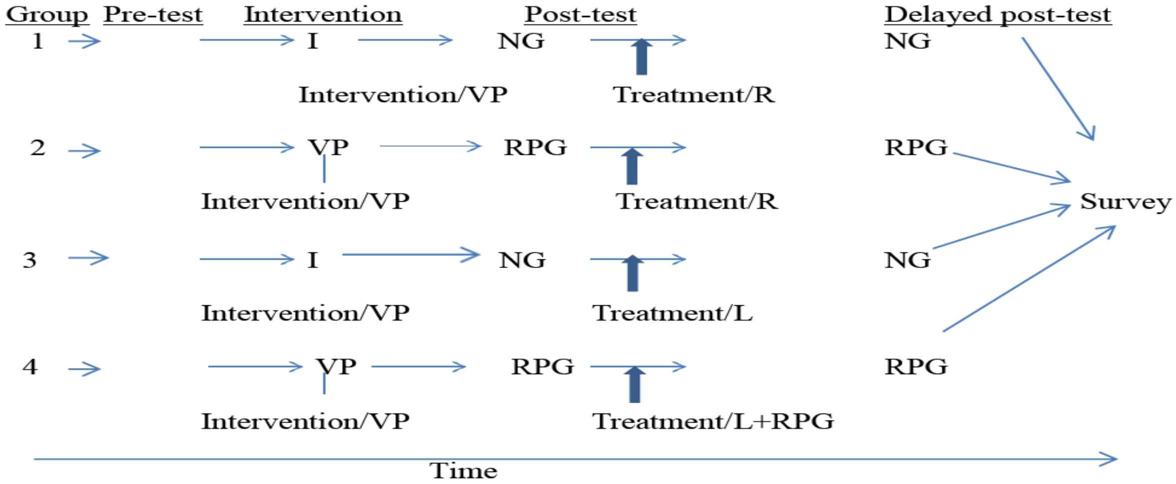
3.2.2.4.1 Grouping

I used a Randomized Comparison Group Design that allows one to select three experimental and one control groups out of the 11th and 12th grades in HiP groups at the high school:

- Group 1: the control group. (N = 31)
- Group 2: An experimental group with the regeneration of ICSCTG only. (N = 32)
- Group 3: An experimental group with Listening practice only. (N = 30)
- Group 4: An experimental group with the regeneration of ICSCTG and Listening practice together. (N = 33)

Figure 4 includes an explanation of the randomized comparison group designs for both oral speech and cloze test in detail:

Oral Speech Test: Randomized Comparison Group Design for Pre-test, Post-test, and Delayed post-test



Cloze test: Randomized Comparison Group Design

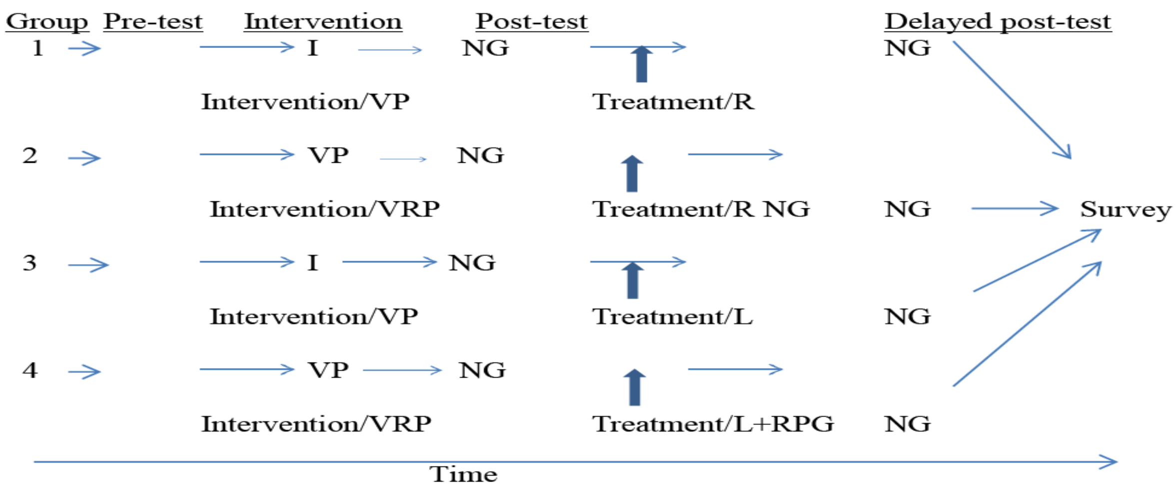


Figure 4. Classroom Interventions.

I: Classroom Intervention

NG: Participants were not asked to perform ICSCTG.

VP: Viewing ICSCTG

R: Reading aloud texts as an assignment

L: Listening practice

RPG: they were asked to reproduce ICSCTG in oral speech condition.

VRP: Intervention with viewing and reproducing ICSCTG

The following demonstrates how a classroom intervention with an instructional design took place in the proposed four groups listed above.

3.2.2.4.2 Classroom Intervention

This section includes an explanation of the theoretical framework of an instructional design, subsequently teaching materials for the classroom interventions is introduced and then a brief outline of an instructional design is described at the selected high school. Finally, the classroom instructions (or interventions) for three lessons are clarified in detail.

The two dimensional four quadrants framework generated by Cummins (1979, 1981, 2000) was used in the classroom intervention of this study. Cummins (2000) argued that this framework would allow participants to "attain greater clarity and precision in the characterization of the construct of language proficiency" (p. 100). Thus this framework can play a role in structuring the classroom interventions to construct Japanese EFL high school students' proficiency in talking about vertical spatial relationships.

- **Instructional Designs**

I distributed three major different types of teaching materials to them for three lessons.

Table 3:

Teaching Materials Given to Students in Each Day

Types	Materials given to students	Grp II & IV	Grp I & III
1	CD, audio material saved in a computer	Yes	No
2	Handout A_1, A_2, and A_3	Yes	Yes
2	Handout B	Yes	Yes
3	Picture cards with ICSTGs	Yes	No
3	Picture cards with images, JPN, ENG	Yes	Yes

Note. Refer to Appendix M for examples of "Handouts A and B", and Appendix N for "ICSCTGs", which explains hand movements of ICSTGs.

Table 4:

Sequence of Distributed Materials for Each Quadrant

Quadrant C (Day 1)	Quadrant A (Day 2)	Quadrant B (Day 3)
Handout A_1	Handout A_2	Handout A_3
Handout B/audio	Picture cards in type 3	Handout B/audio

Note. JPN refers to Japanese and ENG means English. Grp refers to *group.*

Handout A_1, A_2, and A_3 listed in Tables 3 and 4 are the formatted versions of PowerPoint presentation slides to help students learn about differences among *on*, *over, above*, and *under.* They are accompanied by images and corresponding expressions in both English and Japanese. Handout B consists of dialogues that give them an idea about the binary semantic categorization with The SVO word order in context. The picture cards in type 3 were used to play a game in quadrant A for day 2 (see Figure 5).

The brief outline of the structures for teaching vertical axis operations is as follows: literature review in this study suggested that whether gestures helped with comprehending a message had been the topic of various research studies over the previous 35 years and there had been little agreement (Gullberg & De Bot, 2010; Kita, 2016; Hostetter, 2011). However, taking into consideration the degree of difficulties in learning about vertical spatial operations, all of them were allowed to observe gestures which the investigator created.

Classroom interventions with all four groups took place one group at a time. Although all four groups observed ICSCTG to understand vertical spatial relationships, I requested only Groups 2 and 4 to recreate ICSCTG which is one of the experimental treatments, but not

Groups 1 and 3. Figure 5 includes an explanation of the two-dimensional four-quadrants framework that consists of four dimensions:

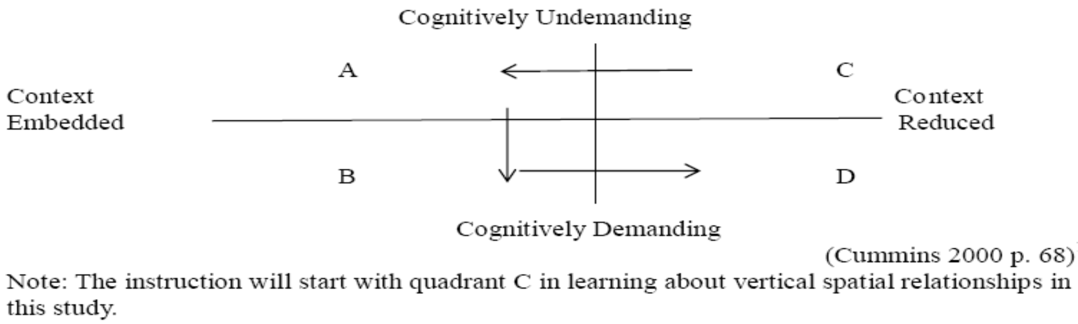


Figure 5. Two-Dimensional Four-Quadrants Framework.
Source: Cummins, J. (2000) *Language, Power and Pedagogy*. Clevedon: Multilingual Matters.

A review of the previous class was given at the beginning of each class except for quadrant C, and subsequently the goal for each quadrant were presented to all the participants, on the basis of literature reviews and theoretical frame work discussed in the previous chapters. Classroom interactions took place among the investigator, students, and their teacher, and PowerPoint presentations in teaching vertical spatial relationships were given in each class with Handout A_1, A_2, A_3 and B. A summary of each class of the day was briefly given to them at the end of each class. Finally, in quadrant B, assignments (treatments) were given to all of them during the last class in this research design. When they had questions, they asked the investigator during class.

- **Activities in Quadrant C**

In quadrant C, cognitively undemanding and context reduced activities took place. One of the goals in this quadrant was to have them be aware of developed meaning of words in EFL with grammatical structures in expressing vertical spatial relationships. Another goal was to have them realize differences between Japanese and English in talking about vertical axis relationships in context to avoid possible errors. The importance of regenerating gestures was explained to Groups 2 and 4 only, and similarly the salience of listening practice was just introduced to Groups 3 and 4.

The differences between the two languages regard describing vertical spatial relationships, including semantic and grammatical structures in English and Japanese (higher concepts in learning about developed meaning of words in EFL). These semantic and grammatical structures were explained by using ICSCTG, PowerPoint Presentation, and Handout A_1 and then all of them were asked to listen to audio material while looking at Handout B. Subsequently, they were requested to verbally repeat each expression in dialogues[70] described in Handout B. The dialogue, pre-recorded by a native English speaker, was about an exchange student from Japan getting help from his friend to organize furniture in a room at his host family's house in the US. The dialogues contained expressions of vertical axis relationships with contact-noncontact distinctions. Each of them was individually given pronunciation practices to confirm whether they could properly sound out each expression in English, and finally at the end of class, a summary of the class was given to them.

- **Activities of Quadrant A**

The goal in this quadrant A was to have participants remember expressions for vertical axis relationships in English with developed meaning of words in EFL, leading to cognitively undemanding and context embedded activities.

---

[70] There are no clear differences between the term dialogue and conversation, according to Oxford Dictionaries. Refer to the following sites: http://www.oxforddictionaries.com/us/definition/american_english/dialogue
http://www.oxforddictionaries.com/us/definition/american_english/conversation
(downloaded on February 14, 2016).

After a brief review, participants in groups 2 and 4 were taught how to express vertical spatial relationships in English with developed meaning of words in EFL through a PowerPoint presentation, Handout A_2, and ICSCTG. Needless to say, the critical point of language acquisition is to have students use visually stimulating expressions and to incorporate fun ---- acceptance and support from other participants and interlocutors. Thus, I organized playing a game with the binary semantic categorization accompanied by contact-noncontact relationships, which was certainly not a threatening activity.

They were given picture cards in a variety of contexts consisting of one sentence per card with developed meaning of words in EFL that has *on, over, above,* and *under.* Besides these, they were provided additional picture cards for English vocabulary that they had already learned or that they would learn for the first time in this class.

All of them were divided into approximately 24 teams which had around five to six members. I assigned a leader to each team to help students in her/his team comprehend and recall descriptions on the cards. After the procedures described above were completed, Groups 2 and 4 had different sessions with different psychological tools from those of Groups 1 and 3 until the game moved on to the next stage.

For instance, 10 picture cards with two different illustrations were given to Groups 2 and 4. The cards depicted vertical spatial relationships with images accompanied by expressions in both Japanese and English on the face of the card. The other side of the cards had the illustrations of hand movements of ICSCTG with a preposition in English and Japanese accompanied by contact-noncontact formations.

Each team had approximately 10 minutes total to learn and remember the expression on each picture card and then each team selected a representative for the team who came to the front of the classroom. The representative was asked to take a card out of 10 picture cards which the investigator held in her hands. Subsequently, the representative explained the vertical axis relationship on the card by using co-thought gestures to all the participants in class. If they felt that they could describe the vertical spatial structure on the card in English with the related co-thought gestures, they raised their hands.

Conversely, Groups 1 and 3 were distributed 10 picture cards which showed vertical axis structures with images and descriptions in both Japanese and English on the face of the card, but the back of it only had a description of a preposition in English and Japanese accompanied by a contact-noncontact configuration without an image of the gesture. Otherwise the cards were the same as those of Groups 2 and 4. Each team in Groups 1 and 3 was also given roughly 10 minutes total to comprehend and recall expressions on each picture card with their team members. When the game started, the investigator showed all of them in the class one picture card with an image which demonstrated vertical spatial relationships by using a computer, a screen and a projector. However, each picture projected on the screen had neither an expression of vertical relationships in English and Japanese nor the illustration of the hand movements of co-thought gestures. If participants thought that they would be able to express in English the vertical relationships on the picture card, they raised their hands.

From that point on, all groups had the same procedures in playing the games. For instance, the participant who raised her/ his hand first had the right to express the relationship. If she/ he described the vertical spatial relationships in English correctly, the team to which she/ he belonged gained a point. The investigator recorded the scores that each team received on a white board in the class. In this way, everybody in the class kept track of points on their team. Five minutes prior to ending the class, the investigator stopped the game, announced the winning team and subsequently summarized the class for that day.

- **Activities of Quadrant B**

The goal of quadrant B was to describe vertical spatial structures in context by using the proper developed meaning of words in EFL with their interlocutors, which leads to cognitively demanding activities with a context embedded dimension. During class, Groups 2 and 4 were asked to regenerate ICSCTG when they practiced dialogues, but not Groups 1 and 3.

After a brief summary of the previous class, all of them were asked to listen to the audio material used in quadrant C while viewing Handout B that was previously given to them, and then orally repeat each sentence in dialogues together.

All of them stood up and walked around in their classroom while music was played. When the music stopped, they identified their interlocutor to dialogue with in Handout B. The investigator instructed them not to look at the handout while they were conversing. However, if it was too difficult to do so, occasionally looking at it was allowed. They changed their interlocutor twice during the game at a minimum. Their teacher and the investigator helped them dialogue with their interlocutor if they had limited abilities to do so.

On that day, an assignment was given to all the groups. For instance, Groups 1 and 2 were asked to read expressions aloud in Handout A_1, A_2, and A_3. Groups 3 and 4 were asked to go to the computer laboratory at the high school to listen to audio material saved in a computer by looking at Handout A_1, A_2, and A_3. The audio materials had the same information as that of Handout A_1, A_2, and A_3. The investigator also provided Groups 3 and 4 with CDs, which had the same information as that of Handout A_1, A_2, and A_3 to allow them to listen to the audio material after school, if necessary. Groups 3 and 4 were instructed to repeat each expression in a dialogue silently when they listened to the audio material. Participants in Groups 1 and 2 were not given an opportunity to listen to the audio material during data collection. The investigator requested Groups 2 and 4 to recreate ICSCTG when they practiced expressions in audio materials and handouts which were given to them, according to their group identities.

- **Activities in Quadrant D**

Activities in quadrant D were omitted because they are too demanding for students who are at a beginners' level. According to Cummins (2000), communications in written speech at a sentence level such as the use of email can be too cognitively demanding and context-reduced for a beginning learner of English. Thus, activities for quadrant D were not organized.

In summary, this general research design was employed on the basis of theoretical and hypothetical framework. I argue that this general research design will help one explore to what extent teaching ICSCTG accompanied by listening practice can help Japanese EFL students learn vertical axis structures in English at a lexicon-grammatical level. Based on this assumption, the following includes a description of how variables in this study were gathered.

3.2.2.4.3 Experimental Treatment[71]

There were two experimental treatments: asking students to reproduce ICSCTG and assigning them to the listening practice homework. For instance, reproducing ICSCTG, was carried out with Groups 2 and 4 during the classroom intervention. After the classroom

[71] Refer to Appendix C for the definition of treatment and intervention in this study.

intervention was completed, Groups 3 and 4 were asked to listen to audio material in the computer lab and/ or outside of the computer lab while Group 4 regenerated the gestures which they learned during the classroom intervention.

The assigned listening practice took 15 minutes for them to complete. The investigator asked students who were not in the listening practice treatment groups to complete silent repetition practices with Handouts A_1, A_2, and A_3 which had already been distributed to them during the classroom interventions.

I argue that assignments should be given to all participants, but otherwise Groups 3 and 4 would have longer study time than Groups 1 and 2. Cooper et al. (1998/ 2006) claim that the more homework they complete, the higher their achievement. Furthermore, they pointed out, "… the optimum benefits of homework for high school students might lie between 1.5 and 2.5 hours [for six different subjects]" a day (Cooper et al., 2006, p. 52).

For this reason, I requested all of them to do the assignments given to them for 15 minutes and three times per week at a minimum until they completed the delayed post-tests in order for all groups to have a similar length of study time. Teachers at the high school have used the method of Read-Texts-Aloud for many years to teach their EFL learners, and thus, they were familiar with the reading aloud homework. In summary, on the basis of theoretical framework and literature review, the present study structures a research design with a quantitative method, including a non-experimental research design.

3.3 Data Collection Procedures

Two alternative stimuli were presented to the participants. These includes listening to audio material recorded by a native speaker of English asking questions in English and viewing the picture book with 20 different images which illustrate vertical axis relationships between two referents at a lexicon-grammatical level.

3.3.1 Stimuli

Two alternative stimuli were presented to them, which were listening to audio material recorded by a native speaker of English asking questions in English and viewing the picture book with 20 different images which illustrate vertical axis relationships between two referents at a lexicon-grammatical level.

3.3.2 Tasks for Oral Speech Tests

Participants were requested to answer 20 questions orally with appropriate developed meaning of words in EFL accompanied by proper SVO word order in English, by formulating oral speech. To produce their answers, they were requested to view one image at a time in the picture book containing the 20 different images, while listening to 20 alternative questions in audio material saved on a computer (See section 3.2.2.4.3). The audio material in the computer generated the questions one by one, while they were viewing the images. The students in Groups 2 and 4 were asked to recreate ICSCTG when they responded to the questions, but not Groups 1 and 3. Needless to say, Groups 1 and 3 were not prohibited from producing gestures because gestures might be spontaneously produced.

Twenty minutes were allowed to complete the tasks described above. Oral responses produced with or without ICSCTG were recorded by a SONY DCR-SX 45 S camcorder in a quiet room. Their performance was video-audio recorded without any identifiable facial feature. All of them were asked to complete tasks described above by taking pre-, post-, and delayed post-tests with or without ICSCTG. Following the completion of the oral tests, their oral responses were transcribed for further analyses. There was no written response for this task.

3.3.3Task for Cloze Tests

Another task for this study was a cloze-type sentence completion test. The cloze test in the current study consisted of a portion of text with prepositions removed. Participants were asked to replace the missing prepositions from a list located at the bottom of each page of the test. The test had 20 images, which showed the vertical spatial relationships of two referents with 20 different sentences in English and a missing preposition in each sentence.

3.3.4 Survey

The survey of the students has already been explained in the previous sections, including the purpose of the use of survey, the examples of independent variables gathered by the survey, and the categorization of independent variables collected by the survey. The survey was distributed to all of the participants in their classroom after the completion of the delayed post-tests. They were given enough time to complete the survey and were encouraged to answer all the questions in the survey which was written in their L1 (Japanese).

3.3.5 Data Collection Procedure

It took five months to complete the data collection, including gathering data for the pilot study. After the completion of the pilot study, I administered pre-, post-, and delayed post-tests for each group to determine participants' knowledge of how to express vertical spatial relationships with developed meaning of words in English at a lexicon-grammatical level. The tests were both oral and written. Post-tests which took place right after the classroom intervention allowed one to examine the effect of the classroom intervention and the experimental treatment of viewing and regenerating ICSCTG, which led to developing

ANCOVA model 1. Delayed post-tests were administered somewhere between 30 and 37 days from the first day of the classroom intervention, which helped this study create ANCOVA model 2.

In oral speech tests, their ID number was recorded by a video recorder, but not their name. Their behavioral data in these tests were digitally recorded by the device described above. All the data were transferred to two different external hard drives (My Book Essential 2 TB, USB 3.0 interface and My Book 500 GB made by Western Digital).

The pre-test took place in a classroom prior to the classroom intervention.[72] All tests were administered in a silent room individually. I asked students in the gesture with listening practice and gesture alone groups to reproduce ICSCTG, but not the control and Listening practice only groups. As explained earlier, I did not prohibit the control and Listening practice only groups from regenerating ICSCTG.

Participants' ID were recorded on each test, but not their names. The pre-test took place in their classroom together prior to the classroom intervention. They were asked to complete post- and delayed post-cloze tests individually in a different room immediately following the oral speech test for post- and delayed post-tests. The investigator did not ask all the participants to regenerate ICSCTG when they took the cloze tests. All tests were administered in a silent room individually except for the pre-test in the cloze tests.

3.4 Data Analysis Procedures

I discuss coding processes in grading their answers to questions in the cloze and oral speech tests generated by the investigator.

[72] Refer to Appendices J and K for the examples of the oral speech tests.

3.4.1 Coding

Complicated mental processes can occur when participants try to shape conceptual structures of vertical spatial operations with developed meaning of words in a certain type of EFL class. To quantify their answers to questions generated by the investigator, I explain correct versus incorrect answers based on the clarification of the coding in this study. Criteria of coding their answers are built upon research presented by Bowerman (1996a), Herskovits (1985), Huddleston and Pullum (2002), Lakoff (1987), Martin (1975), Tyler and Evans (2003), COCA, sociocultural theory proposed by Vygotsky (1987), and studies of positive and negative language transfer in SLA.

Transfer is defined as “influence resulting from the similarities and differences between the target language and any other language that has been previously (and perhaps imperfectly) acquired” (Odlin,1989, p. 27). Studies of language transfer examines how bilingual and/ or multilingual speakers apply their knowledge from one language to another language, including L1 (or L2, L3) interferences and crosslinguistic influences. In this study, discussions are limited to language transfer from Japanese to English. I take into account negative language transfer when coding and grading participants’ expressions in describing vertical spatial relationships, including positive and negative language transfer[73] (cf. Ellis, 2015; Gass & Selinker, 1992; MacWhinney, 2008, 2014; Odlin, 1989). This study examines whether language transfer affects both negative and positive influences on development of meaning of words in EFL (see Odlin, 1989 for negative and positive transfer)

[73] A term *interference*, that has also been used, has a disadvantage of leading one’s thought to the negative consequences of language transfer only. It also has behaviorist implications (Jarvis & Pavlenko, 2008). In this study, the term language transfer is employed.

The concepts of contact-noncontact distinctions are fundamental in the clarifications of a correct versus incorrect use of a prepositional phrase in the present study. ***Over*** is used for both contact and noncontact distinction when a TR is located above a LM and a TR is positioned under a LM. Whereas ***on*** is for contact only regardless of whether a TR is located above or under a LM. Conversely, ***above*** is not used in describing two referent relationships if a TR is located under a LM, even though ***above*** is employed for a contact-noncontact formation in a TRHLM condition. ***Under*** is utilized to express both contact-noncontact structures, when a TR is located under a LM only, but not above a LM. These are the basic differentiations of ***on, over, above,*** and ***under*** (cf. Kunisawa, under review for detail). Crucially, the influence of meaning of words in context is significant in the selection of proper prepositions.

3.4.1.1 Coding of Cloze Test on Correct Versus Incorrect

Participants were asked to fill in the blanks with a proper preposition for each expression matching to an image. They were given five prepositions to choose from: ***under, over, above***, ***in*** and ***on*** (***in*** was used as a distractor). There are various ways to express spatial relationships. However, this study codes an expression of vertical axis operations using published papers, books, and corpora as a model for expressions of beginning EFL students. For instance, of note, ***on*** should be taught to express that "apples are on the tree," suggesting that apples grow on the tree. There is an expression, "There are apples in the tree," which indicates, "There are apples in a hole of the tree" or "Someone placed apples in the tree." Tyler and Evans (2003) suggested that *on* is used for "broad definition of support which potentially includes notion of attachment" (p. 27). (cf. Kunisawa, under review). Bowerman (1996b) states, "apple on twig" (p. 153), by citing Herskovits (1986). "Lakoff (1987) stated, "In English, the basic spatial use of *on* make use of three image schema – CONCTACT, SUPPORT and ABOVE … which form a single conceptual unit" (p. 313). This study employs *on* to express a spatial relationship between a tree and an apple as a default expression in a Japanese EFL beginners' class, even though some English speakers use *in*. I argue that *in* cannot be a default expression in this case. Additionally, an expression, "The mold spread all on the celling," is not considered to be correct in this study, even though some Americans use this expression because COCA does not have an expression "spread all on," suggesting that this expression is not common in American English.

Table 5:

Cloze test

	List of questions	Correct answers
1	Clinton wears a hat ( ) his head.	on
2	He will nail a board ( ) the holes in the ceiling.	over
3	The clouds float ( ) the sun.	over under
4	The dogs are ( ) the tree.	under
5	The spider is walking ( )the ceiling.	on
6	She spread out the paper ( )the massage table.	over on
7	There are two boats ( ) the Golden Gate Bridge.	under
8	A dog jumps ( ) a barrel.	over
9	The airplane is now ( ) Colorado.	over above in
10	The International Space Station floats 200 miles ( ) the Earth's surface.	above over
11	The show was in the beautiful area one floor ( ) the ground level.	above
12	He put the soccer ball ( ) his right foot.	under
13	The water falls ( ) the rocks.	over on
14	A: "Do you want this hotdog?"	
	B: "No. I want the hotdo) it."	above
15	The flag is flying ( ) the White House.	over above
16	The plastic sheet I covering ( )the painted ceiling of the chapel during repair.	over under
17	The mold spreads all ( ) the ceiling.	over on
18	There is the special pad ( ) the wireless mouse.	under
19	The soccer ball is ( )the ground.	on
20	There are apples ( ) the tree.	on

If multiple options were possible, the answer was considered correct. For instance, 65 % can have only one response, 25 % two responses, and 10 % simply three responses. Additionally if they used the same preposition for all the blanks, all the answers were considered to be wrong, and if left blank, the participant' answer was considered to be a missing variable when data analysis was done.

Moreover, predicative expressions proposed by sociocultural theory are not discussed because the cloze tests in this study are not required in oral speech. Note that if negative transfer had occurred, students would have selected only one preposition for all questions (one of ***over, above,*** or ***on***), which did not occur in the cloze tests except for one participant.

3.4.1.2 Coding of Oral Speech Test on Correct and Incorrect

I present explanations of students' correct answers first and then follow with the explanations of incorrect ones in coding their responses. If they expressed vertical spatial relationships with an appropriate contact-noncontact distinction with a proper prepositional phrase that corresponds to a given image in each expression, it was naturally regarded to be a correct answer. I explain eight different categorizations of correct answers and subsequently explain nine alternative classifications of incorrect responses presented by participants in coding their utterances. These processes help with understanding development of meaning of words in EFL that entails TRHLM-TRLLM and contact-noncontact configurations.

3.4.1.2.1 Correct and Incorrect Answers

The following questions are examples of how I coded correct and incorrect answers in oral speech tests in this study. The sample questions in the oral speech tests:[74]

[74] A letter (b) is used for examples of definite versus indefinite articles in section 3.4.1.2.1.2 and letter (d) is for those of ellipses of transitive verb phrase in section 3.4.1.2.1.6.

(a) Where is the tea?

(c) The ceiling in this room does not look good. Are those handprints?

(e) Do you want this pizza?

(f) Do you know if our guests will come to our potluck on time?

(g) What is he doing?

(h) The cloud is on the mountain, isn’t it?

(i) Where is your room?

(j) Where can I get an orange?

(k) Can you take look at my car?

(l) Where is my cat?

(m) Where is your room?

3.4.1.2.1.1 Subject[75] and Copula Ellipses

Subject and copula ellipses in this study should be accepted if vertical axis relationships (TRHLM-TRLLM and contact-noncontact formations) are properly expressed. Two reasons account for this;

First, Vygotsky (1987) suggests that there was the tendency for ellipses in oral speech by citing Dostoevskii’s (1994). Additionally, Vygotsky (1987) states that:

> … the tendency for abbreviation and pure predicativity of expression arises in two circumstances in oral speech -- where the situation being referred to is clear to the interlocutors and where the speaker expresses the psychological context of his expression through intonation (p. 272).

[75] Li and Thompson (1976) proposed that the subject was determined by the predicate in subject-predicate constructions (459).

Second, as stated earlier, Japanese speakers use ellipses, these are common linguistic phenomena in face-to-face-dialogue. Furthermore, subject and copula ellipses are prevalent in face-to-face conversation in Japanese (Hasegawa & Hirose, 2005; Nariyama, 2003; Martin, 1975).

I argue that ellipses in conversational pragmatics demonstrate approximate meanings with a probable syntactic construction through which speakers intend to express a vertical spatial structure. Thus, I propose that an ellipsis to be viewed as important psychological proximity with which speakers can observe a specific referent to other referents. Additionally, an ellipsis psychologically makes a connection to semantics and syntax, which helps one with understanding the proximity of an unmarked linguistic construction.

Examples:

Question (a): Where is the tea?

Possible answers to this question are:

(a.1) *The tea is* ***on*** *the Table.* (correct/support)

(a.2) *The tea* ***on*** *the Table.* (the ellipsis of the copula/a predicative expression/support)

(a.3) ***On*** *the Table.* (the subject and copula ellipsis/a predicative expression/support)

For the reasons explained above, these answers are correct. The subject and copula ellipses are predicative expressions with the proper relationships between the TR and LM and contact between the two referents in describing vertical axis structures.

3.4.1.2.1.2 Non-obligatory Distinction between Singular Versus Plural Nouns, and Definite Versus Indefinite Articles

Even though negative transfer regarding non-obligatory distinction between singular versus plural nouns, and definite versus indefinite articles occurs, participants' answer are treated as an acceptable answer, if the lack of these terms does not affect expressing proper relationships between two referents (TRs & LMs) located in vertical space.

In support of this stance, I present contrastive analyses of English and Japanese on singular versus plural nouns and definite versus indefinite articles proposed by an American linguist of Japanese, Martin (1975).[76] Martin (1975) suggested that:

> In Japanese, as in many other languages of East Asia, nouns are characteristically vague as to number… Japanese is also free of the obligatory distinction of definite vs. indefinite that forces the English speaker…(p. 143).

That is, English usually requires a speaker to clarify whether each noun in her/his expression is singular or plural. For instance, a suffix produces *dogs* from *dog*, a vowel change turns *woman* into *women*, etc.

In Japanese, however, nouns are typically vague in terms of expressing plurality. English speakers will make an immediate distinction between *I need a pen* and *I need* (*some*) *pens*, yet in the same situations, the Japanese will simply say:

(b.1)	*Pen*	*ga*	*iru*
	Pen	NOM	need-NONP

This means that (I) need pen (s) without specifying the exact number of pens for his/her specific needs (Martin, 1972). Nonetheless, they have no difficulty when they wish to express plurality, and thus, an obvious device is that of explicit counting:

(b.2)	*Pen*	*ga*	*ippon*	*iru*
	Pen	NOM	CLS	need: NONP
		'I need one pen.'		
(b.3)	*Pen*	*ga*	*nihon*	*iru*
	Pen	NOM	CLS	need: NONP
		'I need two pens.'		

[76] Samuel Elmo Martin (1924 -2009) was a professor of Far Eastern Languages at Yale University who studied for a PhD in Japanese Linguistics under Bernard Bloch (1907–1965). Today, functional linguists in Japanese employ studies presented by Martin, including Ono and Iwasaki.

There is no singular or plural marker on the term *pen* in examples (b.1) and (b.3) noted above.

Additionally, as stated earlier, Japanese has the non-obligatory distinction of definite versus indefinite unlike English. For instance, English speakers say either I need a pen (some pens) or I need the pen(s), but, Japanese speakers do not always need to express definiteness; accordingly this usually goes untranslated.

An example of negative transfer regarding non-obligatory distinction between singular versus plural nouns, and definite versus indefinite articles in Japanese is:

Question (c): The ceiling in this room is not good. Are those handprints?

Possible answers are:

(c.1) *The handprints spread all* ***over*** *the ceiling.* (correct)

(c.2) *ØHandprintØ is* [sic] *all* ***over*** *the ceiling.* (acceptable)

Example (c.2) is considered to be acceptable because the contact of the two referents and the relationship of the TR and the LM are properly described, even though example (c.2) does not show the obligatory distinction of definite versus indefinite articles and has a vagueness as to number.

Thus, the lack of singular versus plural and definite versus indefinite configurations in participants' answers are treated as correct expressions, provided that the meaning of vertical spatial structures of the two referents is appropriately expressed.

3.4.1.2.1.3 Subject-Verb Agreement Error

Even though negative transfer of subject-verb agreement arises, an participant's answer is treated as acceptable if relationships between TRHLM-TRLLM and contact-noncontact distinctions are appropriately expressed. English has subject-verb agreement (Langacker, 1987).[77] However, Shibatani (1978) states, "there is no phenomenon that is comparable to the subject-verb agreement found in European languages" (p. 57) except for honorification.[78] An example of a subject-verb agreement error can be seen below

(c.3) *Handprint are all **over** the ceiling*. (Acceptable) (*Be* verb shows plurality, but the subject is singular)

In this study, honorification was not used to assess participants' ability of vertical spatial operations, and thus, a subject-verb agreement error can be ignored.

3.4.1.2.1.4 Ungrammaticality that is not Negative Transfer

An ungrammatical expression of the present progressive in this study should be accepted, even though Japanese has the progressive (or continuative) (Martin, 1975):

guest	*ga*	*hashi*	*o*	*aruite-iru*
guest(s)	NOM	bridge	ACC	walk-PROG:NONP

'Guest(s) is (are) walking over (on) the bridge.'

The participants' expression *The guests are walk **over** the bridge* should be accepted because vertical spatial structure for two referents (TR versus LM) is appropriately expressed with the differentiation of contact-noncontact even though the present progressive in the expression noted above lacks an affix, -*ing* for the verb *walk.* I am not testing the proper use of the present progressive. Instead, I am examining an appropriate use of a prepositional phrase in describing vertical spatial structures.

[77] Langacker (1987) stated, "what is traditionally known as subject verb agreement is analyzed as being a part of the grounding predication"(p. 247).

[78] Refer to Kindaichi (2011) for interesting discussions regarding a difference of subject-verb agreement between English and Japanese (p. xi).

3.4.1.2.1.5 Incorrect Pronunciation

The incorrect pronunciation of terms which does not affect expressing a proper vertical spatial relationship should be accepted as in **diractor* instead of *detector.* Vygotsky (1987) suggests, "The spontaneous use of phonetics (what is called pronunciation) is a [sic] extremely difficult for the school child who is learning a foreign language" (p. 221).

3.4.1.2.1.6 Ellipses of Transitive Verb Phrase

An ellipsis of a transitive verb phrase (VP) in this study should be accepted if the meaning of vertical spatial relationships is properly expressed (a correct TR and LM structure with contact-noncontact distinction) because English itself employs an ellipsis of a transitive VP. For instance;

Question (d) *Will you have dinner?*[79]

The answer with the ellipsis of a transitive VP is:

(d.1) *Yes**,** I will (have dinner)**.*** (a predicative expression)

Examples in this study:

Question (e): *Do you want this pizza?*

Possible answers are:

(e.1) *No. I want the pizza* ***above*** *it*. (correct)

(e.2) *No.* ***Above*** *it*. (correct, because even though the transitive VP ellipsis occurs, the expression is considered as correct due to the predicative nature of the expression in both English and Japanese.) Thus, it can be argued that the ellipsis of the transitive VP should be accepted because the unmarked TR (*the pizza*) is psychologically understandable and the predicative expression can be adequate in the expression noted above.

[79] This question is not in the oral speech test in this study.

3.4.1.2.1.7 Ellipsis of Progressive Auxiliary *Be*

Huddleston and Pullum (2002) classify six different ways of using ***be*.** In this study, out of the six categories, only copula ***be*** and progressive ***be*,** and ellipses of these are discussed. The progressive is constructed with the progressive auxiliary ***be*** and the *–ing* form of the verb[80] (Heny, & Richards, 1983). In the present study, an ellipsis of progressive auxiliary ***be*** *in* a vertical axis operation should be considered appropriate because of the predicative nature of oral speech, if the structure of a TRHLM-TRLLM and contact-noncontact alignment is properly expressed. Examples:

Question (f): Do you know if our guests will come to our potluck on time?

Participants' answers:

(f.1) *Guests are walking* ***on/over*** *the bridge.* (correct / the progressive auxiliary ***be*** + the participle. COCA downloaded on September 4, 2016.)

(f.2) *The Guest walk* **on** *the bridge.*[81] (correct)

Even though the ellipsis of the progressive auxiliary ***be***, the lack of the affix, *-ing* in the present participle, and the absence of subject-verb agreement are observed, this expression can be treated as a correct answer in this study because the configuration of a TRHLM-TRLLM and contact-noncontact alignment is properly expressed. Additionally, I am not testing whether they were able to employ the proper use of the progressive ***be*.** Furthermore, as explained earlier, Japanese does not have subject-verb agreement unlike English, and thus, *The guest walk on the bridge* should be accepted as a correct answer (see section 3.4.1.2.2.1 for a discussion about progressive copula ***be*** versus progressive auxiliary ***be***).

---

[80] That is a present participle, but not gerund.

[81] Participants' answer ***Over*** *the bridge* is incorrect. See later in this chapter.

3.4.1.2.1.8 Code-Switching:

In linguistics, code-switching arises when a speaker switches between different languages in a context of a single conversation. Code-switching is a common linguistic phenomenon in everyday life. Researchers have presented studies about code-switching for decades. However, it is largely unknown why code-switching occurs. In this study, if code-switching does not affect a proper expression of vertical spatial relationships, the code-switching is allowed. Examples:

Question (h): The cloud is on the mountain, isn't it?

Participant' answer:

(h.1) *The cloud is on janai (not)* ***over*** *the mountain.* (correct)

Question (c): The ceiling in this room does not look good. Are those handprints?

Participant's answer:

(c.4) *Tegata (handprints) is all* ***over*** *the ceiling.* (correct)

Expressions noted above show the code switching that demonstrates proper TRHLM-TRLLM and contact-noncontact structures, and for this reason the expressions with the code switching are acceptable.

3.4.1.2.2 Incorrect Answers

3.4.1.2.2.1 End-point Focus and Copula *be* Versus Progressive Auxiliary *be*

This section discusses issues of end-point focus that involves the differences between copula ***be***[82] versus progressive auxiliary ***be.*** The structural function of copula ***be*** differs from the use of the progressive auxiliary ***be***, where ***be*** links to *–ing* to qualify the action denoted by a main

[82] Huddleston and Pullum (2005) proposed that ***be*** was an auxiliary. However, traditional grammar did not explain syntactic criteria for determining what verbs were auxiliaries in English and accordingly the membership of the class of ***be*** was ill defined and varied from one grammar to another.

verb and that has ***be*** as an aspect marker (Huddleston & Pullum, 2005). I first explain what end-point focus is as, proposed by Lakoff (1987) and then follow with explanations of progressive auxiliary ***be*** versus copula ***be*** in the current study.

Lakoff (1987) suggested that ***over*** in an expression "*Sausalito is* **over** *the bridge* has the sense of "on the other side of" as a result of end-point focus" (p. 424).

Examples in this study:

Question (f): Do you know if our guests will come to our potluck on time?

Participants' answers:

(f.3) *The guests are walking* ***on/ over*** *the bridge*. (correct)

(f.4) **Guests* <u>*are* ***over***</u> *the bridge*[83] (incorrect because of end-point focus that does not mediate TR-LM configurations).

Question (f) is to ask whether the guests are coming (their action) with the progressive auxiliary ***be*** + the *present participle.* The answer (f.3) has the progressive auxiliary ***be*** + the present participle with the proper TRHLM-TRLLM and contact-noncontact formation and is therefore correct.

However, the response (f.4) has the copula ***be,*** but does not have the progressive auxiliary ***be*** + the present participle, resulting in end-point focus that the subject (the TR) is on the other side of the bridge (the LM), but not walking (***on/ over)*** the bridge, which does not correspond to a given image, and thus, the answer is incorrect. The copula ***be*** versus the progressive auxiliary ***be*** distinctions are crucial to answering the question properly. The following example also demonstrates the crucial factor of the progressive auxiliary ***be*** versus copula ***be*** distinctions.

[83] The asterisk indicates wrong.

Question (g): What is he doing?

Participants' answer:

(g.1) *He is jumping over the fence*. (correct/cf., Tyler & Evans, 2003)

(g.2) **He is over the fence*. (end-point focus)

The answer (g.2) has the copula ***be,*** but does not have the progressive auxiliary ***be*** + the present participle, resulting in end-point focus that the subject (the TR) is on the other side of the fence (the LM), but is not in the process of jumping over the fence (the TR), which does not correspond to a given image, either. Therefore the answer is incorrect. For this reason, both answers (f.4) and (g.2) are incorrect.

3.4.1.2.2.2 Erroneous Vertical Spatial Meaning

This section includes coding criteria regarding questions (f) and (g), and answers to these questions. It focuses on the proper use of *on* to elucidate incorrect uses of *on* and *over*, which involves end-point focus, and progressive auxiliary ***be*** versus copula ***be*** distinctions.

Examples:

Question (f): Do you know if our guests will come to our potluck on time?

Participants' answers:

(f.5) *Yes. The guests are* ***on*** *the bridge*. (correct/the copula ***be*** + the prepositional phrase, but not an ellipsis of the present participle). Although, the question asks them an action of the TR, using copula ***be*** is appropriate in providing the correct answer because it does not create end-point focus in the response.

(f.6) ****Over*** *the bridge*.

This expression shows end-point focus, and thus, is incorrect.

Therefore, (f.5) is correct, but (f.6) is not.

Question (g): What is he doing?

Participants' answers:

(g.3) **He is **on** the fence*.

This expression consists of the copula **be** + the prepositional phrase, which is an improper meaning of the preposition ***on*** (contact). The image (noncontact) does not correspond to this answer, and thus, this can lead example (g.3) to an erroneous expression of TR and LM relationship.

(g.4) ****Over*** *the fence*.

Example (g.4) has endo-point focus resulting in the erroneous answer.

Concisely, example (f.5) is correct, even though it lacks the progressive auxiliary ***be*** and the present participles because it has the proper TRHLM-TRLLM and contact-noncontact distinction, which does not demonstrate end-point focus. Additionally, (f.5) expresses the proper vertical axis relationships of two referents despite the fact that it has copula ***be.*** However, participants' answers (f.6), (g.3), and (g.4) are incorrect because of the lack of proper contact-noncontact distinctions and improper use of end-point focus.

3.4.1.2.2.3 Incorrect Use of Prepositional Phrase

If students express vertical spatial relationships with incorrect contact-noncontact distinctions, it is an error, despite the fact that TRHLM-TRLLM structures of two referents are properly described. Examples:

Question (h): Do you know if our guests will come to our potluck on time?

(h.2) *The cloud is **over/ above** the mountain*. (correct)

(h.3) **The cloud is **on** the mountain*. (a wrong contact-noncontact distinction)

Question (i): Where is your room?

Participants' answer:

(i.1) **My room* ***on*** *the room.*

Above is proper, which expresses "The Next-one-up Sense (3) In this sense, above relates to the next one up in a vertical sequence" (Tyler & Evans, 2003, p. 120), but ***on*** does not. Tyler and Evans (2003) clearly differentiate between ***above*** and ***on***. For this reason, the expressions marked with an asterisk noted above are incorrect.

3.4.1.2.2.4 Wrong Word Order

This section explains negative transfer of word order in a Japanese EFL class. Even though they expressed contact-noncontact configurations, if they were unable to create proper word order (appropriate TRHLM - TRLLM alignments), it is an error. Two reasons account for this. First, wrong word order creates the erroneous meaning of vertical spatial relationships between two referents, which does not correspond to images given for oral speech tests (wrong TRHLM-TRLLM distinctions). Second, Comrie (1989a) stated that English had a very high correlation between grammatical relations and word order, suggesting that word order is a basic grammatical relation as opposed to Japanese in which it is not. Examples:

Question (j): Where can I get an orange?

Participants' answer:

(j.1) **The tree is on the orange.*

Question (f): Do you know if our guests will come to our potluck on time?

Their response:

(f.7) **The bridge on the people*.

Question (k): Can you take look at my car?

Participants' answer:

(k.1) **Car bottom is over the oil.*

(k.2) **Oil is under the car bottom.*

Question (l): Where is my cat?

Their response:

(l.1) **Desk under the cat.*

The answers marked with asterisks do not parallel the images of the test, although contact-noncontact distinctions are described. Therefore, asterisked answers are wrong because of inappropriate word order (erroneous TRHLM-TRLLM alignments). As stated earlier, Japanese has highly flexible word order, whereas word order is crucial in English, and thus, this type of negative transfer can occur in a Japanese EFL class (see Appendix H).

3.4.1.2.2.5 Ellipsis of Preposition

This section explains negative transfer of a postpositional structure in Japanese as a result of lacking prepositions in encoding vertical axis structures (see Appendix H). English speakers usually use a prepositional phrase in describing vertical relationships, although there is a case that they employ an unexpressed preposition as in "He is seated."[84] Japanese also has ellipses (pragmatic inference) that are prevalent throughout face-to-face dialogue. However, in the present study, if the ellipses of the prepositions occur in expressing vertical spatial relationships, it is incorrect inasmuch as I am examining the proper use of prepositional phrases. If students use unexpressed prepositional phrases, the inevitable consequence is that development of meaning of words cannot take place through expressions of vertical spatial structures. Besides, vertical spatial relationships between the two referents cannot be

[84] COCA shows an ellipsis in expressing vertical axis relationships in English. For example, "Please, **<u>be</u>** **<u>seated</u>**" that implies "Please be seated <u>on</u> the chair," "Please be seated <u>on</u> the sofa," or "Please be seated <u>on</u> the bench." (Downloaded on January, 30, 2016).

expressed properly, leading to improper TRHLM-TRLLM and contact-noncontact formations. Examples:

Question (e): Do you want this pizza?

Their response:

(e.3) *I want the pizza.*

Question (k): Can you take look at my car?

Participants' answers:

(k.3) *Oil is bottom the car*.

(k.4) *Oil is bottom of the car.*

Question (m): Where is your room?

Their response:

(m.1) *My room is one floor my room*.

Consequently, the ellipses of the prepositions in the present study are incorrect.

3.4.1.2.2.6 Wrong Subjects

If wrong subjects are used in expressing vertical spatial relationships, answers are regarded to be incorrect because this leads to unclear and/ or leads to erroneous vertical axis relationships between two referents (uncertain TRHLM-TRLLM distinctions). Examples:

Question (g): What is he doing?

Participants' answers:

(g.5) **It is skate board over the fence.* (end-point focus and the unclear subject)

(g.6) **Skateboard is over the fence*. (end-point focus and the wrong subject)

(g.7) **It is over the fence.* (end-point focus and the unclear subject)

The three expressions described above do not explain the vertical relationship between the boy (the TR) and the fence (the LM), but rather the skateboard and the fence. Additionally, an interlocutor does not understand what the subject *it* refers to in this context. Furthermore, the image of this scene does not indicate end-point focus.

Question (c): The ceiling in this room does not look good. Are those handprints?

Their response:

(c.5) **Hand is on the ceiling*.

The expression noted above does not describe a vertical spatial relationship between the ceiling and handprints, instead the expression explains a vertical spatial relationship between someone’s hand and the ceiling of which an interlocutor does not ask participants. Consequently, the expression noted above is incorrect. Therefore, if incorrect subjects are used, it is treated as a wrong answer.

3.4.1.2.2.7 Inconclusive Selection of Preposition

Question (e): Do you want this pizza?

Participants’ answer:

(e.3) **I want over the next above it*.

By using the multiple prepositions in the sentence noted above, the answer becomes unconvincing and incomprehensible for the interlocutor. Additionally, the transitive VP is improperly expressed. Consequently, the sentence noted above is incorrect.

3.4.1.2.2.8 Wrong Meaning

Even though a preposition is properly expressed in explaining a TRHLM-TRLLM and contact-noncontact formation, if a prepositional phrase has a wrong meaning, it is incorrect.

Examples:

Question (m): Where is your room?

Participants' answers:

(m.1) **My room is above you.*

(m.2) *My room is one floor above my friends' room.* (correct)

An image in the test suggests that the speaker's room is located one floor above his/her friend's room, but the room is not located above the interlocutor. Thus, the expression marked with the asterisk is incorrect, wherein a term *you* refers to the interrogator.

3.4.1.2.2.9 Improper Use of Genitive Noun Phrase

An improper use of a genitive noun phrase is not accepted as a correct answer. An example:

Question (k): Can you take look at my car?

Their response:

(k.5) **The oil is all over the bottom's car.*

The expression noted above does not refer to the bottom of the car. It implies the possession of the car by *the bottom*, which does not properly respond to the question in this study and is therefore incorrect.

In summary, based on studies suggested by Bowerman (1996a), Herskovits (1985), Huddleston and Pullum (2002), Lakoff (1987), Martin (1975), Tyler and Evans (2003), COCA, sociocultural theory proposed by Vygotsky (1987), and studies of language transfer (positive and negative transfer in SLA). I coded participants' responses and graded correct and incorrect answers in cloze and oral speech tests. These can help one assess whether students have achieved proper development of meaning of words in this study or were still in the process of learning development of meaning of words. Chapter 5 examines whether the Gesture Listening Higher Concept Approach (a part of support processes in MacWhinney's terms, 2013) can help Japanese EFL students with development of their language skills at a lexical and syntactic level (developed meaning of words in EFL), based on data analyses presented in Chapter 4.

3.5 Statistical Methods

To test the hypotheses and research questions in this study, the investigator transcribed all data obtained from participants' oral speech tests and then coded their answers in all of the oral speech and cloze tests to assess knowledge of developed meaning of words in EFL based on criteria discussed in the previous sections. The results of assessments of their answers in pre-, post-, and delayed post-tests in oral speech and cloze tests, scores in standardized tests, and survey variables were entered in an Excel file to run statistical analyses. Those statistical analyses are ANCOVA, a repeated-measures ANOVA, and T-Tests (see Appendix E for all the variables in this study).

3.5.1Variables without Influence of Classroom Intervention

ANCOVA does not accept variables which were influenced by the classroom intervention (Clark-Carter, 2009). However, I argue that a survey variable such as "students' motivation of learning English grammar to obtain a better job offer in the future", can be acceptable as an independent variable, even though the investigator collected this variable after the classroom intervention. Four reasons account for this:

First, the 150 minutes classroom intervention that the investigator used cannot influence participants' motivation of learning English grammar in obtaining a better job offer in the future. Second, the Ministry of Education, Culture, Sports, Science and Technology in

Japan emphasizes the importance of teaching English grammar.[85] Third, a high standard in English grammar for college admission is crucial, and likewise a majority of well-respected colleges in Japan require students to take English examinations that test their English grammatical knowledge. Fourth, a college degree has become increasingly necessary for success in obtaining a better job offer in Japan.

In short, learning English grammar is linked to college admissions and to better job opportunities, and thus, participants in this study had been in an environment whereby they were motivated to learn English grammar. Therefore, the investigator used their motivation of learning English grammar as a covariate for the two ANCOVA models.[86] The following includes an explanation of how the investigator used variables described above to conduct statistical testing.

3.5.2 ANOVA, Repeated-Measures ANOVA, ANCOVA, Paired Sample T-Test, and Graphical Data Analysis

To assess the effect of the Gesture Listening Higher Concept Approach, the investigator selected ANCOVA model 1 for the post- test and ANCOVA model 2 for the delayed post-tests with higher model accuracies, satisfying no serious multicollinearities/ collinearities, and the statistical assumptions of normality and of constant variance of residuals (see Appendix D for model selection processes).

ANCOVA is frequently used when random assignments of participants to groups are not possible as in educational research (Burns & Burns, 2008). There are multiple steps in

[85] http://www.mext.go.jp/component/a_menu/education/micro_detail/__icsFiles/afieldfile/2010/01/29/1282000_9.pdf (downloaded on October, 30, 2016).

[86] Ordinal variables measured by Likert scales in the survey can be used as covariates (Lesaffre &Senn, 2003; Tsangari, & Akritas, 2004; Schacht, Bogaerts, Bluhmki & Lesaffre, 2008).

which one should examine the proper ANCOVA models, prior to determining the final ANCOVA models (cf. Appendix D for detail).

The ANOVA and the repeated-measures ANOVA can also be used to further examine their test scores for which ANCOVA has a limited ability. For instance, the ANOVA was performed to find differences in initial group means among the four groups regarding their scores in standardized tests in L1 and EFL, and the survey variables which had not been influenced by the classroom intervention. Similarly, the repeated-measures ANOVA was implemented to examine statistically significant difference in scores between pre- and post-test, and post- and delayed post-test in both oral speech and cloze test conditions.

The investigator also used the paired-samples t-test to measure the means in two of the groups when the samples were the matched pairs and the ANOVA did not show a group difference (McMillan & Schumacher, 2009). Additionally, the investigator employed a graphical data analysis technique when statistical analysis was not possible (Unwin, 2015).

3.5.3 Initial Group Difference

The investigator performed ANOVA to identify initial group differences among the four groups. The ANOVA revealed that the means of all four groups significantly differed from each other in their L1 test scores as shown by Table 6 (n=125, M=53.24, SD=14.288, p = .002).

Table 6:

Standardized Test in L1

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	2960.933	3	986.978	5.343	.002*

Within Groups	22351.867	121	184.726
Total	25312.800	124	

*p <.05, **p <.001

However, all independent variables had no statistically significant group difference, except for those of their L1 score. For instance:

1. Participants' EFL standardized test score: (n=125, M=43.224, SD=9.4, p < .560).
2. Their pre cloze test scores: (n=126, M=.48, SD=.13112, p = .681).
3. Their oral speech pre-tests: (n=126, M=.15084, SD=.13194, p = .571).
4. Their motivation of learning English grammar (n=126, M=2.176, SD=.6099, p = .065).

The two ANCOVA models in the present study have three covariates (groups,[87] the pre oral speech test scores, and the motivation of learning English grammar), which do not have the initial group differences with statistical significance. The most common reason for using the ANCOVA models is to control for individual differences related to the outcome measure. Even though the three covariates do not have statistically significant group mean differences, these three covariates had a confounding effect on outcomes in the two ANCOVA models.

3.6 Summary

The methods outlined in this chapter include clarification of a comprehensive framework within which to test the hypotheses and to investigate the research questions addressed in Chapter 1. Accordingly, effects of the Gesture Listening Higher Concept Approach can be investigated at a high school level in Japanese EFL education with inferential analyses on theoretical and experimental bases, which leads to achieving the goals in the current study. Chapter 4 presents results of data analyses.

[87] Groups are classified as both a covariate and a fixed factor.

Chapter 4

Results

4.1 Introduction

This chapter examines Hypotheses 1, 2, and 3, and Research Questions 1 and 2 by analyzing test scores generated by four different groups with statistical techniques. These groups are the control (Group 1), gesture only (Group 2), listening alone (Group 3), and concurrent use of ICSCTGs and listening (Group 4). The statistical techniques are ANCOVA, the repeated-measures ANOVA, T-tests, and the graphical data analysis. Participants' test scores are obtained by pre-, post-, and delayed post-tests both in oral speech and cloze test conditions.

4.2 Outlines of Outcomes

Section 4.2 examines whether the experimental results indicate that teaching students ICSCTG (Research Question 1) and implementing listening practice (Research Question 2) facilitates learning how to express vertical spatial relationships in English for Japanese EFL high school students (Hypotheses 1), as well as whether the combination of these pedagogical interventions is better than the use of ICSCTG alone (Hypothesis 2).

Based on results defined by the statistical techniques noted above, section 4.2.1 examines Research Question 1 and Hypothesis 1 and then investigates Research Question 2 and Hypothesis 2 together, and subsequently explores Hypothesis 3. Lastly, this section includes a statistical analysis of the positive correlation of two test scores which do not directly link to the hypotheses and the research questions for this study.

4.2.1 Hypothesis 1 and Research Question 1 in Pre- and Post-Test Oral Speech Condition

Hypothesis 1 posits that ICSCTG facilitates Japanese EFL students' learning of how to express vertical spatial relationships in English? As shown in Figure 6-2, the gesture groups (Groups 2 & 4) demonstrated significantly better performance on the oral speech task post-test than the non-gesture groups (Groups 1 & 3) despite the fact that all groups performed similarly on the pre-test (see Table 15). The gesture groups scored 59% correct on average on the post-test, whereas the non-gesture groups scored only 45% correct on average [F (3, 119) = 3.302, p = .023, Partial η = .077,[88] see Table 7-1]. Likewise, improvement from the pre-test to the post-test was significantly greater for the gesture groups than for the non-gesture groups [F (1, 124) = 7.435, p = .007, Partial η = .057, see Table 10 and Figure 6-2].

Therefore, ICSCTGs alone facilitated learning about vertical axis structures, which supports Hypothesis 1 and answers Research Question 1 in the post-oral speech test setting and between the pre- and post-oral circumstance (an effect of time).

4.2.2 Hypothesis 1 and Research Question 1 in Post- and Delayed Post-Oral Speech Setting

Between the post- and the delayed post-oral speech setting, outcomes generated by the repeated-measures ANOVA and Figure 6-2 demonstrates that participants' maintenance of the newly developed knowledge of vertical spatial structure in English was evaluated with a delayed post-test. There was no evidence of a significant loss or gain between the post-test and the delayed post-test for the gesture-only group (Group 2). Their average performance

---

[88] Partial η^2 (effect size or ES): 0.01: small, 0.06: medium, 0.14: large (Cohen, 1969/1988; Carlson & Winquist, 2013; Lauson-Hall, 2015; Richardson, 2011). Partial η^2 quantifies the overall effect of the interaction on the dependent variable.

was 57% on both tests (0.3% mean decrease, Wilks' Lambda[89] = 1.0, F (1,31) < 1, n.s.). Accordingly, Hypothesis 1 was also supported, answering Research Question 1 in the post- and the delayed post-oral speech setting.

4.2.3 Hypothesis 2 and Research Question 2

I hypothesized that the concomitant use of ICSCTGs and listening practice would have an effect on learning about vertical spatial structures more than either treatment alone (Hypothesis 2 that corresponds to Research Question 2). The effect of the different intervention types on the oral post-test was evaluated first with an ANCOVA comparing performance of the four groups, and subsequently, with a repeated-measures ANOVA to determine whether there were changes in performance from the post-test to the delayed post-test. Results of both analyses support Hypothesis 2.

First, in a delayed post-oral speech test condition, Group 4 outperformed the other three groups (control, gesture only, and listening groups) [$p < .001$, F (3, 118) = 7.205, Partial $\eta^2 = .155$], controlling for the three covariates. (See to Table 8-1 and Figure 6-2). That is, the concurrent use of these two techniques was the most effective compared to either method alone (gesture only & listening practice alone).

Second, between the post- and the delayed post-oral speech setting, outcomes of the repeated-measures ANOVA suggested that Group 3 enhanced their capability to express vertical spatial operations with statistical and practical significance [$p < .05$, Wilks' Lambda[90] = .851, F (1, 29) = 5.092, Partial $\eta^2 = .149$] (Table 14-3).

---

[89] Wilk's lambda indicated the proportion of variance in the combination of dependent variable; lesser its value more is the variation in the group mean (as cited in Verma, 2015). APA (2010) indicated the use of Wilks' lambda.

[90] Wilk's lambda indicates the proportion of variance in the combination of dependent variable; lesser its value more is the variation in the group mean (Verma, 2015). APA (2010) indicated the use of Wilks' lambda.

Nevertheless, between the post- and delayed post-oral speech condition, outcomes of the repeated-measures ANOVA also suggested that Group 4 was superior to the other three groups as well [$p < .001$, Wilks' Lambda = .639, F (1, 31) = 18.086, Partial $\eta^2 = .368$]. This suggests that Group 4 outperformed the other three groups not only in the delayed post-oral speech test condition, but also made greater gains than all the other groups between the post- and the delayed post-oral speech setting, including Group 3 (Table 14-4).

Thus, Hypothesis 2 was supported, leading to answering Research Question 2 that teaching EFL students in Japan about vertical space by using ICSCTG and listening practice together is more effective than using ICSCTG alone.

4.2.4 Hypothesis 3

A repeated-measures ANOVA was used to investigate Hypothesis 3, namely, to determine whether learning gains made through the use of ICSCTG and listening practice are maintained over time by comparing performance of the four groups on the post-test to a delayed post-test, administered 30 to 37 days after the post-test.

Table 14-4 generated by the repeated-measures ANOVA [$p = .000$, Wilks' Lambda = .632, F (1, 31) = 5.092, Partial $\eta^2 = .368$] and Figure 6-2 demonstrated the continuous enhancement of the capacity of participants in Group 4 in expressing vertical axis operations between the post- test and the delayed post-test. Indeed, Group 4 improved performance between the post-test and the delayed post-test significantly more than all three other groups. For the reason explained above, Hypothesis 3 was supported between the post- and the delayed post-oral speech condition. (See Tables 14 S & 14-4).

4.2.5 All Groups Learned and Maintained Knowledge

The following includes an explanation of how statistical data analyses display all four groups learned and maintained knowledge of developed meaning of words, including the control group that did not directly links to an effect of the concomitant use of ICSCTGs and listening practice, ICSCTGs alone, and listening practice only on learning about developed meaning of words in EFL. This led to discovering an additional effect on learning about it.

The non-gesture groups (Groups 1 & 3) learned about developed meaning of words with statistical and practical significance both between the pre- and post-oral speech settings. For instance, Figure 6 and the repeated-measures ANOVA suggested that there was a significant effect of time on learning about it both between the pre- and post-oral speech test conditions [$p < .001$, Wilks' Lambda = .398, $F(1, 60) = 90.580$, Partial $\eta^2 = .602$] (Table 11-1), and between the pre- and post-cloze tests [$p < .001$, Wilks' Lambda = .641, $F(1, 60) = 33.653$, Partial $\eta^2 = .359$] (Tables 12-1, 12-2, & 12-3).

Beside, between the post- and delayed post-oral speech test circumstance, Group 1 (5.2 % increase of mean cf., Table 14-1-1), Group 2 (0.3 % decrease of mean cf., Table 14-2-2), and Group 3 (7.8 % increase of mean cf., Table 14-3-2) retained knowledge of it without statistical significant memory decay. However, Group 4 outperformed all the other three groups (18.3% increase, $p <.0001$, partial $\eta^2 = .368$) with statistical and practical significance and Group 3 was superior to Groups 1 and 2 ($p < 05$, partial $\eta^2 = .149$) with statistical and practical significance (cf., Tables 14 S, 14-3, & 14-4).

Between the post- and delayed post- cloze test circumstances, all four groups retained knowledge of developed meaning of words in EFL without a statistically significant decay in recall (*p* =.139, 2.3% increase of mean on average) [*p* = .139, Wilks' Lambda = .982, *F* (1, 122) = 2.213, Partial η^2 = .018], including control group. (See Table 13 and Figure 6-1).

Therefore, all four groups learned developed meaning of words in EFL and maintained knowledge of it for at least 30-37 days both oral speech and cloze test conditions, regardless of group identities. Of note, even though all groups learned developed meaning of words in EFL between a pre- and post-cloze test setting with statistical significance and practical significance, an effect of time did not differ among groups (*p* =.451) (cf., Table 13-4**)**. Why did all four groups learn and preserve knowledge of it? A possible answer could be given by the following data analysis.

4.2.6 Positive Correlation

Table 15 and Figure 6 demonstrate that as oral speech test scores in all four groups improved, cloze test scores in the groups also increased, suggesting that there was a positive correlation between the two alternative test scores measured by two different assessments, and thus, a possible reason should be investigated in this regard (See Chapter 5).

In short, based on results determined by the statistical methods in this study, Hypotheses 1, 2, and 3 were supported and Research Questions 1 and 2 were answered in oral speech settings. Additionally, the graphical data analyses, and the repeated-measures ANOVA revealed that there should have been another effect on learning about developed meaning of words in EFL, in addition to the effects of the concurrent ICSCTGs and listening practice, ICSCTGs only, and listening practice alone.

Importantly, the survey variables did not have a statistically significant effect on generating the ANCOVA models both in the oral speech and cloze tests settings, although the ANCOVA models determined the participants' motivation of learning English grammar as a covariate that had a confounding effect. Additionally, there is no significant effect of their L1 and EFL on learning about vertical spatial operations in this study (Table 7-1 & in Table 8-1).[91] Lastly, of note, empirical evidence that supports Hypotheses1, 2, and 3 is unconvincing in the cloze test settings. Thus, the cloze test results do not inform the research questions as expected. The following summarizes the relationships between variables in the two ANCOVA models, the repeated-measures ANOVA, and paired sample t-tests in this experimental study.

Table 7: *Oral Speech Tests*

Table 7-1:

Outcomes Defined by ANCOVA Model 1 Tests of Between-Subject Effects (Dependent Variable: Post-Oral Speech Test)

Source	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial η^2
Corrected Model	2.710[a]	5	.542	7.839	.000***	.248
Intercept	2.857	1	2.857	41.313	.000***	.258
Pre-oral speech	1.386	1	1.386	20.038	.000***	.144
ENG grammar helps with future job	.481	1	.481	6.956	.009**	.055
Group	.685	3	.228	3.302	.023*	.077
Error	8.228	119	.069			
Total	44.935	125				
Corrected Total	10.939	124				

a. R Squared = .248 (Adjusted R Squared = .216)
b. * $p < 0.05$, ** $p < 0.01$ *** $p < 0.001$

[91] Test scores in L1 and EFL did not show up as a covariate in ANCOVA models 1 and 2.

Table 7-2:

Results Determined by ANCOVA Model 1 Pairwise Comparisons (Dependent Variable: Post-oral Speech Test)

(I) Group	(J) Group	Mean Difference (I-J)	Std. Error	Sig. [b]	95% Confidence Interval for Difference [b] Lower Bound	Upper Bound
(1) Control	(2) Gesture	-.154*	.069	.027	-.289	-.018
	(3) Listening	-.027	.069	.699	-.164	.111
	(4) Gesture with listening	-.169*	.067	.012	-.301	-.037
(2) Gesture	(1) Control	.154*	.069	.027	.018	.289
	(3) Listening	.127	.068	.064	-.008	.261
	(4) Gesture with listening	-.016	.066	.813	-.146	.115
(3) Listening	(1) Control	.027	.069	.699	-.111	.164
	(2) Gesture	-.127	.068	.064	-.261	.008
	(4) Gesture with listening	-.142*	.067	.036	-.275	-.010
(4) Gesture with listening	(1) Control	.169*	.067	.012	.037	.301
	(2) Gesture	.016	.066	.813	-.115	.146
	(3) Listening	.142*	.067	.036	.010	.275

4.3 Cloze Tests

This study conducted one-way ANCOVA to test a statistical significance of the difference among group means in post- and delayed post-cloze test situations. The same methods as those of oral speech were employed in selecting an ANCOVA model for cloze tests, however no statistically significant difference in group mean resulted. (See Chapter 5 for a discussion). The summaries of ANCOVA models in cloze tests are explained in Tables 9-1 and 9-2 below by differentiating results in a delayed post-cloze test (Table 9-1) from those of a post-cloze test condition (Table 9-2).

4.4 Summary

In sum, the participants' performance on the oral speech task provided evidence that Japanese students in an EFL classroom not only learned, but also maintained and further enhanced their knowledge of how to express vertical spatial relationships, if they employ ICSCTGs and listening practice together. Additionally, the use of ICSCTG alone and of listening practice alone also facilitated not only learning, but also preserving knowledge of how to describe vertical axis operations.

Furthermore, outcomes defined by these statistical methods and Figure 6 explained above uncovered that all groups learned and retained knowledge of developed meaning of words in EFL both in oral speech and cloze test settings. Additionally, the positive correlation of the two alternative test scores was revealed. However, empirical evidence that supports Hypotheses 1, 2, and 3 is unconvincing in the cloze test situations, and thus, no additional support for the hypothesis was gained from the results.

Table 8: *Oral Speech Tests: Outcomes Defined by ANCOVA Model 2*

Table 8-1:

Tests of Between-Subject Effects (Dependent Variable: Delayed Post-oral Speech Test)

Source	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial η^2
Corrected Model	3.130[a]	5	.626	8.205	.000***	.258
Intercept	3.495	1	3.495	45.809	.000***	.280
Pre-oral speech	.925	1	.925	12.125	.001**	.093
ENG grammar helps with future job	.338	1	.338	4.434	.037*	.036
Group	1.649	3	.550	7.205	.000***	.155
Error	9.002	118	.076			
Total	57.058	124				
Corrected Total	12.131	123				

a. R Squared = .258 (Adjusted R Squared = .227)

Table 8-2:

Results Determined by ANCOVA Model 2 Pairwise Comparisons (Dependent Variable: Delayed Post-oral Speech Test)

(I) Group	(J) Group	Mean Difference (I-J)	Std. Error	Sig. [b]	95% Confidence Interval for Difference [b] Lower Bound	Upper Bound
(1) Control	(2) Gesture	-.096	.072	.187	-.238	.047
	(3) Listening	-.068	.073	.351	-.213	.076
	(4) Gesture with listening	-.307*	.070	.000	-.446	-.167
(2) Gesture	(1) Control	.096	.072	.187	-.047	.238
	(3) Listening	.027	.071	.703	-.114	.168
	(4) Gesture with listening	-.211*	.070	.003	-.349	-.073
(3) Listening	(1) Control	.068	.073	.351	-.076	.213
	(2) Gesture	-.027	.071	.703	-.168	.114
	(4) Gesture with listening	-.238*	.071	.001	-.379	-.098
(4) Gesture with listening	(1) Control	.307*	.070	.000	.167	.446
	(2) Gesture	.211*	.070	.003	.073	.349
	(3) Listening	.238*	.071	.001	.098	.379

*. The mean difference is significant at the .05 level.

b. Adjustment for multiple comparisons: Least Significant Difference (equivalent to no adjustments).

The following reports findings in cloze tests.

Table 9: *Investigations of ANCOVA Models: Results of Cloze Test for Dependent Variable in Delayed Post-Cloze Tests*

Table 9-1:

Results of Delayed Post Cloze-Test with Four Groups

	Model 1		Model 2		Model 3	
	Coefficient	t	Coefficient	t	Coefficient	t
Group 1	.	.	.	.	.	.
Group 2	0.11	(-1.87)	0.09	(-1.50)	0.10	(-1.71)
Group 3	0.07	(-1.31)	0.07	(-1.36)	0.05	(-0.80)
Group 4	0.03	(-0.60)	0.03	(-0.72)	0.00	(-0.04)
TV contribute improvement	0.07	(-2.07)	0.102*	(-2.66)	0.09	(-2.29)
English grammar helps with personal life	-0.06	(-1.62)	-0.04	(-1.10)		
Pre-oral test				(-0.23)	-1.65	
First language				0.00	-1.22	
_cons	0.54	(-7.63)	0.57	(-7.95)	0.46	
N	126.00		125.00		125.00	
adj. R-sq	0.20		0.24		0.28	
Log lik.	33.26		34.76		36.99	
AIC	-56.52		-57.52		-57.97	
Chi-squared						

t statistics in parentheses

* p < 0.05, **p < 0.01 *** p < 0.001

Table 9-2:

Results of Post-Cloze Test with Four Groups

	Model 1		*Model 2*		*Model 3*	
	Coefficient	t	Coefficient	t	Coefficient	t
Group 1	.	.	.	.	.	.
Group 2	-0.05	(-1.16)	-0.04	(-0.92)	-0.03	(-0.79)
Group 3	-0.03	(-0.64)	-0.04	(-0.91)	0.04	(-0.78)
Group 4	-0.01	(-0.34)	-0.03	(-0.67)	0.02	(-0.42)
Pre cloze test	0.44**	(-4.06)	0.42**	(-3.78)	0.38**	(-3.53)
L1			0.00	(-1.77)	0.00248*	(-2.35)
English conversation help with private life			-0.03	(-1.39)	-0.03	(-1.21)
Age					0.02	(-0.64)
_cons	0.59**	(-18.48)	0.56**	(-8.09)	0.18	(-0.34)
N	126.00		125.00		125.00	
adj. R-sq	0.10		0.11		0.15	
Log lik.	53.84		55.35		58.77	
AIC	-97.68		-96.70		-101.50	
Chi-squared						

t statistics in parentheses
* p<0.05 ** p<0.01

Table 10: *Repeated-Measures ANOVA in Gesture versus Non-Gesture Groups in Oral-Speech*

Table 10-1:

Tests of Within-Subject Contrasts

Measure: MEASURE_1

Source	Time	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial η²
Time	Linear	8.452	1	8.452	238.804	.000	.658
Time * group	Linear	.263	1	.263	7.435	.007**	.057
Error(Time)	Linear	4.389	124	.035			

Table 10-2:

Pairwise Comparisons

Measure: MEASURE_1

(I) Time	(J) Time	Mean Difference (I-J)	Std. Error	Sig.[b]	95% Confidence Interval for Difference [b] Lower Bound	Upper Bound
1	2	-.366*	.024	.000	-.413	-.320
2	1	.366*	.024	.000	.320	.413

The following is the repeated-measures ANOVA of pre- versus post-test results in Groups 1 and 3 (non-gesture groups) in detail.

Table 11: *Oral Speech Test*

Table 11-1:

Tests of Within-Subject Contrasts

Measure: MEASURE_1

Source	Time	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial η²
Time	Linear	2.778	1	2.778	90.580	.000***	.602
Error(Time)	Linear	1.840	60	.031			

Table 11-2:

Pairwise Comparisons

Measure: MEASURE_1

(I) Time	(J) Time	Mean Difference (I-J)	Std. Error	Sig. [b]	95% Confidence Interval for Difference [b] Lower Bound	Upper Bound
1	2	-.302*	.032	.000***	-.365	-.238
2	1	.302*	.032	.000***	.238	.365

Based on estimated marginal means.

Table 11-3:

Multivariate Tests

Effect		Value	F	Hypothesis df	Error df	Sig.	Partial η²
Time	Wilks' Lambda	.398	90.580	1.000	60.000	.000***	.602

Table 12: *Cloze Test*

Table 12-1:

Tests of Within-Subject Contrasts

Measure: MEASURE_1

Source	Time	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial η²
Time	Linear	.691	1	.691	33.653	.000***	.359
Error(Time)	Linear	1.232	60	.021			

Table 12-2:

Pairwise Comparisons

Measure: MEASURE_1

(I) Time	(J) Time	Mean Difference (I-J)	Std. Error	Sig. [b]	95% Confidence Interval for Difference [b] Lower Bound	Upper Bound
1	2	-.150*	.026	.000***	-.202	-.099
2	1	.150*	.026	.000***	.099	.202

Based on estimated marginal means

*. The mean difference is significant at the .05 level.

Table 12-3:

Multivariate Tests [a]

Effect		Value	F	Hypothesis df	Error df	Sig.	Partial η^2
Time	Wilks' Lambda	.641	33.653[b]	1.000	60.000	.000***	.359

a. Design: Intercept Within Subjects Design: Time
b. Exact statistic

Table 12-4:

Paired Samples Test

		Paired Differences							
					95% Confidence Interval of the Difference				
		Mean	Std. Deviation	Std. Error Mean	Lower	Upper	t	df	Sig. (2-tailed)
Pair 1	Pre – Post-cloze test	-.15049	.20261	.02594	-.20238	-.09860	-5.801	60	.000
Pair 2	Pre – Post-oral speech test	-.30180	.24767	.03171	-.36523	-.23837	-9.517	60	.000

Table 12-5:

Tests of Between-Subject Effects

Measure: MEASURE_1
Transformed Variable: Average

Source	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial η^2
Intercept	78.275	1	78.275	2844.139	.000***	.958
Group	.018	1	.018	.654	.420	.005
Error	3.413	124	.028			

Table 13: *REPEATED-MEASURES ANOVA: Post- and Delayed Post-Cloze Test Setting in 4 Groups*

Table 13-1:

Tests of Between-Subject Effects

Measure: MEASURE_1
Transformed Variable: Average

Source	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial η^2
Intercept	104.960	1	104.960	2455.965	.000***	.953
group	.122	3	.041	.955	.417	.023
Error	5.214	122	.043			

Table 13-2:

Tests of Within-Subject Contrasts

Measure: MEASURE_1

Source	Time	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial η²
Time	Linear	.032	1	.032	2.213	.139	.018
Time * group	Linear	.038	3	.013	.885	.451	.021
Error(Time)	Linear	1.767	122	.014			

Table 13-3:

Pairwise Comparisons

Measure: MEASURE_1

					95% Confidence Interval for Difference [a]	
(I) Time	(J) Time	Mean Difference (I-J)	Std. Error	Sig. [a]	Lower Bound	Upper Bound
Post-Test	Delayed post-test	-.023	.015	.139	-.053	.007
Delayed post-test	Post-Test	.023	.015	.139	-.007	.053

Based on estimated marginal means

a. Adjustment for multiple comparisons: Least Significant Difference (equivalent to no adjustments).

Table 13-4:

Multivariate Tests

Effect		Value	F	Hypothesis df	Error df	Sig.	Partial η²
Time	Wilks' Lambda	. 982	2.213	1.000	122.000	.139	.018
Time * Group	Wilks' Lambda	.979	.885	3.000	122.000	.451	.021

Table 13-5:

Paired Samples Test

		Paired Differences							
					95% Confidence Interval of the Difference				
		Mean	Std. Deviation	Std. Error Mean	Lower	Upper	t	df	Sig. (2-tailed)
Pair 1	Post-Test - Delayed post-test	-.0223	.16996	.01514	-.05227	.00766	-1.47	125	.143

Table 14 S:

Summary Table of Repeated-Measures ANOVA in Post- versus Delayed Post-Oral Speech Tests of Four Groups

Time2	Time 1	SS	DF	MS	F	Sig.	Partial η^2	Wilks' Lambda
Post-Test	Delayed post-test	(sum of squares)	(degree of freedom)	(mean squares)				
Group 1	Group 1	.042	1	.042	2.395	.132	.074	.926
Group 2	Group 2	.000	1	.000	.011	.917	.000	1.000
Group 3	Group 3	.092	1	.092	5.092	.032	.149	.851
Group 4	Group 4	.535	1	.535	18.086	.000	.368	.632

The following Tables are the repeated-measures ANOVA in test of within-subject contrasts in post- versus delayed post-oral speech test in each group in detail.

Table 14-1: *Group 1. Tests of Within-Subject Contrasts*

Measure: MEASURE_1

Source	Time	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial η²
Time	Linear	.042	1	.042	2.395	.132	.074
Error(time)	Linear	.524	30	.017			

Table 14-1-1:

Pairwise Comparisons

(I) Time	(J) Time	Mean Difference (I-J)	Std. Error	Sig. [a]	95% Confidence Interval for Difference [a] Lower Bound	Upper Bound
1	2	-.052	.034	.132	-.121	.017
2	1	.052	.034	.132	-.017	.121

Table 14-1-2:

Multivariate Tests

Effect		Value	F	Hypothesis df	Error df	Sig.	Partial η²
Time	Wilks' Lambda	.926	2.395[b]	1.000	30.000	.132	.074

Table 14-2: *Group 2.*

Table 14-2-1:

Tests of Within-Subject Contrasts

Measure: MEASURE_1

Source	Time	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial η²
Time	Linear	.000	1	.000	.011	.917	.000
Error(time)	Linear	.543	31	.018			

Table 14-2-2:

Pairwise Comparisons

Measure: MEASURE_1

					95% Confidence Interval for Difference [a]	
(I) Time	(J) Time	Mean Difference (I-J)	Std. Error	Sig. [a]	Lower Bound	Upper Bound
1	2	.003	.033	.917	-.064	.071
2	1	-.003	.033	.917	-.071	.064

Table 14-2-3:

Multivariate Tests

Effect		Value	F	Hypothesis df	Error df	Sig.	Partial η²
Time	Wilks' Lambda	1.000	.011[b]	1.000	31.000	.917	.000

Table 14-2-4:

Descriptive Statistics

	Mean	Std. Deviation	N
Post-oral speech test	.56423611112	.326729498882	32
Delayed post-oral speech test	.56076388888	.338772638433	32

Table 14-3: *Group 3.*

Table 14-3-1:

Tests of Within-Subject Contrasts

Measure: MEASURE_1

Source	Time	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial η^2
Time	Linear	.092	1	.092	5.092	.032*	.149
Error(time)	Linear	.524	29	.018			

Table 14-3-2:

Pairwise Comparisons

Measure: MEASURE_1

					95% Confidence Interval for Difference [b]	
(I) Time	(J) Time	Mean Difference (I-J)	Std. Error	Sig. [b]	Lower Bound	Upper Bound
1	2	-.078*	.035	.032	-.149	-.007
2	1	.078*	.035	.032	.007	.149

Table 14-3-3:

Multivariate Tests

Effect		Value	F	Hypothesis df	Error df	Sig.	Partial η²
Time	Wilks' Lambda	.851	5.092^b	1.000	29.000	.032*	.149

Table 14-4: *Group 4.*

Table 14-4-1:

Tests of Within-Subject Contrasts

Measure: MEASURE_1

Source	Time	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial η²
Time	Linear	.535	1	.535	18.086	.000***	.368
Error(time)	Linear	.917	31	.030			

Table 14-4-2:

Pairwise Comparisons

Measure: MEASURE_1

(I) Time	(J) Time	Mean Difference (I-J)	Std. Error	Sig. [b]	95% Confidence Interval for Difference [b]	
					Lower Bound	Upper Bound
1	2	-.183*	.043	.000	-.270	-.095
2	1	.183*	.043	.000	.095	.270

Table 14-4-3:

Multivariate Tests

Effect		Value	F	Hypothesis df	Error df	Sig.	Partial η²
Time	Wilks' Lambda	.632	18.086^b	1.000	31.000	.000***	.368

Table 15: *Summary Table for Test Scores Generated by 4 Different Groups*

Table 15-1:

Oral Speech Tests

4 Different Groups (mean scores)	Pre-oral speech test	Post-test in oral speech test	Delayed post-oral speech test
Group 1	0.12	0.43	0.48
Group 2	0.15	0.56	0.56
Group 3	0.17	0.47	0.54
Group 4	0.16	0.61	0.79
Average	0.15	0.52	0.59

Table 15-2:

Cloze Test

4 Different Groups (mean scores)	Pre-cloze test	Post cloze test	Delayed post-cloze test
Group 1	0.49	0.64	0.64
Group 2	0.46	0.61	0.61
Group 3	0.49	0.64	0.70
Group 4	0.48	0.65	0.68
Average	0.48	0.63	0.66

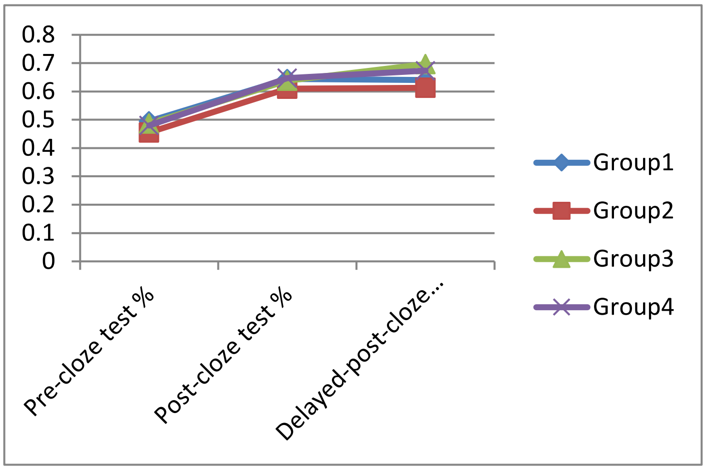


Figure 6-1. Cloze tests.

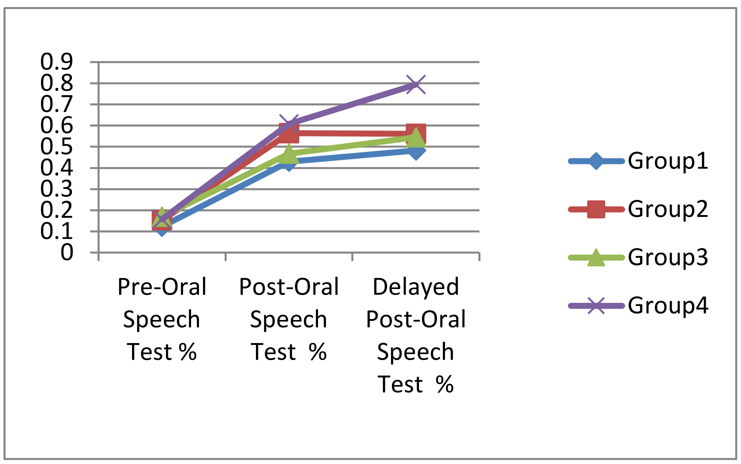


Figure 6-2. Oral speech-tests.

Figure 6. Correlation of Test Results in Four Groups Among Test Scores in Pre-, Post-, and Delayed Post- Tests.

Chapter 5

Discussions and Conclusions

5.1 Introduction

Based on the experimental and theoretical investigations in literature review and results in this study, this final chapter concludes that the Gesture Listening Higher Concept Approach helped with accomplishing three goals in the present study. This chapter revisits the three hypotheses, the two research questions, and the two issues that remained uncertain in experimental gesture studies in an FL to date, which I addressed in Chapter 1 that built upon empirical, hypothetical, and theoretical studies in L1, FL, and bilingualism. Furthermore, discussions in this chapter allow one to provide pedagogical suggestions and implications of future research in experimental gesture studies in FL.

One of the three goals in this experimental study was to respond to the novel research trend that necessitates the incorporation of linguistic relativity into SLA in sociocultural theory insomuch as linguistic relativity does not discuss language acquisition, but sociocultural theory does in SLA (cf. Kunisawa, under revision).[92] Additionally, Vygotsky strengthened diachronic, historical-developmental approach, whereas Whorf' established synchronic, comparative-interpretive approach (Lucy & Wertsch, 1987). It appears that both studies complement each other, and thus, based on experiments in this study and findings in literature review, I argue that this study experimentally and theoretically accomplished the first goal. Second, the current study revealed that the concomitant use of ICSCTGs and listening practice helped Japanese high school students learn about developed meaning of

[92] Kunisawa (under review) indicates that the concept of development of meaning of words in sociocultural theory also can be applicable to the synchronic, comparative-interpretive approach in linguistic relativity in SLA research.

words in EFL more than they would learn from either treatment alone. Third, this study uncovered to what extent the simultaneous use of the two techniques helped students with learning about vertical spatial operations, although literature review indicated that the concurrent iconic gestures and listening practice assisted students with learning artificial words for 5 day (Macedonia et al., 2011).

Accordingly, exploring these three goals confirms the significance of the effects of the Gesture Listening Higher Concept Approach by supporting the hypotheses and responding to the research questions. Additionally, the current study revealed that this approach shares analogous thought to that of the gesture-for-conceptualization hypothesis concerning the effects of reproducing ICSCTGs. Furthermore, this approach uncovered the potential reason why Group 1 also learned a vertical axis operation and preserved knowledge of it without statistically significant memory decay in a level of a higher concept, even though they did not receive any treatment. To discuss results and findings in this study, I revisit the following research questions and hypotheses. Chapter 1 presented:

Research Question 1: The major research question in the current study is: Does teaching students iconic co-speech co-thought gestures (ICSCTG) facilitate learning how to express vertical spatial relationships in English for Japanese EFL high school students?

Research Question 2: A secondary research question in the present study is: Does listening practice with or without ICSCTG facilitate learning how to express vertical spatial relationships in English for Japanese EFL high school students?

The same chapter generated:

> **Hypothesis 1:** Teaching ICSCTG will facilitate learning how to express vertical spatial relationships in English for Japanese EFL high school students,
>
> **Hypothesis 2:** Teaching ICSCTG and listening practice combined will facilitate learning how to express vertical spatial relationships in English for Japanese EFL high school students more than either treatment alone, and
>
> **Hypothesis 3:** Teaching ICSCTG and listening practice combined will help Japanese EFL learners maintain knowledge of how to express vertical spatial relationships in English.

Based on the research questions and hypotheses noted above, the following section 5.2 includes a discussion of findings in literature review and results in this study.

5.2 Findings and Discussions

This final chapter presents discussion of how this study supports Hypothesis 1that corresponds to Research Question 1. Subsequent sections include my arguments for novel discoveries in the experimental gesture studies in FL. The current study investigated how distinct worldview and cognitive development in EFL can co-evolve. Arguments and investigations in this study built upon: (a) the forgetting curve hypothesis; (b) the gesture-for-conceptualization hypothesis; (c) the Gesture Listening Higher Concept Approach; (d) studies of sociocultural theory and of contemporary linguists; (e) findings in literature review; and (f) experimental results in this study. Arguments and investigations supported the hypotheses in answering the research questions in this study. Furthermore, this chapter offers the explanations of current status of comprehending gestures and finally includes a description of the summary of this study. The following section discusses how the present study fully supports Hypothesis 1 and properly responds to Research Question 1.

5.2.1 Primacy: Iconic Co-Speech Co-Thought Gesture and Hypothesis 1

The gesture-for-conceptualization hypothesis argues that speakers generate gestures for the purpose of speaking and that gestures and speech act independently as two distinct systems, interacting during conceptualization. However, later the two systems merge (Chu & Kita, 2016; Hostetter & Alibali, 2007; Kita, 2010; Kita et al., 2017; Kita & Ozyurek, 2003), unlike the lexical retrieval hypothesis proposed by Krauss et al. (2000). Additionally, co-speech gestures helps students with learning and problem solving and assisted them with remembering vocabulary. Moreover, as introduced in Chapter 2, co-verbalized gesture has a semantic connection to oral speech. The following section 5.2.1.1 includes investigation of to what extent a study presented by Mayberry and Nicoradis (2000), Quinn-Allen (1995), Tellier (2008) supports Hypothesis 1 in this study.

5.2.1.1 (Co-Speech) Gestures: Findings in Preschool and High School EFL Class

Although, differences exist between this and Tellier's (2008) study, findings in Tellier's study suggested that teaching (co-speech) gestures helped students learn a single word meaning with statistical significance, supporting Hypothesis 1 and answering Research Question 1to a certain extent.[93] The results in this and findings in Tellier's (2008) study indicated that (co-speech) gestures alone had an effect on learning a single word meaning with statistical significance, which this and Tellier's study share to a certain extent. The following sections address a limitation in Tellier's study and an issue in comparing these two studies noted

[93] Henceforth, the current study replaces a term *active knowledge* with *active selection* and *passive recording* with *passive knowledge* in this chapter.

above, which led to examining a study presented by Quinn-Allen (1995) and Mayberry and Nicoludis (2000).

5.2.1.1.1. Instruction-Assessment Dualism: Silent Gesture Versus Co-Speech Gesture

This section addresses a limitation in an assessment that Tellier's (2008) study, which did not employ co-speech gestures in assessing knowledge of English words, even though Tellier taught participants those words with co-speech gestures. The concept of instruction-assessment dualism proposed by sociocultural theorists is a crucial factor in determining the effectiveness of any instruction. However, Tellier evaluated participants' knowledge with silent gestures (passive recording) in learning English vocabulary, which led to lack of instruction-assessment dualism. Additionally, Ozcaliskan et al. (2016) directly compared silent gesture with co-speech gestures and concluded that silent gestures had no effect on language.

Nonetheless, the gesture group in Tellier's study enriched learning those words with silent gestures. Additionally, non-gesture group in Tellier's (2008) study significantly improved in learning those words without co-speech gestures. The following section includes a discussion of a potential reason why participants in Tellier's study not only learned the words, but also enhanced their knowledge of those words with silent gestures, which lack instruction-assessment dualism.

5.2.1.1.2 Age Difference

Different outcomes in the current study and in the Tellier's (2008) study suggest that the age difference may impact the way that gesture supports learning a single word meaning in FL class. For instance, Figure 6-2 indicates that between the pre- and the post-test setting, gesture groups in this study improved knowledge of developed meaning of words in EFL with statistical and practical significance. However, gesture only group (Group 2) did not advance their ability of it between the post-and the delayed post-oral speech setting, although maintained their knowledge of it without statistically significant memory decay.

Conversely, Tellier's (2008) study suggests that both gesture and non-gesture groups constantly enhance knowledge of English words between Assessment 1 and 3b, although the non-gesture group decreased their test scores between the Assessments 2 and 3 slightly (Figure 1). A comparison of these two studies suggests that students at a high school level did not enhance their ability to learn a single word meaning with co-speech gestures between the post- and the delayed post-test condition. However, in Tellier's study, participants at a pre-school level improved knowledge of English words with statistical significance, despite lacking the instruction-assessment dualism (i.e., the silent gesture), which did not have an effect of language on gestures. Refer to Figure 1 for Tellier's study (2008) in Chapter 2. As explained earlier, Vygotsky (1987) argues, "early age has shown that the mastery of two or three languages does not slow language learning … It cannot even be compared with the memory of the adolescent or adult" (p. 307). Additionally, Dick et al. (2012) conclude that children's cortical activations differed from those of adults when processing co-speech gestures.

Comparable results in the current study and in Tellier's (2008) study of preschoolers indicate that gesture influences learning a single word meaning in a FL class differently. However, based on literature review in gesture research in L1, SLA, and bilingualism, I argue that gestures can facilitate learning wide range of age groups, if gestures are properly employed (Goldin-Meadow et al., 2001).

Concisely, although, differences between the two studies exist, findings in Tellier's (2008) study suggested that Hypothesis 1 was supported and Research Question 1 was answered to a certain extent. The following section 5.2.1.2 includes an explanation that (co-speech) gestures also have an effect on learning FL at a college level. Quinn-Allen (1995) presented the first experimental gesture study in FL in a first year college FFL class[94] (Bergmann & Macedonia, 2013; Macedonia, 2014; Tellier, 2008). Other experimental gesture studies in FL extended these findings in 1995 (Bergmann & Macedonia, 2013).[95]

5.2.1.2 (Co-Speech) Gestures: Findings in FFL First Year College Class

Quinn-Allen's (1995) study claimed that recreating and observing gestures helped students learn the French expressions to a certain extent at a college level (see Figure 7). Thus, I argue that Quinn-Allen's study supported Hypothesis 1 and answered Research Question 1 to some extent by rectifying the issue of the age difference between this and Tellier's (2008) study, although the methodological issues (the instruction-assessment dualism and the use of two different assessment methods in this study) remains as addressed earlier. These studies raise the question of why gestures help students to acquire word meaning. One explanation is that gesture and speech conjointly emerge, helping learners incorporating thinking and speaking. This view is consistent with several prior studies in the literature. For example, Mayberry and Nicoladis (2000) studies the use of gesture by French-English bilingual preschoolers, and

[94] A study published by Carels (1981) proposed systematic use of gestures in FL instructions. Additionally, Mariani (1981), and Schewe and Shaw (1993) suggested gestures involving drama helped students learn FL. However, none of them employed experimental studies, and thus, lacked empirical evidence for effects of gestures on FL learning.
[95] They (2013) did not explain how and whose experimental studies had extended Quinn-Allen's (1995) study in FL education.

found that gesture did not precede spoken language use, but instead was more likely to be used with the dominant language.

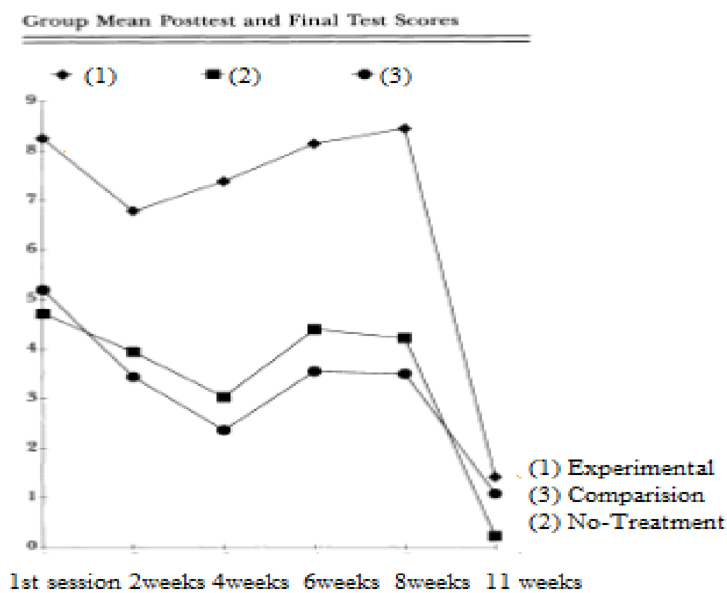


Figure 7. Group Mean Posttest and Final Test Scores.
Source: Based on Quinn-Allen's study, the author added some information to make Figure 7 clearer (cf. Quinn-Allen, 1995, p. 525).

The following section 5.2.1.3 investigates how iconic co-speech gestures CAN help preschoolers with language development accompanied by syntax, which led to fully resolving issues in supporting Hypothesis 1that corresponds to Research Question 1 except for an age difference from that of this study.

5.2.1.3 Co-Speech Gestures: Findings in Study of Young Children

To support Hypothesis 1 regarding co-speech co-thought gestures at a high school level, the present study introduced the study of Mayberry and Nicoladis (2000). Mayberry and Nicoladis (2000) claimed that:

- Generating iconic co-speech gestures could affect oral speech development in bilingual children (preschoolers) accompanied by properly organized semantic and syntax at a sentence level with the instruction-assessment dualism.
- Participants began to create iconic and beat gestures while speaking only after their oral speech became longer than two words. Consequently, their oral speech accompanied by iconic or beat gestures was more linguistically complex than oral speech accompanied with pointing or no gestures at all.

These findings noted above explain that co-speech gestures encoded both semantic and syntax inasmuch as participants generated gestures after their oral speech became longer than two words.

Thus, creating co-speech gestures facilitated oral speech/ language development with linguistically complex expressions, activating lexical items with maintaining their remembering, manipulating, exploring their utterance and gestures, packaging information when constructing their oral speech with multiple words, which entails the gesture-for-conceptualization hypothesis.

5.2.1.4 Summary of Gesture Studies

Gesture studies in L1 and FL raise questions why co-speech gestures help speakers learning. Alibali et al. (2017) argue, “Speakers who cannot gesture should have greater difficulty planning syntactic units that correspond to units for speech production” by citing a study presented by Bock and Cutting (1992). Additionally, the information packaging hypothesis contends, “speakers package spatio-motoric information into units appropriate for verbal encoding … Difficulty in Maintaining Spatio-Motoric Information Triggers Gesture” (Kita et al., 2017, p. 246-247). Moreover, based on an experiment presented by So et al. (2014), Kita et al., (2017) suggest, “Gesture activated, manipulated, packaged or explored spatio-motoric information, and consequently more resources were available for the memory task” (Kita et al., 2017, p. 258). That is, co-speech gesture helps with lightening cognitive load when participants engage in a difficult verbal material (Goldin-Meadow et al., 2001). As explained earlier, the gesture-for-conceptualization hypothesis claims that representational gestures link not only to speaking, but also to learning and problem solving (cf., Kita, 2000 for the information packaging hypothesis; Kita et al., 2017; Nicoladis et al., 2009).

Furthermore, Vygotsky (1993) suggests, “Speech not only fulfills the function of communication between children but also appears as an instrument of thought” (p. 114), and thus, speech is a crucial factor of thinking. “Motor processes linked to speech play a central role in assisting the thinking process. Accordingly, a subject understands a complicated verbal material” (Vygotsky, 1987, p. 44). That is, co-verbalized gestures can enhance speakers’ thinking. The study presented by Mayberry and Nicoladis (2000) demonstrated how iconic co-speech gestures assisted participants with creating linguistically complex speech accompanied by properly organized semantic and syntax. Mayberry and Nicoladis (2000) did not use two different assessments methods as in a study presented by this, Macedonia el al. (2011), and Tellier (2008) because of a natural language learning setting, but not a classroom setting. Therefore, the lack of two assessment tools, which consists of a passive recording and a positive selection, can be ignored.

In sum, based on results in this study and findings in literature review, the present study fully supported Hypothesis 1 and appropriately answered Research Question 1, which reveal the effect of the Gesture Listening Higher Concept regarding ICSCTG. As a result, the gesture-for-conceptualization hypothesis is also supported. Additionally, I contend that studies explained above show that gestures facilitate learning up to college age in minimum.

5.2.2. Gesture Listening Higher Concept Approach

This section presents reasons that this study generates the Gesture Listening Higher Concept Approach, which leads to three novel discoveries in experimental gesture studies in FL. These three innovative discoveries are an effect of concurrent ICSCTGs and listening practice, supporting Hypotheses 2 and 3 and an effect of positive correlations that could induce a higher concept. These discoveries could advance not only an experimental gesture study in FL, but also SLA research.

5.2.2.1 Novelty 1: Hypothesis 2

The present study tested whether the concomitant use of ICSCTGs and listening practice helped Japanese EFL high school students learn about developed meaning of words (a single word meaning) more than they would learn from either treatment alone. This issue remained unclear in experimental gesture studies in FL research to date. Results of this study demonstrate that the simultaneous use of these two techniques assisted students with learning about developed meaning of words in EFL more than either treatment alone with statistical and practical significance, and thus, the current study supported Hypothesis 2, which is one of the three goals in this study. I argue that the Gesture Listening Higher Concept Approach accelerated supporting Hypothesis 2, which is an innovative discovery in an experimental gesture study in FL (Novelty 1). However, researchers have not yet established a theoretical framework in this regard (Hostetter, 2011; Muller et al., 2013), and thus, a theoretical framework remains uncertain in this regard.

The following section 5.2.2.2 investigates to what extent the study presented by Macedonia et al. (2011) can explain that iconic gesture and listening practice helps students maintain knowledge of a single word meaning. I re-analyze data in Macedonia et al. (2011) by using the forgetting curve hypothesis to investigate it. Consequently, Hypothesis 3 was fully supported, which led to another new discovery in an experimental gesture study in FL (Novelty 2).

5.2.2.2 Novelty 2: Hypothesis 3

This section explains how results in the present study differ from those of studies of the forgetting curve hypothesis, of Macedonia et al. (2011), and of Quinn-Allen (1995), although their studies support the results in the current study. That is, the higher retention rate found in this study reported in Chapter 4 can be accounted for by limitations that I identified in their studies explained below.

To what degree the concurrent ICSCTGs and listening practice can help EFL students preserve knowledge of how to express vertical axis operations remained unclear to date. Based on the forgetting curve hypothesis, the gesture-for-conceptualization hypothesis, and sociocultural theory, this section leads to fully supporting Hypothesis 3 that is one of the three goals in this study. The following section 5.2.2.2.1 provides a brief overview of a study reported by Macedonia et al. (2011). Subsequently, I present my arguments about how Hypothesis 3 was supported. To support the arguments, I will explain three alternative theoretical justifications in addressing three different limitations in Macedonia et al. (2011), which led to explanations of re-analyzing their data by using the forgetting curve hypothesis and this study (the Gesture Listening Higher Concept Approach). Then, I provide additional three limitations in Macedonia et al. (2011). I also present a brief summary of studies presented by the forgetting curve hypothesis, leading to the necessitation of a joint analysis of a study presented by Macedonia et al. (2011), the forgetting curve hypothesis, and this study. A final section discusses attrition in FL research.

5.2.2.2.1 Brief Overview of Study Presented by Macedonia et al. (2011)

Macedonia et al. (2011) used 92 artificial nouns (Vimmi), created based on Italian phonotactics for experimental purposes[96] in conducting data analyses with a statistical method 33 native German participants [M age 23.17, SD = 1.61, N =17 females, 16 males] learned 92 artificial nouns of which half were learned with iconic gestures accompanied by listening practice and the other half with meaningless gestures and listening practice together.[97] Daily sessions were120 minutes long (with breaks every 29 minutes) consisting of viewing video shots of a gesture corresponding to a word written in Vimmi accompanied by a subtitle that showed German translation of Vimmi. Participants simultaneously listened to each word produced by audio files in a computer and then reproduced a co-speech gesture corresponding to the word repeatedly after seeing and hearing the word. Assessments took place for 4 days (once a day) at the beginning of each training session and then the final assessments took place roughly 60 days after the last measurements (the fourth assessment).

Additionally, Macedonia et al. (2011) also investigated brain activations of words learned by iconic and meaningless gestures. After completing the entire training sessions,[98] participants' cortical activities were registered by means of fMRI while carrying out a word recognition task. Macedonia et al. (2011) trained participants one additional day to confirm

[96] Refer to Macedonia and Repetto (2016b).
[97] Meaningless gestures by definition are not symbolic, according to Macedonia et al. (2011).
[98] Participants in Macedonia et al.'s study (2011) viewed written words and gestures on a computer screen and they repeated the words, after hearing the words generated by an audio file during training sessions.

reaching a ceiling in both training conditions. The stimuli were the 92 trained Vimmi words (cf. Table I in Macedonia et al., 2011, p. 984) and 23 unknown filler words (cf. Table II in Macedonia et al., 2011, p. 989) that were neither German nor Vimmi, and thus, participants did not know the filler words.

Macedonia et al. (2011) showed written Vimmi words to participants with an LCD projector onto a projection screen and then played the audio file coinciding with the start of the visual stimulus (the written word). Macedonia et al. (2011) asked participants to press a key if the participants detected an unknown word. They (2011) concluded that:

- The “behavioral results clearly demonstrate that performing iconic gestures during learning has a positive impact on memory for new nouns” (p. 898).
- Based on an outcome generated by fMRI, they (2011) claim, “a gesture leads to better memory performance only if the motor image (iconic gestures) matches an internal representation of the concepts of semantics” (p. 994-5).

Based on findings in Macedonia et al. (2011) and literature review, I argue that Macedonia et al.’s (2011) study supports Hypothesis 3 for 5 days in minimum in learning 46 artificial words, but it is uncertain whether their study supports Hypothesis 3 for more than 5 days (see Figure 8), and thus, the following presents my arguments on this issue.

5.2.2.2.1.1 Arguments of Hypothesis 3

When participants in Macedonia et al. (2011) learned 46 artificial words with the concurrent use, their cognitive load in their study was initially higher than that of the forgetting curve hypothesis in one day. However, somewhere between 31 and 60 days, the words learned with the simultaneous use of these two techniques in Macedonia et al. (2011) (M = 34.00%) outperformed the words acquired in the absence of the concurrent use in the forgetting curve hypothesis (M = 17.78 %), although the one-syllable nonsense words are inherently difficult (Vygotsky, 1987) (see Table 16). Therefore, undoubtedly, iconic gestures and listening practice together had a significant effect on learning artificial words.

Nonetheless, after 60 days from the day when the last assessment took place, memory retention of words learned by the two techniques was at a similar level to that of somewhere between the first and the second day, suggesting words learned by the concomitant use deteriorated as time goes by (see Figure 8 & Table 16). Thus, Macedonia et al. (2011) did not explain to what extent the concurrent use helped participants maintain knowledge of the artificial words, although the Gesture Listening Higher Concept Approach does. Based on theoretical justifications, I propose re-analyzing data in Macedonia et al. (2011) to investigate to what extent the concurrent use of iconic gestures and listening practice had an effect on learning artificial words after 60 days, which led to supporting Hypothesis 3.

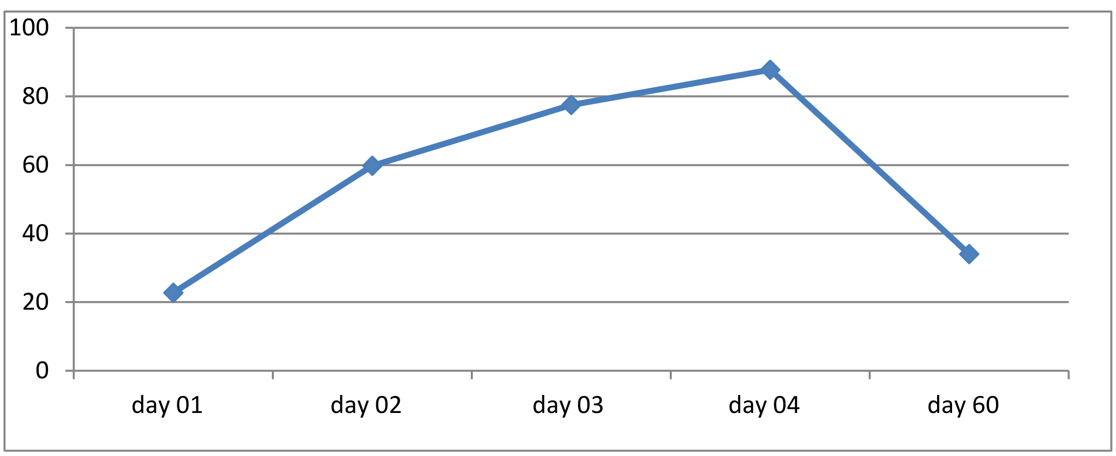


Figure 8. Artificial Nouns Learned by Iconic Gestures and Listening Practice Together (%).

Table 16:

Retention of Words in Forgetting Curb Hypothesis and of Artificial Nouns Learned by Meaningless and Iconic Gestures with Listening Practice in Macedonia et al. (2011)

		Forgetting curve hypothesis (1885/ 1913, 1991, 2015)	Words learned by meaningless gestures: Macedonia et al. (2011)	Words learned by iconic gestures: Macedonia et al. (2011)
1	Mean score after one day	31.00%	10.00%: Translation from German (G) into Vimmi (V)	16.00%: Translation from German into Vimmi
			22.00%: from V to G	29.50%: from V to G
	On average		16 %	22.76 %
2	Mean score after 4 days on average	N/A	81.75 %	87.75 %
	Mean score after 31 days	17.78%	N/A	N/A
3	Mean score of pairing words after 60 days	N/A	8.26%	28.26%
4	Mean score of free recall loose item 60 days	N/A	15.11%	39.78%
5	Average score in Figure 1and 2 listed below	N/A	11.67%	34.00%

In the following sections, I address three limitations in Macedonia et al.'s (2011) study, which leads to theoretical justifications in re-analyzing data in Macedonia et al. (2011). Furthermore, I also present three additional limitations in Macedonia et al. (2011). Subsequently, I briefly summarize the forgetting curve hypothesis and then follow with the explanation of the joint analysis of findings in Macedonia et al. (2011), in the forgetting curve hypothesis and results in this study, and lastly I explain the attrition in FL class.

5.2.2.2.1.1.1 Limitations of Macedonia et al.'s (2011) Study and Theoretical Justifications for Re-analyzing Data in Macedonia et al. (2011)

Revealing three limitations in Macedonia et al.'s (2011) study led to the joint analysis that made this study possible to present convincing arguments of Hypothesis 3. As a result, this section presents theoretical justifications of the collaborative analysis of a study offered by Macedonia et al. (2011), the forgetting curve hypothesis, and this study. Consequently, Hypothesis 3 can be properly discussed.

(1) Limitation 1: Use of Meaningless Gestures

The first limitation of Macedonia et al.'s study (2011) is the inclusion of meaningless gestures in their study. Figure 8 demonstrates a significant effect of iconic gestures and listening practice together on learning artificial words for 5 days in minimum. However, Macedonia et al. (2011) has limitations in claiming to what extent the concurrent use had an effect on maintaining knowledge of the artificial words between the last assessment and 60 days later from the last assessment. Macedonia et al. (2011) contend, "iconic gestures compared with meaningless gestures significantly help to enhance the memorization of foreign language nouns" (pp. 994-995), however, their argument was unpersuasive. Two reasons account for this. First, Macedonia et al. (2011) claim an effect of meaningless gestures[99] on verbal memory by citing studies presented by research in L1.

[99] Meaningless gestures and mismatching gestures should be differentiated. Gesture researchers interpret mismatching gesture differently. For instance,

- Goldin-Meadow (2003) claims, "when gesture-speech mismatch is a step on a child's path to mastery, learning is deep and robust" (p. 54). McNeill et al. (1994) supported Goldin-Meadow's view of gesture-speech mismatching and stated, "The surprising discovery by Goldin-Meadow et al. was that children with such mismatching gestures and speech give independent evidence of being in a transitional knowledge state "(p. 235). However,
- Macedonia et al. (2011) states that mismatching gestures interferes (negative effects on information processes) with the semantic of the words.

However, Cook, Yip, and Goldin-Meadow (2012) concluded, “Speakers recalled significantly more letters when producing movements that coordinated with the meaning of the accompanying speech, i.e., when gesturing, than when producing meaningless movements or no movement” (p. 594).

Second, Goldenberg (2013) argues that meaningless gestures are difficult to assess because those gesture do not have any specific name. Additionally, visuo-imitative apraxia is not a general defect in imitating gestures, but a particular defect of imitating meaningless gestures, which indicates that cognitive processes of regenerating meaningless gestures differs from those of imitating gestures (meaning in them). These factors noted above led to re-analyzing data presented by Macedonia et al. (2011) by comparing those in the forgetting curve hypothesis. As explained earlier, the artificial words learned by using the concurrent use were superior to the nonsense words in the forgetting curve hypothesis after 60 days. However, Macedonia et al. (2011) has a limitation in claiming to what degree the simultaneous use had an effect on preserving knowledge of the artificial words between the last assessment and 60 days later from the last assessment. Therefore, Macedonia et al. (2011)’s argument of an effect of iconic gestures and listening practice together remained unconvincing regarding to what extent the simultaneous use has an effect on maintaining knowledge of artificial words, even though the concomitant use had a significant effect on learning the single words meaning, which is theoretical justification 1.

McNeill (1992) and Goldin-Meadow (2003) did not offer their discussions on meaningless gestures in the books published in 1992 and 2003.

(2) Limitation 2: Artificial Words and Nonsense One-Syllable Words

The second limitation of Macedonia et al.'s (2011) study is their reliance on artificial words. Vygotsky (1987) criticizes the use of artificial words for an experiment and contended that:

> A major deficiency of the method of definition is that the concept is torn from its natural connections. It is isolated in a congealed and static form from the actual processes of thinking where it is encountered. It is isolated from the processes of thinking where it is born and lives…It tells us nothing of how the child operates with the concept in the real-life process of solving a problem, of how he uses it when some real- life need for it arises (p. 123).[100]

Similarly, Whorf (1956) intentionally remained distant from promoting the artificial language (as cited in Levelt, 2012).

Furthermore, Vygotsky (1987) states, "There is a large range of thinking that has no direct relationship to verbal thinking" (p. 115). Even though participants of Macedonia et al. (2011) study reproduced co-speech gestures during training sessions, recreating those gestures had a limitation in generating thinking because of the use of artificial words. Thus, knowledge created by the simultaneous iconic gestures and listening practice lost a linkage to thinking and memory as time goes by. Accordingly, artificial words learned by the concurrent use of these two techniques increasingly experienced memory deterioration of the words, suggesting learning artificial words is extremely difficult.

According to Vygotsky (1987), learning nonsense one-syllable words are inherently difficult and acquiring both artificial and nonsense one-syllable words are extremely difficult, suggesting that the degree of difficulty in learning these words in both studies can be equivalent. These factors led this study to employing data presented by the forgetting curve

[100] The first experimenter of artificial language learning was Fischer (1922) (as cited in Levelt, 2012).

hypothesis that uses the one-syllable nonsense words in re-analyzing data reported by Macedonia et al. (2011), which is theoretical justification 2.

(3) Limitation 3: Spacing Effect

Third, Macedonia et al. (2011) failed to include the necessary spacing in the learning trials to optimize learning outcomes. Participants in Macedonia e al. (2011) received intensive training in learning artificial words for approximately for 464 minutes in total in 4 days (400 minutes for listening practice) at the beginning of their experiment by using the concomitant use of iconic gestures and listening practice. The spacing effect (or graduated interval recall) can make a big difference in learning a single word meaning. Contemporary researchers have supported this effect (Ellis, 1995; MacWhinney, 2013; Pavlik & Anderson, 2005; Pimsleur, 1967). However, after participants completed the fourth assessments, Macedonia e al. (2011) did not ask participants to learn artificial nouns by utilizing the simultaneous use and then assessed their memory of the words after 60 days from the fourth assessment (the lack of the spacing effect). Ebbinghaus (1885/ 1913) suggested that:

> It makes the assumption probable that with any considerable number of repetitions a suitable distribution of them over a space of time is decidedly more advantageous than the massing of them at a single time (p. 89).

Accordingly, the lack of a spacing effect in Macedonia et al. (2011) induced decay in memory after the fourth assessment that took place in the beginning part of their experiment. Researchers in the forgetting curve hypothesis also measured participants' knowledge of nonsense one-syllable words without spacing effects. The lack of spacing effect in both studies noted above led to comparing data offered by Macedonia et al. (2011) to that of the forgetting curve hypothesis, which revealed to what extent the concurrent use can have an effect on learning artificial words, which is theoretical justification 3.

5.2.2.2.1.1.2 Three Additional Limitations of Macedonia et al.'s study (2011)

(4) Limitation4: Knowledge Versus Image

Macedonia et al.'s (2011) explanation of the effect of gesture on learning words in terms of gesture's role in the generation of a motor image is too simplistic. Based on analyses of participants' behaviors (statistical analyses), their interpretations of neurological language processes with fMRI, and the enactment effect, Macedonia et al. (2011) argues, "a gesture leads to better memory performance only if it allows to create a motor image that matches with an internal representation of the concept's semantics" (p. 995). Vygotsky (1987) suggests, "Like any object, the word can be replaced by a mental representation or image in memory" (p. 256) and recognized an importance of image in memory to some extent and stated that thinking connected to memory image. However, as stated earlier, Vygotsky (1998) further indicates, "His memory is no longer a store for single image, but is an archive of knowledge" (p. 88). That is, Vygotsky emphasizes the importance of knowledge, thinking, and memory in pedagogy. Moreover, Baddeley (2007) claims that episodic memory "will in due course contribute to semantic memory, as part of the normal process of accrual of knowledge" (p. 153). Furthermore, Vygotsky (1987) contends that the "accumulation of knowledge leads directly to an increase in the level of scientific thinking" (p. 168).

Accordingly, the experiment conducted by Macedonia e al. (2011) clarified the crucial factor of creating a motor image in the processes of learning the artificial words. However, their argument had the limitation in explaining how learning the words linked to the accumulation of knowledge of artificial words because of the lack of discussions noted above.

(5) Limitation 5: Instruction-Assessment Dualism

Macedonia et al. (2011) lack the concept of the instruction-assessment dualism. Although, Macedonia et al. (2011) used iconic co-speech gestures during training sessions, they did not ask the participants to use co-speech gestures in their assessments. As stated earlier, Vygotsky (1987) argues that, “the motor processes associated with speech play an important role in facilitating the thinking process in particular, in improving the subject’s understanding of difficult verbal material.” (p. 44). I argue that recreating co-speech gestures has more of an effect on improvement of learning a single word meaning than without regenerating them. It appears that Macedonia et al. (2011) lacked the concept of instruction-assessment dualism.

(6) Limitation 6: Lack of Proper Analysis and Discussions

Macedonia et al. (2011) seriously failed to evaluate and discuss an effect of listening practice, despite Macedonia et al. (2011) trained participants by using auditory cues for approximately 400 minutes in 4 days, while participants recreated iconic co-speech gestures. This suggests that Macedonia et al. (2011) essentially evaluated an effect of the simultaneous use of iconic gestures and listening practice on learning a single word meaning (an artificial word). Nonetheless, Macedonia et al. (2011) neither reported nor discussed results of their experiments as an effect of the concomitant use of these two techniques on learning artificial words.

Based on data analysis of the present study, Chapter 4 revealed to what degree (a) the concomitant ICSCTGs and listening practice, (b) ICSCTGs alone, and (c) listening practice only have an effect on learning about a single word meaning (developed meaning of words in EFL) as opposed to Macedonia et al.’s study (2011) that did not. Data analyses in Chapter 4 suggested that an effect of (a) differed from (b) and (c) noted above, and thus, an effect of these three different techniques should be properly analyzed by comparing each effect. Furthermore, as discussed in Chapter 2, “inner speech is an internal plane of verbal thinking which mediates the dynamic relationship between thought and word” (Vygotsky, 1987, p. 279). Therefore, I argue that Macedonia et al. (2011) experimentally and participially support Hypothesis 3. However, Macedonia et al. (2011) failed to provide a relevant theoretical framework of the concurrent use of iconic gesture and listening practice, and lacked assessing and discussing an effect of listening practice alone and of recreating iconic gesture only on learning the single word meanings.

5.2.2.2.1.1.3 Brief Summary of Studies Presented by Forgetting Curve Hypothesis

Ebbinghaus (1885/ 1913), who is the originator of the forgetting curve hypothesis, conceptualized forgetting as a function of time (Vygotsky, 1987). In this section, I discuss the relevance of the use of the forgetting curve hypothesis in FL research because of the nature of forgetting that is common phenomena in any classroom. For instance, Baddeley et al. (2009) suggested that as with other studies shown by Ebbinghaus (1883/ 1913), the results could apply to a wide-range of learning conditions. Additionally, Weltens (1989) stated, “the Ebbinghaus curve does generally fit FL attrition data, except when we are dealing with relatively high levels of proficiency” (p. 12). Thus, the outcomes reported by the forgetting curve hypothesis can be used in FL research to examine an effect of the simultaneous use of these two techniques at a beginners’ level, even though the forgetting curve hypothesis did not provide participants with the meaning of each nonsense word.

The forgetting curve hypothesis was based on the performing of a group comprised of four men whose age range from 22 to 30. Participants in the study memorized eight separate series of 13 different one-syllable nonsense words created by experimenters such as *BUP* and

TOV (see Figure 9). Refer to Ebbinghaus (1885/ 1913), and Murre and Dros (2015) as well. Participants in their studies learned the one-syllable nonsense words by repeating each word in the absence of iconic gestures, of listening practice, and of meaning of each word (see Table 17 for the results of experiments in the forgetting curve hypothesis). Needless to say, teaching students with meaning of each word is crucial in FL class, even if teaching the meaning of words or expressions alone does not assure that they can learn the FL.

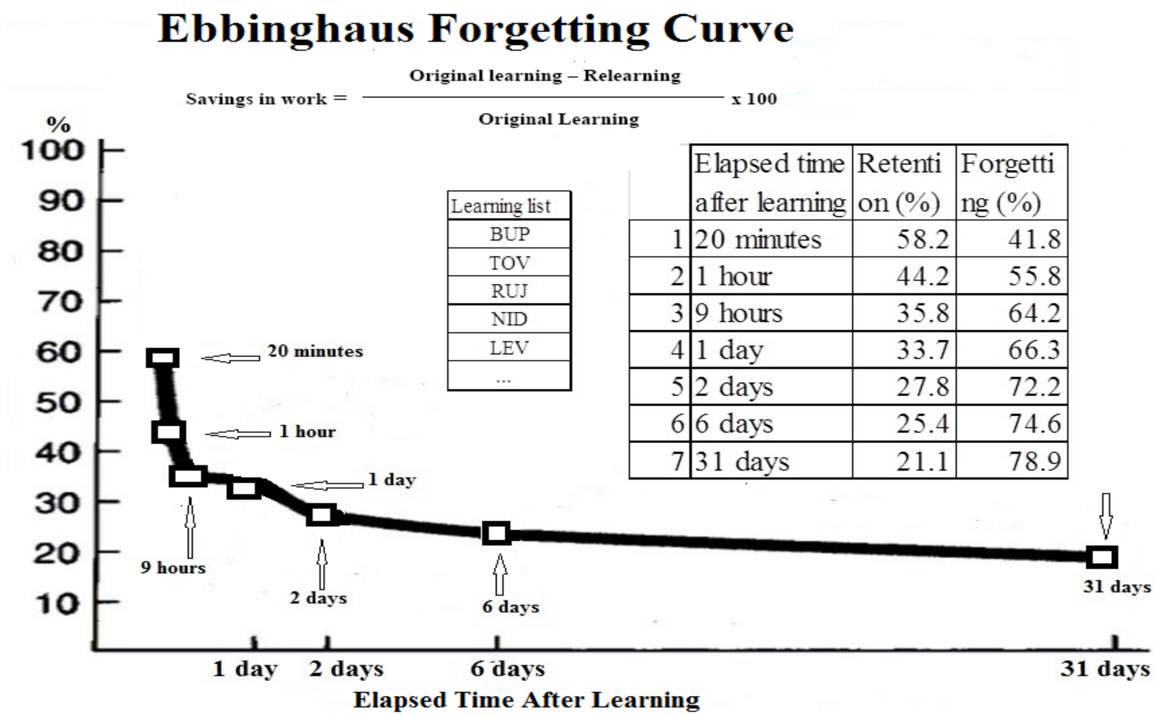


Figure 9. Outcomes presented by Ebbinghaus (1885/ 1913). The author has modified Figure 9 described above based on studies presented by Ebbinghaus (1885/ 1913, p. 66-76), Heller et al. (1991); Murre and Dros (2015)], and Baddeley, Eysenck, & Anderson (2009, p. 234). Scores listed above are based on savings scores.

By contrast, Macedonia et al. (2011) and Quinn-Allen (1995) taught meaning to all participants. When the participants in studies of the forgetting curve hypothesis learned the 13 nonsense words, they remembered 17.78 % of the words on average after 31 days (Table 17 & Figure 9). However, gesture group in Quinn-Allen (1995) remembered 14.2 % of

French expressions in 77 days[101] (cf. Table 18 & Figure 8) and in Macedonia et al.'s (2011) study, participants remembered 11.67% of the words learned with meaningless gestures in approximately 60 days (Table 18). These findings noted above suggested that the French expressions learned by gesture group in Quinn-Allen (1995) and artificial words learned with meaningless gestures in Macedonia et al.'s (2011) study experienced larger memory loss than the participants in Ebbinghaus' study. Thus, teaching meaning of word or of expression alone does not ensure their success in learning FL.

Table 17:

Results Presented by Forgetting Curve Hypothesis (1885/ 1913, 1991, 2015) [Ebbinghaus (1885/ 1913, p. 76); Heller et al. (1991); Murre and Dros (2015)]

	Ebbinghaus	Mack	Seitz	Dros	
Elapsed time after learning	Retention (%)	Retention (%)	Retention (%)	Retention (%)	Average (%)
20 minutes	58.2	54.4	44.2	47.2	51.0
1 hour	44.2	43.2	32.5	37.3	39.3
9 hours	35.8	28.5	27.0	27.6	29.7
1 day	33.7	31.6	27.0	31.7	31.0
2 days	27.8	36.5	28.6	23.0	29.0
6 days	25.4	30.9	20.5	16.8	23.4
31 days	21.1	25.8	20.1	4.1	17.78

Note. The author modified Table 17 noted above based on Murre and Dros (2015, p. 10). Mack and Seitz were participants in Heller et al. (1991) and Dros in Murre and Dros (2015).

[101] Macedonia et al. (2011) assessed students' memory of the artificial nouns approximately 60 days after the last training day and Quinn-Allen (1995) 77 days, whereas the forgetting curve hypothesis measured their participants' memory of one-syllable words after 31days from the day of relearning started. However, these differences between the three studies were not a big issue because newly learned information loss was quite rapid initially, but later slowed down over time, according to the forgetting curve hypothesis (1885/ 1913, 1991, 2015). Therefore, participants of the forgetting curve hypothesis preserved the memory at a similar level in remembering the one-syllable words between 31 and 60 days (cf. Figure 7).

Additionally, no large difference existed among the three studies regarding participants' age. For instance, mean age of participants in Macedonia et al. (2011) was [M age 23.17, SD = 1.61], their age range in the forgetting curve hypothesis was somewhere between 22 and 30, and Quinn-Allen's (1995) participants were the first-semester college students.[102] Furthermore, participants' memory were assessed in both studies of Macedonia et al. (2011) and of the forgetting curve hypothesis at the word level, though Quinn-Allen (1995) assessed memory at the sentence level. [103]

These factors outlined above explain that even though participants in the experiments in the forgetting curve hypothesis were not taught meaning of each word, they learned the one-syllable nonsense words. Taking into considerations of the theoretical justifications and the limitations in Macedonia et al. (2011) explained earlier, the issues addressed in this section, and results in this study, I propose the collaborative analysis of Macedonia et al. (2011), of the forgetting curve hypothesis, and of the Gesture Listening Higher Concept Approach. This joint analysis can lead to a proper investigation of to what extent the concomitant use of ICSCTGs and listening practice can have an effect on maintaining single word meanings (Hypothesis 3).

---

[102] Quinn-Allen (1995) did not specified participants' age.
[103] There were two major differences among the three studies. First, researches who propose the forgetting curve hypothesis tested only 4 participants, whereas Macedonia et al. (2011) had 33 and Quinn-Allen (1995) 112. Second, researchers for the forgetting curve hypothesis used a saving score to assess their memory, but Macedonia et al. (2011) and Quinn-Allen (1995) did not.

Table 18:

French Expressions Learning

	Posttest 1	Final Recall Test
	Mean	Mean
Experimental Group	82.6	14.2
No-Treatment Group	47.1	2.3
Comparison Group	51.9	10.9
Average score (%)	60.5	9.1

Source: Quinn-Allen (1995, p. 526)

5.2.2.2.1.1.4 Joint Analysis of Macedonia et al. (2011), Forgetting Curve Hypothesis, and Gesture Listening Higher Concept Approach

The comparative data analysis of Macedonia et al. (2011), of the forgetting curve hypothesis, and the Gesture Listening Higher Concept Approach allows one to make stronger and more proper arguments of the effect of the concomitant use of iconic gestures and listening practice than Macedonia et al.'s (2011) study alone and the Gesture Listening Higher Concept Approach only. As a result, the joint analysis can provide an additional support for Hypothesis 3.

As argued earlier, the concurrent use of iconic gestures and listening practice in Macedonia et al.'s (2011) study partially supported Hypothesis 3 for 5 days in learning 46 artificial words, however, whether the simultaneous use can support Hypothesis 3 more than 5 days is unclear (see Figure 8). Participants' cognitive load in Macedonia et al. (2011) was initially higher than that of the forgetting curve hypothesis in one day. However, somewhere between 31 and 60 days, the words learned with the simultaneous use of these two techniques outperformed the words acquired in the absence of those two techniques in the forgetting curve hypothesis. Therefore, it appears that the effect of iconic gestures and listening practice together was enormous in learning artificial words.

Nevertheless, after 60 days from the day when the last assessment took place, memory retention of words learned by the two techniques was at a similar level to that of somewhere between the first and the second day, suggesting words learned by the concomitant use deteriorated as time goes by (see Figure 8). Thus, Macedonia et al. (2011) did not explain to what extent the concurrent use helped participants maintain knowledge of the artificial words. I propose re-analyzing data in Macedonia et al. (2011) to investigate to what extent the concurrent use of iconic gestures and listening practice had an effect on learning artificial words between the last assessment and 60 days after the last assessment.

Furthermore, as argued earlier, Macedonia et al. (2011) exposed three limitations, including the reliance of meaningless gestures, the use of artificial words which may have an equivalent difficulty to that of learning the nonsense word in the forgetting curve hypothesis, and the lack of spacing effect that share with the forgetting curve hypothesis. Moreover, as argued earlier, the forgetting curve hypothesis did not provide a meaning of word to participant. Needless to say, teaching students with meaning of each word is crucial in FL class, whereas teaching meaning of words or expressions only does not assure that they can learn FL. Therefore, I argue that the use of the forgetting curve hypothesis is appropriate in re-analyzing data presented by Macedonia et al. (2011).

Additionally, the results in this study revealed that the simultaneous use of ICSCTG and listening practices can have an effect on preserving knowledge of single word meanings for 30-37 days in minimum.

On the basis of the results in this study, the findings in the forgetting curve hypothesis and in Macedonia et al. (2011), I contend that the concurrent use of ICSCTG and listening practice had an effect on learning the developed meaning of words in EFL and maintaining knowledge of it for 30-37 days in minimum.

5.2.3.1.1.5 Forgetting: Discussions in FL Research

Contemporary researchers indicated that students in a beginners' FL class forget easily about what they learned (Bahrick, 1984; Weltens, 1987). Quinn-Allen (1995) claims that even though participants learned meaning of French expressions with emblematic gestures, "none of the three groups improved significantly as the study progressed" (p. 526). As stated earlier, the retention scores in studies by Macedonia et al. (2011) and Quinn-Allen (1995) noted above were worse than those of learning nonsense words in the forgetting curve hypothesis (17.8% retention).

Moreover, the forgetting curve presented by Ebbinghaus (1885/ 1913) shown in Figure 9 displays that newly learned information loss was quite rapid initially (somewhere between 20 minutes and 2 days), but later slowed down over time. As stated earlier, the lack of spacing effect and of the instruction-assessment dualism, and higher cognitive load could lead to attrition. However, given that the circumstances of degree of difficulty in learning FL vary among cultures (FSI) and teaching methods significantly diverge in FL class, it is plausible that forgetting curves in FL instruction and research differ among studies to some extent.

Nonetheless, Figure 7 in Quinn-Allen (1995) and Figure 8 in Macedonia et al. (2011) did not indicate when memory decay in these two studies started slowing down. This finding suggests that participants' cognitive load in the two studies noted above was significant, compared to participants in the forgetting curve hypothesis (1885/ 1913, 1991, 2015), despite the fact that co-speech gesture lightens the cognitive load of speaking (as cited in Kita et al., 2017). Therefore, if the claims presented by Baddeley et al. (2009) and by Weltens (1989) are correct, the issue in the data presented by Quinn-Allen's and in Macedonia et al. (2011) should be appropriately addressed. That is, when instructors make a decision an item that they will teach, they should avoid teaching students by using items which induce a high level of extraneous cognitive load.

In summary, the results of the current study support Hypothesis 3, namely, that the concurrent use of iconic gestures and listening practice not only facilitates learning how to express vertical spatial relationships in English for Japanese EFL high school students, but that it also helps students maintain knowledge long after the instructional intervention. While several prior studies have documented some improvement in FL knowledge retention through the use of concurrent gesture use (Macedonia et al., 2011), no prior study has shown such extensive retention of FL knowledge as the current study. While the simultaneous use of iconic gestures and listening practice generated the best performance among four groups, Group 1 also made some progress in learning how to express vertical spatial relationships in English. The following section 5.2.2.3 provides a potential explanation how learning in Group 1took place.

5.2.2.3 Novelty 3: Higher Concept

The three experimental gesture studies in FL employed two alternative assessments, including Macedonia et al. (2011), Tellier (2008), and this study. In Chapters 2 and 4, I explained positive correlations between the two alternative test scores in these three studies.

Refer to Figure 1 for Tellier (2008), Figure 2 for Macedonia et al. (2011), and Figure 6 for this study.

Based on sociocultural theory, I argue that remembering and knowledge simultaneously improve by creating a higher concept when two different assessment tools evaluate students' knowledge. Refer to Figure 1 for two alternative assessment tools in Tellier (2008) and Figure 2 for them in Macedonia et al. (2011). For instance, memory shifts "from passive recording into a function of active selection and active and intellectual remembering" (Vygotsky, 1999, p. 25). Vygotsky (1998) also argues that memory connects to an archive of knowledge, but not a storehouse of images, and thus, this transformation can create an accumulation of knowledge in a level of scientific thinking (a higher concept). However, the solution of the issue of to what extent the remembering last depends upon pedagogical strategies, including the instruction-assessment dualism and the Gesture Listening Higher Concept Approach, and excluding the limitations as addressed earlier.

I contend that all groups in this study could have obtained accumulated knowledge of semantic and grammatical structures when they took oral speech and cloze tests by creating the positive correlation inasmuch as this positive correlation could trigger a higher concept in learning about vertical axis structures by the transformation explained above. Consequently, all groups learned and maintained developed meaning of words in EFL in the level of a higher concept, including Group 1 that did not receive any treatments. This positive correlation led to simultaneous acquisition of the distinct worldview and of cognitive development in EFL. Importantly, Group 4 acquires distinct worldview and cognitive development in EFL the most effectively because of the effects of the concurrent use and of the positive correlation on learning about vertical axis structures in a higher level of scientific and categorical thinking than the other groups. In this way, this study accomplished one of the three goals, which theoretically and experimentally can respond to the novel research direction in SLA.

5.2.2.4 Distinct Worldview and Cognitive Development in EFL

I argue that distinct worldview and cognitive development in EFL together arises when Japanese EFL students learn vertical spatial structure with the Gesture Listening Higher Concept Approach. Cognitive development in EFL includes a definition as higher mental functions that are not simply a continuation of elementary functions and, “are not their mechanical combination, but a qualitatively new mental formation that develops according to completely special laws and is subject to completely different patterns” (Vygotsky 1998, p. 34). Learning about vertical spatial operations in a Japanese EFL class demands that students create novel psychological formation that develops from the single to binary semantic categorization, leading to generating entirely unique laws. The new psychological formation noted above is subject to absolutely distinct patterns with the use of the Gesture Listening Higher Concept Approach. It appears that gesture is instrumental in helping Japanese EFL students by reducing the level of constraint of linguistic relativity to a certain extent.

Furthermore, Vygotsky (1987) states, “what the child is able to do in collaboration today he will be able to do independently tomorrow” (p. 210), which explains how cognitive development in L1 children takes place in the zone of proximal development that is scaffolding. Bakhurst (2007) argues, “Vygotsky’s dialectical conception is development through qualitative transformation” (p. 68). Additionally, Bein et al. (1993) suggest, “Vygotsky saw the possibility of advancing and improving children’s cognitive activities in the active formation of their higher mental processes” (p. 306). A major theme in Vygotsky’s work is the fundamental role of word meaning (experience or knowledge) in all higher mental processes (as cited in van der Veer, 1997). For this reason, as the meaning of words develops in EFL class, students' mental function and verbal thinking that ties in with word meaning can also develop by creating totally a dissimilar mental pattern from that of their L1. At this moment, gestures triggers generating a different worldview from that of their L1 in Whorfian's sense by orchestrating listening practice and a higher concept, which leads to accelerating verbal thinking.

The simultaneous ICSCTGs and listening practice (helping hands) play a role of generating the zone of proximal development in creating entirely distinct thinking patterns (contact and noncontact distinctions) from that of their L1 when Japanese EFL students express vertical axis structures. Similarly, Vygotsky (1987) also suggests that the higher mental functions develop in the zone of proximal development. Thus, not only Group 1's use of a higher concept,[104] but also Group 2's use of gestures, Group 3's use of listening, and Group 4's use of gestures and listening can advance the accumulated knowledge of developed meaning of words in English (grammatical and semantic structures) in a Japanese EFL class. As a result, students acquired cognitive development (higher mental function) in EFL, leading to generating a different worldview from that of their L1, if they continuously practice in vertical axis operations with developed meaning of words and verbal thinking in their everyday life. Therefore, I argue that distinct worldview and cognitive development in EFL could co-emerge in Japanese EFL education, if word(s) meaning develops in Japanese EFL class.

[104] This higher concept was used in all four groups.

5.2.3 Issue: Comprehending Gestures

Existing studies in L1 showed that language comprehension improved when listeners observe co-speech gestures (Dick et al., 2009; Grisoni, Dreyer & Pulvermuller, 2016; Guenther, 2006; Mashal et al., 2012; Skipper, Goldin-Meadow, Nusbaum & Small, 2009). Sueyoshi and Hardison (2005) reported that the presence of gestures helped students improved the comprehension of bilinguals whose proficiency was low in their L2, but not for high proficiency learners at a college level.

However, Gullberg and De Bot (2010) suggested that no theory or hypothesis concerning viewing gestures had yet been undertaken to explain developmental processes in either children or adults. Similarly, in experimental gesture studies in L1 and FL, the issue whether gestures help with comprehending a message "has been the topic of many research studies over the previous 35 years, and there has been little consensus" (Hostetter, 2011, p. 297). Likewise, the gesture research regarding an effect on comprehending information had not progressed yet since Hostetter's (2011) review paper. Quinn-Allen (1995) claimed that viewing gestures helped students learn language. However, Quinn-Allen did not discuss developmental processes of learning FL by observing co-speech gestures. Interestingly, preschoolers who had not observed gestures developed their memory of English words between second and fourth week as in Tellier's (2008) study. Thus, an effect of viewing gesture on developmental processes in language learning remains unconvincing in this study as well. In the following section, I summarize this experimental gesture study in EFL.

5.2.4 Summary

In conclusion, the fundamental discovery of this study is that the results in the present study supported Hypothesis 3, "Teaching ICSCTG and listening practice combined will help

Japanese EFL learners maintain knowledge of how to express vertical spatial relationships in English."

Furthermore, the outcome in this study also supported Hypothesis 2, "Teaching ICSCTG and listening practice combined will facilitate learning how to express vertical spatial relationships in English for Japanese EFL high school students more than either treatment alone." Despite the fact that theoretical and hypothetical framework has not yet been established in this regard, the findings can have implication for a study of multimodality that is a growing area in gesture studies.

Also, the findings in the current study supported Hypothesis 1, "Teaching ICSCTG will facilitate learning how to express vertical spatial relationships in English for Japanese EFL high school students." This hypothesis is consistent with the prior studies in gesture studies in L1, L2, and FL. Furthermore, the results revealed the possible reason why all groups also learned a vertical spatial relationship and maintained knowledge of it without statistically significant memory decay in a level of a higher concept, which linked to categorical thinking (semantic and grammatical structures) and to the use of two alternative assessment tools.

Consequently, this study theoretically, hypothetically, and experimentally demonstrates the innovative research direction by incorporating the principle of linguistic relativity into SLA, leading to achieving the three goals in this study. Additionally, the present study reveals that the Gesture Listening Higher Concept Approach is analogous to the gesture-for-conceptualization hypothesis regarding the effects of recreating ICSCTGs that help students with learning, problem solving, and remembering. On the basis of results in this experimental study and findings in literature review, I argue that distinct worldview and cognitive development in EFL simultaneously emerge. Importantly, this study demonstrates that not only gestures and listening practice alone, but also the Gesture Listening Higher Concept Approach helps overcome the constraints imposed by the thinking-for-speaking hypothesis.

Furthermore, this approach offers the support process proposed by MacWhinney (2013). Acquiring vertical axis operations with a prepositional phrase in English is the entrenchment[105] for EFL students, specifically if their L1 does not have the obligatory contrast and typologically differs from English in syntax when students learn about vertical spatial operations. Nonetheless, this approach assists students with not only their cognitive development in EFL, but obtaining a different worldview from their L1, if they continuously practice expressing vertical spatial operation. The arguments of developed meaning of words in a certain type of EFL class build upon studies presented by linguistic relativity, sociocultural theory, studies presented by Kabata (2000), Kunihiro (1986), Kunisawa (under review), Munnich et al. (2001), and Tanaka (1997).

Finally, results of data analyses from this study suggested that participants' knowledge in L1 and in EFL in the standardized tests administrated by the prefecture board of education did not have a statistical significant effect on learning about vertical spatial operations. It has to be noted, however, that empirical proof that supports Hypotheses 1, 2, and 3 is inconclusive in cloze test conditions. Thus, the cloze test results do not inform the research questions as expected. In experimental gesture studies in L1 and FL, researchers have little consensus regarding the issue whether gestures help with comprehending a

---

[105]Placing "(someone or something) in a very strong position that cannot easily be changed" http://www.learnersdictionary.com/definition/entrench (Merriam-Webster, downloaded on November 9, 2017).

message. On the basis of discussions and findings in this study, the following section 5.3 includes an explanation of pedagogical implication and suggestions for future research.

5.3 Implications

Based on findings, results, and discoveries in the present study, in the following sections, I present pedagogical implications in FL acquisition first and then follow with suggestions of future research in experimental gesture studies in FL, which may solve issues addressed in previous chapters.

5.3.1 Pedagogical Implications

An implication from this empirical study lies in the importance of the simultaneous use of ICSCTGs and listening practice, when a student's L1 lacks the obligatory contrast for the vertical axis structures and their L1 typologically differs from English in syntax, rather than the use of one of them alone. Besides, the accumulation of knowledge that increases in a level of scientific and categorical thinking is also crucial. Thus, teaching student grammatical and semantic structural differences between their L1 and a target language will help them increase knowledge of FL with the higher concept as in learning about vertical axis operations in this study. Additionally, the use of different types of assessment tools is important to help them advance their abilities of FL.

Needless to say, studies in SLA indicated that individual differences were significant and each teacher differed. Just ing pedagogy in the current study may or may not work. The point is, teachers, instructors, and language professionals may make an adjustment to fit their class by integrating findings of the current study with their teaching experience, their own pedagogy, and taking into consideration their students' age, and language ability in their L1 and ESL (or EFL). Even though the results of the current study suggested participants' abilities in L1and EFL did not influence test scores in both oral speech and cloze test conditions, degree of difficulty differs in learning different languages, according to FSI. Additionally, cultural differences are significant in multicultural classrooms.

My arguments of this pedagogical implication build upon findings in this study and in the literature review as well as nearly 40 years of teaching and research experience in the United States, Canada, and Japan. However, I remain in the process of learning how language professionals can help students in an FL class. The following outlines implications for future research.

5.3.2 Implications for Future FL Experimental Research

5.3.2.1 Long-Term Learning

It still remains unclear whether regenerating gestures can have an effect on long-term learning in FL class (Gullberg, 2014). Future research may investigate the effect of simultaneous ICSCTGs and listening practices with the higher concept on learning about developed meaning of words in EFL for over 30-37 days, including two semesters long study. Through my teaching experience, the concurrent ICSCTGs and listening practices with a higher concept can have an effect on learning developed meaning of words for 7 months at least. The definition of long-term and short-term learning is unclear. However, since Gullberg (2014) used the term, in this section, the present study employs the term to avoid any potential confusion in this regard.

5.3.2.2 Passive Recording

No statistically significant difference was found in group means between the gesture and the non-gesture group(s) in cloze test conditions as observed in this study and a passive recording in Tellier’s (2008) study. Future research may test whether recreating either

ICSCTGs with listening practice, iconic co-thought gestures only, or ICSCTGs alone can have an effect on group mean difference in learning about developed meaning of words in EFL in a cloze test setting or other types of passive recording conditions as in Tellier's study (2008).

A cloze test was invented by theoretically criticizing the discrete-point tests (Oller, 1972; Spolsky, 1969), and thus, a cloze test is the examined method widely employed not only in SLA, but also in L1 acquisition. Furthermore, HI (Kita & Ozyurek, 2003); AGH (Chu & Kita, 2016), and sociocultural theory (Vygotsky, 1987) suggest that co-thought gestures facilitate difficult verbal materials and that speech links to thinking. Therefore, administering a cloze test and/ or a passive recording type of assessment with a demanding task (without any image corresponding to each expression) and co-speech gestures may be considered.

5.3.2.3 Viewing Gestures

Future research may study to what extent observing co-speech gestures has an effect on learning. In Tellier's (2008) study, preschoolers significantly improve their ability to express the words without viewing gestures. However, in Quinn-Allen's (1995) study, observing gestures had an effect on learning French expressions with statistical significance at a college level. Thus, the both studies have a conflict regarding an effect of observing co-speech gestures, even though participants' age differs between the two studies as noted above (Sueyoshi & Hardison, 2005).

5.3.2.4 Higher Concept

Chapter 4 revealed that both Groups 1 and 2 learned developed meaning of words in EFL and maintained this learning without statistically significant memory decay. Future research may study to what extent ICSCTGs alone without a higher concept can help students learning vertical spatial structures.

5.3.2.5 Incorporation of ICSCTG and Listening Practice

The importance exists to establish a theoretical framework with regard to an effect of the simultaneous use of ICSCTGs and listening practice. Future research may investigate a clear theoretical framework in this regard.

References

Agar, M. (1994). *Language Shock: Understanding the Culture of Conversation*. New York, William Morrow and Company.

Alibali, M. W. & Goldin-Meadow, S. (1993). Gesture-speech mismatch and mechanisms of learning: What the hands reveal about a child's state of mind. *Cognitive Psychology 25*: 468-523.

Alibali, M. W., Heath, D. C., & Myers, H. J. (2001). Effects of visibility between speaker and listener on gesture production: Some gestures are meant to be seen. Journal of Memory and Language 44(2): 169-188.

Alibali, M. W., 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.

Alibali, M. W., Yeo, A., Hostetter, A. B., & Kita, S. (2017). Representational gestures help speakers package information for speaking. In R. B. Church, Alibali, M. W. & Kelly, S. D. (Eds.), *Why gesture? How the hands function in speaking, thinking and communicating*. Amsterdam, the Netherlands, John Benjamins: 15-38.

American Council on the Teaching of Foreign Language (ACTFL) (2012). ACTFL Provisional Proficiency Guidelines. See https://www.actfl.org/publications/guidelines-and-manuals/actfl-proficiency-guidelines-2012/english/speaking (accessed October, 13, 2017).

An, Z. (2014). The ambiguity of the meaning of Japanese spatial nouns 'ue' and 'shita'. *Papers from the National Conference of the Japanese Cognitive Linguistics Association 14*: 336-348.

Athanasopoulos, P., Damjanovic, L., Krajciova, A. & Sasaki, M. (2011). Cognitive restructuring in bilingualism. Thinking and Speaking in Two Languages. In A. Pavlenko (Ed.), Bristol, Multilingual Matters: 29-66.

Babaii, E. & Shahri, S. (2010). Psychometric rivalry: The c-test and the cloze test interacting with test takers' characteristics. In R. E. Grotjahn (Ed.), The *C-test: Contributions from Current Research*. Germany, Peter Lang: 41-56.

Baddeley, A. (1986). *Working memory*. Oxford, England, Oxford University Press.

Baddeley, A. (2000). The episodic buffer: A new component of working memory. *Trend in Cognitive Sciences 4*: 417- 423.

Baddeley, A. (2007). *Working Memory, Thought, and Action*. New York, Oxford University Press.

Baddeley, A., Eysenck, M. W. & Anderson, M. C. (2009). *Memory*. New York, Psychology Press.

Baddeley, A., Gathercole, S. & Papagno, C. (1998). The phonological loop as a language learning device. *Psychological Review 105*: 158-173.

Baddeley, A. D. (1997). *Human Memory: Theory and Practice*. Hove, UK, Psychology Press.

Baek, M. (2015). English Preposition over and its Counterparts in Korean. *Journal of Humanities Studies 54(4)*: 431-462.

Bahrick, H. P. (1984). Semantic memory content in permastore: Fifty years of memory for Spanish learned in school. *Journal of Experimental Psychology: General 113(1)*: 1-29.

Bahrick, H. P., Bahrick, L. E., Bahrick, A. S. & Bahrick, P. E. (1993). Maintenance of foreign language vocabulary and the spacing effect. *Psychological Science 4(5)*: 316-321.

Bakhurst, D. (2007). Vygotsky's Demons. In H. Daniels, Cole, M. & Wertsch, J. V. (Eds.), *The Cambridge Companion to Vygotsky*. New York, Cambridge University Press: 50-76.

Bartlett, F. (1932). *Remembering*. Cambridge, Cambridge University Press.

Bate, E. & MacWhinney, B. (1982). Functionalist approaches to grammar. In E. Wanner & Gleitman, L. R. (Eds.), *Language Acquisition: The State of the Art*. New York, Cambridge University Press: 173-218.

Beattie, G. & Coughlan. J. (1999). An experimental investigation of the role of iconic gestures in lexical access using the tip-of-the tongue phenomenon. *British Journal of Psychology 90*: 35–56.

Beattie, G. & Shovelton, H. (1999). Do iconic hand gestures really contribute anything to the semantic information conveyed by speech? An experimental investigation. *Semiotica 123*: 1–30.

Bein, E. S., Vlasova, T. A., Levina, R. E., Morozova, N. G. & Shif, Zh. I. (1993). Afterword. In R. W. E. Rieber & Carton A. S. (Eds.), *The Collected Works of L.S. Vygotsky: Volume 2: Fundamentals of Defectology*. New York, Plenum Press: 302-314.

Bergmann, K. & Macedonia, M. (2013). A Virtual Agent as Vocabulary Trainer: Iconic Gestures Help to Improve Learners' Memory Performance. In Aylett, R., Krenn, B., Pelachaud, C. & Shimodaira, H. (Eds.), *The 13th International Conference on Intelligent Virtual Agents (IVA 2013)*, Edinburgh, UK, Springer: 139-48.

Berman, R. & Slobin, D. (1994). *Relating events in narrative: A crosslinguistic developmental study*. Hillsdale, NJ, Lawrence Erlbaum Associates.

Block, D. (2014). Moving beyond "Lingualism": Multilingual Embodiment and Multimodality in SLA. In *The Multilingual Turn: Implications for SLA, TESOL, and Bilingual Education*. S. E. May (Ed.), New York, Routledge, Taylor & Francis Group: 54-77.

Boas, F. (1911). *The Mind of Primitive Man*. London, UK, Forgotten Books.

Bock, K. & Cutting, J. C. (1992). Regulating mental energy: Performance units in language production. *Journal of Memory and Language 31*: 99-127.

Boers, F. & Demecheleer, M. (1998). A cognitive semantic approach to teaching prepositions. *EFL Journal 52(3)*: 197-204.

Boggess, L. C. (1978). *Computational interpretation of English spatial prepositions*. Unpublished doctoral dissertation, University of Illinois, Urbana, IL.

Bowerman, M. (1996a). Learning how to structure space for language. In *Language and space*. P. Bloom, Peterson, M. A., Nadel, L. & Garrett, M. F. (Eds.), Cambridge, MA, The MIT Press: 385-436.

Bowerman, M. (1996b). The origin of children’s spatial semantic categories: Cognitive versus linguistic determinants. In J. Gumperz & Levinson, S. (Eds.), *Rethinking Linguistic Relativity*. New York, Cambridge University Press: 145-176.

Bowerman, M. & Choi, S. (2001). Shaping meanings for language: universal and language-specific in the acquisition of spatial semantic categories. In M. Bowerman & Levinson, S. (Eds.), *Language acquisition and conceptual development*. Cambridge, UK/New York, USA, Cambridge University Press: 475-511.

Bowerman, M. & Choi, S. (2003). Space under construction: Language-specific spatial categorization in first language acquisition. In D. Gentner & S. Goldin-Meadow (Eds.), *Language in mind. Advances in the study of language and thought*. Cambridge, MA, The MIT Press: 387-428.

Brinton, L. J. & Traugott, E. (2005). *Lexicalization and Language Change*. New York, Cambridge University Press.

Broders, S. C., Cook, S.W., Mitchell, Z.A. & Goldin-Meadow, S. (2007). Making children gesture brings out implicit knowledge and leads to learning. *Journal of Experimental Psychology: General 136*: 539-540.

Bruner, J. (1987). Prologue to the English Edition. In R. W. Rieber & Carton, A. S. (Eds.), *The Collected Works of L.S. Vygotsky: Volume 1: Problems of General Psychology, Including the Volume Thinking and Speech*. New York, Plenum Press: 1-16.

Burns, R. P. & Burns, R. (2008). *Business Research Methods and Statistics Using SPSS*. Los Angeles, SAGE.

Butterworth, B. & Hadar, U. (1989). Gesture, speech and computational stages: A reply to McNeill. *Psychological Review 96*: 168-174.

Bybee, J. (2010). *Language, Usage and Cognition*. New York, Cambridge University Press.

Bybee, J., Perkins, R. & Pagliuca, W. (1994). *The Evolution of Grammar: Tense, Aspect, and Modality in the Languages of the World*. Chicago, IL, University Of Chicago Press.

Carels, P. E. (1981). Pantomime in the foreign-language classroom. *Foreign Language Annals 14:* 407–411.

Carlson, K. & Winquist, J. R. (2018). *An Introduction to Statistics: An Active Learning Approach*. Thousand Oaks, CA, SAGE Publications.

Celce-Murcia, M., Larsen-Freeman, D. & Williams, H. A. (1999). *The grammar book: An ESL/EFL teacher's course*. Rowley, MA, Newbury House.

Chilton, P. (2015). *Language, Space and Mind: The Conceptual Geometry of Linguistic Meaning*. New York, Cambridge University Press.

Chu, M. & Kita, S. (2008). Spontaneous gestures during mental rotation tasks: Insights into the microdevelopment of the motor strategy. *Journal of Experimental Psychology: General 137*: 706-723.

Chu, M. & Kita, S. (2011). The Nature of Gestures' Beneficial Role in Spatial Problem Solving. *Journal of Experimental Psychology: General 140(1)*: 102-116.

Chu, M. & Kita, S. (2012). The Role of Spontaneous Gestures in Spatial Problem Solving. In E. Efthimiou, Georgios, K. & Fotinea, S. E. (Eds.), *Gesture and Sign Language in Human-Computer Interaction and Embodied Communication*. Berlin, Springer: 57-68.

Chu, M. & Kita, S. (2016). Co-thought and Co-speech Gestures Are Generated by the Same Action Generation Process. *Journal of Experimental Psychology: Learning, Memory, and Cognition 42(2)*: 257-270.

Cipris, Z. & Hamano, S. (2002). *Making Sense of Japanese Grammar: A Clear Guide Through Common Problems*. Honolulu, University of Hawaii Press.

Clancy, P. M. (1985). The acquisition of Japanese. In D. I. E. Slobin (Ed.), *The Crosslinguistic Study of Language Acquisition*. Hillsdale, NJ, Lawrence Erlbaum Associates: 373-524.

Clark, E. V. (1973). Non-linguistic strategies and the acquisition of word meanings. *Cognition 2(2)*: 161-82.

Clark, H. (1973). Space, Time, Semantics and the Child. In T. E. More (Ed), *Cognitive Development and the Acquisition of Language*. New York, Academic Press: 27-62.

Clark-Carter, D. (2009). *Quantitative Psychological Research: The Complete Student's Companion*. New York, Psychology Press.

Cohen, J. (1969/1988a). *Statistical power analysis for the behavioral sciences*. Hillsdale, NJ, Lawrence Erlbaum Associates.

Cohen, J. (1988b). Set correlation and contingency tables. *Applied Psychological Measurement 12*: 425-434.

Cohen, J. (1992). A power primer. *Psychological Bulletin, 112(1)*: 155-159.

Cohen, J., Cohen, P., Stephen, G. W. & Aiken, L. S. (2003). *Applied Multiple Regression/Correlation Analysis for the Behavioral Sciences*. New York, Routledge.

Cohen, R. L. (1981). On the generality of some memory laws. *Scandinavian Journal of Psychology 22*: 267–281.

Comrie, B. (1985). *Tense*. New York, Cambridge University Press.

Comrie, B. (1989a). *Language Universals and Linguistic Typology: Syntax and Morphology*. Chicago, IL, University of Chicago Press.

Comrie, B. (1989b). Some general properties of reference-tracking systems. In D. Arnold, Atkinson, M., Durand, J., Grover, C. & Sadler, L. (Eds.), *Essays on grammatical theory and universal grammar*. Gloucestershire, UK, Clarendon Press: 37-51.

Cook, S. W., Yip, T. K. & Goldin-Meadow, S. (2012). Gestures, but not meaningless movements, lighten working memory load when explaining math. *Language and Cognitive Processes 27*: 594–610.

Cook, V. (2003). *Effects of the Second Language on the First*. Clevedon/ Buffalo, Multilingual Matters.

Cooper, H., Nye, B., & Greathouse, S. (1998). Relationships Among Attitudes About Homework, Amount of Homework Assigned and Completed, and Student Achievement. *Journal of Educational Psychology 90(1)*: 70-83.

Cooper, H., Robinson, J. C. & Patall, E. A. (2006). Does Homework Improve Academic Achievement? A Synthesis of Research, 1987-2003. *Review of Educational Research 76(1)*: 1-62.

Coventry, K., Guijarro-Fuentes, P. & Valdes, B. (2012). On the First and Second Language Acquisition of Spatial Language. *Spatial Cognition & Computation 12(4)*: 219-230.

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

Craik, F. I. M. & Watkins, M. J. (1973). The role of rehearsal in short-term memory. *Journal of Verbal Learning and Verbal Behavior 12*: 599-607.

Croft, W. (2003). *Typology and Universals*. New York, Cambridge University Press.

Croft, W. & Cruse, D. A. (2004). *Cognitive Linguistics*. New York, Cambridge University Press.

Cummins, J. (1979a). Cognitive/academic language proficiency, linguistic interdependence, the optimum age question and some other matters. *Working Papers on Bilingualism* 19: 121-9.

Cummins, J. (1979b). Linguistic Interdependence and the Educational Development of Bilingual Children. *Review of Educational Research 49*: 222-251.

Cummins, J. (1980/1981a). Age on arrival and immigrant second language learning in Canada: A reassessment. *Applied Linguistics 1(1)*: 132-149.

Cummins, J. (1981b). The role of primary language development in promoting educational success for language minority students. *Schooling and language minority students: A theoretical framework*. Los Angeles, California State University, National Evaluation, Dissemination and Assessment Center: 3-49.

Cummins, J. (1991a). Interdependence of first- and second-language proficiency in bilingual children. In E. Bialystok (Ed.), *Language processing in bilingual children*. Cambridge, England, Cambridge University Press: 70-89.

Cummins, J. (1991b). The development of bilingual proficiency from home to school: A longitudinal study of Portuguese-speaking children. Journal of Education *173*: 85-98.

Cummins, J. (2000). *Language, Power and Pedagogy: Bilingual Children in the Crossfire*. Clevedon, UK; Buffalo NY, Multilingual Matters.

Cummins, J. (2001). *An introductory reader to the writings of Jim Cummins*. Clevedon, UK, Multilingual Matters.

Cummins, J. (2005). Teaching for Cross-Language Transfer in Dual Language Education: Possibilities and Pitfalls. *TESOL Symposium on Dual Language Education: Teaching and Learning in Two Languages in the EFL Setting. Istanbul, Turkey: Bogazici University*.

Cummins, J. (2009). Fundamental psychological and sociological principles underlying educational success for linguistic minority students. In T. Skutnabb-Kangas, Phillipson, R., Mohanty, A. K. & Panda, M. (Eds.), *Social Justice Through Multilingual Education*. Bristol, UK, Multilingual Matters: 19-35.

Daniels, H. (2007). Pedagogy. In H. Daniels, Cole, M. & Wertsch, J. V. (Eds.), *The Cambridge Companion to Vygotsky*. New York, Cambridge University Press: 307-331.

Davis, M. (2008). The Corpus of Contemporary American English (COCA), Brigham Young University.

de Groot, A. (2011). *Language and Cognition in Bilinguals and Multilinguals: An Introduction*. New York, Psychology Press.

de Groot, A. & Van Hell, J.G. (2005). The learning of foreign language vocabulary. In J. Kroll & de Groot, A. (Eds.), *Handbook of Bilingualism: Psycholinguistic Approaches*. New York: 433-453.

de Guerrero, M. C. M. (2006). *Inner Speech - L*2: *Thinking Words in a Second Language*. New York, Springer.

de Ruiter, J. P. (2000). The production of gesture and speech. In D. McNeill (Ed.), *Language and Gesture*. New York, Cambridge University Press: 284-311.

Dempster, F. N. (1981). Memory span: Sources of individual and developmental. *Psychological Bulletin, 89*: 63-100.

Dick, A. S., Goldin-Meadow, S., Hasson, U., Skipper, J.I. & Small, S. L. (2009). Co-speech gestures influence neural activity in brain regions associated with processing semantic information. *Human Brain Mapping 30(11)*: 3509-26.

Dick, A. S., Goldin-Meadow, S., Solodkin, A. & Small, S. L. (2012). Gesture in the developing brain. *Developmental Science 15(2)*: 165-180.

Dostoevsky, F. (1994/ 1873-1876). *A Writer's Diary*. Evanston, IL, Northwestern University Press.

Ebbinghaus, H. (1885/ 1913). *Memory: A Contribution to Experimental Psychology*. New York, Bureau of Publications, Teachers College, Columbia University.

Elkman, P. & Friesen, W. V. (1969). The Repertoire of Nonverbal Behavior: Categories, Origins, Usage, and Coding. *Semiotica 1(1)*: 49-98.

Ellis, N. (1995). The Psychology of Foreign Language Vocabulary Acquisition: Implications for CALL. *Computer Assisted Language Learning*. *8(2-3)*: 103-128.

Ellis, N. (2001). Memory for language. In P. Robinson (Ed.), *Cognition and Second Language Instruction*. New York, Cambridge University Press: 33-68.

Ellis, N. & Robinson, P. (2008). An Introduction to Cognitive Linguistics, Second Language Acquisition, and Language Instruction. In P. Robinson & Ellis, N. (Eds.), *Handbook of Cognitive Linguistics and Second Language Acquisition*. New York, Routledge: 3-24.

Ellis, N. & Sinclair, S. G. (1996). Working Memory in the Acquisition of Vocabulary and Syntax. *The Quarterly Journal of Experimental Psychology 49A*: 234-250.

Ellis, N. B. & Beaton, A. (1993). Factors affecting the learning of foreign language vocabulary: Imagery keyword mediators and phonological short-term memory. *Quarterly Journal of Experimental Psychology* 46A: 533-558.

Ellis, R. (2015). *Understanding Second Language Acquisition: Second Edition (Oxford Applied Linguistics*). London, UK, Oxford University Press.

Engelkamp, J. (1980). Some studies on the internal structure of propositions. *Psychological Research 41(4)*: 355–371.

Engels, F. (1883/ 1934). Dialectics of Nature. Mosco, Progress publisher.

Feng, F. & Su, X. (2009). Chuugoku-go no hookoo shi no "shita" no imi ni tsuite (On meaning of Chinese directional complement "xia". *Aichi Shukutoku Knowledge Archive 9*: 21-32.

Feyereisen, P. (1998). Le rôle des gestes dans la mémorisation d'énoncés oraux (The role of gestures in the memorization of oral statements. *Oralité et gestualité. Communication multimodale, interaction (Orality and gestuality. Multimodal communication, interaction)*, Paris.

Feyereisen, P. (2006). Further investigation on the mnemonic effect of gestures: Their meaning matters. *European Journal of Cognitive Psychology 18(2)*: 185-205.

Field, A. P. (2013). *Discovering Statistics Using IBM SPSS Statistics: and Sex and Drugs and Rock 'n' Roll*, Sage Publications.

Fiest, M. I. (2000). *On In and On: An investigation into the linguistic encoding of spatial scenes*. Doctoral Dissertation. Northwestern University, Evanston, IL. Retrieved from http://www.ucs.louisiana.edu/~mif8232/thesis.pdf

Flynn, S. (1989). The role of the head-initial/head-final parameter in the acquisition of English relative clauses by adult Spanish and Japanese speakers. In Gass, S. & Schachter, J. (Eds.), *Linguistic Perspectives on Second Language Acquisition*. New York, Cambridge University Press: 89-108.

Foley, W. A. & Van Vallin, R. D. (1984). *Functional Syntax and Universal Grammar*. New York, Cambridge University Press.

Fujii, N. & Ono, T. (2000). The Occurrence and Non-occurrence of the Japanese Direct Object Marker o in Conversation. *Studies in Language and Culture 24(1)*: 1-39.

Gaonac'h, D. (1995). La mémoire dans l'apprentissage des langues vivantes: fonctionnement, acquisitions, utilisation. *Les langues modernes 2*: 9-24.

Gass, S. & Seilinker, L. (1992). Introduction. In S. Gass & Seilinker, L. (Eds.), *Language Transfer in Language Learning*. Amsterdam/ Philadelphia, J. Benjamins Pub. Co.**:** 1-17.

Goldberg, A. (1995). *Constructions: A Construction Grammar Approach to Argument Structure*. Chicago, IL, University of Chicago Press.

Goldenberg, G. (2013). *Apraxia: The Cognitive side of motor control*. Oxford, UK, Oxford University Press.

Goldenberg, G. & Hagmann, S. (1997). The meaning of meaningless gestures: A study of visuo-imitative apraxia. *Neuropsychologia 35*: 333-341.

Goldenberg, G. & Strauss, S. (2002). Hemisphere asymmetries for imitation of novel gestures. *AAN Enterprises, 2*(2): 893-897.

Goldin-Meadow, S. (2003). *Hearing Gesture: How Our Hands Help Us Think*. Cambridge, MA, Belknap Press of Harvard University Press.

Goldin-Meadow, S. (2006). Meeting other minds through gesture: How children use their hands to reinvent language and distribute cognition. In N. J. Enfield & Levinson, S. C. (Eds.), *Roots of Human Sociality: Culture, Cognition and Interaction*. Oxford/ New York, BERG: 353-374.

Goldin-Meadow, S. (2015). From action to abstraction: Gesture as a mechanism of change. *Developmental Review 38*: 167-184.

Goldin-Meadow, S., Cook, S. W. & Mitchell, Z. A. (2009). Gesturing Gives Children New Ideas About Math. *Psychological Science 20*(3): 267-272.

Goldin-Meadow, S., Levine, S., Zinchenko, E., KuangYi Yi, T., Hemani, N. & Factor, L. (2012). Doing gesture promotes learning a mental transformation task better than seeing gesture. *Developmental Science 15*(6): 876-884.

Goldin-Meadow, S., Nusbaum, H., Kelly, S. & Wagner, S. (2001). Explaining Math: Gesturing Lighten the Load. *Psychological Science 12*(6): 516-522.

Gravetter, F. & Forzano, L. (2011). *Research Methods for the Behavioral Sciences*. Belmont, CA, CENGAGE learning.

Greenberg, J. (1972). Numeral classifiers and substantial number: Problems in the genesis of a linguistic type. Working Papers on Language Universals *9*: 1-39.

Greenberg, J. H. (1990). *On Language: Selected Writings of Joseph H. Greenberg*. Stanford, CA, Stanford University Press.

Grisoni, L., Dreyer, F. R. & Pulvermuller, F. (2016). Somatotopic Semantic Priming and Prediction in the Motor System. *Cerebral Cortex 26*: 2353-2366.

Grotjahn, R. (2013). Cloze test. In M. Byram, & Hu, A. (Eds.), *Routledge Encyclopedia of Language Teaching and Learning.* London/ New York, Routledge: 118-9.

Guenther, F. H. (2006). Cortical interactions underlying the production of speech sounds. *Journal of Communication Disorders 39*: 350-360.

Guilford, J. P. (1967). *The nature of human intelligence*. New York, McGraw-Hill.

Gullberg, M. (2009). Why Gestures are Relevant to the Bilingual Lexicon. In A. Pavlenko (Ed.), *The Bilingual Mental Lexicon: Interdisciplinary Approaches*. Bristol, UK, Multilingual Matters: 161-184.

Gullberg, M. (2010). Methodological reflections on gesture analysis in second language acquisition and bilingualism research. *Second Language Research 26(1)*: 75–102.

Gullberg, M. (2014). Gestures and second language acquisition. In C. Muller, Cienki, A., Fricke, E., Ladewig, S. H., McNeill, D. & Bressem, J. (Eds.), *Body - Language - Communication*, *Volume 2*. Berlin, De Gruyter Mouton: 1868-1875.

Gullberg, M. & De Bot, K. (2010). *Gestures in language development*. Amsterdam, Philadelphia, John Benjamins Pub. Co.

Gullberg, M. & McCafferty, S. G. (2008). Introduction to gesture and SLA: Toward an integrated approach. *Studies in Second Language Acquisition 30(2)*: 133-146.

Gumperz, J. J. & Levinson, S. (1996). Introduction: Linguistic relativity re-examined. In J. J. Gumperz & Levinson, S. (Eds.), *Rethinking linguistic relativity*. New York, Cambridge University Press: 1-36.

Hagoort, P. (2005). On Broca, brain, and binding: A new framework. *Trend in Cognitive Sciences 9(9)*: 416-423.

Hagoort, P. & Poeppel, D. (2013). The Infrastructure of the Language-Ready Brain. In M. A. Arbib (Ed.), *Language, Music, and the Brain: A Mysterious Relationship*. Cambridge, MA, The MIT Press: 233-256.

Hamano, S. (1998). *The Sound Symbolic System of Japanese*. Stanford, CA/Tokyo, Japan, CSLI Publications/ Kuroshio.

Hamano, S. (2006). Palatalization in Japanese sound symbolism. In L. Hinton, Nichols, J. & Ohala, J. J. (Eds.), *Sound Symbolism*. New York, Cambridge University Press: 148-160.

Hamers, J. H. M. & Resing, W. C.M. (1992). Learning Potential Assessment: Introduction. In J. H. M. Hamers, Ruijssenaars, A.J.J.M. & Sijtsma, K. (Eds.), *Learning Potential Assessment*. Lisse/ Abingdon/ Exton (PA)/ Tokyo, Swets & Zeitlinger Publishers: 23-42.

Hansen, L., Umeda, Y. & McKinney, M. (2002). Savings in the relearning of second language vocabulary: the effects of time and proficiency. *Language Learning 52(4)*: 653-678.

Hasegawa, Y. H., Y. (2005). What Japanese language tells us about the alleged relational self. *Australian Journal of Linguistics, 25*: 219-251.

Haun, D. B. M., Rapold, C. J., Janzen, G. & Levinson, S. C. (2011). Plasticity of human spatial cognition: Spatial language and cognition covary across cultures. *Cognition 119*: 70-80.

Hawkins, B. W. (1993). On universality and variability in the semantics of spatial adpositions. In Zelinsky-Wibbelt, C (Ed.), *The Semantics of prepositions: From mental processing to natural language processing*. New York, Mouton de Gruyter: 327-350.

Hawkins, J. & Filipovic, L. (2012). *Criterial Features in L2 English: Specifying the Reference Levels of the Common European Framework (English profile Studies).* Cambridge, UK, Cambridge University Press.

Hawkins, J. A. (1990). Seeking motives for change in typological variation. In W. A. Croft, Kemmer, S. & Denning, K. (Eds.), *Studies in Typology and Diachrony: Papers presented to Joseph H. Greenberg on his 75th birthday*. Amsterdam/ Philadelphia, John Benjamins B. V.: 95-128.

Hayashi, O. & Ikegami, A. (1979). Kokugoshi Jiten (A Dictionary of the History of the Japanese Language. Tokyo, Tokyodoo.

Hayward, W. & Tarr, M. (1995). Spatial language and spatial representation. *Cognition 55*: 39-84.

Hegarty, M. & Waller, D. (2005). Individual differences in spatial abilities. In P. Shah & Miyake (Eds.), *Handbook of visuospatial thinking*. Cambridge, England, Cambridge University Press: 121-169.

Heller, O., Mack, W. & Seitz, J. (1991). Replikation der Ebbinghaus'schen Vergessenskurve mit der Ersparnis-methode: "Das Behalten und Vergessen als Function der Zeit" (Replication of the Ebbinghaus forgetting curve with the saving method - retention and forgetting as function of time. *Zeitschrift für Psychologie 199*: 3-18.

Hergenhahn, B. R. & Henley, T. (2013). *An Introduction to the History of Psychology*. Belmount, CA, Wadsworth Publishing.

Herskovits, A. (1986). *Language and spatial cognition: An interdisciplinary study of the prepositions in English*. New York, Cambridge University Press.

Hoiting, N., & Slobin, D. I. (2007). From Gestures to Sign in the Acquisition of Sign Language. In S. D. Duncan, Cassell, J. & Levy, E. T. (Eds.), *Gesture and the Dynamic Dimension of Language: Essays in Honor of David McNeill*. Amsterdam/ Philadelphia, John Benjamins Publishing Co.**:** 51-65.

Holland, D. & Lachicotte, W. (2007). Vygotsky, Mead, and New Sociocultural Studies of Identity. In D. Harry, Cole, M. & Wertsch, J. V. (Eds.), *The Cambridge Companion to Vygotsky*. New York, Cambridge University Press: 101-135.

Horikawa, T. (1988). Kakujoshi ni no imi nituie no ichishiron (A working theory about the meanings of the particle ni). *Tokyo University Linguistics Papers 88*: 321-333.

Hostetter, A. B. (2008). *Mind in Motion: The Gesture as Simulated Action Framework*. Doctoral Dissertation. University of Wisconsin, Madison, WI. Retrieved from https://catalog.hathitrust.org/Record/007967296

Hostetter, A. B. (2011). When Do Gestures Communicate? A Meta-Analysis. *Psychological Bulletin*, *137(2)*: 297–315.

Hostetter, A. B. & Alibali, M. W. (2008). Visible embodiment: Gestures as simulated action. *Psychonomic Bulletin & Review 15(3)*: 495-514.

Hostetter, A. B., Alibali, M. W. & Kita, S. (2007). I see it in my hand's eye: Representational gestures reflect conceptual demands. *Language and Cognitive Processes 22(3)*: 313-336.

Hostetter, A. B., & Boncoddo, R. (2017). Gestures highlight perceptual motor representations in thinking. In R. B. Church, Alibali, M. W. & Kelly, S. D. (Eds.), *Why gesture? How the hands function in speaking, thinking and communicating.* Amsterdam, the Netherlands, John Benjamins: 155-174.

Houston, J. P. (1986). *Fundamentals of learning and memory*. New York, Harcourt, Brace, Jovanovich.

Huddleston, R. & Pullum, G. K. (2002). *The Cambridge Grammar of the English Language*. Cambridge, UK, Cambridge University Press.

Hulstijn, J. H. (2010). Measuring second language proficiency. In E. Blom & Unsworth, S. (Eds.), *Experimental Methods in Language Acquisition Research*. Philadelphia, John Benjamins Pub. Co.: 185-200.

Humboldt, W. F. v. (1836/1999). *Humboldt: 'On Language': On the Diversity of Human Language Construction and its Influence on the Mental Development of the Human Species*. New York, Cambridge University Press.

Ijaz, H. (1986). Linguistic and cognitive determinants of lexical acquisition in a second language. *Language Learning 36(4)*: 401-451.

Iwabuchi, E. & Muraishi, S. (1976). *Yoji no yougo* (*Vocabulary of Children)*. Tokyo, Nihon Hosou Shuppan Kyokai.

Iwasaki, S. (2013). *Japanese: Revised Edition*. Philadelphia, PA, John Benjamins Publishing Company.

Iwasaki, S. & Ono, T. (2002). "Sentence" in spontaneous spoken Japanese discourse. In J. Bybee & Noonan, M. (Eds.) *Complex Sentences in Grammar and Discourse: Essays in Honor of Sandra A. Thompson*. Amsterdam/ Philadelphia, John Benjamins Publishing Co.: 175-203.

Jarvis, S. & Odlin, T. (2000). Morphological type, spatial reference, and language transfer. *Studies in Second Language Acquisition 22*: 535-556.

Jarvis, S. & Pavlenko, A. (2008). *Crosslinguistic Influence in Language and Cognition*. New York, Routledge.

John-Steiner, V. (2007). Vygotsky on Thinking and Speaking. In D. Harry, Cole, M. & Wertsch, J. V. (Eds.), *The Cambridge Companion to Vygotsky*. New York, Cambridge University Press: 136-152.

Johnston, J. R. & Slobin, D. I. (1979). The development of locative expressions in English, Italian, Serbo-Croatian and Turkish. *Journal of Child Language* 6: 529-45.

Kabata, K. (2000). *Japanese ni: A cognitive analysis of a lexically complex particle*. Doctoral Dissertation. University of Alberta, Edmonton, AB, Canada.

Kabata, K. (2014). Interchangeability of so-called interchangeable particles. In K. Kabata & Ono, T. (Eds.), *Usage-based approaches to Japanese grammar: towards the understanding of human language*. Amsterdam/ Philadelphia, John Benjamins Publishing Company: 171-192.

Kabata, K. & Rice, S. (1997). Japanese ni: The Particulars of a somewhat Contradictory Particle. In M. H. Verspoor, Lee, K. D. & Sweetser, E. (Eds.), *Lexical and Syntactical Constructions and the Construction of Meaning*. Amsterdam/ Philadelphia, John Benjamins Publishing Company: 108-127.

Kaiser, S., Ichikawa, Y., Kobayashi, N. & Yamamoto, H. (2013). *Japanese: A Comprehensive Grammar*. New York, Routledge.

Kapatsinski, V. (2014). What is grammar like? A usage-based constructionist perspective. *Linguistics Issues in Language Technology 11(1)*: 1-41.

Kelly, S., Creigh, P. & Bartolotti, J. (2010). Integrating speech and iconic gestures in a stroop-like task: Evidence for automatic processing. *Journal of Cognitive Neuroscience 22(4)*: 683–694.

Kelly, S. D., Barr, D. J., Church, R. B. & Lynch, K. (1999). Offering a hand to pragmatic understanding: The role of speech and gesture in comprehension and memory. *Journal of Memory and Language, 40*: 577-592.

Kelly, S. D., Kravitz, C. & Hopkins, M. (2004). Neural correlates of bimodal speech and gesture comprehension. *Brain Lang 89*: 253–260.

Kelly, S. D., McDevitt, T. & Esch, M. (2009). Brief training with co-speech gesture lends a hand to word learning in a foreign language. *Language and Cognitive Processes 24(2)*: 313-334.

Kendon, A. (1994). Do gestures communicate? A review. *Research on Language and Social Interaction, 27*: 175-200.

Kendon, A. (2004). *Gesture: Visible action as utterance*. Cambridge, U.K, Cambridge University Press.

Kess, J. F. & Miyamoto, T. (1994). *Japanese Psycholinguistics: A Classified and Annotated Research Bibliography.* Amsterdam/ Philadelphia, John Benjamins Publishing Company.

Kim, K. H. S., Relkin, N. R., Lee, K.M. & Hirsch, J. (1997). Distinct cortical areas associated with native and second languages. *Nature 388*: 171-4.

Kindaichi, H. (2011). *The Japanese Language*. Tokyo, Turtle Publishing.

Kirk, E. & Lewis, C. (2016). Gesture Facilitates Children's Creative Thinking. *Psychological Science* 28(2): 225-232.

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

Kita, S. (2014). Production of speech-accompanying gesture. In V. Ferreira, Goldrick, M. & Miozzo, M. (Eds.), *The Oxford Handbook of Language Production*. Oxford, UK, Oxford University Press: 451-459.

Kita, S., Alibali, M. W., & Chu, M. (2017). How Do Gestures Influence Thinking and Speaking? The Gesture-for-Conceptualization Hypothesis. *Psychological Review 124* (3): 245–266.

Kita, S. & Ozyurek, A. (2003). What does cross-linguistic variation in semantic coordination of speech and gesture reveal? Evidence for an interface representation of spatial thinking and speaking. *Journal of Memory and Language 48*: 16-32.

Kita, S. & Ozyurek, A. (2007). How does spoken language shape iconic gestures? In S. Duncan, Cassel, J. & Levy, E. (Eds.), *Gesture and the dynamic dimension of language*. Amsterdam, Benjamins: 67-74.

Kita, S., Ozyurek, A., Allen, S., Brown, A., Furman, R. & Ishizuka, T. (2007). Relations between syntactic encoding and co-speech gestures: Implications for a model of speech and gesture production. *Language and Cognitive Processes 22*(8): 1212-1236.

Kitahara et al. (1981). Nihonbunpoo Jiten (Encyclopedia of Japanese Grammar. Tokyo, Yuuseidoo.

Klammer, T. P., Schulz, M. & Volpe, A. D. (2006). *Analyzing English Grammar*. Harlow, UK, Pearson/Longman.

Ko, R. (2002). Gendei chuugoku-go "Shang" ga shimesu kuukan saikoo (Re-thinking of space "Shang" in Modern Chinese. *Gengo bunka ronshu 23*(2): 81-91.

Koda, K. (1990a). Factors affecting second language text comprehension. In J. Zutell, & McCormick, S. (Eds.), *Literacy theory and research: Analyses from multiple paradigms.* Chicago, IL, National Reading Conference**:** 419-427.

Koda, K. (1990b). The use of L1 reading strategies in L2 reading. *Studies in Second Language Acquisition 12*(4): 393-407.

Koda, K. (2005). *Insights into Second Language Reading: A Cross-Linguistic Approach (Cambridge Applied Linguistics)*. New York, Cambridge University Press.

Kondo, T. & Yamashita, H. (2011). Why Speakers Produce Scrambled Sentences: An Analysis of a Spoken Language Corpus in Japanese. In H. Yamashita, Kondo, T. & Packerd, L. (Eds.), *Processing and Producing Head-final Structures*. New York, Springer: 195-218.

Krauss, R., Chen, Y., & Chawla, P. (1996). Nonverbal Behavior and Nonverbal Communication: What Do Conversational Hand Gestures Tell Us? In M. E. Zanna (Ed.), *Advances in experimental social psychology*. San Diego, Academic Press: 389-450.

Krauss, R., Chen, Y., & Gottesman, R. (2000). Lexical Gestures and Lexical Access: A Process Model. In D. McNeill (Ed.), *Language and gesture*. New York, Cambridge University Press: 261-283.

Krauss, R. & Robert, M. (1998). Why Do We Gesture When We Speak? *Current Directions in Psychological Science. 7*(2): 54-60.

Kumashiro, T. (1994a). How the goal can be the source: The semantics of the Japanese dative marker ni. *The twentieth LACUS forum*: 401-417.

Kumashiro, T. (1994b.) On the conceptual definitions of adpositions and case markers: A case for the conceptual basis of syntax. *Papers from the 30th Regional Meeting of the Chicago Linguistic Society, 1994: The main session*, Chicago Linguistic Society.

Kumashiro, T. (2000). *A Cognitive approach lo Japanese clausal structure.* Doctoral dissertation. University of California. San Diego, CA. UMI Dissertation Services.

Kumashiro, T. (2016). *A cognitive grammar of Japanese clause structure*. Amsterdam/ Philadelphia, John Benjamins Publishing Company.

Kunihiro, T. (1967a). 'And' to 'to, ni, ya, mo'-Nichiei ryogo goi no hikaku ('And' and 'to, ni, ya, and mo' - A Contrastive Analysis of Meaning in Japanese and English. *Gengo Kenkyuu (Journal of the Linguistic Society of Japan) 50*: 34-49.

Kunihiro, T. (1967b). *Kozoteki imiron (Structural Semantics.* Tokyo, Taishukan.

Kunihiro, T. (1986). *Imiron nyuumon (Introduction to Semantics. Gengo 15(12)*: 194-202.

Kunisawa, T. (Under review). Development of Meaning of Words in EFL While Talking About Space. *Language and Sociocultural Theory.*

Kuno, S. (1973). *The structure of the Japanese language*. Cambridge, MA, The MIT Press.

Lakoff, G. (1987). *Women, Fire and Dangerous Things: What Categories Reveal about the Mind*. Chicago, IL, University Of Chicago Press.

Lakoff, G. J., M. (1980). *Metaphors We Live By*. Chicago, University Of Chicago Press.

Lakoff, G. J., M. (1999). *Philosophy in the Flesh: The Embodied Mind and Its Challenge to Western Thought*. New York, Basic Books.

Langacker, R. (1976). Semantic Representations and the Linguistic Relativity Hypothesis. *Foundations of Language 14(3)*: 307-357.

Langacker, R. (1988). A view of linguistic semantics. In B. Rudzka-Ostyn (Ed.), *Topics in Cognitive Linguistics*. Berlin/Philadelphia, John Benjamins Publishing Company: 49-90.

Langacker, R. W. (1987). *Foundations of Cognitive Grammar: VOLUME I Theoretical Prerequisites*. Stanford, CA, Stanford University Press.

Langacker, R. W. (1993). Reference-Point Constructions. *Cognitive linguistics 4(1)*: 1-38.

Langacker, R. W. (2008). *Cognitive Grammar: A Basic Introduction*. New York, Oxford University Press.

Langacker, R. W. (2009). *Investigations in Cognitive Grammar*. Berlin/New York, Mouton de Gruyter.

Lantolf, J. & Thorne, S. L. (2006). *Sociocultural theory and the genesis of second language development*. New York, Oxford University Press.

Larson-Hall, J. (2015). *A Guide to Doing Statistics in Second Language Research using SPSS*. New York, Routledge.

Laufer, B. & Goldstein, Z. (2004). Testing Vocabulary Knowledge: Size, Strength, and Computer Adaptiveness. *Language Learning 54(3)*: 399-436.

Laufer, B. & Nation, I.S.P. (2012). In S. Gass & Mackey, A. (Eds.), Vocabulary. *The Routledge handbook of second language acquisition*. New York, Routledge: 163-176.

Lesaffre, S. & Senn, S. (2003). A note on non-parametric ANCOVA for covariate adjustment in randomized clinical trials. *Statistics in Medicine 22(23)*: 3583-3596.

Levelt, W. (2012). *A History of Psycholinguistics: The Pre-Chomskyan Era*. Oxford, UK, Oxford University Press.

Levinson, S. (1992a). Prime for the field investigation of spatial description and conception. *Pragmatics 2(1)*: 5-47.

Levinson, S. (1992b). Language and cognition: the cognitive consequences of spatial description in Guugu Yimithirr. *Working paper No. 13 Nijmegen: Cognitive Anthropology Research Group*.

Levinson, S. (1996). Relativity in spatial conception and description. In J. J. Gumperz & Levinson, S. (Eds.), *Rethinking Linguistic Relativity*. New York, Cambridge University Press: 177-202.

Levinson, S. (1997). From outer to inner space: linguistic categories and non-linguistic thinking. In J. P. Nuyts, E. (Eds.), *Language and conceptualization*. New York, Cambridge University Press: 13-45.

Levinson, S. (2003a). Language and Mind: Let's get the issues straight! In Gentner, D. & Goldin-Meadow, S. (Eds.), *Language in mind: Advances in the study of language and thought*. Cambridge, MA, The MIT Press: 25-46.

Levinson, S. (2003b). *Space in Language and Cognition: Explorations in Cognitive Diversity*. New York, Cambridge University Press.

Levinson, S. & Wilkins, D. P. (2006a). Patterns in the data: towards a semantic typology of spatial description. In S. Levinson & Wilkins, D. P. (Eds.), *Grammars of Space: Explorations in Cognitive Diversity*. New York, Cambridge University Press: 512-552.

Levinson, S. & Wilkins, D. P. (2006b). The background to the study of the language of space. In S. Levinson & Wilkins, D. P. (Eds.), *Grammars of Space: Explorations in Cognitive Diversity*. New York, Cambridge University Press: 1-23.

Li, C. N. & Thompson, S. A. (1976). Subject and topic: A new typology of language. In C. N. Li (Ed.), *Subject and topic*. London, UK, *Academic Press*: 457-489.

Li, J. & Cai, J. (2015). L1 transfer in Chinese Learners' Use of Spatial Preposition. In T. Odlin & Yu, L. (Eds.), *New Perspectives on Transfer in Second Language Learning*. Bristol, Buffalo, Toronto, Multilingual Matters: 63-75.

Li, P. & Gleitman, L (2002). Turning the tables: Language and spatial reasoning. *Cognition 83*: 265-294.

Loewen, S. & Plonsky, L. (2015). *An A–Z of Applied Linguistics Research Methods*. London, UK, Palgrave Macmillan.

Lotto, A. & Holt, L. (2011). Psychology of auditory perception. *WIREs Cognitive Science 2:* 479–489.

Lucy, J. (1985). Whorf's View of the Linguistic Mediation of Thought. In E. Mertz & Parmentier, R. J. (Eds.), *Semiotic mediation: sociocultural and psychological perspectives*. New York, Academic Press: 73-97.

Lucy, J. (1996). The scope of linguistic relativity: An analysis and review of empirical research. In J. J. Gumperz & Levinson, S. (Eds.), *Rethinking linguistic relativity*. New York, Cambridge University Press: 37-69.

Lucy, J. & Wertsch, J. (1987). Vygotsky and Whorf: A Comparative Analysis. In M. E. Hickmann (Ed.), *Social and functional approaches to language and thought*. FL, Academic Press: 67-86.

Luria, A. R. (1972). *The Man with a Shattered World: The History of a Brain Wound*. New York, Basic Books.

Luria, A. R. (1973/ 2011). *Basic Problems of Neurolinguistics*. The Hague-Paric, De Gruyter Mouton.

Luria, A. R. (1987). Afterword to the Russian Edition. In R. W. Rieber & Carton, A. S. (Eds.), *The Collected Works of L.S. Vygotsky: Volume 1: Problems of General Psychology, Including the Volume Thinking and Speech*. New York, Plenum Press: 359-383.

Macedonia, M. (2003). Sensorimotor enhancing of verbal memory through 'Voice Movement Icons' during encoding of foreign language (German: Voice Movement Icons. Sensomotorische Encodierungsstrategie zur Steigerung der quantitativen und qualitativen Lerneffizienz bei Fremdsprachen. Salzburg, University of Salzburg.

Macedonia, M. (2014). Bringing back the body into the mind: gestures enhance word learning in foreign language. *Frontiers in Psychology 5*: 1-6.

Macedonia, M. & Muller, K. (2016). Exploring the Neural Representation of Novel Words Learned through Enactment in a Word Recognition Task. *Front Psychology 7*: 1-14.

Macedonia, M., Muller, K. & Friederici, A. D. (2010). Neural Correlates of High Performance in Foreign Language Vocabulary Learning. *Mind Brain and Education 4(3)*: 125-134.

Macedonia, M., Muller, K. & Friederici, A. D. (2011). The Impact of Iconic Gestures on Foreign Language Word Learning and Its Neural Substrate. *Human Brain Mapping 32*: 982-998.

Macedonia, M. & Repetto, C. (2016). Brief Multisensory Training Enhances Second Language Vocabulary Acquisition in Both High and Low Performers. *International Journal of Learning, Teaching and Educational Research 15(3)*: 42-53.

Macedonia, M. & von Kriegstein, K. (2012). Gestures Enhance Foreign Language Learning. *Biolinguistics 6(3-4)*: 393-416.

Mackey, A. & Gass, S. (2005). *Second Language Research: Methodology and Design*. Mahwah, NJ, Lawrence Erlbaum Associates.

MacWhinney, B. (1989). Competition and lexical categorization. In M. Noonan, Eckman, R. & Corrigan, R. (Eds.), *Linguistic Categorization.* Amsterdam, John Benjamins Publishing Company: 195-242.

MacWhinney, B. (2008). A unified model. In P. Robinson & Ellis, N. (Eds.), *Handbook of Cognitive Linguistics and Second Language Acquisition*. Mahwah, NJ, Lawrence Erlbaum Associates**:** 341-370.

MacWhinney, B. (2013). The Logic of the Unified Model. In Gass, S. & Mackey, A. (Eds.), *The Routledge Handbook of Second Language Acquisition*. New York, Routledge: 211-227.

MacWhinney, B. (2014). Introduction. In B. MacWhinney. *The Handbook of Language Emergence*. Hoboken, NJ, Wiley-Blackwell**:** 1-32.

Makino, S. (1978). Kotoba to kuukan (Language and Space. Tokyo, Tookai daigaku shuppan kai.

Malt, B. C., Gennari, S. & Imai, M. (2010). Lexicalization Patterns and the World-to-Words Mapping. In B. Malt & Wolff, P. M. (Eds.), *Words and the Mind: How words capture human experience*. New York, Oxford University Press: 29-57.

Mariani, L. (1981). English through mime. *English Teaching Forum 19*: 36-7.

Martin, S. (1975). *A Reference Grammar of Japanese*. University of Hawaii Press, Honolulu, HI.

Maruo, M. (2004). A Study of Locative Words in Chinese: From the Viewpoint of a Place as Thing or Location. *Studies in Language and Culture 25(2)*: 151-166.

Maruo, M. (2014). Extended Meanings of the Directional Complements “xia(lai/qu)” in Chinese: Relations with the Concept of Quantity. *Studies in Language and Culture 35(2)*: 83-97.

Mashal, N., Andric, M. & Small, S. (2012). Motor and Nonmotor Language Representations in the Brain. In M., Faust (Ed.), *The Handbook of the Neuropsychology of Language, Volume 1*. M. Faust. Hoboken, NJ, Wiley-Blackwell: 276-293.

Masuda, K. (2002). *A Cognitive Approach to Japanese Locative Postpositions ni and de a Case Study of Spoken and Written Discourse*. Interdisciplinary Program in Second Language Acquisition and Teaching. The University of Arizona, Tucson, AZ.

Masuda, K. (2007). Japanese postpositions ni and de: A cognitive linguistic approach. *CLS 39-2. The panels: papers from the 39th annual meeting of the Chicago Linguistic Society*, Chicago, AuthorHouse.

Masuda, K. & Arnett, C. (2015). Cognitive linguistics, Sociocultural Theory and language teaching: Introduction 1. In K. Masuda, Arnett, C. & Labarca, A. (Eds.), *Cognitive Linguistics and Sociocultural Theory Applications for Second and Foreign Language Teaching*. Berlin/ Boston, De Gruyter Mouton: 1-24.

Masuda, K. & Labarca, A. (2015). Schematic diagram use and languaging quality in learning Japanese polysemous particles ni and de. In K. Masuda, Arnett, C. & Labarca, A. (Eds.), *Cognitive Linguistics and Sociocultural Theory Applications for Second and Foreign Language Teaching*. Berlin/ Boston, De Gruyter Mouton: 203-232.

Matsumoto, Y. (2003). *Ninchi Imi-Ron* (*Cognitive Semantics)*. Tokyo, Taishu kan sho-ten.

Matyushkin, A. M. (1997). *Epilogue*. In R. W. Rieber (Ed.), *The Collected Works of L.S. Vygotsky: Volume 4: The History of the Development of Higher Mental Function*. New York, Plenum Press: 261-278.

Mayberry, R. & Nicoladis, E. (2000). Gesture Reflects Language Development: Evidence from Bilingual Children. *Current Directions in Psychological Science 9(6)*: 192-196.

Mayer, M. (1969). *Frog, Where are You?* New York, Dial Books for Young Readers.

McCafferty, S. G. & Stam, G. (2008). Gesture Studies and Second Language Acquisition: A Review. In S. G. McCafferty & Stam, G. (Eds.), *Gesture: Second Language Acquisition and Classroom Research (Esl and Applied Linguistics).* New York, Routledge: 3-24.

McCraw, M. (2010). *First-person singular pronouns in Japanese: How do they work in conversation?* Doctoral dissertation. The University of New Mexico, Albuquerque, NM.

McGeoch, J. A. & Irion, A. L. (1952). *The psychology of human learning*. New York, McKay.

McMillan, J. H. & Schumacher, S. (2009). *Research in Education: Evidence-Based Inquiry*. Boston, MA, Pearson.

McNeill, D. (1992). *Hand and mind: what gestures reveal about thought*. Chicago, IL, University of Chicago Press.

McNeill, D. (2005). *Gesture and Thought*. Chicago/London, University of Chicago Press.

McNeill, D. (2008). Imagery for speaking. In J. Guo, Lieven, E., Ervin-Tripp, S., Nakamura, K., & Ozucaliskan, S. (Eds.), *Crosslinguistic Approaches to the Psychology of Language: Research in the Tradition of Dan Isaac Slobin*. London, UK, Psychology Press: 517-530.

McNeill, D. (2012). *How Language Began: Gesture and Speech in Human Evolution*. New York, Cambridge University Press.

McNeill, D., Cassell, J. & McCullough, K-E. (1994). Communicative effects of speech-mismatched gestures. *Research on Language Interaction 27(3)*: 233-237.

McNeill, D. & Duncan, S. (2000). Growth Points in Thinking-for-Speaking. In D. McNeill (Ed.), *Language and Gesture*. New York, Cambridge University Press: 141-161.

Miller, G. A. (1956). The Magical Number Seven, Plus or Minus Two: Some Limits on our Capacity for Processing Information. *Psychological Review 63*: 81-97.

Miller, G. A. (1969). The organization of lexical memory: are associations sufficient? In G. Talland & Waugh, N. (Eds.), *The Pathology of Memory*. Cambridge, MA, Academic Press: 223-236.

Milton, J., Richards, B., Mantyla, K. Treffers-Daller, J., Bell, H.Miralpeix, I. & Daller, M. (2010). Essential Readings in Vocabulary Studies Vocabulary Studies SIG Colloquium: *Applied Linguistics, Global and Local: Proceedings of the 43rd Annual Meeting of the British Association for applied Linguistics 9-11 September 2010*. R. M. D. Millar, M. Aberdeen, UK.

Mittelberg, I. (2008). Peircean semiotics meets conceptual metaphor: Iconic modes in gestural representations of grammar. In C. Muller & Cienki, A. (Eds.), *Metaphor and Gesture.* Amsterdam/ Philadelphia, John Benjamins Publishing Company: 115-154.

Mittelberg, I. & Waugh, L. R. (2009). Metonymy first, metaphor second: A cognitive-semiotic approach to multimodal figures of thought in co-speech gesture. In C. Forceville & Urios-Aparisi, E. (Eds.), *Multimodal Metaphor*. Berlin/New York, M. de Gruyter: 329-356.

Miyajima, T. (1972). *Doshi no imi Yoohou no kijutsu-teki kenkyuu (Descriptive studies on meanings of verbs and its usage.* Tokyo, Shuei shuppan.

Mol, L. & Kita, S. (2012). Gesture structure affects syntactic structure in speech. In N. Miyake, D. Peebles, & R. P. Cooper (Eds.), *Proceedings of the 34th Annual Conference of the Cognitive Science Society*: 761–766, Austin, TX, Cognitive Science Society.

Morgan, G. A., Leech, N. L., Gloeckner, G. W. & Barrett, K. C. (2012). *IBM SPSS for Introductory Statistics: Use and Interpretation, Fifth Edition*. New York, Routledge.

Morrison, R. (1882). *Dictionary Chinese Language*. Macao, China, East India Company's Press.

Moskowitz, G. (1976). The classroom interaction of outstanding foreign language teachers. *Foreign Language Annals 9*: 135-157.

Munnich, E. & Landau, B. (2003). The Effects of Spatial Language on Spatial Representation: Setting Some Boundaries. In D. Gentner & Goldin-Meadow, S. (Eds.), *Language in mind: advances in the study of language and thought*. Cambridge, MA, The MIT Press: 113-155.

Munnich, E. & Landau, B. (2010). Developmental Decline in the Acquisition of Spatial Language. *Language Learning and Development 6(1)*: 32-59.

Munnich, E., Landau, B. & Dosher, B. (2001). Spatial language and spatial representation: Across-linguistic comparison. *Cognition 81*: 171-207.

Murre, J. M. J. & Dros, J. (2015). Replication and Analysis of Ebbinghaus' Forgetting Curve. *PLoS ONE. 10(7)*: 1-23.

Myers, C. E. (2008). *Development of the Trauma Play Scale: Comparison of Children Manifesting a History of Interpersonal Trauma with a normative sample*. North Texas.

Nariyama, S. (2000). Markedness theory and ellipsis: ellipsis as the unmarked representation of 'sameness'. *Melbourne Papers in Linguistics and Applied Linguistics 1*: 55-68.

Nariyama, S. (2003). *Ellipsis and Reference Tracking in Japanese*. Philadelphia, PA, John Benjamins Pub Co.

Neilson, K. (1990). Essay Review: Development of meaning and meaning of development in the single word period: Early Lexical Development. *First Language 10:* 61-73.

Nicoladis, E., Pika, S. & Marentette, P. (2009). Do French-English Bilingual Children Gesture More Than Monolingual Children?" *Journal of Psycholinguistic Research 38(6)*: 573-585.

Niimura, I. (1969). *Koji en*. Tokyo, Iwanami shoten.

Nilpel, C. (2013). Distinctive functions of quotative markers: Evidence from Meidai Kaiwa Corpus. Gengo bunka ronshu *35(1)*: 87-103.

Nilsson, L. G., Cohen, R.L. & Nyberg, L. (1989). Recall of enacted and nonenacted instructions compared: Forgetting functions. *Psychological Research 51*: 188-193.

Norman, D. A. & Bobrow, D.G. (1975). On data-limited and resource-limited processes. *Cognitive Psychology 7*: 44-64.

Novack, M. A., Congdon, E. L., Hemani-Lopez, N. & Goldin-Meadow, S. (2014). From action to abstraction: Using the hands to learn math. *Psychological Science 25*(*4*): 903-910.

Odlin, T. (1989). *Language Transfer: Cross-Linguistic Influence in Language Learning*. New York, Cambridge University Press.

Ohso, M. (2003). Nagoya University Conversation Corpus, National Institute for Japanese Language and Linguistics. 2016.

Oller, J. W. (1972). Assessing competence in ESL: Reading. *TESOL Quarterly 6(4)*: 313-323.

Ono, T. & Thompson, S. (1997). Deconstructing "Zero Anaphora" in Japanese. *Proceedings of the Twenty-Third Annual Meeting of the Berkeley Linguistics Society*: 481-491.

Ono, T., Thompson, S. & Suzuki, R. (2000). The pragmatic nature of the so-called subject marker ga in Japanese: evidence from conversation. *Discourse Studies 2(1)*: 55-84.

Onodera, N. (2004). *Japanese Discourse Markers: Synchronic and Diachronic Discourse Analysis*. Amsterdam/ Philadelphia, John Benjamins Publishing Company.

O'Reilly, R. C., Braver, T. S. & Cohen, J. D. (1999). A biologically based computational model of working memory. In A. Miyake & Shah, P. (Eds.), *Models of working memory: Mechanisms of active maintenance and executive control*. New York, Cambridge University Press: 375–411.

Ortega, L. (2015). Usage-based SLA: A Research Habitus Whose Time Has Come. In T. Cadierno & Eskildsen, S. W. (Eds.), *Usage-based perspectives on second language learning*. Berlin, Morton de Gruyter: 353-376.

Ozcaliskan, S., Lucero, C. & Goldin-Meadow, S. (2016). Does language shape silent gesture? *Cognition 148*: 10-18.

Paggio, P. (2012). Towards an emplically-based grammar of speech and gesture. In P. Bergmann, Brenning, J., Pfeiffer, M. & Rebe, E. (Eds.), *Prosody and Embodiment in Interactional Grammar*. Berlin/Boston, de Gruyter: 281-314.

Papagno, c., Valentine, T. & Baddeley, A. D. (1991). Phonological short-term memory and foreign-language vocabulary learning. *Journal of Memory and Language 30*: 331-47.

Pavlenko, A. (2000). L2 influence on L1 in late bilingualism. *Issues in Applied Linguistics 11*: 175-205.

Pavlenko, A. (2005). Bilingualism and Thought. In J. Kroll & De Groot, A. M. B. (Eds.), *Handbook of Bilingualism: Psycholinguistic Approaches*. New York, Oxford University Press: 433-453.

Pavlenko, A. (2011a). Introduction: Bilingualism and Thought in the 20th Century. In A. Pavlenko. (Ed.), *Thinking and Speaking in Two Languages*. Bristol, UK, Multilingual Matters: 1-28.

Pavlenko, A. (2011b). Thinking and Speaking in Two Languages: Overview of the Field. In A. Pavlenko. (Ed.), *Thinking and Speaking in Two Languages*. Bristol, UK, Multilingual Matters: 237-257.

Pavlenko, A. (2014). *The Bilingual Mind: And What It Tells Us about Language And Thought*. New York, Cambridge University Press.

Pavlik, P. & Anderson, J. (2005). Practice and forgetting effects on vocabulary memory: An activation-based model of the spacing effect. *Cognitive Science 29*: 559-586.

Pederson, E., Danziger, E., Wilkins, D., Levinson, S., Kita, S. & Senft, G. (1998). Semantic Typology and Spatial Conceptualization. *Language 74(3)*: 557-589.

Perry, M., Church, R. B. & Goldin-Meadow (1992). Is Gesture-Speech Mismatch A General Index of Transitional Knowledge? *Cognitive Development 7*: 109-122.

Petrick-Steward, E. (1995). *Beginning Writers in the Zone of Proximal Development*. Hillsdale, N.J., L. Erlbaum Associates.

Pettito, L. A. (1992). Modularity and constraints in early lexical acquisition: Evidence from children's first words/signs and gestures. In M. R. Gunnar & Maratsos, M. (Eds.), *Modularity and Constraints in Language and Cognition.* Hillsdale, NJ, L. Erlbaum Associates: 25-58.

Pharkit, A. (2015). *Experimental Research Methods in Language Learning*. New York, Bloomsbury Academic.

Piaget, J. (1927/ 2005). *La Causalite Physique chez l'Enfant (The Child's Conception of Physical Causality.* Paris, Alcon/Routledge.

Piaget, J. (1970). *Psychology and epistemology: towards a theory of knowledge*. London, Penguin.

Pimsleur, P. (1967). A memory schedule. *The Modem Language Journal 51*: 73-75.

Ping, R., Goldin-Meadow, S. & Beilock, S. L. (2014). Understanding gesture: Is the listener's motor system involved?" *Journal of Experimental Psychology: General 143*(1): 195-204.

Platt, E. & Brooks, F. B. (2008). Embodiment as Self-Regulation in L2 Task Performance. In S. G. McCafferty & Stam, G. (Eds.), *Gesture: Second Language Acquisition and Classroom Research (Esl and Applied Linguistics).* New York, Routledge: 66-87.

Poehner, M. E. (2008). *Dynamic Assessment: A Vygotskian Approach to Understanding and Promoting L2 Development (Educational Linguistics)*. Berlin/ New York, Springer.

Poeppel, D. & Overath, T. (2012). Introduction: Why Human Auditory Cortex? In D. Poeppel, Overath, T., Popper, A. N. & Fay, R. R. (Eds.), *The Human Auditory Cortex*. New York, Springer: 1-12.

Porter, A. (2012). A helping hand with language learning: teaching French vocabulary with gesture. *The Language Learning Journal 44(2)*: 236–256.

Quinn-Allen, L. (1995). The effect of emblematic gestures on the development and access of mental representations of French expressions. *Modern Language Journal 79(4)*: 521-529.

Quinn-Allen, L. (2000). Nonverbal accommodations in foreign language teacher talk. *Applied Language Learning 11*: 155-176.

Radden, G. & Dirven, R. (2007). *Cognitive English Grammar*. Berlin, John Benjamins Publishing.

Rantner, C. (1991). *Vygotsky's sociohistorical psychology and its contemporary applications*. New York, Plenum Pres.

Rantner, C. (1998). Prologue. In R. W. Rieber (Ed.), *The Collected Works of L.S. Vygotsky: Volume 5: Child Psychology*. New York, Plenum Press: v-xv.

Rauschecker, J. P. (2012). Processing streams in auditory cortex. In Y. E. Cohen, Popper, A. N. & Fay, R. R. (Eds.), *Neural Correlates of Auditory Cognition*. New York: Springer: 7-44.

Rauscher, F. H., Krauss, R. M. & Chen, Y. (1996). Gesture, Speech, and Lexical Access: The Role of Lexical Movements in Speech Production. *Psychological Science 7*(4): 226-231.

Ravizza, S. (2003). Movement and lexical access: Do noniconic gestures aid in retrieval? *Psychon Bull Rev 10*(3): 610-615.

Richardson, J. T. E. (2011). *Eta squared and partial eta squared as measures of effect size in educational research*. Educational Research Review *6*: 135-147.

Rimè, B., Shiaratura, L., Hupet, M. & Ghysselinckx, A. (1984). Effects of relative immobilization on the speaker’s nonverbal behavior and on the dialogue imagery level. *Motivation and Emotion 8*: 311-325.

Robbins, D. (1999). Prologue. In R. W. Rieber (Ed.), *The Collected Works of L. S. Vygotsky Volume 6: Scientific Legacy*. New York, Plenum Press: v-xii.

Robinson, P. & Ellis, N. (2008). Conclusion: Cognitive linguistics, second language acquisition and L2 instruction issues for research. In P. Robinson & Ellis, N. (Eds.), *Handbook of cognitive linguistics and second language acquisition*. New York, Routledge: 489-545.

Rohlfing, K. J. (2006). Facilitating the acquisition of UNDER by means of IN and ON – a training study in Polish. *Journal of Child Language 33*: 51-69.

Sadler, M. (2007). Grammar in Use Across Time and Space: Deconstructing the Japanese 'dative subject' construction. Amsterdam/ Philadelphia, John Benjamins Publishing Company.

Saito, M. (1985). *Some asymmetries in Japanese and their theoretical implications.* Cambridge, MA, MIT.

Saltz, E. & Donnenwerth-Nolan, S. (1981). Does motoric imagery facilitate memory for sentences? *Learning and Verbal Behavior 20*: 322–332.

Salvato, G. (2015). *Looking Beyond Words: Gestures in the Pedagogy of Second Languages in Multilingual Canada*. Newcastle upon Tyne, UK, Cambridge Scholars Publishing.

Sapir, E. (1921). *Language: An introduction to the study of speech*. New York, Harcourt, Brace & Company.

Sapir, E. (1949/1985). *Selected writings of Edward Sapir in Language, Culture, and Personality*. Berkeley/Los Angeles, CA, University of California Press.

Schacht, A., Bogaerts, K., Bluhmki, E. & Lesaffre, E. (2008). A New Nonparametric Approach for Baseline Covariate Adjustment for Two-Group Comparative Studies. *Biometrics, 64*(4): 1110-1116.

Schaller, S. (2012). *A Man Without Words*. Berkeley, CA, University of California Press.

Schewe, M. & Shaw, P. (1993). *Towards Drama as a Method in the Foreign Language Classroom*. Frankfurt am Main, Peter Lang.

Schumann, J. (1986). Locative and directional expressions in basilang speech. *Language Learning, 36*: 277–94.

Sekine, K., Stam, G., Yoshioka, K., Tellier, M. & Capirci, O. (2015). Cross-linguistic Views of Gesture Usage. *Vigo International Journal of Applied Linguistics VIAL 12*: 91-106.

Service, E. (1992). Phonology, working memory, and foreign-language learning. *The Quarterly Journal of Experimental Psychology Section A: Human Experimental Psychology, 45*(1): 21-51.

Service, E. & Kohonen, V. (1995). Is the relation between phonological memory and foreign language learning accounted for by vocabulary acquisition? *Applied Psycholinguistics, 16*, 155-172.

Shepard, R. N. & Metzler, J. (1971). Mental rotation of three-dimensional objects. *Science 171*: 701–703.

Shibatani, M. (1978). Mikami Akira and 'Subject' in Japanese grammar. In J. Hinds & Howard, I. (Eds.), *Problems in Japanese syntax and semantics*. Tokyo, Kaitakusha: 52-67.

Shibatani, M. (2001/2013). *The Languages of Japan*. New York, Cambridge University Press.

Sinha, C. (1983). Background Knowledge, Presupposition, and Canonicality. In W. Wannenmacher & Seiler, T. B. (Eds.), *Concept Development and the Development of Word Meaning*. Berlin, New York, Springer-Verlag: 269-296.

Sinha, C., Thorseng, L. A., Hayashi, M. & Plunkett, K. (1999). Spatial Language Acquisition: In Danish, English and Japanese. In P. M. Broeder, J. & Murre, J. (Eds.), *Language and Thought in Development*. Württemberg, Germany, G. Narr Verlag: 95-126.

Skipper, J. I., Goldin-Meadow, S., Nusbaum, H. C. & Small, S. L. (2009). Gestures orchestrate brain networks for language understanding. *Current Biology 19*: 661-667.

Slobin, D. (1987). Thinking for speaking. *Proceedings of the Thirteenth Annual Meeting of the Berkeley Linguistics Society*: 435-444.

Slobin, D. (1991). Learning to think for speaking: Native language, cognition, and rhetorical style. *Pragmatics 1(1)*: 7-25.

Slobin, D. (1996). From "thought and language" to "thinking for speaking". In J. J. Gumperz & Levinson, S. (Eds.), *Rethinking Linguistic Relativity*. New York, Cambridge University Press: 70-96.

Slobin, D. (1997). The universal, the typological, and the particular in acquisition. In D. E. Slobin (Ed.), *The crosslinguistics study of language acquisition Volume 5*: *Expanding the Contexts*. Mahwah, NJ, Lawrence Erlbaum Associates, Publishers: 1-39.

Slobin, D. (2000). Verbalized Events A Dynamic Approach to Linguistic Relativity and Determinism. In S. Niemeier & Dirven, R. (Eds.), *Evidence for linguistic relativity*. Amsterdam, John Benjamins: 107-138.

Slobin, D. (2001). Form-function relations: How do children find out what they are? In M. Bowerman & Levinson, S. (Eds.), *Language acquisition and conceptual development*. Cambridge, UK, Cambridge University Press: 406-449.

Slobin, D. (2003). Language and thought online: Cognitive consequences of linguistic relativity. In D. Gentner & Goldin-Meadow, S. (Eds.), *Language in mind: Advances in the study of language and thought*. Cambridge, MA, The MIT Press: 157-192.

Slobin, D. (2004). The many ways to search for a frog: Linguistic typology and the expression of motion events. In S. Stromqvist & Verhoeven, L. (Eds.), *Relating events in narrative: Vol. 2. Typological and contextual perspectives*. Mahwah: NJ, Lawrence Erlbaum Associates: 219-257.

Slobin, D. (2006). Relating Narrative Events in Translation. In D. Ravid & Shyldkrot, H. B. (Eds.), *Perspectives on language and language development: Essays in honor of Ruth A. Berman*. Dordrecht, Kluwer: 115-130.

So, W. C., Ching, T. H-W., Lim, P. E., Cheng, X. & Ip, K. Y. (2014). Producing gestures facilitates route learning. *PLoS ONE* 9(11): 2-21.

Sokolov, A. N. (1972). *Inner Speech and Thought*. New York, Plenum Press.

Spolsky, B. (1969). Reduced Redundancy as a Language Testing Tool. In G. E. Perren & Trim, J. L. M. (Eds.), *Applications of Linguistics*. New York, Cambridge University Press: 383-390.

Spolsky, B. (1973). What does it mean to know a language or how do you get someone to perform his competence? In J. W. E. Oller (Ed.), *Focus on the learner: pragmatic perspectives for the language teacher*. New York, Newbury House Publishers: 164-176.

Stone, A., Webb, R. & Mahootian, S. (1992). The generality of gesture-speech mismatch as an index of transitional knowledge: Evidence from a control-of-variables task. *Cognitive Development 6*: 301-313.

Sudharshana, N. P. (2015). Encoding 'support' and 'containment' relations in Kannada. *The EFL Journal 6(2)*: 15-30.

Sueyoshi, A. & Hardison, D. M. (2005). The role of gestures and facial cues in second language listening comprehension. *Language Learning 55*: 661-699.

Suvorov, R. S. (2008). *Context Visuals in L2 Listening Tests: The Effectiveness of Photographs and video vs. audio-only format*. Ames: IW, Iowa State University.

Svorou, S. (1994). The Grammar of Space. Amsterdam/ Philadelphia, John Benjamins Publishing Company.

Talmy, L. (1985). Lexicalization patterns: Semantic structure in lexical forms. In T. Shopen (Ed.) *Language Typology and Syntactic Description. Vol. III: Grammatical Categories and the Lexicon*. Cambridge, New York. Cambridge University Press: 57-149.

Talyer, J. (1999). Cognitive semantics and structural semantics. In P. Kochi & Blank, A. *Historical Semantics and Cognition*. Berlin, De Gruyter Mouton: 17-48.

Tanaka, S. & Matsumoto, Y. (1997). *Kuukan to idoo no hyogen* (*Expressions on space and movement*. Tokyo, Kenkyusha.

Tayler, W. L. (1953). Cloze procedure: a new tool for measuring readability. *Journalism Quarterly 30*: 415-433.

Tellier, M. (2008). The effect of gestures on second language memorisation by young children. *GESTURE 8*(2): 219-235.

Terazawa, T. (2013). *Gendai chuugoku go no hooi shi "Shan" to" Li" ni kansuru kenkyuu* (*A study of orientational marker "Shan" and "Li" in Modern Chinese language*. Doctoral Thesis. Nagoya, Nagoya University.

Tsangari, H. & Akritas, M. G. (2004). Nonparametric ANCOVA with two and three covariates. *Journal of Multivariate Analysis 88(2)*: 298-319.

Tulviste, P. (1991). *The cultural-historical development of verbal thinking*. Commack, N.Y., Nova Science Publishers.

Tyler, A. & Evans, V. (2003). *The Semantics of English Prepositions: Spatial Scenes, Embodied Meaning, and Cognition*. New York, Cambridge University Press.

Unwin, A. (2016). *Graphical Data Analysis*. Boca Raton, FL, Chapman and Hall/ CRC Press.

Ushakova, T. N. (1994). Inner speech and Second Language Acquisition: An Experimental - Theoretical Approach. In J. Lantolf & Appel, G. (Eds.), *Vygotskian Approaches to Second Language Research*. Norwood, NJ, Ablex Publishing Corporation: 135-156.

van der Veer, R. (1996). The Concept of Culture in Vygotsky's thinking. *Culture & Psychology 2*: 247-263.

van der Veer, R. & Valsiner, J. (1991). *Understanding Vygotsky: A Quest for Synthesis*. Oxford, Blackwell.

van Heuven, W., Schriefers, H., Dijkstra, T. & Hagoort, P. (2008). Language Conflict in the Bilingual Brain. *Cerebral Cortex 18(11)*: 2706-2716.

Verma, J. P. (2015). *Repeated Measures Design for Empirical Researchers*. Hoboken, New Jersey, John Wiley & Sons.

Vygotsky, L. S. (1978). *Mind in Society: The Development of Higher Psychological Processes*. Edited by Cole, M., John-Steiner, V., Scribner, S, & Souberman, E. Cambridge, MA, Harvard University Press.

Vygotsky, L. S. (1987). *The Collected Works of L.S. Vygotsky Volume 1: Problems of General Psychology*. Edited by R.W. Rieber & Carton, A. S. New York, Plenum Press.

Vygotsky, L. S. (1993). *The Collected Works of L.S. Vygotsky Volume 2: Fundamentals of Defectology*. Edited by R.W. Rieber & Carton, A. S. New York, Plenum Press.

Vygotsky, L. S. (1997a). *The Collected Works of L.S. Vygotsky Volume 3: Problems of the Theory and History of Psychology*. Edited by R.W. Rieber & Wollock, J. New York, Plenum Press.

Vygotsky, L. S. (1997b). *The Collected Works of L.S. Vygotsky Volume 4: The History of the Development of Higher Mental Function*. Edited by R. W. Rieber Ed. New York, Plenum Press.

Vygotsky, L. S. (1998). *The Collected Works of L.S. Vygotsky Volume 5: Child Psychology*. Edited by R.W. Rieber. New York, Plenum Press.

Vygotsky, L. S. (1999). *The Collected Works of L. S. Vygotsky Volume 6: Scientific Legacy*. Edited by R.W. Rieber Ed. New York, Plenum Press.

Vygotsky, L. S. (2001). Shikoo to Gengo (Thinking and Speech). Edited by Shibata, Y. Tokyo, Shin Dokusho Sha.

Wagner, D. A. (1998). Putting second language first: Language and literacy learning in Morocco. In *Literacy Development ill a Multilinona! Context: Crose-Culturat Perspectives*. Mahwah, NJ, Lawrence Erlbaum Associates: 169- 183.

Wang, F. F. (2002). *Mandarin Chinese Dictionary: Chinese-English*. Mineola, New York, Dover Publications.

Wang, Z. Y. (2006). A Study of Chinese "xia" and "-xia". *Journal of the Faculty of Humanities and Social Sciences 8*: 13-24.

Wannenmacher, W. & Seiler, T. B. (1983/ 2012). Historical Trends of Research in Concept Development and the Development of Word Meaning. In T. B. Seiler & Wannenmacher, W. (Eds.), *Concept Development and the Development of Word Meaning*. Berlin, New York, Springer-Verlag: 1-14.

Washburn, G. N. (1994). Working in the ZPD: Fossilized and Nonfossilized Nonnative Speakers. In J. Lantolf & Appel, G. (Eds.), *Vygotskian Approaches to Second Language Research.* Westport, CT, Ablex Publishing: 69-82.

Weltens, B. (1989). *The attrition of French as a foreign language*. The Netherlands: Foris, Dordrecht.

Weltens, B. & Grendel, M. (1993). Attrition of vocabulary knowledge. In R. Schreuder & Weltens, B. (Eds.), *The bilingual lexicon*. Amsterdam, John Benjamins Publications: 135-156.

Wertsch, J. (1985). *Vygotsky and the Social Formation of Mind*. Cambridge, MA, Harvard University Press.

Wertsch, J. & Stone, C. A. (1978). Microgenesis as a tool for developmental analysis. *Quarterly The Laboratory of Comparative Human Cognition 1*: 8-10.

Whorf, B. L. (1956). *Language, thought, and reality; selected writings of Benjamin Lee Whorf*. Edited by J. Carroll. Cambridge, MA, The MIT Press.

Willems, R. M. & Hagoort, P. (2007). Neural evidence for the interplay between language, gesture, and action: A review. *Brain and Language 101(3)*: 278-289.

Woods, A., Fletcher, P. & Hughes, A. (2008). *Statistics in language studies*. New York, Cambridge University Press.

Yamada, S. (1981). *Kinou go no imi no hikaku* (*Comparison of the meaning of the functional word.* In T. E. Kunihiro (Ed.), *Nichi-Eigo hikaku koza Imi to goi* (*Course of comparative analysis between Japanese and English: Meaning and Vocabulary*. Tokyo, Taishukan shoten: 53-100.

Yamada, S. (1994). Nichijogengo no ninchikaku moderu (Ninchikaku model in Cognitive linguistics. *Gengo 23(6)*: 104-109.

Yano, Y. & Long, M. (1994). The Effects of Simplified and Elaborated Texts on Foreign Language Reading Comprehension. *Language Learning 44(2)*: 189-219.

Yaroshevsky, M. G. (1999). Epilogue. In R. W. Rieber (Ed.), *The Collected Works of L. S. Vygotsky Volume 6 Scientific Legacy*. New York, Plenum Press: 245-266.

Yoshioka, K. (2008). Linguistic and Gestural Introduction of Ground Reference in Ll and L2 Narrative. In S. G. McCafferty & Stam, G. (Eds.), *Gesture: Second Language Acquisition and Classroom Research (Esl and Applied Linguistics).* New York, Routledge: 211-230.

Appendices

Appendix A

Acronyms in Linguistics

Acronyms Adapted from Iwasaki, 2013, p. XXI, Hanamo, 1998. & Martin, 1975

		Acronyms in Linguistics
1	ACC	Accusative particle.
2	ALL	Allative.
3	CLS	Numeral classifier.
4	COND	Conditional form.
5	COP	Copula.
6	GA	Marker commonly considered as nominative.
7	GEN	Genitive particle.
8	LOC	Locative particle.
9	NEG	Negative.
10	NONPAST	Non past.
11	NOM	Nominative.
12	NP	Noun phrase.
13	PAST	Past tense.
14	PP	Pragmatic particle.
15	Qes	Question marker.
16	Q	Stem final.
17	S	Subject.
18	TE	-*te* (conjunctive) form.
19	TOP	Topic marker.
20	V	Verb.
21	VP	Verb phrase.
22	Ø	Ellipsis.

Appendix B

Acronyms in General

		Acronyms in General
1	AIC	Akaike information criterion.
2	AICC	Akaike's information criterion corrected.
3	ANCOVA	Analysis of covariance.
4	BIC	Bayesian information criterion.
5	cf.	Refer to.
6	CIs	Confidence intervals.
7	e.g.	For instance.
8	EFL	English as a foreign language.
9	ES	Effect size.
10	η^2	Eta squared.
11	FFL	French as a foreign language.
12	HiP	Higher proficiency.
13	I	Intervention with viewing ICSCTG.
14	i.e.	That is.
15	I/G	Intervention with viewing and performing ICSCTG.
16	ICSCTG	Iconic co-speech co-thought gesture.
17	IH	Interface hypothesis.
18	IP	Intervention with viewing and treatment with performing ICSCTG.
19	IPH	Information packaging hypothesis.
20	IRB	Institutional Review Board.
21	JFL	Japanese as a Foreign Language.
22	L	Listening practices.
23	L1	First language.
24	L2	Second language.
25	LM	Landmark.
26	LoP	Lower proficiency.
27	M	Mean age.
28	MCC	Meidai Conversational Corpus.
29	NG	Participants will be asked not to perform ICSCTG.
30	PF	Performing of ICSCTG.
31	Q-Q plots	Quantile-quantile plots.
32	R	Reading aloud text assignment to participants.
33	RP	Participants will be asked to reproduce ICSCTGs in oral speech.
34	SLA	Second language acquisition.
35	SOV	Word order of subject-object-verb.

36	SVV	Word order of subject-verb-object.
37	T	Taking a test.
38	TFS	Thinking for speaking.
39	TR	Trajector.
40	VIV	Variance inflation factor.

Appendix C

Definitions of Major Key Terms

Bilingualism: There are different views of bilingualism. One view is “native-like control of two or more languages” (Bloomfield, 1933, p. 56). A second view is “The term bilingual [as] … used to refer to speakers who use two languages in their daily lives, be it simultaneously (in language contact situations) or consecutively (in the context of transnational migration), regardless of respective levels of proficiency in the two” (Pavlenko, 2006, p. 2). The present study adopts Pavlenko’s definition in this regard.

Classroom Intervention and experimental treatment: Some researchers use the term intervention and treatment interchangeably. This study has not found clear differences between intervention and treatment. However the current study differentiates both terms. Intervention refers to *an act of intervening*. Experimental treatment is defined as a treatment “administered over time to a single individual or a small number of individuals” (Creswell, 2013, p. 12) and “ensuring that the characteristics essential to the conceptualization of treatment are the ones, and the only ones, causing the effect” (Krathwohl & Smith, 2005, p. 146).

Developed meaning of words in EFL: Developed meaning of words refers to the meaning of words with nonobligatory contrast in Japanese which develops from the single to the binary semantic categorization by generalizing vertical spatial conceptualizations in a Japanese EFL class. (cf. Kunisawa under review for development of meaning of words in L1 and EFL).

***Discovery, Finding,* and *Result*:** I differentiate between these three terms based on the literature review in gesture studies.

(a) The term *discovery* is used for something new or novel findings such as "gestural discovery of novel ideas" (Kita et al., 2017, p. 253) and "Producing gestures can help learners to discover new conceptualizations of problems" (Kita et al., 2017, p. 260).

(b) The term *finding* is utilized for a finding through literature review. I have found that Kita et al. (2017) used term in the entire article except for two pages.

(c) The term *result* is defined as "to proceed or arise as a consequence, effect, or conclusion" (Merriam-Webster). https://www.merriam-webster.com/dictionary/result

Chu and Kita (2012) described that "Our results are compatible with the findings that gesture plays an active role in children's learning process [12], [13], [29]" (p. 63), and thus, I consider that the term result can be used for a result of an experiment in my study. Kita suggested, "I think your definitions sound right. Finding and result are very similar to each other and can be used interchangeably" (S. Kita, personal communication, September 4, 2017).

External validity: "*External validity* is the extent to which the results of an experiment can be generalized to people and environmental conditions outside (of) the context of the experiment" (McMillan & Schumacher, 2009, p. 265).

Foreign Language learning: "Foreign language learning is generally differentiated from second language acquisition in that the former refers to the learning of a non-native language in the environment of one's native language" (Gass, 2013, p. 4).

Higher concept: Vygotsky (1987) stated that a "higher concept presupposes both a hierarchical system and concepts subordinate and systematically related to the given concept" (p. 192). In this study, a higher concept refers to scientific and categorical thinking, used as synonymous terms

Internal validity: The internal validity of a study is a judgment that is made concerning the confidence with which plausible rival hypotheses can be ruled out (McMillan & Schumacher, 2009).

Lexicon: Lexicons are organized during vocabulary development (Li, Farkas, & MacWhinney, 2004). There is no distinct boundary between lexicon and grammar (J. Bybee, personal communication, November. 16, 2014).

Lexicon-grammar: There is an interconnected network between lexicon and grammar. The concept of lexicon-grammar in cognitive linguistics (Goldberg, 1995; Langacker, 2009; Lakoff, 1987) is similar to that of functional linguistics (Bybee, Perkins, & Pagliuca, 1994). In other words, it is a cognitive-functional approach. The cognitive-functional linguistics framework is widely accepted (Wilcox, 1999). See also Heyvaert (2003).

Linguistic relativity:

There are three different versions of linguistic relativity. Those are:

Version 1: Slobin (1996) suggests, “Here I am following a tradition in anthropological linguistics that has taken a less deterministic approach in the face of linguistic diversity, as exemplified by the thinking of Franz Boas” (p. 71). Slobin (1996) also stated, “In making this claim, I wish to present a new version of the von Humboldt-Wharf position on linguistic relativity and determinism” (p. 75). Herbert Clark (1996) consider that “Whorf argued for two proposals. One was linguistic relativity…the other…was linguistic determinism” (p. 324). Kita and Ozyurek (2003) stated, “the current results that language can shape non-linguistic spatial representation in thinking-for-speaking opens the door to the possibility of language shaping thinking-in-general under certain circumstances (p. 27). Levinson (2003) stated, “Whorfianism, understood as a limited linguistic determinism” (p.

315). Therefore, I assume that Whorf, Slobin, Levinson, and Kita support a weak version of linguistic relativity.

Version 2: Slobin (1996) also stated, “Again, we see that thinking for speaking is not a Whorfian straightjacket”(p. 86), which is a moderate version of linguistic relativity.

Version 3: Linguistic determinism that refers to the Sapia-Whorf hypothesis. The term *Sapir–Whorf hypothesis* is regarded as a misnomer by linguists for some reasons: Edward Sapir and Benjamin Lee Whorf have never co-authored any work. The distinction between a weak and a strong version of this hypothesis is a later creation (Gumperz & Levinson, 1996).

Preposition and Particle: Crystals (1992) defined a preposition and stated, “An item that typically precedes a noun phrase to form a single constituent of structure” (p. 312). This study adapts the interpretation of preposition described by Lakoff. Lakoff (1987) treats terms *on*, *over* and *above* as prepositions. Crystals (1992) noted that “The term is especially used for a form which does not readily fit into a standard classification of parts of speech” (p. 291). However, Tyler and Evans (2003) defined *on, over, above* and *under* as “an important subset of which are prepositions” (p. 246). Goldberg (2014) suggested that the use of term *preposition* for *on, over, above* and *under* is proper, but not a particle (A. Goldberg, personal communication, May 8, 2015). In British English, there are differences between the two. Refer to Huddleston and Pullum (2002) for further discussion.

Speech: Speech has a symbolic function (Vygotsky, 1987). “ … speech plays an important role in the organization of higher mental functions” (Vygotsky, 1999, p. 17). “Speech included in the operation was the system of psychological signs that acquired a very special functional significance and resulted in a complete reorganization of behavior” (Vygotsky, 1999, p. 27).

Second Language Acquisition: “SLA refers to the process of learning another language after the native language has been learned … regardless of whether it is the second, third, fourth, or fifth language” (Gass, 2013, p. 4).

Sign: Language is one of many psychological tools (Vygotsky,1997a). Gesture is a visual sign (Vygotsky, 1997b). Mimicry is gestured language (Vygotsky, 1993). Sound in human speech is a sign (Vygotsky, 1987). Word is a higher form of sign (Vygotsky, 1999). Vygotsky (1987) stated, “the symbolic function of the word represents thinking activity in the true sense of the word” (p. 94). Thus, “[…] the word represents thinking activity in the true sense of the word” (Vygotsky, 1987, p. 94). However, Vygotsky differentiates the sign from the tool and states (1997b) that:

> A more substantial difference of the sign from the tool and the basis of the real divergence of the two lines is the different purpose of the one and the other. The tool serves for conveying man’s activity to the object of his activity, it is directed outward, it must result in one change or another in the object, it is the means for man’s external activity directed toward subjugating nature. The sign changes nothing in the object of the psychological operation, it is a means of psychological action on behavior, one’s own or another’s, a means of internal activity directed toward mastering man himself; the sign is directed inward (p. 62).

In short, a way of use of a sign (a word) is crucial to generate a higher psychological operation, leading to an internal activity.

Single and Binary Semantic Categorization: A binary semantic categorization refers to an obligatory contrast when English speakers describe vertical spatial structures accompanied by contact and noncontact differentiation. Conversely, a single semantic categorization refers to a nonobligatory contrast when Japanese speakers express vertical axis operations without contact and noncontact differentiation (Munnich et al., 2001; Munnich & Landau, 2003).

Single word meaning: Vygotsky (1987) suggested that in languages, "several words or an entire phrase can carry the functional meaning of a single word" (p. 277). English structures a prepositional phrase[106] by using multiple different words (*on the table* and *above the table*). Japanese constructs a noun phrase by using multiple alternative words (*teeburu no ue ni* or *teeburu ni* / *on, over,* or *above the table*). Based on Vygotsky's claim noted above, it appears that both English and Japanese speakers express vertical space by using a single word meaning such as contact and/or noncontact. Therefore, a term *a single word meaning* is used when discussing existing experimental gesture studies in FL and this study. The terms *single word meaning, developed meaning of words in EFL*, *lexicon and grammar*, and *a prepositional phrase* are synonymous in this study.

[106] Macedonia (2000) published a study that described the use of gesture when students learned a prepositional phrase. However, it seems that she published it in German and thus the investigator has no access to this thesis. Besides, after completion of her study in 2000, her publications indicated that her research focus was learning vocabulary with iconic gestures, but not a prepositional phrase.

Appendix D

Brief Outline of ANCOVA Model Selection Processes

The purpose of this appendix is to explain multiple steps in which one should examine proper analysis of covariance (ANCOVA) models, prior to determining final ANCOVA models. The present study identifies better ANCOVA models with higher model accuracy and satisfying no serious multicollinearity/collinearity, and statistical assumptions of normality and of constant variance of residuals. These multiple steps can lead to testing Hypotheses 1, 2, and 3, and responding to Research Questions 1 and 2 correctly in the current study. The multiple steps are:

First of all, I selected four subsets of covariates by using information criterion in SPSS and STATA. These are Subsets A, B, C, and D explained below. Subsets A and B were used to identify two temporal ANCOVA models (models 1_1 and 1_2) for the post oral speech test.

Additionally, Subset C and D were employed to identify another two provisional ANCOVA models (models 2_1 and 2_2) for the delayed post-oral speech. Covariates[107] in each subset for these temporal ANCOVA models are as follow:

1. Subset A: Consisting of the four groups and the pre-oral speech test
2. Subset B: Consisting of the four groups, the pre-oral speech test, and the participants' motivation of learning English grammar
3. Subset C: Consisting of the four groups and the pre-oral speech test
4. Subset D: Consisting of the four groups, the pre-oral speech test, and the participants' motivation of learning English grammar.

[107] Groups are classified as both a covariate and a fixed factor.

The following table demonstrates the structures of possible final ANCOVA models 1 and 2.

Table 1:

		Final ANCOVA model 1 for post-oral speech		Final ANCOVA model 2 for delayed post oral speech	
1		Temporal ANCOVA model 1_1	Temporal ANCOVA model 1_2	Temporal ANCOVA model 2_1	Temporal ANCOVA model 2_2
2	Subset A	Pre-oral test/4 groups			
3	Subset B		Pre-oral test/mortivation of learning ENG grammar/4 groups		
4	Subset C			Pre-oral test/4 groups	
	Subset D				Pre-oral test/mortivation of learning ENG grammar/4 groups

Model accuracy, no serious multicollinearity/ collinearity, and satisfying statistical assumptions should be examined to select the final ANCOVA models.

For the reasons explained above, model accuracy of these four temporal ANCOVA models is examined first. Second, multicollinearity/ collinearity and the variance inflation factor (VIF) are assessed, and subsequently, tested for statistical assumptions of normality and of constant variance of residuals are explained.

- **Each Step in Determining Two Final ANCOVA Models**

Table 2 was created using STATA[108] and demonstrates that relative model accuracy of each temporal ANCOVA model, and thus, the temporal ANCOVA model 1_2 (adj. R-sq = 0.216; AIC = 26.65) was selected for the post-oral speech test due to the higher model accuracy.

---

108 STATA version 12 was used.

Similarly, the temporal ANCOVA model 2_2 (adj. R-sq = 0.227; AIC = 38.66) was chosen for the delayed post-oral speech test because of the higher model accuracy.

Table 2: Possible final ANCOVA models

			Final ANCOVA model 1				Final ANCOVA model 2	
Dependent variables	Post-test		Post-test		Delayed post test		Delayed post test	
	Temporal ANCOVA model. 1_1		Temporal ANCOVA model. 1_2		Temporal ANCOVA model. 2_1		Temporal ANCOVA model. 2_2	
	coefficient	t	coefficient	t	coefficient	t	coefficient	t
1.group	.	.	.	.	.	.	.	.
2.group	0.109	-1.6	0.154*	-2.24	0.057	0.03	0.0956	-1.33
3.group	-0.00397	(-0.06)	0.0269	-0.39	-0.8	-0.39	0.0683	-0.94
4.group	0.148*	-2.19	0.169*	-2.54	0.0937	0.07	0.307**	-4.36
pre-oral test	0.869**	-4.78	0.808**	-4.48	0.741**	-3.88	0.660**	-3.48
English grammar helps with future job.			-0.106**	(-2.64)			-0.0888*	(-2.11)
_cons	0.323**	-6.08	0.539**	-5.57	0.391**	-7	0.575**	-5.66
N	126		125		125		124	
adj. R-sq	0.182		**0.216**		0.212		**0.227**	
Log lik.	-10.49		-7.323		-16.46		-13.33	
AIC	30.99		**26.65**		42.92		**38.66**	
="* p<0.05	** p<0.01"							

Additional data analyses should be further performed because an unacceptable model selection can lead to improper significance tests, including the actual alpha level of the significance tests can be too high or too low, if large differences in the residual variances are obtained (Cohen et. al. 2003), and thus, additional data analyses are performed. Those are as follows:

The Pearson correlation coefficient in this study suggests that all independent variables in the present study have less than moderate correlation levels (r <.526). Similarly, the variance inflation factor (VIF) for all the models in the current study showed less than 1.1. Cohen et al. (2003) suggest that any VIF of 10 or more shows evidence of serious multicollinearity/collinearity. For the reasons explained above, proper outcomes can be obtained by conducting ANCOVA and ANOVA.

Cohen et al. (2003) noted, "Homoscedasticity and normality of residuals are required for inference in linear MR[109]" (p. 253). Similarly, they (2003) suggested, "When the assumption of constant variance of the residuals regardless of the value of X is met, this condition is termed *homoscedasticity*" (p. 120). *Homoscedasticity*, that refers to the variance of the residuals within each group, which can lead to an invalid model selection.

Likewise, they (2003) recommend the graphical examination of the distribution of the residuals and the use of 95% confidence intervals (CIs) to analyze residuals. Therefore, those methods listed above were employed to examine the temporarily selected ANCOVA models. Additionally, Loess[110]curve, that is "central methodology in nonparametric regression" (Cohen et al., 2003, p. 252), was used in assessing the detrended Q-Q plots in the each ANCOVA model, but 95% CIs are not.

The processes of the two final ANCOVA model selections are simultaneously presented by employing the graphical methods explained above, in addition to the values of adjusted *R* squared, and those of AIC.

---

[109] ANCOVA is the combination of regression and ANOVA, and thus, testing the statistical assumptions of normality and constant variance of residuals (homomoscedasticity) is crucial in this study MR is acronym for multiple regression.

[110] Loess is an acronym for *locally weighted* scatterplots smoother that is an nonparametric method that generates a smooth regression line or curve. Loess describes the trend of the X-*Y* relationship in a scatterplot (Cohen et al. 2003). A loess curve is another important exploratory graphic aid, which is a method for fitting a smooth curve between two variables by adding a smooth curve to a scatter plot to provide better perception of the pattern of dependence.

- **Q-Q Plots and Residual Plots**

Diagnostic Q-Q plots examining all four temporary ANCOVA models indicate that there was no fundamental deviation from the statistical assumptions of the linear regression models with 95% CIs. Three reasons account for this.

First, the residuals approximate straight lines in Q-Q plots in these four models. Second, in detrended Q-Q plots, residuals for these four models do not form any specific form such as U or W shape and major data points relatively cluster y = 0. Third, residual plots that examine the constant variance of residuals in these four models hold relatively rectangular distributions with 95% CIs.

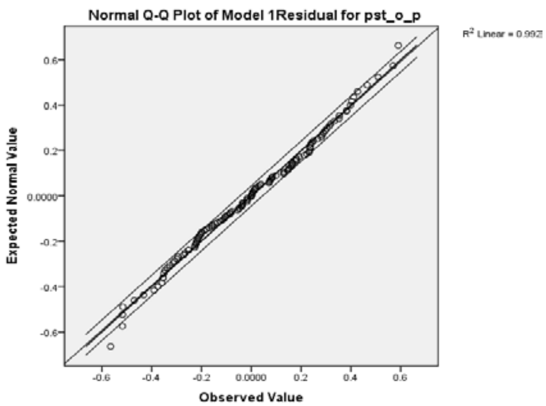
The statistical assumptions of normality and of constant variance of the residuals in these four ANCOVA models are met. Furthermore, the temporal ANCOVA models 1_2 and 2_2 have better values in adjusted *R* squared and AIC than others. Additionally, the three figures explained below demonstrate that the statistical assumptions of normality and of constant variance of residuals for two ANCOVA models 1_2 and 2_2 are met.

Concisely, the ANCOVA model 1_2 was selected as the final ANCOVA model 1 for the post-tests. The ANCOVA model 2_2 was also chosen as the final ANCOVA model 2 for the delayed post-tests. Thus, I argue that these final ANCOVA models can help one produce proper outcomes in conducting data analyses of this study with higher model accuracy, satisfying no serious multicollinearity/collinearity, the statistical assumptions of normality, and constant variance of residuals. This leads to testing Hypotheses 1, 2, and 3, and answering Research Questions 1 and 2 appropriately in the current study.

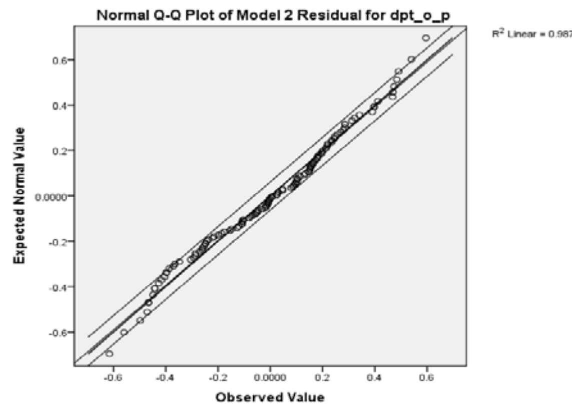
Plots to Evaluate ANCOVA Models 1 and 2

FIGURE 1

ANCOVA Model 1



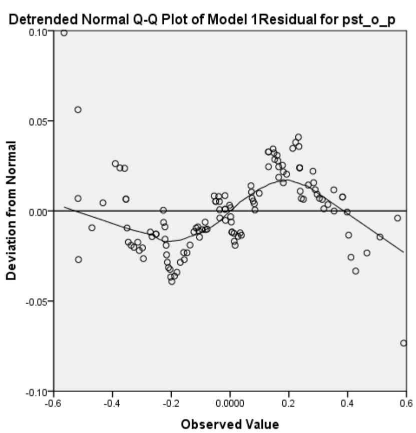
ANCOVA Model 2



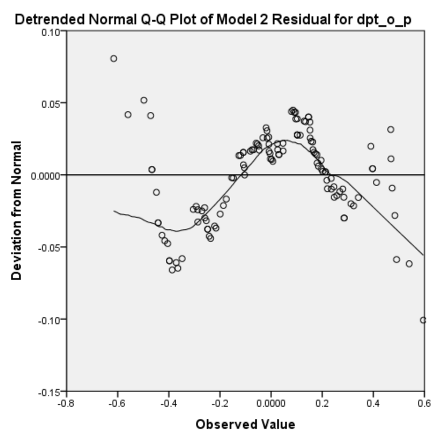
Q-Q PLOTS to evaluate the normality of the residuals in ANCOVA models

FIGURE 2

ANCOVA Model 1



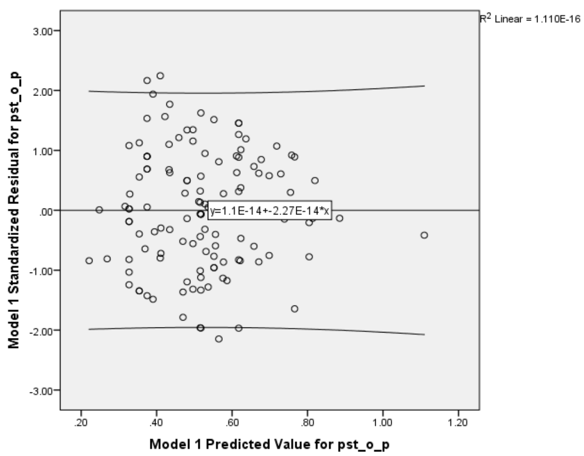
ANCOVA Model 2



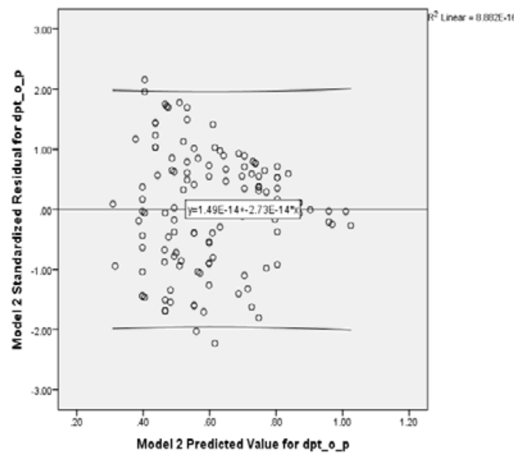
Detrended Q-Q plots to assess normality of the residuals in ANCOVA models

FIGURE 3

ANCOVA Model 1



ANCOVA Model 2



Plots to evaluate constant variance of residuals in the ANCOVA models

Appendix E

Variables

1. Dependent Variable

Students completed pre-, post-, and delayed post-tests. The post-test and the delayed post-test were used to evaluate learning at a lexical and grammatical level (developed meaning of words in EFL/ a single word meaning). Each student had a test score reflecting the appropriate use of developed meaning of words and word order, and thus this study has four dependent variables, which are:

1. Percentage of oral speech test score in a delayed post-test,
2. Percentage of oral speech test score in a post-test,
3. Percentage of cloze test score in a post-test, and
4. Percentage of cloze test score in a delayed post-test.

2. Independent Variables

I obtained 48 independent variables.

(a) Test scores:

(i) Participants' pre-test scores for the both ANCOVA models 1 and 2, and post-test scores for ANCOVA model 1 in oral speech and cloze tests.

(ii) Standardized test scores in EFL (grammar, reading, listening, and writing) and Japanese as L1.

(b) Survey variables: A survey conducted with 126 participants, which contained 114 questions: Their answers for 114 questions were categorized into 27 survey variables to analyze the data.

Survey Variables

	Name of variables
1	The use of audiovisual devices contributing to the improvement of their English abilities.
2	Participants' age of learning EFL for the first time.
3	Age. Participants' age at the time of data collection
4	The number of audiovisual devices which students use to study English.
5	The duration of their EFL learning.
6	Watching digital movies in English contributes to the improvement of their English abilities.
7	Whether English abilities of their family, relatives, and friends contribute to the improvement of their English abilities.
8	Whether students are applying for four year, two year college, vocational school, or applying for a job.
9	Whether talking with a friend helps with the improvement of their English abilities.
10	Learning English grammar helps them with getting a job offer in the near future.
11	Gender.
12	The ratio of submission of homework in their regular English class.
13	Using the Internet, which contributes to the improvement of their English abilities.
14	Using an iPad, which contributes to the improvement of their English abilities.
15	Using a smartphone, which contributes to the improvement of their English abilities.
16	Listening to music in English, which contributes to the improvement of their English abilities.
17	Help from people outside of school such as private tutors and teachers at other schools.
18	Learning English conversation helps them with their private life.
19	Learning English grammar helps them with their private life.
*20	Whether PowerPoint presentation was helpful in understanding class.
*21	Whether participants are willing to participate in future research.
22	Listening to the radio in English, which contributes to the improvement of their English abilities.
*23	Whether students repeated after the CD whenever the CD spoke to them.

24	Whether students learned for the first time about the characteristics of spatial terms (*on, over, above*, and under) in the investigator's class.
25	Going to a movie theater to watch movies in English, which contributes to the improvement of their English abilities.
*26	The length of study time in learning about vertical space.
27	Watching TV programs in English with subtitles, which contributes to the improvement of their English abilities?

Note: Survey variables marked by an asterisk indicate that the variables may be influenced by the classroom intervention. Thus, these variables were not used as independent variables.

Independent variables obtained from test scores:

1. Standardized test scores in their EFL
2. Standardized test scores in their L1
3. Pre-test scores in oral speech tests
4. Pre-test scores in cloze tests
5. Frequency of gesture generation in delayed post-test scores in oral speech tests
6. Frequency of gesture generation in post-test scores in oral speech tests
7. Frequency of gesture generation in pre-test scores in oral speech tests

Appendix F

Erroneous Expression

The purpose of this appendix is to address four potential problems, to explain probable reasons why the problems can occur, and includes an explanation of possible solutions when students learn about how to verbalize vertical axis relationships in a Japanese EFL class. The four major problems which Japanese EFL students will face in learning about vertical axes operations are: First, some of students will not make a contact-noncontact distinction that lacks proper developed meaning of words because of the confusion of an appropriate use of developed meaning of words.

Second, students will employ wrong SVO word order because Japanese has both SOV and OSV word order even though Japanese is considered to be “an SOV language” (Iwasaki, 2013, p. 11), but not SVO word order. A preposition includes classification as lexicon and grammar (Brinton & Traugott, 2005). The concept of TRHLM-TRLLM distinctions is inseparable in discussions of development of meaning of words inasmuch as improper word order can result in erroneous TRHLM-TRLLM relationships, hindering proper development of meaning of words. Examples:

(1) My eyeglasses are on the table.

(2) The table is on the eyeglasses.

Expressions (1) and (2) noted above have different meaning because of the totally distinct word order. Comrie (1989a) stated that English has a very high correlation between grammatical relations and word order, which suggests that word order is a basic grammatical relation. Thus, appropriate word order is fundamental for development of meaning of words in expressing proper meaning of vertical axis relationships (lexicon-grammar) in English for a particular type of EFL education, including Japanese, Korean, Kannada,[111] and others (Sudharshana, 2015). These languages are classified as SOV languages with the nonobligatory contrast when speakers of these languages express vertical spatial operations. The nonobligatory contrast refers to non-differentiation between contact and non-contact when speakers describe vertical axis structures.

Thus, one can argue that the improper TRHLM- TRLLM distinctions can lead to erroneous vertical spatial relationships of two referents, which results in incomplete development of meaning of words. For this reason, word order with lexicon and grammar should be simultaneously and explicitly taught in the EFL class to avoid the erroneous meaning of vertical spatial relationships of two referents.

Third, some students will express vertical axes relationships without a preposition when describing vertical axes relationships because Japanese lacks a preposition because of the postpositional structures in oral speech conditions in Japanese. Additionally, Japanese has ellipses (pragmatic inference) prevalent throughout face-to-face dialogue (Fujii & Ono, 2000; Nariyama, 2000, 2003; Ono & Thompson, 1997), whereas, generally speaking, English is not a zero anaphora (no ellipsis) language (Kibrik, 2001), even though English speakers occasionally use an ellipsis in a referential form (Appendix H).

Lastly, improper choice of a preposition can occur because Japanese does not have a preposition in oral speech, although Japanese has a preposition in a written form as in discussion in Appendix H. Furthermore, Greenberg (1990) stated “Universal 4: With overwhelmingly greater than chance frequency, languages with normal SOV order are postpositional” (p. 45). Thus, Japanese is postpositional, which tends not to develop a

[111] Kannada language is Dravidian language spoken primarily by people in South India.

prepositional structure. This postpositional nature of Japanese results in improper word choice of a preposition.

For the reasons explained above, it is important to note that instructors not only educate Japanese EFL students the linguistic differences between the two languages, but also create an effective and evidence-based teaching method, including the Gesture Listening Higher Concept Approach.

Appendix G

Generalization and Development of the Meaning of Words in L1

As the meaning of words develops, generalizations also can develop (Vygotsky 1997a). For this reason, development of the meaning of words and generalizations concurrently can occur in learning about vertical axis structure. Psychologically, generalization in L1 is elementary inasmuch as L1 speakers naturally generalize a vertical axis structure without efforts. Conversely, EFL students must consciously generalize a vertical spatial operation by reconceptualizing a vertical axis operation in talking about vertical relationships, which is a higher form of generalization.

Meaning of words in language can be a vehicle for the development of higher order mental function (van der Veer, 1997). Vygotsky (1987) suggested that learning a foreign language involves a higher mental function. Development of the meaning of words can lead to the generalization of the meaning of words in learning about vertical spatial relationships in an EFL class with the binary semantic categorization in a certain type of EFL education. Furthermore, as discussed in Appendix F (erroneous expressions), meaning relates to grammatical structures (head-initial SVO word order), and thus as the meaning of words develops, generalization of grammatical structures in English can also occur in learning about vertical axis operations. In other words, if an instructor employs an effective method, EFL students will then not only come to understand, but also express developed meaning of words using the higher form of generalization.

Therefore, I argue that development of the meaning of words will lead to a higher form of generalization, which can create development of verbal thinking for the reason that verbal thinking connects with the meaning of words as noted earlier (Vygotsky, 1987).

Appendix H

Major Linguistic Differences Between Japanese and English

The purpose of this appendix is to explain the markedly distinct typological dissimilarity in syntax and morphology, which leads to a different way in describing vertical spatial structures with contact-noncontact relationships between Japanese and English (see Appendix F). Robinson and Ellis (2008) state that the relative difficulty experienced by individuals who speak typologically distinct languages should be investigated in SLA research. The following section 1 includes discussion of major characteristics of Japanese when Japanese speakers express vertical spatial structures, followed by an explanation of key characteristics of English.

1. Major Characteristics in Japanese

SOV languages such as Japanese have a postpositional structure (cf. Comrie, 1989a; Greenberg, 1990), although "7% of SOV languages have preposition" (Hawkins, 1990, p. 119). SOV and agglutinative (synthetic) languages such as Japanese share a variety of features with Ural-Altaic languages[112] that is an absolute language-family uniting the Uralic and Altaic languages, including the use of postpositions rather than prepositions (Shibatani, 1990). In other words, Japanese lacks prepositions to encode vertical axes relationships. When talking about vertical spatial relationships, Japanese speakers use postpositions and two relational nouns (***ue*** and ***shita***) singularly, and the combination of both a relational noun and a postposition.[113]

---

[112] Japanese (Japonic) is one of the Altaic languages, including Turkic, Mongolic, Tungusic, Korean (Koreanic) (Janhunen, 2009).

[113] Other postposition (*de*) and grammaticized *no-tokoro* are used in vertical spatial operations in Japanese. However, I will not discuss a postposition *de* and *no-tokoro* in this thesis.

When expressing vertical axis relationships, crucial characteristics in Japanese are four-fold:

First, Japanese has a nonobligatory contrast between contact and noncontact when describing space in three dimensions (lateral, sagittal, and vertical orientations), even if the distinctions are made perceptually (the single semantic categorization = the nonobligatory contrast). Refer to Kunisawa under review for detail.

Second, linguistic typology[114] suggests that Japanese is a head-final language with a high degree of flexibility in an SOV word order (Kess & Miyamoto, 1994; Kondo & Yamashita, 2011; Kuno, 1973[115]), which influences an expression that has developed meaning of words when students learn about vertical axis operations in a certain types of EFL class. Additionally, even though Japanese has an SOV word order as a default, it also has an OSV word order (Saito, 1985).

Kuno (1973) showed examples about SOV and OSV with highly flexible head-final constructions in Japanese:

a. John	ga	Mary	o	Cambridge	de	mita
NOM	particle	NOM	particle	NOM	particle	in saw
<‘John saw Mary in Cambridge.’>						[SOV][116]

b. John ga Cambridge de Mary o mita. [SOV]
c. Mary o John ga Cambridge de mita. [OSV]
d. Mary o Cambridge de John ga mita. [OSV]
e. Cambridge de John ga Mary o mita. [SOV]
f. Cambridge de Mary o John ga mita. [OSV]

(Kuno, 1973, pp. 351-352).

---

[114] Studies of linguistic typology include discussion of a variety of domains of linguistics, including typology of morphology and syntax, general typology of reference-tracking systems, and others (Comrie, 1998a; Nariyama, 2003). The goal of this study is to focus on syntactical and morphological typology, but not others.
[115] Kuno (1973) stated, “Japanese is a left-branching language” (p. 4). “Left-branching” refers to head-final (Flynn, 1989).
[116] The author inserted these in each sentence.

Therefore, Japanese has a highly flexible head-final elliptic structure, which influences an appropriate expression of meaning of words in a Japanese EFL class.

Third, SOV languages such as Japanese have a postpositional structure (Comrie, 1989; Greenberg, 1990[117] ; Kuno, 1973), while “only 7% of SOV languages have prepositions” as stated earlier (Hawkins, 1990, p. 119). SOV languages such as Japanese share a variety of features with Ural-Altaic languages, including the use of postpositions rather than prepositions (Shibatani, 1990). However, Japanese has prepositions in written text and in a noun. Examples:

(1) *至名古屋*: *Toward (to)* Nagoya-city

(2) *於名古屋大学*: *At* the University of Nagoya (Tsunoda, 2009)

Nonetheless, when Japanese speakers verbally express space as in the examples below, the prepositions disappear and postpositions appear instead as in:

(1)-1: *Nagoya* ***e***
Nagoya ALL[118]
‘Toward (to) Nagoya-city’

(2)-1: *Nagoya-daigaku* ***de***
Nagoya-University LOC
‘At the University of Nagoya.’

It is proper to claim that Japanese can have a preposition in a written text and in a noun not used in expressing vertical space, but not in a verbal communication and a noun used in describing vertical space. The current study addresses a verbal communication only in

[117] Greenberg (1990) stated “*Universal* 4: With overwhelmingly greater than chance frequency, languages with normal SOV order are postpositional” (p. 45).
[118] ALL is an acronym for allative “to bring to” Iwasaki (2002, p. 324).

expressing vertical space. Therefore, those prepositions and nouns in Japanese do not influence discussions of the postpositional structure in Japanese.

When talking about vertical spatial relationships, Japanese speakers use postpositions and two relational nouns (***ue****: on, over, above*, and ***shita***: *under*) singularly, the combination of a relational noun and a postposition.[119] ***Ue*** and ***shita*** (relational nouns) include the common translation as a preposition in English. These relational nouns indicate a position relative in space or in time (Kaiser, Ichikawa, Kobayashi, & Yamamoto, 2013; Martin, 1975).

Fourth, when Japanese speakers talk about vertical space, they use ellipses, which are common linguistic phenomena in a face-to-face-dialogue (Fujii & Ono, 2000; Nariyama, 2000, 2003; Ono & Thompson, 1997). An ellipsis is “a cross-linguistic tendency for the unmarked choice of a trajector to correspond to the entity that would normally be construed as the figure on general perceptual grounds” (Langaker, 1987, p. 233). Ellipses in conversational pragmatics can show approximate meanings and a possible syntactic construction which speakers intend to express through vertical axis relationships. I propose that an ellipsis can have a psychological proximity with which Japanese speakers’ view on a specific referent to other referents. Therefore, I argue that an ellipsis psychologically links to semantics and syntax which leads to a proximity of unmarked linguistic structure.

[119] Japanese speakers use other postposition (*de*) and grammaticized *no-tokoro* in vertical spatial operations. However, I will not discuss the postposition *de* and the grammaticized *no-tokoro* in this thesis.

2. Major Features in English

English is typical of the SVO and analytic languages with head-initial word order (Biberauer, 2014; Hawkins & Filipović, 2012), which tends to develop prepositions (Bubenik, 2006), as opposed to synthetic language as in Japanese. Additionally, in general, English is a relatively less elliptic language, even though English speakers occasionally use an ellipsis in a referential form.[120] Analytic languages share common fundamentals, including "an abundance of grammatical words (auxiliaries and prepositions) and reliance on word order" (van Geleren, 2014, p. 282).

English is one of the Germanic languages, which is a branch of the Indo-European language family. The preposition ***above*** "traces its origins to Anglo-Saxon [Old English] *be* + *ufan*, which is related to the same Sanskrit root, 'upari', as ***over***" or higher (Tyler & Evens, 2003, p. 110). Sanskrit is a member of the Indo-Iranian subfamily of the Indo-European family of languages, and thus, Sanskrit has influenced English. Tyler and Evans (2003) also stated that:

> In Anglo-Saxon, the antecedent of *over* was *ufa*, the comparative form of *above*. In Gothic, *be + fan* was interpreted roughly as 'being in an up position'. Thus, *over* and *above* have long been closely associated semantically (p. 110).

Over and *above* are synonymous to some extent, but *on* is not.[121] English clearly distinguishes between contact and noncontact positioning with prepositional phrases (the binary semantic categorization = the obligatory contrast) with the SVO word order by using

[120] Corpus of contemporary American English (COCA) shows an ellipsis when English speakers express vertical axis relationships. For example, "Please, **be** **seated**" [2015 SPOK] that implies "Please be seated on the chair," "Please be seated on the sofa," or "Please be seated on the bench." (Downloaded on January 30, 2016).
[121] Tyler and Evans (2003) do not describe the etymology of a preposition *on.*

prepositions *on, over, above,* and *under*, unlike Japanese. Additionally, English speakers do not use morphemes and relational nouns as in Japanese in expressing vertical space.

Consequently, markedly distinct typological dissimilarities in syntax and morphology subsist in grammatical and semantic structures, leading to a different way in describing contact-noncontact relationships between Japanese and English, which can create problems when Japanese EFL students learn about vertical axis operations.

Appendix I

Cloze Test

空間についての文法 [122]
I. 名前は実名は記入せずに、与えられたアルファベットと数字の組み合わせ（ID）を使ってください。あなたの ID:___________________________
II. 最も適切な前置詞を下記の中から選択し、（　　　　　）の中に適語を入れなさい。
(under, over,　above, in, on) 新出単語は 2 ページ目に日本語訳があります。わからない単語があれば、手を上げてください。

1. The man wears the hat (　　　) his head.　2.He will nail a board (　　　) the holes in the ceiling.

3. The clouds float (　　　) the sun.　4. The dogs are (　　　) the tree.

5. The spider is walking (　　　) the ceiling.　6. She spread out the cloth (　　　) the massage table.

7. There are two boats (　　　) the Golden Gate Bridge.　8. The dog jumps (　　　) a barrel.

9. The airplane is now (　　　) Colorado.　10. The International Space Station floats 200 miles (　　　) the Earth's surface.

---

[122] This closed test was administered by using images corresponding to each expression. However, the images were removed in this thesis because of an issue related to a copyright of the images.

11. The show was in the beautiful area one floor (   ) the ground level.
12. He put the soccer ball (   ) his right foot.

13. The water falls (   ) the rocks.
14. A: “Do you want this hotdog?”
B: “No. I want the hotdog just (   ) it.”

15. The flag is flying (   ) the White House.
16. They put a transparent plastic sheet (   ) the painted ceiling of the chapel during repair.

17. The mold spreads all (   ) the ceiling.
18. There is the special pad (   ) the wireless mouse.

19. The soccer ball is (   ) the ground.
20. There are apples (   ) the tree.

単語

nail:固定する/ceiling:天井/ float:ただよう/spider:クモ/spread:広がる/barrel:ドラム缶/ surface:表面/chapel:礼拝堂/　remove:取り除く/wireless:無線の/ pad:あて物/mold:カビ/ repair: 修理

Appendix J

Speech Tests with Directions[123]

英会話テスト (English conversation test)

- グループA (Group A)

ジェスチャーを使用し、絵を見て、質問に答えてください。自分の顔より低い位置でジェスチャーをしてください。
(Please look at these pictures and answer questions, while generating gestures. Please gesture at a location lower than your face.)

テープに録音されたアメリカ人の男性による英語での説明と質問を聞いて、その質問に対し、英語で答えてください。答えるときは前置詞 on, over, above, under, in のいずれか一つを選択して、絵で表現されているものの位置関係を英語で答えてください。質問は20(18)問あります。答えがわからないときは、"I don't know"と言ってください。

(Please listen to explanations and questions in this auditory material that is recorded by the American male. Please select an appropriate spatial preposition among on, over, above, under, and in to express a relationship between two objects (or between a person and an object) described in these pictures. After you listen to the questions and select one of the terms, please answer the questions in English. There are twenty [eighteen] questions which I expect you to answer. If you do not know the answers, please tell me "I don't know.)

英会話テスト (English conversation test)

グループ B (Group B)

テープに録音されたアメリカ人の男性による英語での説明と質問を聞いて、その質問に対し、英語で答えてください。答えるときは前置詞 on, over, above, under, in のいずれか一つを選択して、絵で表現されているものの位置関係を英語で答えてください。質問は20(18)問あります。答えがわからないときは、"I don't know"と言ってください。

(Please listen to explanations and questions in this auditory material that is recorded by the American male. Please select an appropriate spatial preposition among on, over, above, under, and in to express a relationship between two objects (or between a person and an object) described in these pictures. After you listen to the questions and select one of the

[123] Images have been removed because of copyright issues.

terms, please answer the questions in English. There are twenty [eighteen] questions which I expect you to answer. If you do not know the answers, please tell me "I don't know.)

英会話テスト (English conversation test)

- グループ C (Group C)

テープに録音されたアメリカ人の男性による英語での説明と質問を聞いて、その質問に対し、英語で答えてください。答えるときは前置詞 on, over, above, under, in のいずれか一つを選択して、絵で表現されているものの位置関係を英語で答えてください。質問は20(18)問あります。答えがわからないときは、"I don't know"と言ってください。

(Please listen to explanations and questions in this auditory material that is recorded by the American male. Please select an appropriate spatial preposition among on, over, above, under, and in to express a relationship between two objects (or between a person and an object) described in these pictures. After you listen to the questions and select one of the terms, please answer the questions in English. There are twenty [eighteen] questions which I expect you to answer. If you do not know the answers, please tell me "I don't know.)

英会話テスト (English conversation test)

- グループ D (Group D)

ジェスチャーを使用し、絵を見て、質問に答えてください。自分の顔より低い位置でジェスチャーをしてください。

(Please look at these pictures and answer to questions, while generating gestures. Please gesture at the location lower than your face.)

テープに録音されたアメリカ人の男性による英語での説明と質問を聞いて、その質問に対し、英語で答えてください。答えるときは前置詞 on, over, above, under, in のいずれか一つを選択して、絵で表現されているものの位置関係を英語で答えてください。質問は20(18)問あります。答えがわからないときは、"I don't know"と言ってください。

(Please listen to explanations and questions in this auditory material that is recorded by the American male. Please select an appropriate spatial preposition among on, over, above, under, and in to express a relationship between two objects (or between a person and an object) described in these pictures. After you listen to the questions and select one of the terms, please answer the questions in English. There are twenty [eighteen] questions which I expect you to answer. If you do not know the answers, please tell me "I don't know.)

Directions for English conversations

① お茶とテーブルの位置関係を英語で言ってください。

Where is the tea relative to the table?

② 猫とテーブルの位置関係を英語で言ってください。

Where is the cat in relation to the Table?

③ 犬とソファーの位置関係を英語で言ってください。

Please describe the spatial relationship between the dogs and the sofa.

④ 鳥とテーブルの位置関係を英語で言ってください。

Please express the spatial relationship between the bird and the table.

⑤ 手形と天井の位置関係を英語で言ってください。

What is the spatial relationship between the handprints and the ceiling?

⑥火災探知機と天井の位置関係を英語で言ってください。

What is the spatial relationship between the ceiling and the smoke detector?

⑦ 雲と空の位置関係を英語で言ってください。

Please describe the spatial relationship between the clouds and the sky.

⑧ポトラックに招かれたゲストは何をしているのか、橋とゲストの位置関係を考えて英語で言ってください。

What are the guests who were invited to potluck doing? Please express the spatial relationship between the guests and the bridge.

⑨テーブルクロスとテーブルの位置関係を英語で言ってください。

Please state the spatial relationship between the tablecloth and the table.

⑩車の底とオイルの位置関係を英語で言ってください。

Please express the spatial relationship between the bottom of the car and the oil.

⑪あなたの友人はあなたがピザBがほしいと思っています。しかし、あなたはピザAがほしいです。ピザAとピザBの位置関係を英語で言ってください。

Your friend considers that you want pizza B. However, you want pizza A. Where is pizza A relative to pizza B?

⑫ナプキンとスプーンの位置関係を英語で言ってください。

Where is the spoon in relation to the napkin?

⑬オレンジと木の位置関係を英語で言ってください。

Describe the spatial relationship between the oranges and the tree.

⑭あなたの部屋（ルームA）とあなたの友人の部屋（ルームB）の位置関係を英語で言ってください。

Express the spatial relationship between your room (room A) and your friend room (room B).

⑮少年とフェンスの位置関係を英語で言ってください。

Where is the boy in relation to the fence?

⑯ソファとスキーをしている写真の位置関係を英語で言ってください。

Where is the picture relative to the sofa?

⑰雲と山の位置関係を英語で言ってください。

Where is the cloud in relation to the mountain?

⑱ヨガボールとあなた(黄色スポーツウェアーを着ている人はあなたと考えてください)の 位置関係を英語で言ってください。

What is the spatial relationship between you (the person wearing yellow sportswear) and the ball? Please use the term "under."

⑲あなたは窓から湖の上を見ています。オーロラと湖の位置関係を英語で言ってください。

You are looking at the Aurora from a window. Please express the position of the Aurora in relation to the lake.

⑳雲と満月の位置関係を英語で言ってください。

Where is the moon in relation to the cloud?

Appendix K

Oral Speech Test

Potluck	**Score**
IT: You and your friend are having a potluck at your and your friend's house this evening.	
IT: Question 1: Where is the tea?	
Answer 1:	
IT: Thank you.	
IT: I do not see my cat, bird and dogs.	
Q2: Where is the cat?	
Answer 2:	
Q3: Do you know where my dogs are?	
Answer 3:	
Q4: How about the bird?	
Answer 4:	
IT: Thank you. I will take them to my room.	
Q5: The ceiling in this room is not good. Are those handprints?	
Answer 5:	
IT: We will clean them later.	
Q6: Is there a smoke detector at this house?	
Answer 6:	
IT: It's good.	
Q7: There are a lot of clouds in the sky. Do you see the sun?	

Answer 7:	
IT: I hope it will not rain.	
Q8: Do you know if our guests will come to our potluck on time?	
Answer 8:	
Q9: That's good. Where is my pink tablecloth?	
Answer 9:	
(One of the guests arrives at the potluck. But she has a problem with her car.)	
Q10: Can you take a look at my car?	
Answer 10:	
IT: I will take it to my mechanic. Thank you, though.	
(The potluck finally starts.)	
Q11: Do you want this pizza?	
Answer 11:	
IT: Okay. No problem.	
Q12: Where is a napkin?	
Answer 12:	
IT: Thank you.	
Q 13: I want an orange. Where can I get it?	
Answer 13:	
IT: That's good. I can get it there.	
Q14: Where is your room?	
Answer 14:	
IT: Look at that boy!	

Q15: What is he doing?	
Answer 15:	
Q16: You have a snow skiing picture. Where is it?	
Answer 16:	
Q17: I have a picture of a cloud. The cloud is on the mountain, isn’t it?	
Answer 17:	
IT: It is the interesting picture.	
Q18: What are you doing with the Yoga ball?	
Answer 18:	
(The potluck is over. It is now 11 pm.)	
Q19: What are you doing?	
Answer 19:	
IT: Wow, it is so lovely. But the Aurora will go away soon.	
That’s okay. We will see the full moon.	
Q20: Do you see the moon?	
Answer 20:	
IT: It is time to go to bed. Have a good night.	

End of interview.

Appendix L

Survey in English

Thinking-for Speaking and the Bilingual Mind: Face-to-Face Dialogue to Talk about Space.

Questionnaire for EFL students at Kochi Higashi High School, Kochi, Japan

Fall 2012

Cover letter

Dear Students,

I would like to invite you to participate in this questionnaire, which is part of the requirements for my dissertation studies in the USA. The goal of the questionnaire is to collect information about your experience with learning English as a foreign language. It may take you up to 45 minutes to complete it.

You will not be asked to provide your name, yet we do need your Student ID in order to link your responses to language proficiency scores. Importantly, no information about your identify will be disclosed. Responses to the questionnaire will be confidential and only group results will be reported. If you have any concern or question about the questionnaire, please contact me.

Tae Kunisawa
Educational Linguistics program at the University of New Mexico
MSC03 2130, Department of Linguistics
Humanities 526
University of New Mexico
Albuquerque, NM 87131, United States of America
US phone #: 505-277-6353

email: tkuni@unm.edu (Phone in Japan: 088-832-2527)

Student ID:________________________________ Date: / / 2012 .

Please select one answer to the following questions and mark your answer with (√), unless you are requested to select multiple answers.

Example 1: You are √□①high school student □②college student

Note:

1. English grammar in this questionnaire pertaining to that which you learned in Ms. Kunisawa's class, including *on, over, above,* and *under*.
2. English conversation in this questionnaire pertaining to that which you were tested in Ms. Kunisawa's class.

I. Please let me know about you.

1. What is your gender?

□①Female

□②male

2. What is your age?

□①Under 15 yrs

□②16 yrs

□③17 yrs

□④18 yrs

□⑤over 19 yrs

3. What is your citizenship?

□① Japan

□②Korea

□③ China

□④ Other ( )

4. Please select your first language.

□①Japanese

□②Chinese

□③Korean

□④other (Specify it:)

5. Please specify your foreign language abilities.

□①Bilingual (Specify languages which you have learned about four skills) :

□②Multilingual (Specify languages which you have learned about four skills) :

6. In what country were you born?

□①Japan

□②Unknown

□③A country other than Japan

7. How many years have you studied English as a foreign language?

□①0-1

□②2-3

□③4-5

□④-6

□⑤more than 7

8. When did you start learning English?

□①Before middle school (describe the age:)

□②At middle school

□③At high school

9. Have you ever been to English speaking countries?

□①Yes □②No ⇨ **PROCEED TO QUESTION 10**

⇩

1) If yes, please specify the countries.______________________________

2) **If yes,** please specify length of your stay.

□①Less than 1 month

□②2-3 months

□③4-11 months

□④More than 12 months

3) **If yes**, how old were you when you were in the country(ies)?

□①0-3

□②4-12

□③13-15

□④16-18

10. Is there a TV which you can watch TV programs at your home?

□①Yes □②No ⇨ **PROCEED TO QUESTION 11**

⇩

1) If yes, do you watch TV programs in English?

□①Yes □②No ⇨ **PROCEED TO QUESTION 11**

2) If yes, do you view TV programs with subtitles?

□①Yes □②No ⇨ **PROCEED TO QUESTION 11**

⇩

3) If yes, watching TV programs in English with subtitles has contributed to improving your English abilities.

□①Strongly Agree

□②Agree

□③Disagree

□④Strongly disagree

11. Do you have devices which you can access the Internet outside your high school?

□①Yes □②No **PROCEED TO QUESTION 12**

1) **<u>If yes,</u>** do you view sites in English?

□ ①Yes □②No **PROCEED TO QUESTION12**

<u>2) If yes,</u> how often do you access the Internet?

□①Every day

□②Several times a week

<u>3) If yes,</u> how many hours do you use English language sites a day?

□①0-1

□②2-3

□③4+

4) **<u>If yes,</u>** accessing the English language sites has contributed to improving your English speaking abilities.

□①Strongly Agree

□② Agree

□③ Disagree

□④ Strongly disagree

12. Do you listen to music in English?

□①Yes □No② **PROCEED TO QUESTION 13**

<u>1) If yes</u>, listening to the music contributed to improving your English.
□①Strongly Agree

□②Agree

□③Disagree

□④Strongly disagree

13. Do you watch digital movies in English with DVD?

□①Yes □②No **PROCEED TO QUESTION 14**

<u>1) If yes</u>, watching the movies has contributed to improving your English.
□①Strongly Agree

□②Agree

□③Disagree

□ ④Strongly disagree

<u>2) If yes</u>, do the movies have subtitles in Japanese?

□①Yes □②No

14. Do you go to a movie theater to watch movies in English?

□①Yes □②No **PROCEED TO QUESTION 15**

<u>1)</u> **<u>If yes</u>**, do the movies have subtitles in Japanese?
□①Yes □②No

<u>2)</u> **<u>If yes,</u>** watching the movies has contributed to improving your English.
□①Strongly Agree

□②Agree

□③Disagree

□④Strongly disagree

15. Do you listen to radio in English?

□①Yes □②No ⇒ **PROCEED TO QUESTION 16**

<u>1) If yes,</u> listening to the radio in English has contributed to improving your English.

□①Strongly Agree

□②Agree

□③Disagree

□④Strongly disagree

16. Do you use an iPad?

□①Yes □②No ⇒ **PROCEED TO QUESTION 17**

<u>1) If yes</u>, how often do you use it?

□①Every day

□②Several times a week

□③Several times a month

□④Several times a year

<u>2) If yes</u>, using an iPad has contributed to improving your English.

□①Strongly Agree

□②Agree

□③Disagree

□④Strongly disagree

17. Do you use an iPhone?

□①Yes □②No **PROCEED TO QUESTION 18**

<u>1) If yes</u>, how often do you use it?

□①Every day

□②Several times a week

□③Several times a month

□④Several times a year

<u>2) If yes</u>, using iPhone has contributed to improving your English.

□①Strongly Agree

□②Agree

□③Disagree

□④Strongly disagree

18. Do you use email in English?

□①Yes □②No **PROCEED TO QUESTION 19**

<u>1) If yes</u>, how often do you use the email?

□①Every day

□②Several times a week

□③Several times a month

□④Several times a year

<u>2) If yes</u>, using email in English has contributed to improving your English.

□①Strongly Agree

□②Agree

□③Disagree

□④Strongly disagree

19. Do you talk with people over the phone in English?

□①Yes □②No **PROCEED TO QUESTION 20**

<u>1) If yes,</u> how often do you speak in English over the phone?

□①Every day

□②Several times a week

□③Several times a month

<u>2) If yes</u>, talking in English over the phone has contributed to improving your English.

□①Strongly Agree

□②Agree

□③Disagree

□④Strongly disagree

20. Do you have a chance to talk with a foreigner(s) in English?

□①Yes □②No **PROCEED TO QUESTION 21**

<u>1) If yes,</u> how often do you talk with him/her?

□①Every day

□②Several times a week

□③Several times a month

<u>2) If yes,</u> talking with them has contributed to improve your English conversation skill.

□①Strongly Agree

□②Agree

□③Disagree

□④Strongly disagree

21. Do you have a tutor who teaches English to you?

□①Yes □②No **PROCEED TO QUESTION 22**

1) **<u>If yes,</u>** how often do you learn English from him/her ?

□①Every day

□②Several times a week

□③Several times a month

2) **If yes**, did you learn *on*, *above*, *over*, and *under* from him/her?

□①Yes □②No **PROCEED TO QUESTION 22**

3) **<u>If yes,</u>** the tutor was helpful in understanding differences among *on, above, over,* and *under*.

□①Strongly Agree

□②Agree

□③Disagree

□④Strongly disagree

4) **<u>If yes,</u>** the tutor was helpful in taking grammar (cloze) tests in Ms. Kunisawa's research.

□①Strongly Agree

□②Agree

□③Disagree

□④Strongly disagree

5) **<u>If yes,</u>** the tutor was helpful in taking English conversation tests in Ms. Kunisawa's research.

□①Strongly Agree

□②Agree

□③Disagree

□④Strongly disagree

22. Do you go to a private tutoring school to learn English?

□①Yes □②No ⇨ **PROCEED TO QUESTION 23**

⇩

<u>1) If yes,</u> how often do you study English at the school?

□①Every day

□②Several times a week

□③Several times a month

2)If yes, did you learn *on*, *above*, *over*, and *under* at the school?

□①Yes □②No ⇨ **PROCEED TO QUESTION 23**

3) If yes, learning English at the school was helpful in understanding differences among *on, above, over,* and *under.*

□①Strongly Agree

□②Agree

□③Disagree

□④Strongly disagree

4) If yes, learning English at the school was helpful in taking the grammar (cloze) tests in Ms. Kunisawa's research.

□①Strongly Agree

□②Agree

□③Disagree

□④Strongly disagree

5) If yes, learning English at the school was helpful in taking the English conversation tests in Ms. Kunisawa's research.

□①Strongly Agree

□②Agree

□③Disagree

□④Strongly disagree

23. Do you have homework in your English classes?

□①Yes □②No ⇨ **PROCEED TO QUESTION 24**

1) **If yes,** how many times per one week do you have it?
①1

②2

③3

④4

⑤5

2) **<u>If yes,</u>** do you do homework?

□①Yes □②No **PROCEED TO QUESTION 24**

<u>a)</u> **<u>If yes</u>,** what % of the homework do you submit your teacher?

①100%

②80%

③50%

④30%

⑤Less than 29%

24. Do you use visual and auditory devices to study English outside of your high school?

□①Yes □②No **PROCEED TO QUESTION 25**

1) **If yes,** what device(s) do you use to study it? Multiple selections are allowed.

□①CD player

□②Computer without the Internet

□③Software

□④Internet

□⑤ipad

□⑥iPhone

□⑦TV

□⑧Radio

□⑨Electric dictionary

□⑩Tape recorder

□⑪IC recorder

□⑫Others (Specify them:)

<u>2) If yes,</u> how often do you use it (them)?

□①Every day

□②Several times a week

□③Several times a month

<u>3) If yes,</u> using it (them) has contributed to improving your English.

□①Strongly Agree

□②Agree

□③Disagree

□④Strongly disagree

25. Are you are going to apply to college?

□①Yes □②No ⇒ **PROCEED TO QUESTION 26**

1) **<u>If yes,</u>** learning English grammar will be helpful in being an attractive candidate for acceptance by the college.

□①Strongly Agree

□②Agree

□③Disagree

□④Strongly disagree

2) **<u>If yes,</u>** learning English conversation will be helpful in being an attractive candidate for acceptance by the college.

□①Strongly Agree

□②Agree

□③Disagree

□④Strongly disagree

26. Are you are going to apply to junior college?

□①Yes □②No ⇨ **PROCEED TO QUESTION 27**

⇩

1) **<u>If yes</u>,** learning English grammar will be helpful in being an attractive candidate for acceptance by the college.

□①Strongly Agree

□②Agree

□③Disagree

□④Strongly disagree

2) **<u>If yes</u>**, learning English conversation will be helpful in being attractive candidate for acceptance by the college.

□①Strongly Agree

□②Agree

□③Disagree

□④Strongly disagree

27. Are you applying to vocational school?

□①Yes □②No ⇨ **PROCEED TO QUESTION 28**

⇩

1) **<u>If yes,</u>** learning English grammar will be helpful in being an attractive candidate for acceptance by the school.
   - □①Strongly Agree
   - □②Agree
   - □③Disagree
   - □④Strongly disagree

2) **<u>If yes</u>**, learning English conversation will be helpful in being an attractive candidate for acceptance by the school.
   - □①Strongly Agree
   - □②Agree
   - □③Disagree
   - □④Strongly disagree

28. Are you applying for a job after graduating from high school, including a public servant?

□①Yes □②No ⇒ **PROCEED TO QUESTION 29**

1) **<u>If yes</u>**, learning English grammar will be helpful in being an attractive candidate for getting a job offer.
   - □①Strongly Agree
   - □②Agree
   - □③ Disagree
   - □④ Strongly disagree

2) **<u>If yes</u>**, learning English conversation will be helpful in being an attractive candidate for getting a job offer.
   - □①Strongly Agree
   - □②Agree
   - □③Disagree
   - □④Strongly disagree

29. Your future occupation and learning English.

1) Learning English grammar is helpful in getting a better job offer.

□①Strongly Agree

□②Agree

□③Disagree

□④Strongly disagree

2) Learning English conversation is helpful in getting a better job offer.

□①Strongly Agree

□②Agree

□③Disagree

□④Strongly disagree

30. Your private life and English

1) Learning English grammar is helpful when you communicate with English speakers, including when you go overseas.

□①Strongly Agree

□②Agree

□③Disagree

□④Strongly disagree

2) Learning English conversation is helpful when you communicate with English speakers, including when you go overseas.

□①Strongly Agree

□②Agree

□③Disagree

□④Strongly disagree

31. Please select areas of improving skills about English that you are interested in. (Multiple selections are allowed.)

□①English conversation

□②English grammar

□③Writing in English

□④Reading in English

□⑤Listening to English

II. Please let me know about your opinions on my research.

1. PowerPoint presentation was helpful in understanding Ms.Kunisawa's class.

□①Strongly Agree

□②Agree

□③Disagree

□④Strongly disagree

2. Did you learn **gestures** in Ms. Kunisawa's class?

□①Yes □②No ⇨ **PROCEED TO QUESTION 3**

⇩

1) If yes, learning gestures was helpful in understanding differences between *on* and *over*.

□①Strongly Agree
□②Agree
□③Disagree
□④Strongly disagree

2) If yes, learning gestures was helpful in understanding differences between *on* and *above.*

□①Strongly Agree
□②Agree
□③Disagree
□④Strongly disagree

<u>3) If yes,</u> learning gestures was helpful in understanding differences between *above* and *over*.

□①Strongly Agree

□②Agree

□③Disagree

□④Strongly disagree

<u>4) If yes,</u> learning gestures was helpful in grammar (cloze) tests**.**

□①Strongly Agree

□②Agree

□③Disagree

□④Strongly disagree

<u>5) If yes,</u> learning gestures was helpful in English conversation tests**.**

□①Strongly Agree

□②Agree

□③Disagree

□④Strongly disagree

3. Have you received a CD for Ms. Kunisawa's class before her class started? (If you received it after the class, please select "no").

□①Yes □②No ⇨ **PROCEED TO QUESTION 4**

⇩

1) Did you listen to the CD by the time of completion of a delayed post-test?

□①Yes □②No ⇨ **PROCEED TO QUESTION 4**

⇩

<u>a) If yes,</u> listening to the CD was helpful in understanding the differences and commonalities between *on* and *over*.

□①Strongly Agree
□②Agree
□③Disagree
□④Strongly disagree

<u>b) If yes,</u> listening to the CD was helpful in understanding the differences and commonalities between *on* and *above.*

□①Strongly Agree
□②Agree
□③Disagree
□④Strongly disagree

<u>c) If yes,</u> listening to the CD was helpful in understanding the differences and commonalities between *above* and *over*.

□①Strongly Agree
□②Agree
□③Disagree
□④Strongly disagree

<u>d) If yes,</u> listening to the CD was helpful in taking grammar (cloze) tests**.**

□①Strongly Agree

□②Agree

□③Disagree

□④Strongly disagree

<u>e) If yes,</u> you repeated after the CD whenever the CD spoke to you.

①Yes □②No

e)-**<u>1. Only a person who selects NO on letter e) described above:</u>**

You did not repeat after the CD, but in your mind you repeated what the CD spoke to you.

①Yes □②No

<u>f) ①If yes</u>, did you go to a computer lab to listen to the CD at Kochi Higashi High School?

□①Yes □②No ⇨ **PROCEED TO QUESTION 4**

⇩

<u>f)-1. If yes</u>, how many times did you go to the computer lab to listen to the CD until the day when you completed the delayed post conversation tests?

□①1

□②2

□③3

□④4

□⑤5

□⑥6

□⑦7

□⑧8

□⑨9

<u>f)-2. If yes in question f): Listening to the CD in the computer lab</u> was helpful in taking grammar (cloze) tests**.**

□①Strongly Agree

□②Agree

□③Disagree

□④Strongly disagree

f)-3. If yes in question f): Listening to the CD in the computer lab was helpful in taking conversation tests**.**

□①Strongly Agree

□②Agree

□③Disagree

□④Strongly disagree

2) if yes : Did you listen to the CD outside Kochi Higashi High School?

□①Yes □②No ⇨ **PROCEED TO QUESTION 4**

⇩

a)If yes, how long did you listen to the CD before you completed the all the delayed post-test?

□①5 min

□②10 min

□③15 min

□④30min

□⑤45min

□⑥60 min

□⑦120min

□⑧180min

□⑨240min

If you studies more than 241 min, please let me know how long you studied (__________ hours and __________ minutes)

4. You did not receive the CD before Ms. Kunisawa's class. In addition, gestures were not taught in her classes.

□①Correct □②Incorrect ⇨ **PROCEED TO QUESTION 5**

⇩

<u>1) If correct, you understood</u> the differences and commonalities between *on* and *over*.

□①Strongly Agree

□②Agree

□③Disagree **PROCEED TO QUESTION 4-2)**

□④Strongly disagree **PROCEED TO QUESTION 4-2)**

<u>a)if agree or strongly agree</u>, understanding the differences and commonalities was helpful in taking grammar (cloze) tests.

□①Strongly Agree

□②Agree

□③Disagree

□④Strongly disagree

<u>b) if agree or strongly agree</u>, understanding the differences and commonalities was helpful in taking conversation tests.

□①Strongly Agree

□②Agree

□③Disagree

□④Strongly disagree

<u>2) If correct, you understood</u> the differences and commonalities between *on* and *above.* This was helpful in taking grammar (cloze) tests.

□①Strongly Agree

□②Agree

□③Disagree ⇨ **PROCEED TO QUESTION 4-3**

□④Strongly disagree ⇨ **PROCEED TO QUESTION 4-3**

a) If agree or strongly agree, understanding the differences and commonalities between *on* and *above* was helpful in taking grammar (cloze) tests.

□①Strongly Agree

□②Agree

□③Disagree

□④Strongly disagree

b) If agree or strongly agree, understanding the differences and commonalities between *on* and *over* was helpful in taking English conversation tests.

□①Strongly Agree

□②Agree

□③Disagree

□④Strongly disagree

3) If correct, you understood the differences and commonalities between *above* and *over*.

□①Strongly Agree

□②Agree

□③Disagree ⇨ **PROCEED TO QUESTION 5**

□④Strongly disagree

a) If agree or strongly agree, understanding the differences and commonalities was helpful in taking grammar (cloze) tests.

□①Strongly Agree

□②Agree

□③Disagree

□④Strongly disagree

b) If agree or strongly agree, understanding the differences and commonalities was helpful in taking English conversation tests.

□①Strongly Agree

□②Agree

□③Disagree

□④Strongly disagree

5. Please select appropriate answers about pre-tests that you took before Ms. Kunisawa's class.

1) The grammar (cloze) test was not difficult for you.

□①Strongly Agree

□②Agree

□③Disagree

□④Strongly disagree

2) The English conversation test was not difficult for you.

□①Strongly Agree

□②Agree

□③Disagree

□④Strongly disagree

6. Please select appropriate answers about the post-tests that you took after Ms. Kunisawa's class

1) The grammar (cloze) test was not difficult for you.

□①Strongly Agree

□②Agree

□③Disagree

□④Strongly disagree

2) The English conversation test was not a difficult for you.

□①Strongly Agree

□②Agree

□③Disagree

□④Strongly disagree

7. Please select appropriate answers about the delayed post-tests which you took approximately three to five weeks later from the Ms. Kunisawa's first class.

1) The grammar (cloze) test was not difficult for you.

□①Strongly Agree

□②Agree

□③Disagree

□④Strongly disagree

2) The English conversation test was not difficult for you.

□①Strongly Agree

□②Agree

□③Disagree

□④Strongly disagree

8. Did you study *on*, *over*, *above*, and *under* after Ms. Kunisawa's classes? (Please exclude the time when you studied with CD in a computer lab at your high school and outside your school.)

□①Yes □②No ⇨ **PROCEED TO QUESTION 9**

1) If yes,

a) Select how many minutes in total you studied those without the CD listening between Ms. Kunisawa's first class and the delayed post conversation test, which you took.

□①5 min

□②15 min

□③30 min

□④45min

□⑤60min

□⑥120 min

□⑦ If you studied over 121 min, how long did you study?
(   ).

b) If yes, with whom did you study these?

□①by yourself

□②with your friend (s)

□③with your tutor(s)

□④with your private tutoring school teacher(s)

□⑤others (Specify them：   )

9. Spatial terms which have characteristics of contact and no contact

1) You learned for the first time that a term *ue* does not differentiate between contact and no contact in Ms. Kunisawa's class.

□①Strongly Agree

□②Agree

□③Disagree

□④Strongly disagree

2) Spatial terms *on, over,* and *above:*

a) You learned for the first time that the terms *over* and *above* are used to describe both contact and no contact in Ms. Kunisawa's class.

□①Yes □②No

b) You understand how to use the term *over* in describing both contact and no contact.

□①Strongly Agree

□②Agree

□③Disagree

□④Strongly disagree

c) You understand how to use the term *above* in describing both contact and no contact.

□①Strongly Agree

□②Agree

□③Disagree

□④Strongly disagree

d) You understand how to use the term *on* in describing contact only.

□①Strongly Agree

□②Agree

□③Disagree

□④Strongly disagree

10. Your comprehension abilities in English conversation of delayed post-tests.

1) You understood a majority of questions in these tests described above.

□①Strongly Agree

□②Agree

□③Disagree

□④Strongly disagree

2) You understood questions in these tests noted above, but not all of them.

□①Strongly Agree

□②Agree

□③Disagree

□④Strongly disagree

3) You were unable to understand most of the questions in these tests described above.

□①Strongly Agree

□②Agree

□③Disagree

□④Strongly disagree

11. You wish to participate again in future research, if you would be given an opportunity.

□①Strongly Agree

□②Agree

□③Disagree

□④Strongly disagree

III. Please let me know about your family, relatives and friends.

1. What languages do people whom you live with speak? Multiple selections are allowed.

□①Japanese

□②English

□③Spanish

□④Chinese

□⑤Korean

□⑥Others (Specify them:)

2. Do your relatives speak in English?

□①Yes □②No ⇨**PROCEED TO QUESTION 3**

⇩

<u>1) If yes,</u> talking with them in English has been helpful in improving your English.

□①Strongly Agree

□②Agree

□③Disagree

□④Strongly disagree

3. Do you have a friend(s) who can speak English fluently?

□①Yes □②No ⇨ **PROCEED TO QUESTION 4**

⇩

<u>1) If yes,</u> talking with them in English has been helpful in improving your English.

□①Strongly Agree

□②Agree

□③Disagree

□④Strongly disagree

4. Please select your parents or guardians' highest degree. (If you have multiple guardians, multiple selections are allowed.)

□①Bachelor

□②Associate

□③High school education

□④Vocational School

□⑤Middle School

□⑥You do not know about that

□⑦Others

5. Do you have elder brothers and sisters?

①Yes ②No

<u>1) If yes</u>, select their highest degrees. (If you have multiple older siblings, multiple selections are allowed.)

□①Bachelor

□②Associate

□③High School

□④Vocational education

□⑤Middle School

□⑥You do not know about that

Thank you for your cooperation

Tae Kunisawa: PhD candidate at the University of New Mexico

Appendix M

Handout B for Day 1 and Day 3[124]

Classroom activity

Practice 1:

Conversation with your partner to help Ken organize his room.
ケンの引越しを手伝っているメアリーとの会話

Ken:	Hi Mary.
ケン:	こんにちは、メアリー。
Mary:	Hi Ken.
メアリー:	こんにちは、ケン。
Ken:	Thank you for your help.
ケン:	手伝ってくれるんですね、ありがとう。
Mary:	Sure, no problem. How can I help you?
メアリー:	どういたしまして。何を手伝ったらいいですか。
Ken:	Can you put the desk there?
ケン:	そこに机を置いてくれますか。
Mary:	Sure. Is this okay?
メアリー:	いいですよ。これでいいですか。
Ken:	Fine.
ケン:	それでいいです。
Mary:	Where can I put the monitor?
メアリー:	どこにモニターを置きましょうか。

[124] Images have been removed because of copyright issues.

Ken: On the desk, please.

ケン: 机の上にお願いします。

Mary: For sure. Is this from Japan?

メアリー: はい。日本から持ってきたんですか。

Ken: Yes, it is.

ケン: そうです。

Mary: It is so nice.

メアリー: とてもいいですね。

Ken: Thank you.

ケン: ありがとう。

Mary: How about the computer?

メアリー: コンピューターはどこに置きますか。

Ken: Under the desk, please.

ケン: 机の下にお願いします。

Mary: Okay. Is it a Mac or a PC?

メアリー: はい。それ、マックですか。それとも PC？

Ken: It’s a PC.

ケン: PC です。

Mary: I like PCs. Where can I put the keyboard?

メアリー: PC 私好きです。キーボードはどこにしましょうか。

Ken: On the desk, please. Can you put the desk lamp on the desk?

ケン: 机の上にお願いします。机の上にスタンドを置いてくれますか。

Mary:	Sure.
メアリー:	いいですよ。
Ken:	Can you put the tablecloth over the table?
ケン:	テーブルの上にテーブルクロスを広げてくれますか。
Mary:	Once more, please?
メアリー:	もう一度、お願いします?
Ken:	Can you put the tablecloth over the table?
ケン:	テーブルの上にテーブルクロスを広げてくれますか.
Mary:	I got you. Thank you. No problem. I will put it over the table. What is next?
メアリー:	わかりました。ありがとう。いいですよ。テーブルの上にそれを広げておきます。次は何ですか。
Ken:	I bought the pizza at Pizza Hut. Please put the pizza on the table.
ケン:	ピザハットでピザを買ったんですけど、テーブルの上にピザを置いてくれますか。
Mary:	Sure.
メアリー:	わかりました。
Ken:	Do you like snowboarding?
ケン:	スノーボード好きですか。
Mary:	For sure. I go snowboarding in the winter!
メアリー:	ええ、好きですよ。冬はスノーボードをしに行くんです!
Ken:	Where do you go snowboarding during the winter?
ケン:	どこに行きますか?

Mary: There is a ski resort over the mountains. I go there every winter. It's fun. Do you like it?

メアリー: 山の向こうにスキー場がありますけど、毎年、冬にはそこに行きます。楽しいですよ。スノーボード好きですか。

Ken: Yes, I do. My brother took this picture for me. Look at this.

ケン: はい。兄が写真を撮ってくれました。見てください。

Mary: It 's so cool! I'll put the picture above the bed for you. Is that okay?

メアリー: かっこいいですね。ベッドの上にその写真を置きますけど、いいですか。

Ken: Thank you.

ケン: ありがとう。

When Mary looks out at a window, she sees a bird.

メアリーが窓の向こうに、鳥を見つけます。

Mary: Look at that! A humming bird is flying over the fence.

メアリー: 見て! ハミングバードがフェンスの上を飛んでいます。

Ken: It's so sweet!

ケン: かわいいですね!

Mary: I love the bird, too. There are a lot of clouds in the sky today.

メアリー: あの鳥、私も大好きなんです。今日は雲が多いですね。

Ken: Clouds are moving slowly over the sun.

ケン: 雲が、太陽の上をゆっくり動いていますね。

When Mary looks up at the ceiling, she finds there is a water stain.

メアリーが天井を見上げたとき、しみを見つけます。

Mary: There is a water stain on the ceiling.

メアリー: 天井に water stain があるんですけど。

Ken: What does the water stain mean?

ケン: water stain はどういう意味ですか。

Mary: The color comes from dripping water. You can paint over the stain on the ceiling.

メアリー: 水漏れのため色がつくという意味です。ペンキでそこのところを塗っておけばいいんですけど。

Ken: I will ask my host-family. Thank you so much for helping me.

ケン: ホストファミリーに聞いて見ます。手伝ってくれてありがとう。

Mary: No problem. Any time. Bye now.

メアリー: どうも。いつでも声かけてください。じゃあ、今日は、これで。

Practice 2:

ケンの部屋

A: インフォメーションギャップ:

インフォメーションギャップ:

Question1: Do you see the monitor in your picture? コンピュータのモニターは絵にありますか

Answer1: Yes, I do. はい。

Q2: Where is the monitor? それはどこにありますか。

A2: It is on the bed. ベッドの上にあります。

Q3: Where do you put the pizza? ピザはどこにおいていますか。

A3: I put it on the desk. 机の上です

Q4: Do you see the computer under the desk? 机の下にはコンピューターがありますか。

A4: No, I don’t. いいえ.

Q5: Where is a cloud moving? 雲はどこに移動していますか。

A5: The cloud is moving over the sun. 雲は太陽をおおうように移動しています。

Q6: Do you see the water stain on the ceiling? 天井にしみがありますか。

A6: No, I don’t see it. いいえ、しみはみあたりません。

Q7: Where is the ski resort? スキー場はどこにありますか。

A7: It is over the mountain. A7: 山の向こうにあります。

- **プリント#1３(A)を見ながら、プリント#13(B)を見ている生徒に質問 (例)**

Q1: Do you see the tablecloth in your picture? 絵にはテーブルクロスがありますか。

A1: Yes, I do. はい。

Q2: Where is the tablecloth? テーブルクロスはどこにありますか。

A2: It is over the table. テーブルの上にあります。

Q3: Where do you see the picture of the snowboarding boy?

スノーボード少年の絵はどこにありますか。

A3: It is above the window. 窓の上にあります。

Q4: Do you see the hummingbird? ハミングバードがみえますか。

A4: Yes, I do. はい。

Q5: What is the bird doing? その鳥は何をしていますか

A5: The bird is flying over the fence. 鳥はフェンスの上を飛んでいます。

Q6: Do you see the water stain on the ceiling? 天井には、しみがありますか。

A6: Yes, I do. はい。

Q7: What will you do to hide it? それをどうやって隠すつもりですか。

A7: I will paint over it. ペンキで塗るつもりです。

Appendix N

Examples of ICSCTGs