A Practical Evaluation of the Effects of Oral Sentence Building Training

—Possible Improvements in Speaking Ability—

A Thesis

Submitted to

Department of English Education

Osaka Kyoiku University

In Partial Fulfillment of

the Requirements for the Degree of

Master of Education

---

by

SHINOZAKI Fumiya

January 2013

Acknowledgements

The life that I spent at Osaka Kyoiku University was the most valuable, experiencing many beneficial things.

I would like to express my cordial gratitude to everyone in the department of English education at Osaka Kyoiku University. First of all, I am grateful to Ms. Ikuma, who provided me with a lot of valuable advice about conducting the study and developing the thesis. In addition, I am really thankful to Mr. Malcolm for editing the thesis. Dr. Yoshida and Dr. Kagata were also available to give me a helping hand at all times.

I would also like to recognize the participants of the study, who took the trouble to spare plenty of time for the sake of my research.

Lastly, I would like to express my appreciation to all the people around me, including my precious friends and my family. They supported me whenever I was in need of their help.

Thanks to all of them, I could complete this study to the end.

SHINOZAKI Fumiya

Osaka Kyoiku University
January 2013

Contents

Figures

Tables

Appendices

1. Introduction

Currently, there is a growing concern about the linkage between elementary, junior high and high schools, along with the official commencement of foreign language activities at elementary school in 2011 (MEXT 2008a). Because of this, more and more emphasis is now put on enhancing communication skills, and teachers are obliged to make more effort to accommodate themselves to this tendency so that the students will be able to learn to communicate with other English speakers around the world in English more effectively.

Nevertheless, it is undeniable that the purpose of learning English at school is still, for most students, to get through high school or university entrance examinations. Despite the fact that, just as the Ministry of Education, Culture, Sports, Science and Technology (MEXT) states in the course of study (MEXT 2008b; 2009), integrating the four English skills of listening, speaking, reading, and writing is said to be of a great importance, it seems that inconsistency in and ambiguity of the goals of English classes conducted from elementary school through high school prevents students from improving these respective skills effectively. Therefore, it is crucial to

change this present situation and provide them with educational continuity where they can not only continue learning English but come to be able to use English in more practical ways.

In order to vary and increase students' sense of purposefulness in learning English, the ongoing English education at university must be revamped and thereby cater to social needs. In this era of expanding globalization, how efficiently they can develop their abilities to communicate or negotiate with others, while attending school, will be a key factor for Japanese society to flourish and survive in the world. In this respect, the entire system of English education in Japan needs considerable and rapid development.

Above all, speaking skills should receive a more focused attention since it has long been said that Japanese people are not adept enough at speaking, compared with the other three skills, which would be attributed to the lack of time for speaking practice during class (Ohtaka, 1998). Therefore, there is a compelling need for efficient methods of speaking practice.

This study explores the effects of a practical method of speaking practice called Oral Sentence Building (hereafter OSB). In OSB, a learner listens to three separate recorded language chunks, memorizes them temporarily, and rearranges them in a correct order so that the chunks can make a meaningful sentence. The sentence needs to be produced orally. Eventually, how much the participants' speaking skills were improved through the practice will be analyzed and discussed. In Chapter 2, previous studies discuss the current situations of English education in Japan, English proficiency levels of Japanese people, and drills for speaking practice. Chapters 3 and 4 contain details of this study: the purpose of this study, its research questions, the participants, the procedure, and analyses and discussion. Finally, Chapter 5 answers the research questions, and Chapter 6 refers to the limitations of this study so as to make suggest ions for future research in similar fields.

2. Previous Studies

2.1 The Background of English Education in Japan

First of all, just as referred to in the previous chapter, English education in Japan needs changing, as reflected in the transition period due to the beginning of foreign language activities at elementary school. All teachers will be obliged to ponder over what changes may occur as to how they teach English. In order to make English learning more efficient, it is necessary to discuss how students learn English at school. Therefore, the curricula of English education at each level (elementary, junior high, and high school) are going to be the center of the discussion here.

In the first place, why was English education at the elementary school level begun? Takanashi & Takahashi (2007) and Butler (2005) summarized the sequence of the introduction of English education to elementary school. In the early Meiji era, approximately 150 years ago, English was learned as a tool that enabled people in those days to learn advanced technologies or skills from Europe. Until recently, or perhaps still now in some areas, English was or is regarded as something to become more sophisticated, or as one of the subjects studied for entrance examinations. However, it has been

an urgent issue to foster people who can use English as a means of communication because Japan is experiencing expanding globalization. According to Erikawa (2007), as a matter of fact, cultivating communication ability started being focused on after the publication of the revised course of study in 1989.

In addition, Japan has to be prepared for globalization, keeping an eye on the movement of other countries. Butler (2005) maintains that many Asian countries launched English education at elementary school around 2000. Since globalization has both advantages and disadvantages, it could be a great threat for a nation if it fails to move with the times.

On account of this background, Japanese people have been expected to use English in a more practical way. Now, there is a question of how proficient they should be in English. MEXT set goals in the strategic concept called "Eigo ga Tsukaeru Nihonjin no Ikusei no tameno Kodokeikaku (The action plan for improving English proficiency of Japanese people) in 2003. The strategic concept says that junior high school students should be proficient enough to pass STEP EIKEN Grade 3, and similarly high school students STEP EIKEN Grade Pre-2 or 2, at the time of graduation. These

tests assess vocabulary, grammatical knowledge, reading, listening, and speaking skills. Thus, comprehensive English ability needs to be learned. In a similar manner, university students should have a good command of English, particularly in their specialized fields, before graduation (MEXT 2003). Therefore, English education at elementary school should be based on these goals and supportive of the curricula of junior high and high school.

Just as the trend of the times is trying to make Japan do thus, MEXT has created the new versions of the course of study. Below are the overall objectives of English education at elementary, junior high, and high school that presumably English teachers should note.

> Foreign Language Activities (Elementary School)
> To form the foundation of pupils' communication abilities through foreign languages while developing the understanding of languages and cultures through various experiences, fostering a positive attitude toward communication, and familiarizing pupils with the sounds and basic expressions of foreign languages.
>
> Foreign Languages: English (Junior High School)
> To develop students' basic communication abilities such as listening, speaking, reading and writing, depending their understanding of language and culture and fostering positive attitude toward communication through foreign languages.

Foreign Languages: English (High School)

To develop students' communication abilities such as accurately understanding and appropriately conveying information, ideas, etc., deepening their understanding of language and culture, and fostering a positive attitude toward communication through foreign languages.

(MEXT, 2010a; 2010b; 2010c)

These objectives each share the term "communication," and students are supposed to improve their communication abilities in stages. At elementary school, the most attention will be paid to input by sounds so that pupils get accustomed to English phonemes. Junior high school students usually begin learning how to write English with grammatical knowledge gained and accumulated through classes. From this stage, teachers are driven by the need to teach the four skills with balance. Education for international understanding should be adopted to English classes, supposing that English is used in such a global society, but at the same time, smoother communication requires students to develop their abilities in the English language itself.

According to Canale and Swain (1980), there are four types of communicative competence: grammatical competence, discourse competence, sociolinguistic competence, and strategic competence. In the explanation of

the five important principals for communicative approach, they admit that one competence does not outweigh the others in importance. That is to say, all the four types of competence are requisite for successful communication. As Nakabachi (2000) acknowledges that sociolinguistic competence and discourse competence would make it possible to have better communication, developing both kinds can be an ultimate goal in learning how to communicate in English. However, it is also true that numerous junior high and high school students take entrance examinations, which ask them questions about English grammatical structures.

In the nature of OSB tasks (details will be discussed in Chapter 2.3.3.), it seems that grammatical competence is the most relevant to those because rearranging chunks demands some knowledge of syntax and semantics. Hence, OSB plays a role in making progress toward smoother communication in this respect. It is from junior high school that grammar is demonstratively taught, so it would be most suitable to teach OSB after elementary school.

To summarize the above points, Japan has been trying to change its educational system. Along with the changes, teachers have to know more

effective and efficient ways of teaching so that the students will be able to keep pace with the globalization. For the sake of achievement, they need to improve their communication abilities by becoming able to use English as a tool, just as people did over a hundred years ago. Therefore, the courses of study were revised accordingly, and development in communicative competence is the pressing issue at school. Communicative competence can be divided into four types, all of which are integral to prosperous communication. Of these four competence fields, practice through OSB tasks will enable students to enhance grammatical competence. Thus, it should be introduced at junior high school or after that.

2.2 English Proficiency Required for Japanese People

It could not be better than having full command of English, with enormous capabilities of listening, speaking, reading, and writing. As discussed in the last section, there are certain goals that students at school should attain, which are indicated with the grades of STEP EIKEN. In practice, how well should Japanese people be able to use English? Here, speaking ability will be mainly conferred.

What standard does a person with a good command of English meet? Nakabachi (2000) defines such a person as one who is able to state his or her own opinions and participate in a discussion actively, assuming that this level of ability is the minimum to keep up with classes at university or graduate school in the U.S. or to negotiate professionally in English as equals in a business setting in internationalized society.

Then, Shiraishi (2010) conducted research into what elements of abilities Japanese companies in which employees have opportunities to use English should weigh heavily and require. According to the research, 89.2% (N = 83) value ability to debate and 90.3% (N = 83) emphasize the importance of ability to make a presentation, when hiring people. Similarly, 86.5% (N = 74) appraise people if they have high ability to debate and 86.6% (N = 75) appreciate them for high ability to make a presentation, in business situations where English needs to be used. Both debate and presentation involve advanced speaking skills, so it is obvious that improving such skills before acquiring a job is essential.

Nevertheless, the fact is that many Japanese people do not meet these criteria and nor do they have adequate speaking skills. Educational

Testing Service or ETS discloses the Test and Score Data Summary for TOEFL iBT Tests and TOEFL PBT Tests. Taking a look at the data will lead underscores of how significant teaching and learning speaking skills at school are. Figure 1 shows the change in the TOEFL scores of people in Japan from 2005 through 2011 (ETS, 2007; 2008; 2009; 2010; 2011; 2012a). The data shown here were analyzed according to geographic regions and native countries.

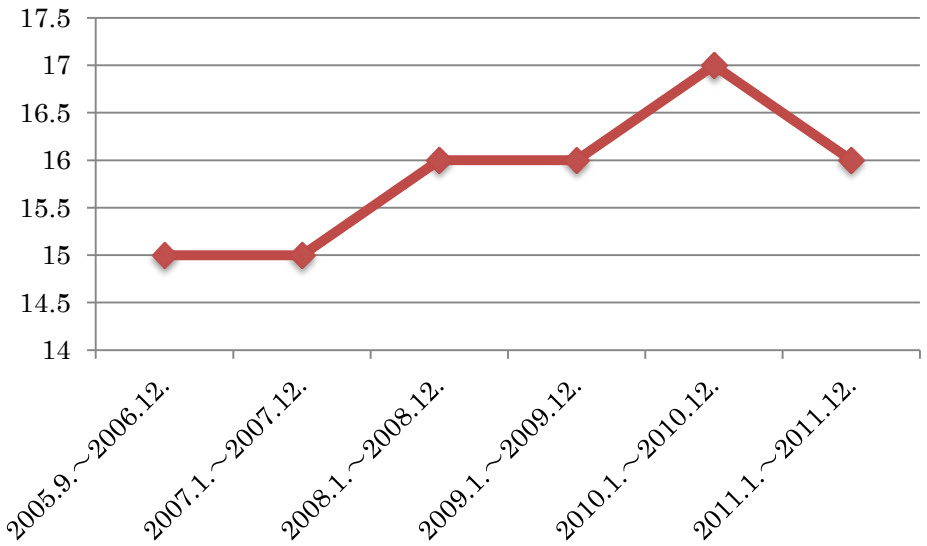


Figure 1. Japanese TOEFL iBT Speaking Score Means from 2005 through 2011 Classified by Geographic Region and Native Country.

The TOEFL iBT speaking score is assessed with the scale of points, 30 being the maximum and 0 being the minimum. According to ETS, the score scale between 10 and 17 shows that the test taker has limited speaking

ability (ETS, 2012b). Figure 1 denotes that Japanese people do not have adequate ability to debate or make presentations in English, and do not cater to the standard defined by Nakabachi. Besides, it seems that this has not changed much in the past six years.

As to the scores of Japanese people, there are some controversies. Japan has low scores on TOEFL over years compared with other countries around the world, and the language distance between English and Japanese can be one of the factors that prevent Japanese people from obtaining higher scores (Manto, 1995). Brown (1993) points out that in Japan English is usually only learned at school and not used outside of classes, as opposed to countries such as Singapore and India, where English is used as an official language. This contention suggests that it is no use comparing EFL settings and ESL settings. According to the data by Foreign Service Institute (1973), it will take Americans whose native language is English 2760 hours to become able to speak Japanese, Korean, or Chinese to the general level of mastery, 720 hours for French or Spanish, 1320 hours for Bulgarian or Indonesian, and 1500 hours for Finnish or Thai. The opposite will be true and it can be also referred from this data that language distance

influences Japan's scores in a negative way, so Japanese people suffer from a handicap in learning English. However, even though Japan is compared with some other Asian countries, its score still seems relatively low (Torikai, 2002). More and more foreign people live in Japan, which accelerates what is called internationalization at home, so English is necessary not only for business but also for daily life (Japan Center for Economic Research, 2011). From the above-noted research, it is clear that there is a strong and immediate demand for Japanese people to improve their English to a level that can be accepted in global society.

2.3 Methods of Speaking Practice

Leaners are able to improve their speaking ability through numerous methods. Usually, they integrate multiple methods of speaking practice so as to gain more effects toward learning. This section will look at a few methods through which learners make progress in their speaking ability, including shadowing, repeating, and OSB. Then, each of them will be compared and contrasted with the others.

2.3.1 Types of Drills

There are several types of drills; the representative examples are mechanical drills, meaningful drills, and communicative drills. A Guide to English Language Teaching Terminology Revised Edition (2009) defines each term as shown below (See Table 1).

Table 1
Types of Drills

Mechanical Drill	This drill focuses on a repetitive practice. It is to promote memorization of learnings and habit-formation with a lot of repetitive practice of mimicry, repetition, substitution, and transformation. The teacher always controls how the learners should respond to the stimuli in a repetitive practice.
Meaningful Drill	Learners respond to the stimuli which the teacher provides, thinking its meaning. The learners' response to the stimuli is controlled to some extent, but they are required to comprehend its meaning in order to determine how they respond.
Communicative Drill	Learners freely respond to the stimuli which the teacher provides.

An example of a mechanical drill is that a learner has to say "I am going to study English" if the teacher says "I am going to study English" as a stimulus. Mimicry-and-memorization and pattern practice are classified into this category of drill. Therefore, learners are not required to

understand the stimuli semantically. In contrast, they need to analyze the meaning of the stimuli so that they can respond to them successfully. For instance, if the teacher asks a learner a question, saying, "May I have your name, please?" and shows him or her three choices: a) I am 18 years old. b) I am Ken. c) I am a student, then the learner has to figure out what it means and choose the choice b) I am Ken. Finally, a communicative drill is different from the other two in that learners' response to the stimuli is not controlled. If the teacher asks a learner the question "What did you do last weekend?" then the learner has to comprehend its meaning and answer, saying, "I went to a museum with my friends," for example. Hence, in order of increasing difficulty, mechanical drills come first, meaningful drills next, and communicative drills last.

Ohtaka (1998) divides speaking practice into two kinds. One is practice through which a learner tries to acquire knowledge of vocabulary, grammar, expressions, and pronunciation. The other is practice through which a learner actually tries to speak English using the knowledge of these things. Needless to say, the former is the basis of the latter, so it is undoubtedly necessary to build up knowledge to improve English speaking

ability. However, as Ohtaka insists, English education in Japan tends to focus on expanding knowledge only, mostly on account of the entrance examinations. Therefore, more and more attention should be paid to a practice which implements the use of English, namely, the latter.

2.3.2 Shadowing and Repeating

OSB is quite similar to shadowing and repeating in that a learner listens to English and vocalizes something based on what he or she has listened to. According to Lambert (1988), shadowing is "a paced, auditory tracking task which involves the immediate vocalization of auditorily presented stimuli, i.e. word-for-word repetition in the same language, parrot-style, of a message presented through headphones." About a decade later, Tamai (2002) tried to clarify this definition and redefined it thus: "Shadowing is an act or a task of listening in which the learner tracks the heard speech and repeats it as exactly as possible while listening attentively to the in-coming information."

One the other hand, repeating is "sentence-level repeating during a pause after the original utterance" (Hiramatsu, 1999), so shadowing and

repeating are often confused with each other but actually different. Kadota (2007) emphasizes that the only difference between them is whether a pause is made before vocalizing. He regards shadowing as an online activity and repeating as an offline activity because, in repeating, a learner can have much more time to process the information gained by listening to the sounds. In other words, repeating involves more semantic analyses before producing words. Even so, there is a possibility that highly proficient speakers process sounds semantically at a much faster pace. From these, both shadowing and repeating will be classified as meaningful drills to a greater or lesser extent.

The effects of shadowing have been reported. Tanaka (2007) acknowledges that shadowing can be beneficial to both learners with high proficiency and to those with low proficiency, as long as instructions are provided according to their English levels. In addition, Eguchi (2007) carried out an experiment in which technical college students had some constant training with shadowing, and made them take TOEIC Bridge, suggesting that it might be possible that shadowing could improve learners' communication ability.

Shiki, et al. (2010) discovered several differences between shadowing and repeating in terms of reproduction rates and types of reproduced words. Even though there were no differences in the improvement of reproduction rates between those two kinds of training, it seems that in either shadowing or repeating only four or five repetition trials could lead to a ceiling point of the reproduction rate. Moreover, they discovered that shadowing could help learners become more sensitive to English sounds, whereas repeating could impregnate learned language items in their brain.

If OSB really has something in common with shadowing and repeating, there should be some positive effects on the improvement in English ability. Before moving to what has been revealed from this study, it will be necessary to look into what OSB actually is and how it relates to the brain activity.

2.3.3 Oral Sentence Building

OSB may not be a breakthrough idea, but there is no impression that it is recognized widely in Japan. Instead, a written version of sentence

building has often been employed in workbooks for students and appeared on exams at school. Figure 2 shows an example of this.

Q1. Make a sentence by using all the words below.
(is / my / to / buy / going / house / brother / new / a)
Answer My brother is going to buy a new house.

Figure 2. A Written Version of Sentence Building.

Usually in this kind of sentence building question, all the words are shown separately, and a learner tries to rearrange those to make a meaningful sentence. It goes without saying that reading the question many times is allowed, providing sufficient time to think about it for the learner. Japanese translations are sometimes given simultaneously as well.

By contrast, in OSB, the words used are normally chunked. According to Negishi (2007), OSB is a way of speaking practice that a sentence is divided into some units and a learner listens to the sound of units, orally presenting the original sentence by rearranging them. He also admits that this kind of question can be created relatively easily and used for all the students in class all together, so it is very practical. Yoshida (2012) argues that it is a very effective form of training to make a sentence by rearranging chunks that have been divided in terms of meanings.

In Versant, an English speaking test created by Pearson, the same type of question is set. This test is taken by phone and takes no longer than twenty minutes. Referring to the sample from the website (Pearson Kirihara, 2012), the procedure is the one shown in Table 2.

Table 2

Procedure of Oral Sentence Building

Question	was reading / my mother / her favorite magazine
Phase 1	You hear “was reading”.
Phase 2	One second later, you hear “my mother”.
Phase 3	One second later, you hear “her favorite magazine”.
Phase 4	Within five seconds, you answer “My mother was reading her favorite magazine” by rearranging the three chunks.

As shown, there are three chunks, and in principle no words are chunked unnaturally. For example, a question like (was reading her / my mother / favorite magazine) is not used. Through this kind of standard test, it can be speculated that OSB will become more widely-recognized in the near future. The next section will explore and explain how a learner’s brain works while doing OSB.

2.4 The Brain Activity during Oral Sentence Building

When people speak a language, certain parts of their brains are activated, and they are always interacting with one another. Figure 3 shows which parts of a brain are responsible for a language use.

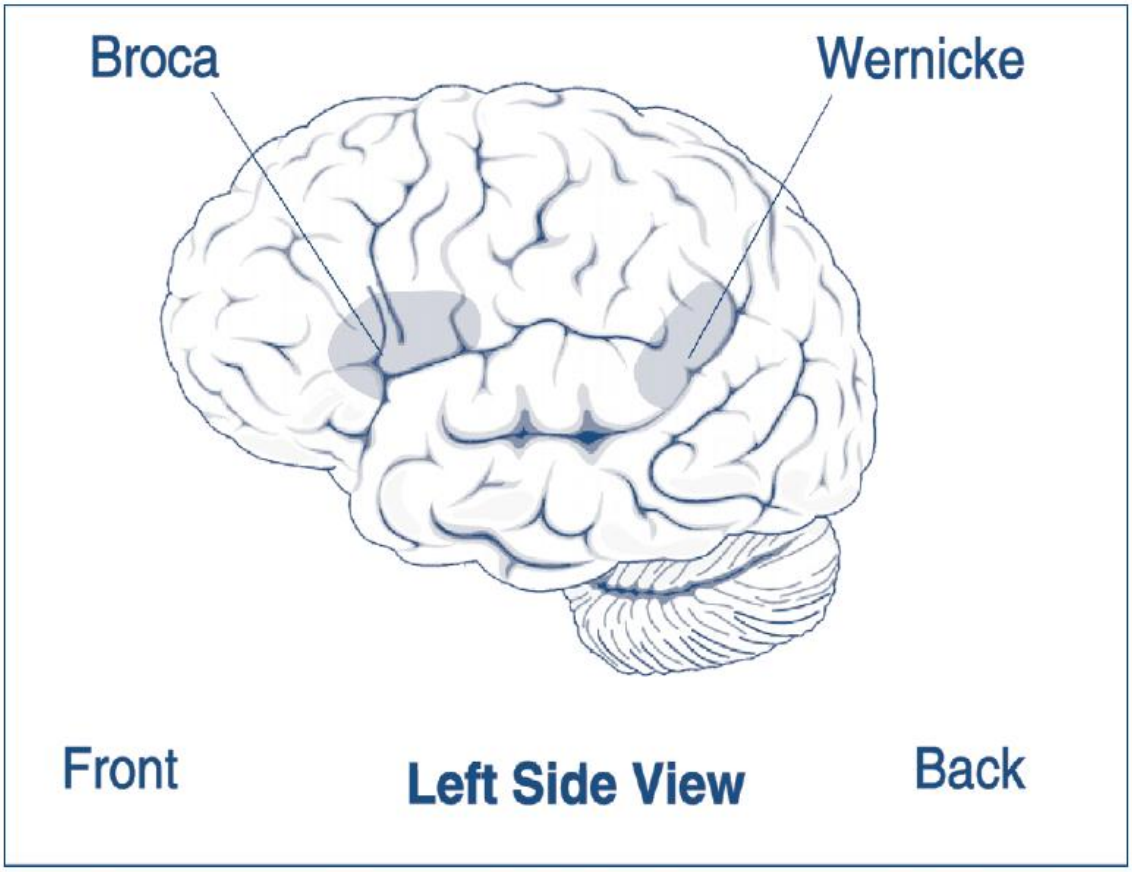


Figure 3. Locations of Broca's Area and Wernicke's Areas.

(NIDCD, 2010)

It is well known that people suffer from aphasia if their Broca's area or Wernicke's area is damaged, though how differently the damage affects

them depends upon the areas impaired; this was discovered from some cases of accidents (NIDCD, 2010). From these, it turned out that Broca's area and Wernicke's area are intimately bound up with language use. Therefore, these two areas are always activated when people use a language; however, these are not the only parts of the brain that are working during speech or conversation.

2.4.1 Automatization and Level of Processing

It is said that fluent speakers and automaticity are closely related to each other. Automaticity is the ability to process input automatically without any conscious effort. It includes the ability to use learned grammatical items instantaneously (A Guide to English Language Teaching Terminology Revised Edition, 2009). The Attention-Processing Model says that lower-level processing relevant to vocabulary or grammar can be automatized with accumulated learning and more attention can be paid to upper-level cognitive processing. Additionally, controlled processing, in which a learner has to pay attention to given tasks consciously and take time for language processing, can make the transition to automatic processing,

wherein a learner pays attention to given tasks unconsciously and can process language input or output relatively speedily (McLaughlin, 1987; 1990). As discussed in Section 2.3.2., both shadowing and repeating require a learner to process information rather quickly and to vocalize immediately thereafter. The processing speed varies according to how much automatization has progressed. It is supposed that simultaneous interpreters are able to interpret very smoothly because they have experienced a great deal of training where automaticity is developed.

The term "level of processing" here indicates how deeply input aural information such as words and sentences is processed. Craik & Lockhart (1972) discuss depth of processing, maintaining that "greater 'depth' implies a greater degree of semantic or cognitive analysis." Furthermore, they allege that input is likely to be retained longer when stimuli that need semantic and cognitive analyses are given, than when stimuli that do not need those are given. This is because more attention is devoted to the former. As a result, they postulate that "analysis proceeds through a series of sensory stages to levels associated with matching or pattern recognition and finally to semantic-associative stages of stimulus enrichment." Thus,

in these phases, semantic analyses cause the deepest level of processing, which means that elaboration reaches its maximum at this point. For these reasons, repeating is more likely to induce semantic analyses than shadowing, particularly during an activity in which the content will be heard for the first time, because a learner has some more time before starting to vocalize. Therefore, repeating requires a deeper level of processing, though shadowing will entail quicker processing without much semantic coding.

Hence, it seems that OSB is more similar to repeating than shadowing. Nonetheless, they will still be different in terms of dual task. Just as Baddeley & Hitch (1974) did in their experiment, OSB demands two kinds of tasks, one being the primary task and the other being the secondary task. In repeating, a learner has to "listen to the sounds of words or sentences," "analyze them semantically," "retain the words and structures used," and "vocalize the original sentences restructured from his or her memory." In short, it is only necessary to repeat the original sentences. On the other hand, in OSB, a learner has to rearrange the chunks (the primary task) and vocalize the original sentences almost at the same time (the secondary task). Therefore, it imposes more cognitive load, which

signifies that a learner will analyze information more semantically in OSB than in repeating. Thus, because of its deeper level of processing that allows a longer retention of information, there is a high possibility that repetitive training with OSB could contribute to improvement in English ability in some way.

2.4.2 The Process of Memorization and Working Memory

At present, numerous studies on human memory have been conducted and some researchers have produced their own models that hypothesize how information is dealt with inside the brain. According to the Atkinson's & Shiffrin's model (1971), any kind of information input through the five senses is divided into three types of memory (See Figure 4). Sensory registers process any environmental input containing images, sounds, textures, tastes, and smells. Then, Short-term Memory (STM) holds limited information that is judged to be necessary by the selective attention. Peterson & Peterson (1959) revealed that STM retains information only for approximately 15 seconds at maximum and if nothing is done with this information approximately 90% of it is irrevocably lost.

Therefore, Atkinson & Shiffrin (1971) assert that rehearsal is the key to retaining information longer.

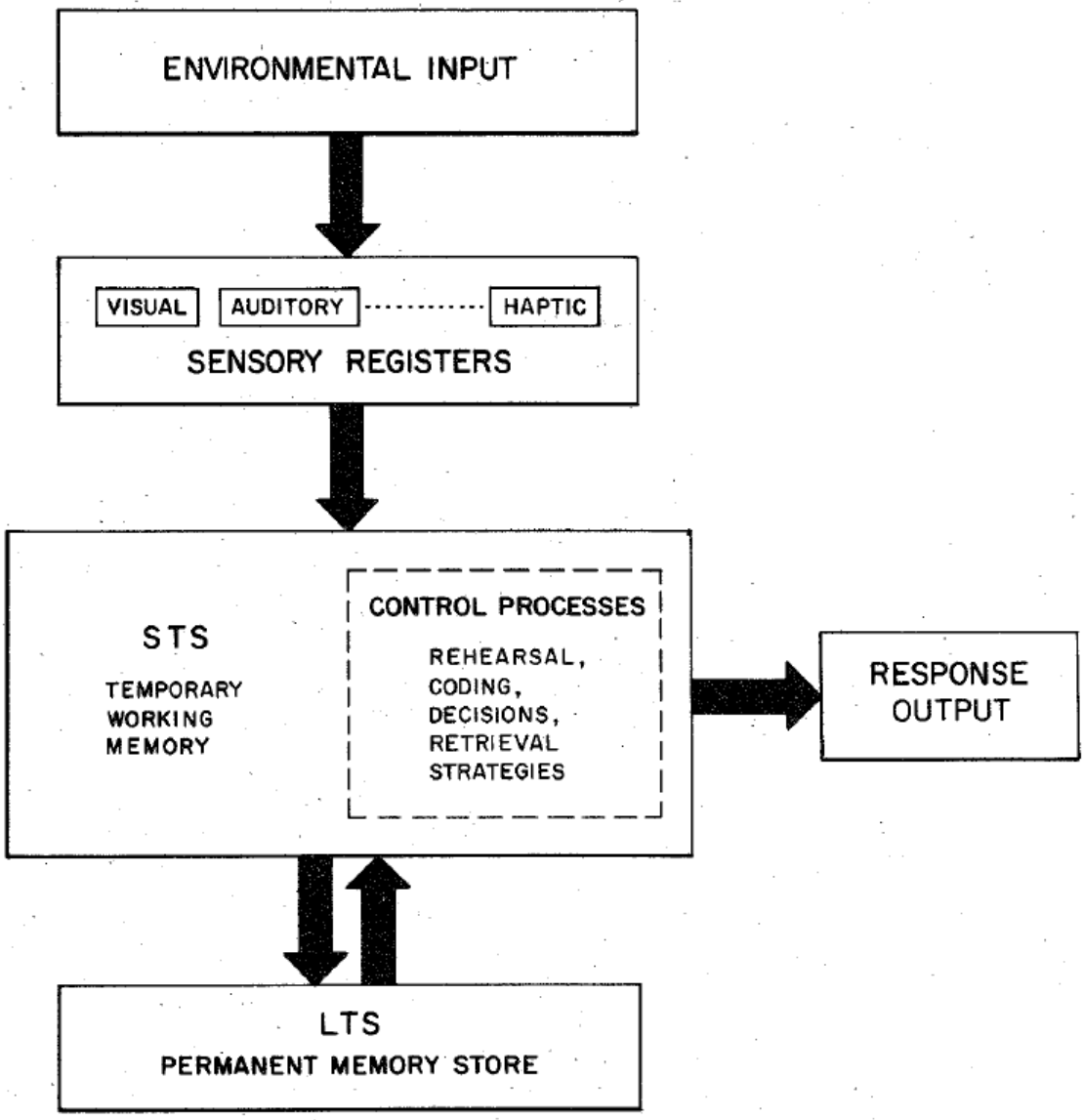


Figure 4. The Atkinson's and Shiffrin's Model of Memories.

(Atkinson & Shiffrin, 1971)

With the repetition of rehearsal, the information in STM is more likely to shift to Long-term Memory (LTM). Therefore, it is important to move what a learner has learned into LTM.

Several years later, a new notion was added to their model by Baddeley & Hitch (1974). They had another look at the mechanism of STM and adopted the idea Working Memory (WM), whose term is said to have been coined by Miller, Galanter and Pribram, according to Baddeley (2003). WM is "a brain system that provides temporary storage and manipulation of the information necessary for such complex cognitive tasks as language comprehension, learning, and reasoning" (Baddeley, 1992). It actually refers to the same realm as STM, but with additional view that in STM information is not just retained for a short time but is also processed. In addition, STM and WM are different from each other in that STM retains information passively, while WM retains actively (Osaka, 2000). For instance, you remember the phone number that you have just heard, and dial the number, or you do a mental calculation on the total amount of expenses when shopping. It is generally believed that WM is deeply linked

with the prefrontal cortex as Figure 5 illustrates, but is not localized in that area and involves other parts of the brain extensively (Osaka, 2000).

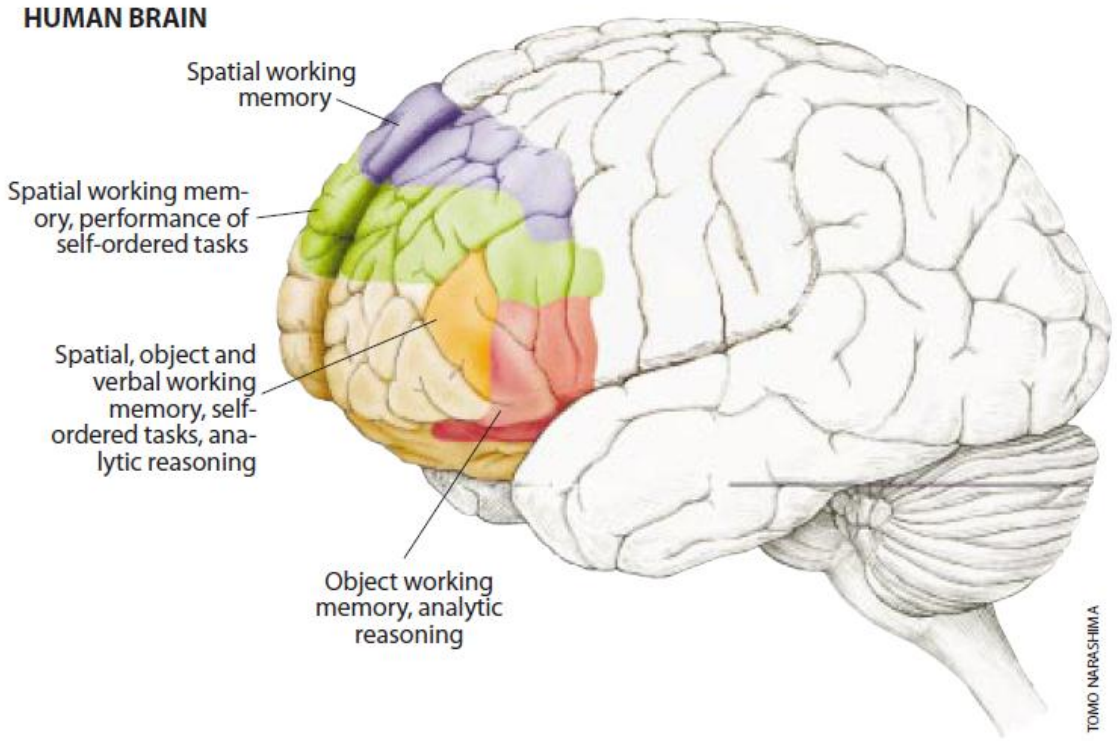


Figure 5. Prefrontal Cortex Deeply Relevant to Working Memory.

(Beardsley, 1997)

Here, the "level of processing," which was already discussed, relates to WM very closely. The degree of how much information can be memorized depends on the depth of processing, so the information can more likely be stored in LTM if some semantic analyses are repeated while it remains in WM.

2.4.3 Phonological Loop and Oral Sentence Building

With years of research, it has been revealed that WM has four slave systems: phonological loop, visuospatial sketchpad, episodic buffer, and central executive (Baddeley, 2003). Each of them interrelates with the others, so they cannot be separated (See Figure 6).

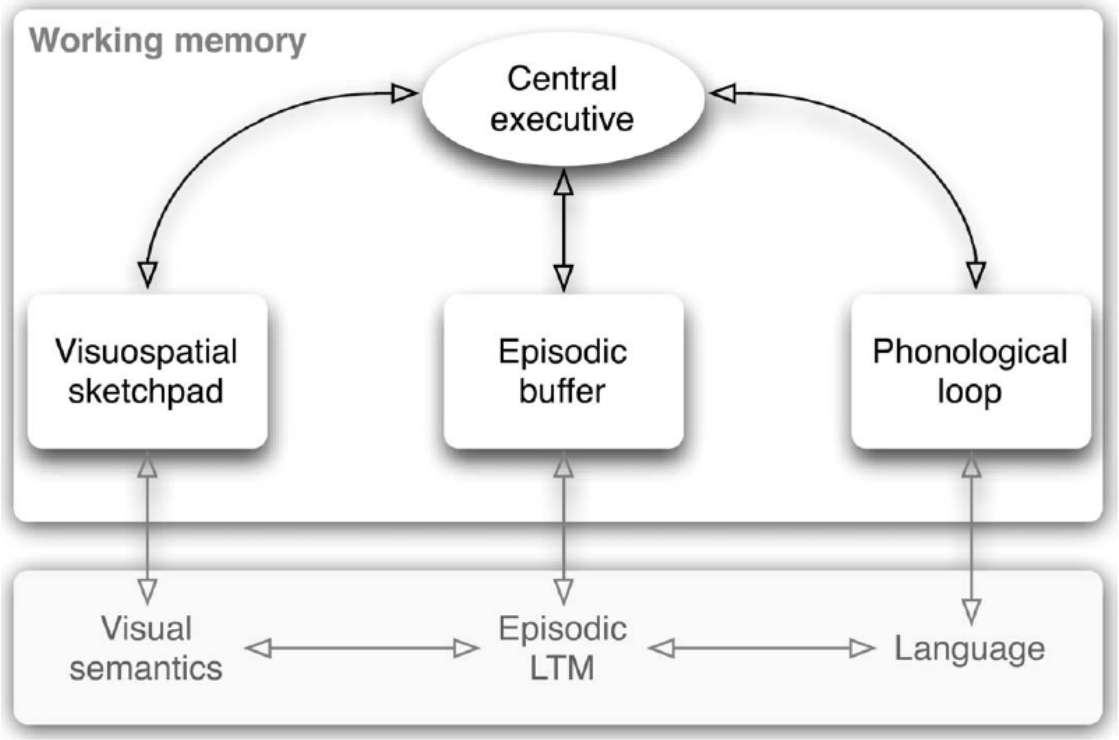


Figure 6. Baddeley's Multi-component Model of Working Memory.

(Repovs & Baddeley, 2006)

Kadota (2007) summarizes the main functions of each of the slave systems. The phonological loop is the system that processes verbalizable stimulus inputs that are retrieved from the various knowledge databases in

LTM, retaining those inputs temporarily. The visuospatial sketchpad is the system that retains and processes unverbalizable stimulus inputs. The episodic buffer is the system that integrates inputs processed both in the phonological loop and the visuospatial sketchpad with information in LTM. Finally, the central executive is the system that regulates the information processing by dividing cognitive resources between the phonological loop and the visuospatial sketchpad in collaboration with the episodic buffer.

Thus, it can be surmised that phonological loop is the system that relates to OSB the most profoundly. It is composed of two components: a phonological store and an articulatory rehearsal process. "The function of the articulatory rehearsal process is to retrieve and re-articulate the contents held in this phonological store and in this way to refresh the memory trace" (Repovs & Baddeley, 2006). In other words, auditory input is retained in the phonological store first, and is repeated in the mind as subvocal rehearsal so as not to lose the information, which allow it to stay in the phonological store longer; without subvocal repetition, it disappears in approximately two seconds (Baddeley, 2002).

WM capacity is controversial. According to the magical number seven, plus or minus two, which Miller advocated, the number of targets that you can remember varies depending on what you try to remember, and approximately five chunks can be retained if the target to be remembered is words (Miller, 1956). Also, it is said that the pure amount of memory is about four chunks if unnecessary elements are excluded (Cowan, 2000).

Furthermore, Baddeley et al (1975) revealed that "subjects can generally remember about as many words as they can say in 2 seconds." Therefore, WM capacity is influenced by what language you use. An experiment administered by Ellis and Hennelly (1980) demonstrated that the memory span of bilinguals who spoke Welsh and English was smaller when they used Welsh because the word length of each language is different. Thus, the languages that you use affect the numbers of words that you can remember.

2.5 Reading Span Test and Listening Span Test

As discussed so far, the functions of WM are essential to complete OSB training that contains a dual task. Reading Span Test (RST) was first

created by Daneman & Carpenter (1980) in order to measure the individual differences in working memory (the processing and storage functions). In their original test, the subjects were shown a card and asked to read a sentence aloud and remember its terminal word. There was only one sentence per card, and each terminal word was different. The number of cards per set shown to the subjects was augmented by the experimenters. The test comprised three sets per level, and it began from two-sentence level, where two cards were shown, to six-sentence level, where six cards were shown. The subjects had to pass two sets out of three so as to go on to the next level. The level that they could pass was treated as a part of the measure of their reading span. They reported that there were few college students who could pass the six-sentence level. Thus, RST is a dual task that demands both reading aloud and remembering and recalling terminal words, so it is used to measure your WM capacity, particularly concerning your reading span.

Alptekin & Ercetin (2011) examined how WM capacity affected second language reading. The participants were Turkish undergraduate students. They used RST in their experiment; the test started at the

two-sentence level and increased in difficulty up to the five-sentence level, with five trials included per level. The tasks that the participants performed while sentences were being shown were a grammaticality judgment task and memorization of the terminal words of the sentences. The experiment was conducted on computers, and seven-second intervals were set between each question. The participants were asked to choose either “grammatical” or “ungrammatical” for each sentence, and to type in the final words in a field box that popped up after all the sentences in the set were shown. In this way, RST is often employed to measure WM capacity.

In addition to RST, Listening Span Test (LST) is also used to measure your WM capacity. According to Osaka (2002), LST is a task in which subjects listen to sentences and remember certain words contained in the sentences. Generally, the subjects are required to judge if the sentences that they hear are grammatically correct.

The evaluation value of RST is called span. Osaka (2002) explained how it is evaluated. Supposing that subjects are required to conduct five trials, succeeding more than three times is necessary to pass a given level. If they pass the two-sentence level, their RST span is 2.0. If they cannot

pass the level but succeed twice, 0.5 points are added. For example, when they can pass the three-sentence level but fail the four-sentence level with two successes, their RST span is 3.5.

Kadota (2007) points out some possible problems of such span tests. It is predictable that scores can increase greatly if subjects take time to perform the tasks. Therefore, as a solution, he states that some time limitation should be necessary for the subjects to do the tasks. Moreover, he brings up several problems of LST. The problems include: at what speed the sentences in tests should be read, how many times the subjects can listen, and who records the sentences, American or British English speakers. Besides, there is a risk that experimenters may have hard time identifying the causes of why the subjects cannot produce the terminal words because they just might have had difficulty in understanding the words, even though they can by looking with their eyes.

3. Methodology of the Present Study

3.1 Purpose of This Study

The purpose of this study is to explore and attempt to circumstantiate the effects of continual English speaking practice through OSB training.

3.2 Research Questions

i. Will the grammatical accuracy of the participants be improved with the continual practice through OSB training?

ii. Will the fluency of the participants be developed with the continual practice through OSB training?

iii. How will WM capacity affect results of each test?

3.3 Participants

There were a total of 30 participants in this study. All of them had experience of learning English over long periods of time. Out of those 30 participants, seventeen were undergraduate students who majored in English education at Osaka Kyoiku University, twelve being freshmen, one a

sophomore, two juniors, and two seniors. Also, eleven were graduate students (ten were English education majors and the other one was a major of another field), another was a Chinese research student, and the other one was receiving general credits not associated with any department.

3.4 Procedure

First, the participants took the pretest, including RST, LST, picture description test (PDT), and OSB test. While they were taking the PDT and the OSB test, their voice was recorded for analysis. Second, they practiced speaking through OSB training as many times as possible (thirty times were the maximum) accessing YouTube, where the sound files had been uploaded by the experimenter beforehand. A few months later, they took the posttest that included the same tests in the pretest (the order of the questions in each test were rearranged, and the picture employed for the PDT was replaced with another one). Similarly, their voice was recorded during the PDT and the OSB test. Then, three participants were selected arbitrarily for an interview. Finally, the collected data were analyzed in terms of vocabulary, grammar, WM capacity, and the interview. Figure 7 charts the procedure.

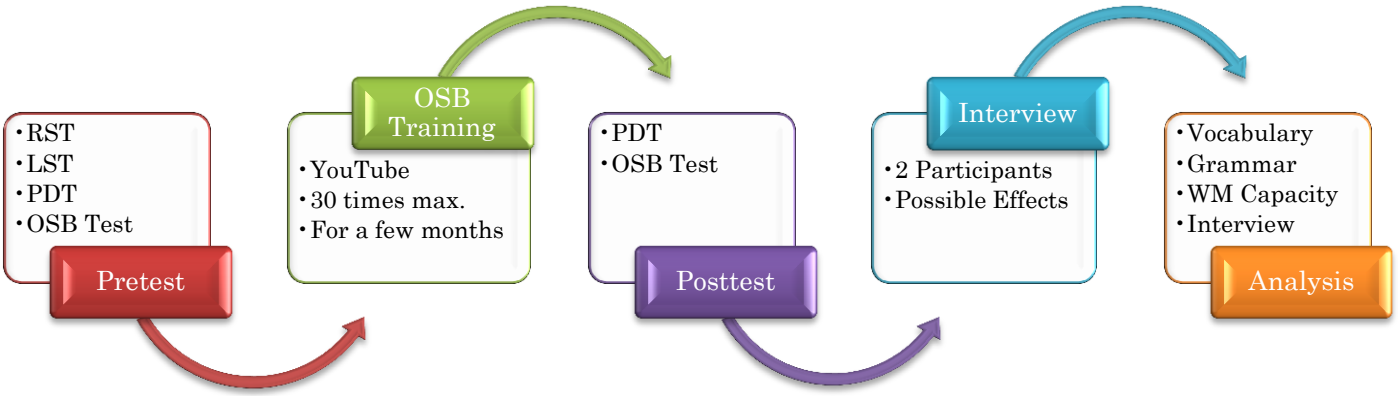


Figure 7. Procedure of This Study.

3.4.1 Reading Span Test

In the present study, an RST was created and used as a measurement of WM capacity. The RST made for the experiment was conducted on computers in order to minimize accidental errors caused by the participants, simulating the fact that the timing of answering or moving on to the next question could not be equal every time for every participant. It contained four sets per level, from two-sentence level to five-sentence level. All the sentences employed were in English, with different terminal words respectively.

There were two kinds of tasks in this test: judging the grammaticality and remembering the final words of each sentence. The

participants had to do these tasks at the same time. The test began with the two-sentence level, so they were supposed to read two sentences and remember two final words, noting that only one sentence was shown at a time. When the target sentence in English orthography appeared on the computer screen, simultaneously two radio buttons to judge its grammaticality with were also shown below the sentence. The left button was "文法 OK" which means "grammatically correct," and the right button was "文法 NG" which means "grammatically incorrect." The participants were given seven seconds per question to choose from the two alternatives by clicking one of those buttons and to memorize the final word. After seven seconds, the test sentence and two choices automatically disappeared and the next question and other two buttons of the same were presented. In the case of the two-sentence level, after seven seconds passed, the text box in which the participants attempted to type the final words that they had memorized appeared on the screen. It disappeared after 15 seconds passed (See Figure 8). This process continued four times. The time limit to type in the final words was made longer as the sentence level increased.

As to evaluation, the full points were given if the participants could pass more than twice out of the four trials, and 0.5 points were given only if they could pass once out of the four trials. This is because those who could not pass once and those who could needed to be discerned. An example of a score is that if you could pass the three-sentence level but pass only once at the four-sentence level, your score will be 3.5. If you only could pass once at the two-sentence level, your score will be 0.5.

3.4.2 Listening Span Test

In addition to the RST, an LST was also made and conducted especially because OSB requires a lot of listening skills as well as speaking skills. Basically, its structure and procedure are the same as those of the RST; however, the participants have to judge the grammaticality and catch the terminal words only by listening to sentences. This is true of the RST, but the terminal words used in the LST were not too difficult for college students because there was a strong necessity of avoiding cases in which the participants could not answer the questions because of lack of English ability.

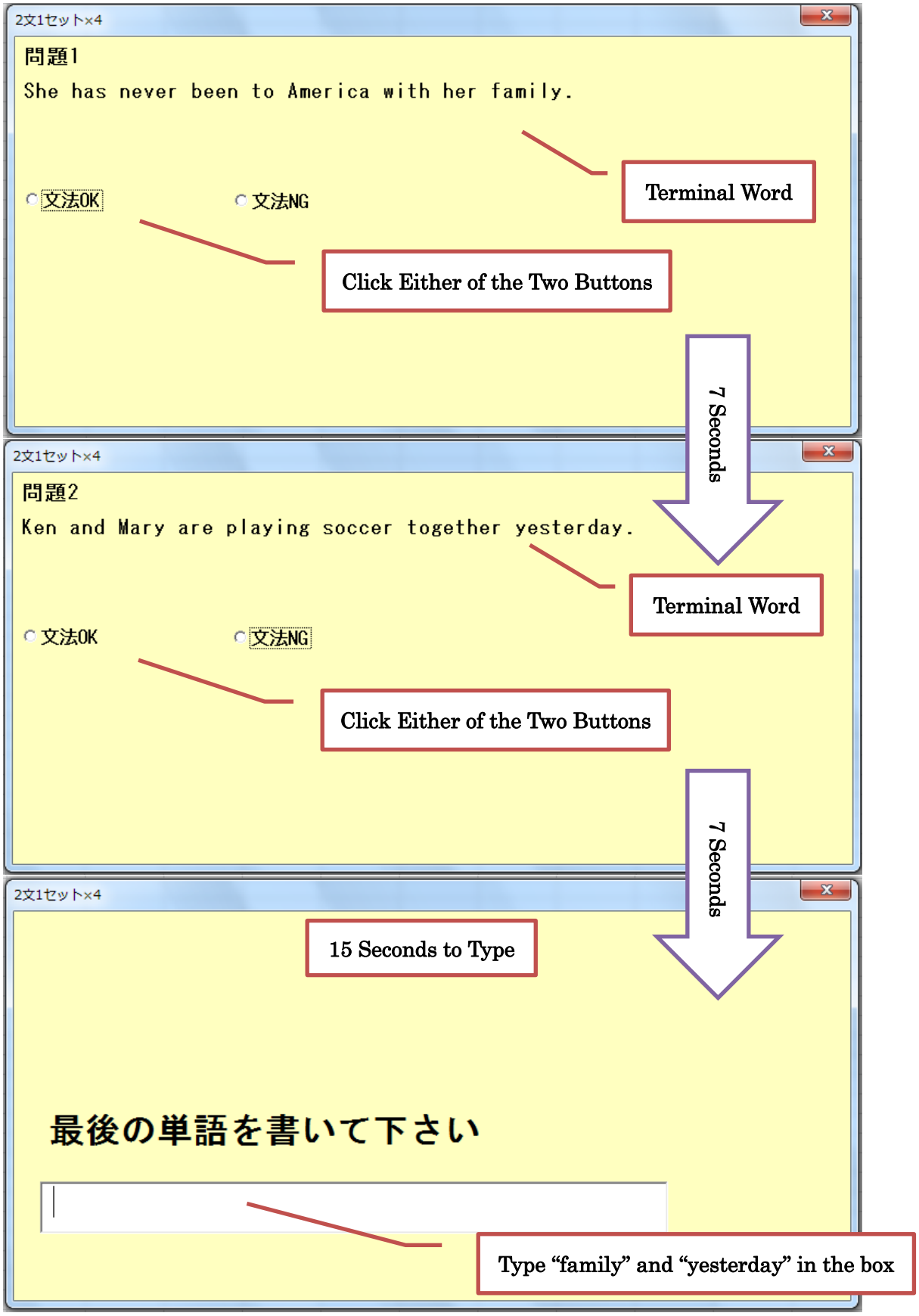
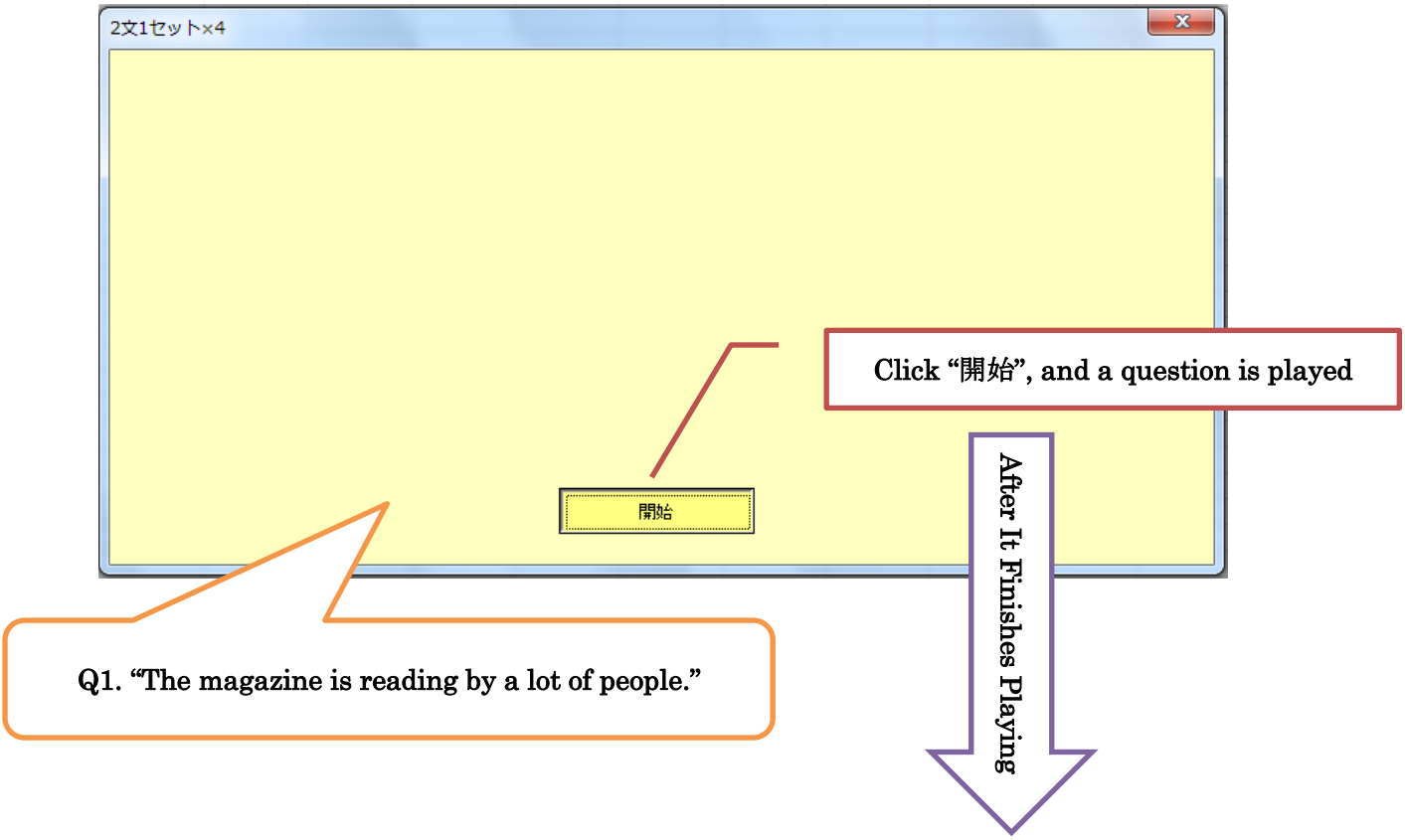


Figure 8. One Set of RST (Two-sentence Level).

The same evaluation system was employed to quantify scores of the participants' achievement on the LST. Figure 9 shows the procedure for a two-sentence level set.



(Continued)

(Continued)

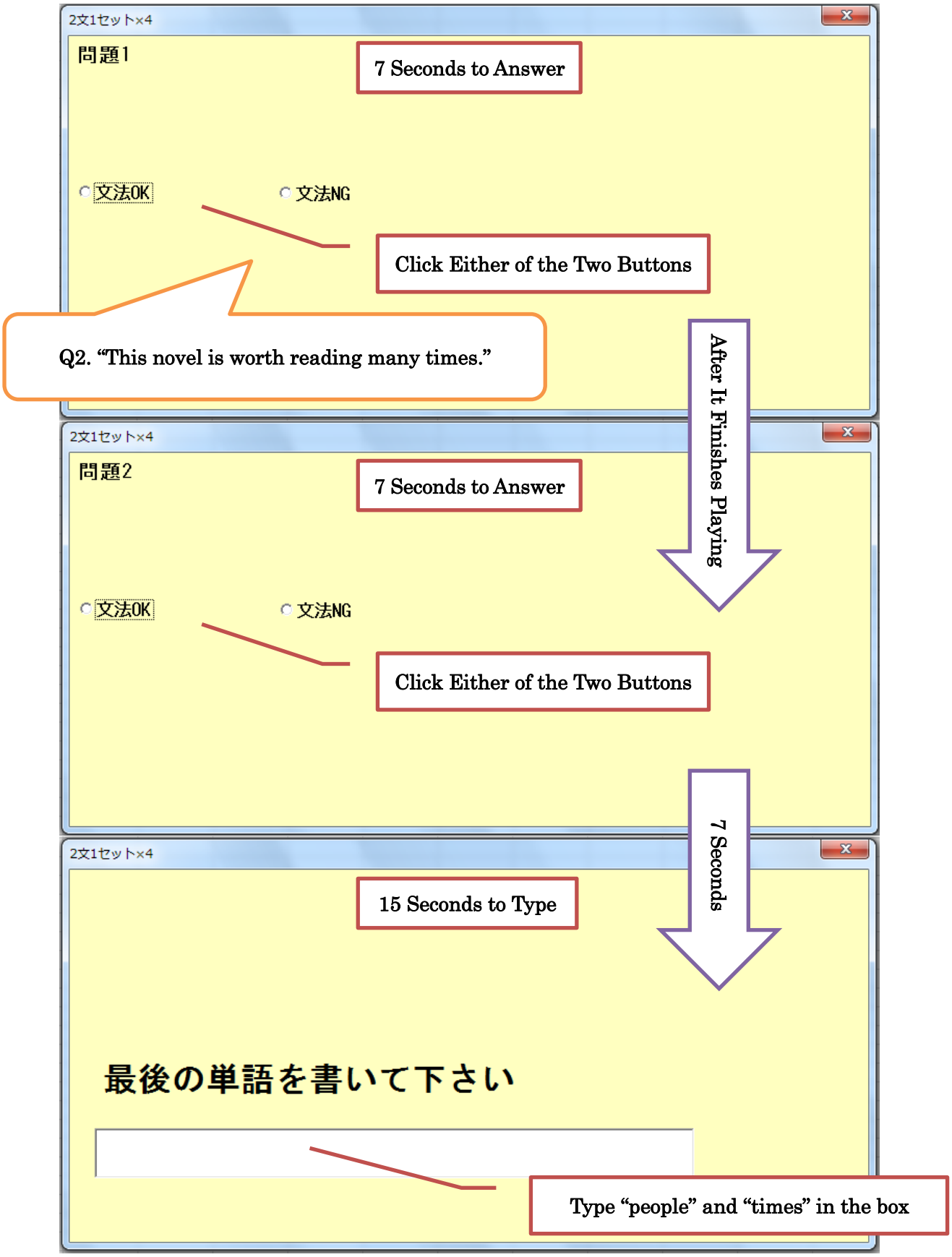


Figure 9. One Set of LST (Two-sentence Level).

3.4.3 Picture Description Test

A PDT was conducted in both pretest and posttest so as to see the participants' speaking ability. The reason why open questions were not chosen for the present study is because it was necessary to avoid situations in which they could not conceive ideas of what to state about given topics.

The participants were given one minute to take a look at the four-frame comic (Appendix III for pretest and Appendix VII for posttest) in order to prepare themselves for the next phase. Subsequently, they were asked to describe it in English in one minute. While describing it, they were allowed to see the comic. For analytical purposes, the voice of the participants was recorded during the test.

3.4.4 Oral Sentence Building Test

The OSB tests used for pretest and posttest respectively were generated by the experimenter (Appendix IV and VIII). One test contained ten questions with diverse grammatical items. Just as discussed in Section 2.3.3., the sentences were divided into three chunks, which were played out of order. There were five patterns of sequences, including 1/3/2, 2/1/3, 2/3/1,

3/1/2, and 3/2/1. There was approximately a one-second pause between chunks, when presumably the participants were trying to do the rehearsal in their brain. Then, five seconds were given to them to build a meaningful sentence using those three chunks. While the participants were taking the test, their voices were recorded for analysis.

The test questions in both pretest and posttest were recorded by a female native speaker of English.

3.4.5 Oral Sentence Building Training

The experimenter created 400 sentences with many kinds of grammatical items such as progressive, passive, auxiliary verb, perfective, infinitive, relative, and so forth. Then, all the sentences were divided into three chunks, and those chunks were arranged randomly. At first, the sentences were sorted according to the types of grammatical items, but they were eventually all mixed together so that a set of questions could have diverse items. Furthermore, since the set of a practice comprises ten questions, all the sentences were split into 40 sets of questions. In the present study, the participants were given 30 opportunities for practice at

maximum. Therefore, the first 30 sets out of those 40 were chosen for the OSB training. Only for these practice exercises, an example answer came five seconds after the final chunk was played for the purpose of providing the participants with feedback.

At the phase of receiving the declaration of their participation in the experiment, the experimenter requested participants' e-mail addresses. As a matter of fact, it was assumed that it would be very difficult to ask them to meet for practice 30 times, so the sound files that contained the recorded questions were uploaded onto YouTube so that the participants would be able to access them whenever they wanted using their computers or cell phones. In principle, the URLs of the OSB questions were sent to them three times a week. That is, ten weeks were used for the training. The participants were supposed to send back an e-mail to the experimenter when they had finished the practice every time. The reply mails were counted as participation on that day.

When it comes to who should record the questions, only Americans were asked in order to minimalize the influence of the speaker's accents; a

man and two women recorded 100 questions each. The playback speed was adjusted to be as homogeneous as possible.

3.4.6 Interview

In order to search for more concrete effects of the OSB training, two participants were chosen arbitrarily for an interview. They were selected because their test results were deemed to be influenced positively and prominently by the continual OSB training. In the interview, the following questions were posed.

1. What do you think the possible effects of a continual OSB training would be?
2. Could the possible effects gained through the OSB training be assessed by the tests (OSB test and PDT)?
3. What other English learning did you do during the period of the OSB training?

4. Analyses and Discussion

This chapter explores the effects of the OSB training based on the analyses of each test result. The discussion will be conducted in terms of the number of OSB training times and the scores on RST and LST. Moreover, the results of the interview will help to answer the research questions as concretely and clearly as possible.

4.1 Statistical Analysis Based on the Number of OSB Training Times

The participants were requested to practice speaking through the OSB training that lasted for ten weeks. There were several who did not practice even once and some who did it all the thirty times. In scoring the OSB tests, two kinds of evaluation criteria were used. Criterion A is where no mistakes are allowed, namely all or nothing. For example, if the answer is "There was a big supermarket near my house," then you need to pronounce the same thing word for word. Criterion B is where a few minor mistakes are allowed as long as the chunk orders are correct and important words such as verbs are not skipped. For example, an answer "There is a supermarket near my house" (big is slipped) is regarded as correct because it

is grammatically accurate, but an answer "There is a big…near my house" is not taken as correct because an important word that should not be omitted to be a meaningful sentence is skipped.

A *t* – test was first conducted between the OSB pretest and the posttest with both Criterion A and Criterion B. Even though the analysis does not necessarily guarantee it will show whether or not the OSB training was effective, there were significant differences in them, as the data shows: $t = -3.587$, $df = 29$, and $p = .001$ (Criterion A), and $t = -2.923$, $df = 29$, and $p = .007$ (Criterion B). Therefore, this denotes a possibility of demonstrating the effects of the OSB training.

Table 3

The Correlation between the Number of OSB Training Times and the OSB Test Scores (Criterion A)

		OSB Test Scores (A)
The Number of OSB Training Times	Pearson Correlation	.366*
	Sig. (2-tailed)	.046

Note. $N = 30$

* Correlation is significant at the 0.05 level (2-tailed).

In order to see the details of the effects of the training, several correlation analyses were carried out. Table 3 shows the correlation

between the number of OSB training times and the increase-and-decrease in the scores on the OSB tests. Criterion A was employed for the evaluation. The average of the number of training times is 11.47, and that of the increase-and-decrease in the test scores is 1.30. The SD of the former is 9.709, and that of the latter is 1.985. From the analysis, the correlation between them is $r = .366$ ($df = 28$, $p = .046$), so it seems that they moderately correlate with each other. The more frequently you practice speaking through OSB training, the better an OSB test result you are likely to obtain.

Table 4
The Correlation between the Number of OSB Training Times and the OSB Test Scores (Criterion B)

		OSB Test Scores (B)
The Number of OSB Training Times	Pearson Correlation	.155
	Sig. (2-tailed)	.414

Note. $N = 30$

Table 4 shows the correlation between those, but criterion B was used. The average of the increase-and-decrease in the test scores is 1.07, and its SD is 1.999. This data reports no correlation and significant difference between these two variables. From Tables 3 and 4, it is possible to imagine that the participants who practiced a lot became able to pay more attention

to function words such as prepositions and articles without concentrating excessively on rearranging the chunks.

Table 5

The Correlation between the Number of OSB Training Times and Increases and Decreases in Token and Type

		Token	Type
The Number of OSB	Pearson Correlation	-.130	-.176
Training Times	Sig. (2-tailed)	.492	.352

Note. $N = 30$

A statistical analysis was executed to see if there was a correlation between the number of OSB training times and the PDT results. The recorded voice data was transcribed, and increase-and-decrease in the number of words (token) and kinds of words (type) was counted. Table 5 shows the result of analysis. There is no correlation between the number of OSB training times and increase-and-decrease in token or type. Irrespective of whether the participants actually practiced, different words are often used in the case that more words are uttered, just as the data $r = .841$ ($df = 28$, $p < .001$) indicates.

Table 6

The Correlation between the Number of OSB Training Times and the Percentage of Uttered Words Divided by JACET List of 8000 Basic Words

		LV 1	LV 4	Others
The Number of OSB	Pearson Correlation	.323	-.384*	-.321
Training Times	Sig. (2-tailed)	.082	.036	.084

Note. $N = 30$

* Correlation is significant at the 0.05 level (2-tailed).

More specific investigation into variation in vocabulary use seemed necessary, so the transcribed data was reanalyzed in reference to JACET List of 8000 Basic Words (JACET, 2003). The list is sorted by the frequency of words, composed of eight levels. Here, only remarkable variables are focused on. Table 6 shows the correlations between the number of OSB training times and the level 1, 4 and Others in the list. The label Others includes numbers such as one and two and words that do not belong to any level. As the data $r = .323$ ($df = 28$, $p = .82$), the frequency of the words in Level 1 might have increased marginally, even though there is no obvious significant difference. In contrast, there is a significant difference in the correlation between the number of OSB practice times and the percentage of the uttered words classified into Level 4. In fact, this does not denote a decrease in the words in Level 4 because almost none of the participants

used any words in Level 4 either in the pretest or posttest. Lastly, there is no significant difference in the correlation between the number of OSB training times and the percentage of the uttered words contained in Others; however, it is meaningful to note that some of the active participants might have used pronouns or other words instead of numbers because some tended to repeat the same expressions such as "one man" and "two boys." Therefore, continual OSB training may not have a great influence on your word-use directly, but there seems to be some impact on the increase-and-decease in the number of words used, so follow-up research is naturally desirable.

The most outstanding analysis in this study will be the correlation between the number of OSB training times and the decrease in dysfluency in the PDT. Table 7 shows the correlation with the data r = -.402 (df = 28, p = .028). The average of the increase-and-decrease in dysfluency is -1.90, and its SD is 4.459. The data demonstrates that the number of dysfluency errors would taper off with more OSB training. Especially, repetition seems to have diminished.

Table 7

The Correlation between the Number of OSB Training Times and the Increase-and-Decrease in Dysfluency Errors

		Dysfluency Errors
The Number of OSB Training Times	Pearson Correlation	-.402*
	Sig. (2-tailed)	.028

Note. $N = 30$

* Correlation is significant at the 0.05 level (2-tailed).

4.2 Statistical Analysis Based on the RST and LST Results

Aside from the direct possible effects of the OSB training, how does WM span affect the test results? In this study, both RST and LST were conducted. Table 8 shows a correlation between the RST or LST scores and the OSB test scores, which indicates a very strong correlation between the scores on both of the tests. Those who gained high scores on one span test also received high scores on the other span test. Does this mean that the scores of each span test will relate equivalently to the results of the OSB test and the PDT?

Even though there are no correlations between either the RST scores or LST scores and the increase-and-decrease in the OSB scores with both of the criteria, several significant differences can be seen in the variables, the pretest and posttest of OSB.

Table 8

The Correlation between the RST or LST Scores and the OSB Test Scores

		RST	LST	PRE (A)	POST (A)	Dif. (A)	PRE (B)	POST (B)	Dif. (B)
RST	Pearson Correlation	1	.623**	.382*	.356	-.020	.223	.134	-.065
	Sig. (2-tailed)		.000	.037	.054	.916	.237	.479	.732
LST	Pearson Correlation	.623**	1	.653**	.602**	-.042	.503**	.351	-.108
	Sig. (2-tailed)	.000		.000	.000	.826	.005	.057	.572

Note. $N = 30$

* Correlation is significant at the 0.05 level (2-tailed).

** Correlation is significant at the 0.01 level (2-tailed).

Compared with the data of RST, it is obvious that the correlations with LST are stronger, as the analysis reveals that the correlations between the LST scores and the pretest (Criterion A), the posttest (Criterion A) or the pretest (Criterion B) are $r = .653$ ($df = 28$, $p < .001$), $r = .602$ ($df = 28$, $p < .001$), and $r = .503$ ($df = 28$, $p = .005$) respectively. This is presumably because the ability required for the OSB tests is more similar to that for LST, that is, the listening skills. Therefore, it can be inferred from this data that those who have a larger WM capacity are more likely to be able to perform well on OSB tests.

Moreover, the participants who gained high RST or LST scores uttered more words and used a wider range of vocabulary. Table 9

illustrates that there are strong correlations between the RST or LST scores and the number of uttered words or kinds of uttered words. The average of the scores of the RST (pretest) is .87, and its SD is 1.121. The average of the scores of the LST (pretest) is .60, and its SD is .995. The average of token (pretest, dysfluency errors omitted) is 42.43, and its SD is 13.523. The average of type (pretest) is 28.97, and its SD is 7.686. Thus, those with a larger listening span are slightly more likely to utter more words and use a variety of words.

Table 9

The Correlations between the Scores of RST or LST and Token or Type (Pretest)

		RST	LST	Token (PRE)	Type (PRE)
RST	Pearson Correlation	1	.623**	.418*	.464**
	Sig. (2-tailed)		.000	.022	.010
LST	Pearson Correlation	.623**	1	.525**	.562**
	Sig. (2-tailed)	.000		.003	.001

Note. $N = 30$

* Correlation is significant at the 0.05 level (2-tailed).

** Correlation is significant at the 0.01 level (2-tailed).

4.3 Findings from the Interview

For a deeper analysis, two participants, who were arbitrarily chosen, were interviewed by the experimenter. Their answers to the questions will contribute to the revelations about the effects of the OSB training. Here, a discussion will follow for each question.

1. What do you think the possible effects of continual OSB training would be?

They both have a common view that continual OSB training notably improved their English speaking proficiency on a syntactical level. Moreover, the interviewees felt less hesitant to speak or became able to take less time to translate Japanese into English, which is not limited to speaking but actually applies to writing. Additionally, they admitted that the training contributes to the development of their listening skills because they listened actively to English at least three times a week. As to reading ability, there is a possibility that learners can process long sentences more quickly because of the improvement in their syntactical ability.

2. Could the possible effects gained through the OSB training be assessed by the tests (OSB test and PDT)?

Their answer to the question was yes. The OSB tests are identical to what they did during the training period, and the reliability between them would be very high. One acknowledged that the training wielded a positive effect on one's syntactical ability, which caused the interviewee to feel development of the test scores from the pretest to the posttest. Furthermore, the other interviewee felt more confident in speaking English in the posttest of picture description as long as appropriate words to be uttered came up because the tests assessed comprehensive English speaking ability, especially vocabulary use and grammatical knowledge. This kind of grammatical use is intimately related to the OSB training.

3. What other English learning did you do during the period of the OSB training?

One of the interviewees did not do any English learning particularly other than the OSB training. The other tried to expand vocabulary using a wordbook for TOEIC and sometimes read online English newspapers.

Therefore, they did not study or review grammar. Above all, they did not have sufficient opportunities to speak English with someone else.

5. Conclusion

This study conducted a practical evaluation on the effects of the OSB training. Thirty participants tried to practice speaking with provided material as many times as possible. This paper concludes by attempting to discover the possible effects of the OSB training in terms of the findings from statistical analyses and the interview, that is, by answering the research questions.

◘ Possible Improvements in Grammatical Accuracy in Speaking

In this study, the possibility that continual OSB training can help you to improve your grammatical accuracy was found. Even though no assured statistical evidence was found to show that the training could make progress in your syntactical ability, the interview results support that it may be developed. In addition, the participants' scores on the OSB tests significantly increased, most likely on account of the fact that many of them could more clearly focus on function words including prepositions and articles.

□ Possible Improvements in Fluency in Speaking

Continual OSB training did not contribute to the improvement in fluency in speaking as the statistics shows, although several participants uttered more words in the pretest of picture description. Nevertheless, a negative correlation between continual OSB training and the number of dysfluency errors was found. Therefore, the training does not positively affect fluency but does improve accuracy. The interview results endorse this since the interviewees realized that they improved their syntactical ability or became less likely to translate Japanese into English in mind.

□ Relations between WM Capacity and the Test Results

The statistical analysis indicates that the scores of RST and LST are interrelated with each other. The further analysis reveals that LST is more closely related to OSB than RST, supposedly because each demands listening attention. Thus, those with a larger WM capacity, especially concerning listening span, are more likely to perform better on OSB tests, paying more attention to function words. In addition, they are more apt to speak more words and use more diverse words.

6. Further Studies

In conducting this experiment and collecting the data, there were several limitations that could hamper effects of the OSB training.

The first limitation is that a total of 400 sentences including 40 grammatical items were created and randomized to be 30 sets of the OSB training, resulting in 300 sentences with mixed grammatical items. The reason this was done is that it was necessary to take into consideration the risk that participants could not complete all the 30 sets of the OSB training. Ideally, one set of training should contain one target grammatical item so that its practice can lead learners to mastery more efficiently by the augmentation of exposure to it at one time. Based on this idea, it might have been better to reduce the number of target grammatical items for more effective learning, particularly in this short period of training, as long as this limitation exists.

Another limitation is that in the present study a reply to the email with a YouTube URL was regarded as achievement of the training, but there was no other better way of checking whether or not each participant actually did the OSB training. It was impossible to meet all the participants three

times a week. Therefore, there were several participants who did not try the OSB training even once, although this had been anticipated. Besides, this experiment did not have any compelling force toward the participants because it was not a part of a class, which would have contributed to the decrease in their extrinsic motivation and the number of training times.

Moreover, there was a concern about how to provide the participants with feedback after each training session. An example answer to each question was played immediately after answering time so as to present quick feedback; however, the participants were just instructed to do the training once per set. More effects of the OSB training could have appeared if they were instructed to solve questions that they had not been able to for the first time as many times as necessary.

For future research, participants should be divided into a control group and an experimental group. By providing them with equal opportunities to practice English through the OSB training, more substantial results would appear. Additionally, a larger number of participants are necessary to see any tendency of the effects of this training.

Bibliography

Alptekin, C., & Ercetin, G. (2011). Effects of working memory capacity and content familiarity on literal and inferential comprehension in L2 reading. *TESOL Quarterly*, *45*, 2, 235-266.

Atkinson, R. C., & Shiffrin, R. N. (1971). The control processes of short-term memory." *Technical Report 173*. Psychology Series. Institute for Mathematical Studies in the Social Sciences, Stanford University.

Baddeley, A. (1992). Working Memory. *Science*, *255*, 556-559.

Baddeley, A.D. (2002). Is working memory still working? *European Psychologist*, *7*, 2, 85-97.

Baddeley, A. (2003). Working Memory: Looking back and looking forward. *Nature Reviews*: *Neuroscience*, *4*, 829-839.

Baddeley, A. D., & Hitch, G. J. (1974). Working memory. In G.A. Bower (Ed.), *Recent advances in learning and motivation, 8*, 47-90. New York: Academic Press.

Baddeley, A., Thomson, N., & Buchanan, M. (1975). Word length and the structure of short-term memory. *Journal of Verbal Learning and Verbal Behavior*, *14*, 575-589.

Beardsley, R. (1997). The machinery of thought. *Scientific American*, August issue, 78-83.

Brown, J.D. (1993). Language test hysteria in Japan? *The Language Teacher*, *17*, 12, 41-43.

Cowan, N. (2000). The magical number 4 in short-term memory: A reconsideration of mental storage capacity. *Behavioral Brain Science*, *24*, 87-185.

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

Daneman, M., & Carpenter, P.A. (1980). Individual differences in working memory and reading. *Journal of Verbal Learning and Verbal Behavior*, *19*, 450-466.

Ellis, N.C., & Hennelly, R.A. (1980). A bilingual word-length effect: Implications for intelligence testing and the relative ease of mental calculation in Welsh and English. *British Journal of Psychology*, *71*, 43-51.

ETS. (2007). Test and Score Data Summary for TOEFL Internet-based Test. Retrieved October 26, 2012, from http://www.ets.org/Media/Research/pdf/TOEFL-SUM-0506-iBT.pdf

ETS. (2008). Test and Score Data Summary for TOEFL Internet-based and Paper-based Tests. Retrieved October 26, 2012, from http://www.ets.org/Media/Research/pdf/71943_web.pdf

ETS. (2009). Test and Score Data Summary for TOEFL Internet-based and Paper-based Tests. Retrieved October 26, 2012, from http://www.ets.org/Media/Tests/TOEFL/pdf/test_score_data_summary_2008.pdf

ETS. (2010). Test and Score Data Summary for TOEFL Internet-based and Paper-based Tests. Retrieved October 26, 2012, from http://www.ets.org/Media/Tests/TOEFL/pdf/test_score_data_summary_2009.pdf

ETS. (2011). Test and Score Data Summary for TOEFL Internet-based and Paper-based Tests. Retrieved October 26, 2012, from http://www.ets.org/Media/Research/pdf/TOEFL-SUM-2010.pdf

ETS. (2012a). Test and Score Data Summary for TOEFL Internet-based and Paper-based Tests. Retrieved October 26, 2012, from http://www.ets.org/s/toefl/pdf/94227_unlf

ETS. (2012b). TOEFL Score Scales. Retrieved October 26, 2012, from http://www.ets.org/toefl/ibt/scores/understand

Foreign Service Institute. (1973). Expected Levels of Absolute Speaking Proficiency in Language Taught at the Foreign Institute. Retrieved October 26, 2012, from http://www.asdk12.org/depts/world_lang/advocacy/ForeignServiceInst.pdf

Heaton, J.B. (1975). *Beginning composition through pictures*. Essex: Longman.

Hiramatsu S. (1999). Inferencing tasks as a process-oriented approach to second language acquisition. *Chugoku Tanki Daigaku Kiyou*, *30*, 173-184.

Lambert, S. (1988). A human information processing and cognitive approach to the training of simultaneous interpreters. In Hammond, D. L. (Ed.), *Languages at Crossroads, Proceedings of the 29th Annual Conference of the American Translators Association*. 379-388.

McLaughlin, B. (1987). *Theories of Second-language Learning*. London: Edward Arnold.

McLaughlin, B. (1990). Restructuring. *Applied Linguistics*, *11*, 2, 113-128.

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.

National Institute on Deafness and Other Communication Disorders (NIDCD). (2010). Aphasia. Retrieved November 7, 2012, from http://www.nidcd.nih.gov/staticresources/health/voice/FactSheetAphasia.pdf

Pearson Kirihara. (n.d.). The test sample of Versant. Retrieved November 7, 2012, from http://www.versant.jp/index.html

Peterson, L. R., & Peterson, M. J. (1959). Short-term retention of individual verbal items. *Journal of Experimental Psychology*, *58*, 193-198.

Repovs, G., & Baddeley, A. (2006). The multi-component model of working memory: Explorations in experimental cognitive psychology. *Neuroscience*, *139*, 5-21.

Shiki, O., Mori, Y., Kadota, S., & Yoshida, S. (2010). Exploring differences between shadowing and repeating practices: An analysis of reproduction rate and types of reproduced words. *ARELE*, *21*, 81-90.

江口誠. (2007).「CALL を活用したシャドーイングの実践」『呉工業高等専門学校研究報告』, 第 69 号, pp.25-31.

大高博美. (1998). 「英語のスピーキング指導―理論と実践―」『言語と文化』関西学院大学言語教育研究センター紀要委員会［編］.

苧阪直行 (2000).「ワーキングメモリと意識」『脳とワーキングメモリ』苧阪直行 (編), 1-18. 京都：京都大学学術出版会.

苧阪満里子. (2002). 「脳のメモ帳ワーキングメモリ」東京: 新曜社.

門田修平. (2007). 「シャドーイングと音読の科学」東京: コスモピア.

白畑知彦, 冨田祐一, 村野井仁, 若林茂則. (2009). 『改訂版英語教育用語辞典』東京: 大修館書店.

白石典義. (2010). 「企業が求める英語力調査」
http://gp-portal.jp/material/refMaterial/769_%E8%8B%B1%E8%AA%9E%E5%8A%9B%E5%A0%B1%E5%91%8A%E8%B3%87%E6%96%99.pdf

大学英語教育学会(JACET). (2003).『大学英語教育学会基本語リスト JACET8000』. 大学英語教育学会(JACET)基本語改訂委員会(編).

高梨庸雄, 高橋正夫. (2007).「新・英語教育学概論」東京: 金星堂.

田中誠. (2007).「授業改善のためのアクション・リサーチ」『長崎国際大学論叢』, 第 7 巻, pp.105-113.

玉井健. (2002).「リスニング力向上におけるシャドーイングの効果について」日本通訳学会第 3 回年次大会講演. 9 月 23 日.

鳥飼玖美子. (2002). 『TOEFL・TOEIC と日本人の英語力』東京: 講談社.

中鉢惠一. (2000). 「日本人に必要な英語力とは」『経営論集』, 第 52 号, pp.99-107.

日本経済研究センター. (2011). 「英語力のビハインドが招く国際競争力の低下」
http://www.jcer.or.jp/report/econ100/pdf/econ100bangai20110711data.pdf

根岸雅史. (2007).「スピーキング・テストをどう作る」『TEACHING ENGLISH NOW』, 8 巻.

バトラー後藤裕子. (2005). 「日本の小学校英語を考える―アジアの視点からの検証と提言」東京: 三省堂.

文部科学省(MEXT). (2003).『「英語が使える日本人」の育成のための行動計画』
http://warp.ndl.go.jp/info:ndljp/pid/286794/www.mext.go.jp/b_menu/houdou/15/03/03033102.pdf

文部科学省(MEXT). (2008a). 『小学校学習指導要領案』

文部科学省(MEXT). (2008b). 『中学校学習指導要領』

文部科学省(MEXT). (2009). 『高等学校学習指導要領』

文部科学省(MEXT). (2010a). 『Chapter 4 Foreign Language Activities』
http://www.mext.go.jp/component/a_menu/education/micro_detail/__icsFiles/afieldfile/2010/10/20/1261037_12.pdf

文部科学省(MEXT). (2010b). 『Section 9 Foreign Languages』
http://www.mext.go.jp/component/a_menu/education/micro_detail/__icsFiles/afieldfile/2011/04/11/1298356_10.pdf

文部科学省(MEXT). (2010c). 『Section 8 Foreign Languages』
http://www.mext.go.jp/a_menu/shotou/new-cs/youryou/eiyaku/__icsFiles/afieldfile/2011/04/11/1298353_9.pdf

萬戸克憲. (1995). 「日本人学生の英語力―TOEFL による国際比較」『和歌山大学教育学部教育実践研究指導センター紀要』, 5 号, pp.165-171.

吉田晴世. (2012). 「『スピーキング力』を磨く: Versant テストを利用して」『英語教育』9 月号, 61 巻, 6, pp.28, 29.

Appendix I. Questions for Reading Span Test (Pretest)

Examples

We are going to study abroad in Canada this month.

My friend don't like to eat tomatoes, broccoli, and onions.

Two-Sentence Level

Set 1

She has never been to America with her family.

Ken and Mary are playing soccer together yesterday.

Set 2

They will be able to came here by noon.

Students have to study a lot of subjects every night.

Set 3

Yesterday the trains stopped because of the heavy rain.

You are really supposed to be on time then.

Set 4

It is interesting for he to climb mountains.

She has many friends to talk at her office.

Three-Sentence Level

Set 1

It is very cold, so it is going to snow.

We had not better get angry at the person.

The gentleman did not know what to talk about.

Set 2

You don't remember the day when we first met.

If I were you, I wouldn't do such as thing.

That is the girl which I saw at the shop.

Set 3

The population of Japan is larger than Korea.

She left the kitchen with the water running.

I am really looking forward to see you again.

Set 4

Now we are going to discuss about this issue.

You should have come to the meeting earlier.

He suggested that you should take a medicine.

Four-Sentence Level

Set 1

My parents didn't allow me to buy a new car.

She is the kindest lady in her class.

Our teacher made us to study harder than before.

My mother was very kind to cook me lunch.

Set 2

My brother is as tall as my father.

I am not sure if he is healthy or not.

We know that there are no rule without exceptions.

I think that your father is very proud of you.

Set 3

Please take a look at this article of the newspaper.

Each students have to bring their own notebooks.

The nurse is busy taking care of her patients.

His Chinese was too difficult for me to understand.

Set 4

He finally decided to get married to his girlfriend.

We have just finished to read the long passage.

Generally speaking, mastering a language is quite difficult.

I would you like to lend me the new book.

Five-Sentence Level

Set 1

The boy likes to listen music very much.

He has a friend whose father is a cook.

Anyone will not go swimming in the small pool.

I found it difficult to solve the problem.

I heard that both of your sisters was very smart.

Set 2

That old lady was spoken by a stranger.

The novelist is going to quit his job.

He enjoys playing the guitar in his band.

Water pollution has been a serious problem for ages.

I want to drink something to hot drink.

Set 3

He will soon get used to the new place.

When young, I often go out to drink alcohol.

I am afraid he will not come to the party.

It is very kind for you to carry my bag.

It seems that this printer is out of ink.

Set 4

The computer is the one my father gave me.

The story of the book is very attractive.

She is having many pens in her pencil case.

I am very exciting because the movie looks fun.

Not only my friend but also I are studying abroad.

Appendix II. Questions for Listening Span Test (Pretest)

Examples

The girl running over there is my sister.

That window has to leave open all the time.

Two-Sentence Level

Set 1

The magazine is reading by a lot of people.

This novel is worth reading many times.

Set 2

She regrets going to the crowded place.

My father has already had lunch.

Set 3

This town is where we first met.

Please give me warm something to wear on.

Set 4

It is unnatural of him to say so.

I am listening to the radio at that time.

Three-Sentence Level

Set 1

Her opinion made me really angry.

My mother has to take care of herself.

The lady was spoken by a strange man.

Set 2

Her story might not be true.

I feel like to sing a song loud.

There used to be a bank near here.

Set 3

Having more free time is almost impossible.

Nobody could understand his question.

I have just finished to do my homework.

Set 4

My friend arrived the airport.

This house was built by my cousin.

This is how he got to the conclusion.

Four-Sentence Level

Set 1

He is the person which I like the best.

Nothing is more important than your health.

I decided not to meet him again.

It is very kind of her to help that person.

Set 2

He is thinking about his future.

What is important is being honest.

I should have gone to the party.

She has to finish her work until tomorrow.

Set 3

The girl don't know how to cook.

We have to discuss about his problem.

Some people are against the new law.

My sister is two centimeters taller than me.

Set 4

The student didn't bring his notebook.

Our office is standing near the station.

She is busy when her husband came home.

I have never watched that movie.

Five-Sentence Level

Set 1

He is a famous scientist for his study.

This computer is superior than that one.

My father likes to eat boiling eggs.

A cat is lying on the bed.

Japanese is not speaking in many countries.

Set 2

It was careless of you to leave your bag.

Please remember to post this letter.

They boy looks very sleepness.

Playing sports are a lot of fun.

I think her as a nice lady.

Set 3

He might have told a lie.

My teacher let me to take a break.

He likes playing both tennis or baseball.

I don’t like the class because of it is boring.

The man is familiar with the town.

Set 4

She is said that she is rich.

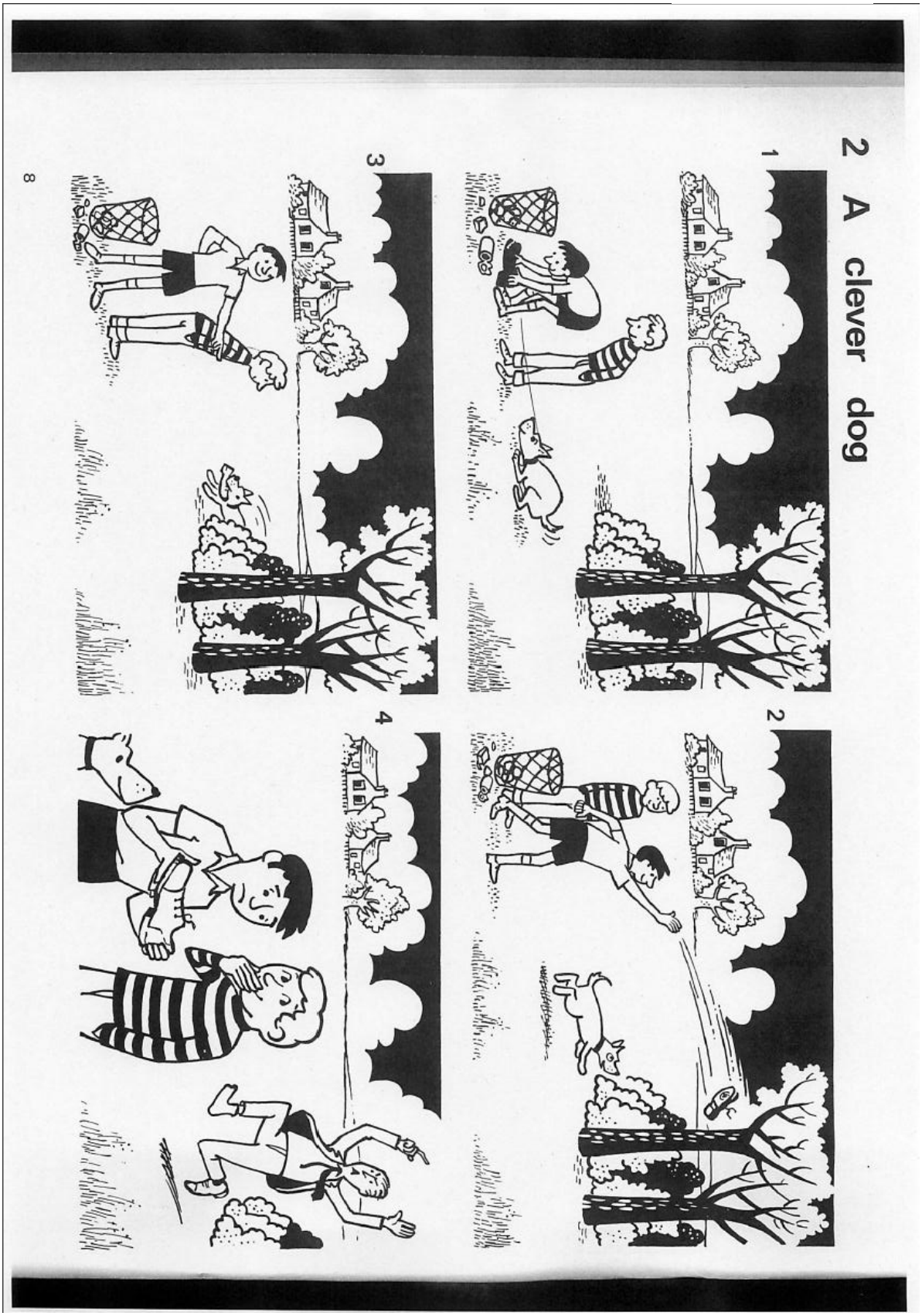
This desk is made of wood.

I didn’t know what to say.

You have to drive safety.

Her mother is going to the school.

Appendix III. Picture Description Test (Pretest) (Heaton, 1975)



Appendix IV. Questions for Oral Sentence Building Test (Pretest)

Original Sentences

1. I expected him / to carry / those heavy books.
2. He was / surprised at / the cat.
3. It is / kind of you / to help me.
4. Studying hard / is often / very difficult.
5. That man / should have kept / the secret.
6. These are / the novels / that he wrote.
7. Are you / going to / bring your friends?
8. This is / the company / where my brother works.
9. There was / a big supermarket / near my house.
10. I am / looking forward to / meeting you.

Actual Questions

1. to carry / those heavy books / I expected him
2. he was / the cat / surprised at
3. to help me / it is / kind of you
4. is often / very difficult / studying hard
5. that man / the secret / should have kept
6. that he wrote / the novels / these are
7. are you / bring your friends / going to
8. the company / this is / where my brother works.
9. there was / near my house / a big supermarket
10. meeting you / looking forward to / I am

Appendix V. Questions for Reading Span Test (Posttest)

Examples

We are going to study abroad in Canada this month.

My friend don't like to eat tomatoes, broccoli, and onions.

Two-Sentence Level

Set 1

That is the girl which I saw at the shop.

She is the kindest lady in her class.

Set 2

I heard that both of your sisters was very smart.

He suggested that you should take a medicine.

Set 3

I am afraid he will not come to the party.

Not only my friend but also I are studying abroad.

Set 4

Ken and Mary are playing soccer together yesterday.

We had not better get angry at the person.

Three-Sentence Level

Set 1

I think that your father is very proud of you.

I found it difficult to solve the problem.

We have just finished to read the long passage.

Set 2

I am really looking forward to see you again.

The story of the book is very attractive.

His Chinese was too difficult for me to understand.

Set 3

The novelist is going to quit his job.

It is very kind for you to carry my bag.

She is having many pens in her pencil case.

Set 4

She has never been to America with her family.

You are really supposed to be on time then.

If I were you, I wouldn't do such as thing.

Four-Sentence Level

Set 1

My mother was very kind to cook me lunch.

The nurse is busy taking care of her patients.

I want to drink something to hot drink.

You should have come to the meeting earlier.

Set 2

They will be able to came here by noon.

You don't remember the day when we first met.

We know that there are no rule without exceptions.

He finally decided to get married to his girlfriend.

Set 3

Water pollution has been a serious problem for ages.

The computer is the one my father gave me.

He has a friend whose father is a cook.

Each students have to bring their own notebooks.

Set 4

Please take a look at this article of the newspaper.

My parents didn't allow me to buy a new car.

Anyone will not go swimming in the small pool.

She has many friends to talk at her office.

Five-Sentence Level

Set 1

It is interesting for he to climb mountains.

Our teacher made us to study harder than before.

I am not sure if he is healthy or not.

I would you like to lend me the new book.

I am very exciting because the movie looks fun.

Set 2

When young, I often go out to drink alcohol.

Generally speaking, mastering a language is quite difficult.

He enjoys playing the guitar in his band.

The gentleman did not know what to talk about.

Yesterday the trains stopped because of the heavy rain.

Set 3

She left the kitchen with the water running.

The boy likes to listen music very much.

It seems that this printer is out of ink.

The population of Japan is larger than Korea.

Students have to study a lot of subjects every night.

Set 4

It is very cold, so it is going to snow.

He will soon get used to the new place.

That old lady was spoken by a stranger.

My brother is as tall as my father.

Now we are going to discuss about this issue.

Appendix VI. Questions for Listening Span Test (Posttest)

Examples

The girl running over there is my sister.

That window has to leave open all the time.

Two-Sentence Level

Set 1

Her opinion made me really angry.

My friend arrived the airport.

Set 2

The girl don't know how to cook.

Playing sports are a lot of fun.

Set 3

This desk is made of wood.

A cat is lying on the bed.

Set 4

Having more free time is almost impossible.

I am listening to the radio at that time.

Three-Sentence Level

Set 1

There used to be a bank near here.

She has to finish her work until tomorrow.

I should have gone to the party.

Set 2

She is busy when her husband came home.

The man is familiar with the town.

Please remember to post this letter.

Set 3

They boy looks very sleepness.

He likes playing both tennis or baseball.

I think her as a nice lady.

Set 4

It was careless of you to leave your bag.

I didn't know what to say.

You have to drive safety.

Four-Sentence Level

Set 1

My teacher let me to take a break.

Her mother is going to the school.

Some people are against the new law.

Our office is standing near the station.

Set 2

I have never watched that movie.

My father likes to eat boiling eggs.

Japanese is not speaking in many countries.

She is said that she is rich.

Set 3

This is how he got to the conclusion.

I have just finished to do my homework.

This house was built by my cousin.

I decided not to meet him again.

Set 4

It is very kind of her to help that person.

My mother has to take care of herself.

Her story might not be true.

The magazine is reading by a lot of people.

Five-Sentence Level

Set 1

My father has already had lunch.

This town is where we first met.

It is unnatural of him to say so.

What is important is being honest.

The student didn't bring his notebook.

Set 2

This computer is superior than that one.

He is a famous scientist for his study.

Nothing is more important than your health.

My sister is two centimeters taller than me.

The lady was spoken by a strange man.

Set 3

He is thinking about his future.

Please give me warm something to wear on.

This novel is worth reading many times.

I feel like to sing a song loud.

Nobody could understand his question.

Set 4

I don't like the class because of it is boring.

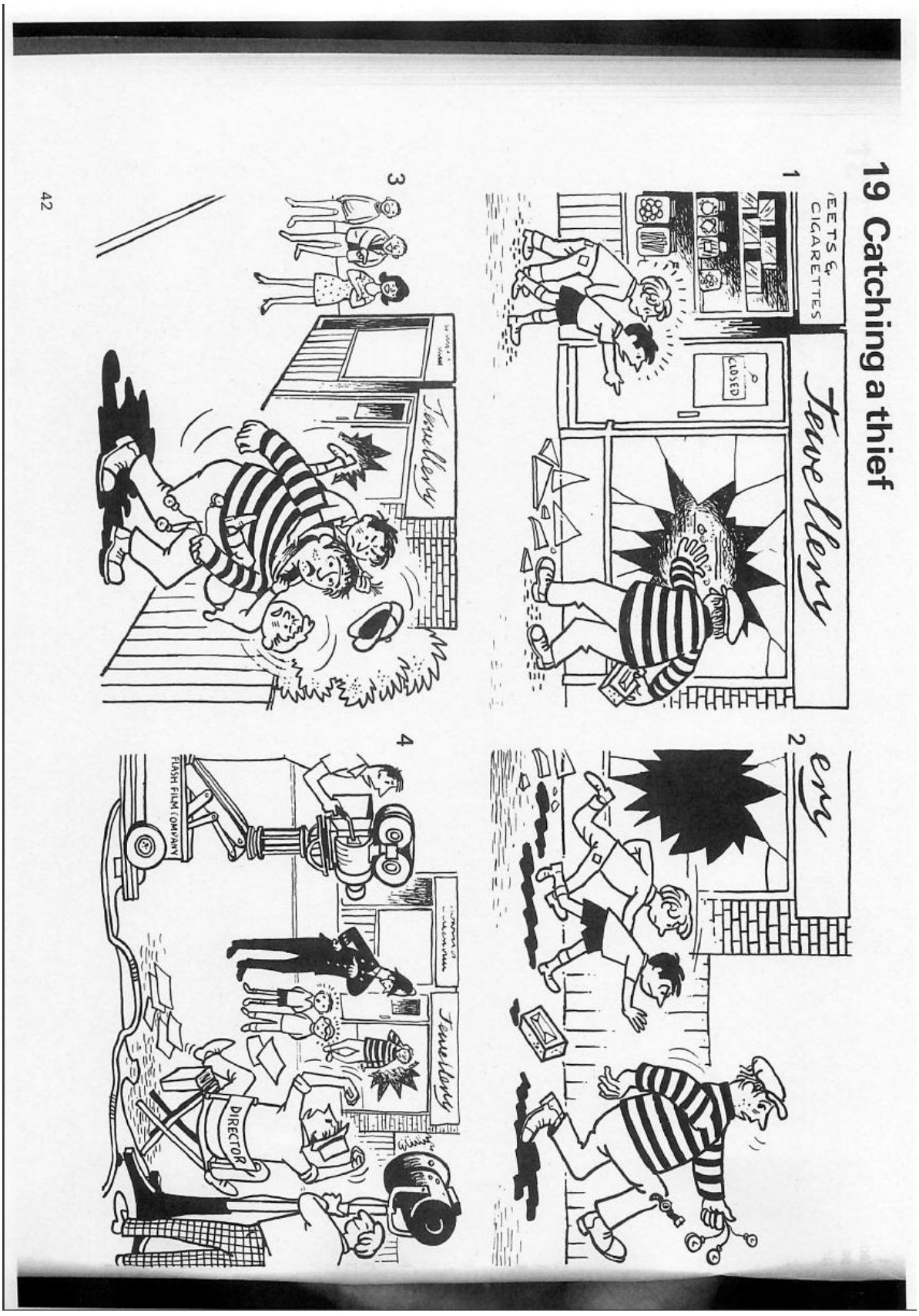
We have to discuss about his problem.

He might have told a lie.

She regrets going to the crowded place.

He is the person which I like the best.

Appendix VII. Picture Description Test (Posttest) (Heaton, 1975)



Appendix VIII. Questions for Oral Sentence Building Test (Posttest)

Original Sentences

1. Are you / going to / bring your friends?
2. I am / looking forward to / meeting you.
3. It is / kind of you / to help me.
4. He was / surprised at / the cat.
5. I expected him / to carry / those heavy books.
6. This is / the company / where my brother works.
7. That man / should have kept / the secret.
8. These are / the novels / that he wrote.
9. There was / a big supermarket / near my house.
10. Studying hard / is often / very difficult.

Actual Questions

1. are you / bring your friends / going to
2. meeting you / looking forward to / I am
3. to help me / it is / kind of you
4. he was / the cat / surprised at
5. to carry / those heavy books / I expected him
6. the company / this is / where my brother works.
7. that he wrote / the novels / these are
8. that man / the secret / should have kept
9. there was / near my house / a big supermarket
10. is often / very difficult / studying hard