Conceptual and lexical buildup in strategic oral production

Chiaki Iwai
Hiroshima City University

Reference data:

Iwai, C. (2009). Conceptual and lexical buildup in strategic oral production. In A. M. Stoke (Ed.), *JALT2008 Conference Proceedings*. Tokyo: JALT.

Several empirical studies about teaching communication strategies (CSs) have shown that teaching CS improves learners' communicative competence; however, the mechanisms for this have not been clarified adequately. Based on Levelt's (1989) psycholinguistic model, an attempt was made in this study to examine them by narrowing its focus only on a Conceptualizer component in his model, while controlling another component, Formulator. Based on the results obtained by analyzing empirical data collected from 39 participants, of whom 19 took part in a 5-week lexical CS training program, the study discusses that teaching CS possibly facilitates learners' processing in the Conceptualizer component by changing their linguistic knowledge of the target language to procedural knowledge.

コミュニケーション方略（CS）の指導が学習者の発話能力を助長することは多くの先行研究によって示されているが、その理由については十分に解明されているとは言いがたい。本研究は　Levelt　(1989)　の心理言語学的発話モデルに基づいて、このモデルで示された概念形成処理過程（Conceptualizer）と文法的変換処理過程　(Formulator)　のうち、前者のみに焦点をあてて実証研究を行ったものである。研究は39名から集めたデータに基づいており、このうち19名の実験群学習者に対しては5週間の語彙に関するCS指導を行った。収集データ分析の結果に基づいて、CS指導は学習者に内在化された言語知識を手続き知識（procedural knowledge）に変えることに貢献する可能性のあることを論じている。

Among various types of strategies used by second language (L2) learners, the present study deals with communication strategies (CSs). The theoretical growth of CSs in the 1980s and 1990s has generated two contrastive approaches (see below) and raised the controversial issue of teaching CSs to L2 learners. The issue, which is now referred to as *CS teachability*, has been challenged empirically by many researchers, especially after the value of CS instruction was straightforwardly rejected by such scholars as Bialystok (1990) and Kellerman (1991). Despite the ample positive empirical evidence that has been found against those researchers' rejections, the exact mechanisms for desirable effects, psycholinguistic mechanisms in particular, have not been adequately clarified yet. Thus, the main objective of this study is to seek possible psycholinguistic accounts for positive effects obtained by teaching CSs.

Literature review

Two major roles have been focused on in CS studies, namely, a *compensatory* role to fill the gap between what speakers can say and what they want to say and a *facilitating* role in their interactive negotiation of meaning. The former has yielded an intrapersonal, psycholinguistic approach, and the latter an interpersonal, sociolinguistic approach (Kasper & Kellerman, 1997). Beyond the theoretical argument on their importance or relevancy (e.g., Yule & Tarone, 1997; Kellerman, 1998), these two approaches are symbolic in CS studies since their theoretical assumptions are directly linked to a CS teachability issue; that is, whether CSs are worth teaching or not, whether CS instruction improves L2 learners' communication skills, and/or whether CS instruction improves their language itself (Dörnyei, 1995).

An outright denial of teaching CS by such psycholinguists as those mentioned above grew into a pro-con debate. Against those cons, the pros countercharged that CS training would facilitate L2 learners' better use of not only strategies but also language itself (Dörnyei, 1995) and that "performance creates competence" (Yule & Tarone, 1997). Researchers in-between the pros and the cons have been a little cautious: Skehan (1998), for example, claims that more direct evidence for interlanguage change and its development over time is needed to justify the value of teaching CSs [for similar views, see Cook (1993), Iwai (2006), and Konishi & Tarone (2004)].

Following these arguments, several researchers have conducted experimental studies to verify effects of teaching CSs. These studies also reflect the two CS perspectives. To mention just a few typical examples, studies based on the psycholinguistic intrapersonal perspective include Dörnyei (1995: about Hungarian EFL learners' use of such strategies as topic replacement, circumlocution and fillers); Kitajima (1997: Japanese college EFL learners' use of message reduction and lexical strategies for unknown words); Salomone and Marsal (1997: American French learners' use of circumlocution); and Iwai (2006: Japanese college EFL learners' use of lexical strategies). Empirical studies from an interpersonal perspective, on the other hand, are in fact not so many, but two of the typical examples are a study by Nakatani (2005: Japanese college EFL learners' interaction in oral communication); and Iwai and Gobel (2004: also Japanese EFL learners' interaction in oral communication).

Putting the findings of these studies together, one plausible generalization is that CS training can improve L2 learners' strategy use in local linguistic problems (e.g., lexical strategies), global communication problems (e.g., message adjustments), and active interactions (e.g., clarification requests). It is, however, by no means clear why such desirable effects are brought about by CS training.

Main purpose of this study

For this reason, the present study was conducted to clarify facilitating mechanisms of CS training. Of various types of CSs, the study narrows its focus only to lexical CSs (or *compensatory strategies* in Poulisse, 1990), since dealing with all different types of strategies in one study could obscure the goal. Lexical CSs are used when a speaker does not know an exact word for a certain target item or concept (e.g., "a talk bird" for "a parrot": Poulisse, 1990).

The theoretical grounding for this study is based on Levelt's (1989) psycholinguistic accounts of speech production and his speech model. According to his model, two psycholinguistic mechanisms are involved in our speech production, represented as *Conceptualizer* and *Formulator*. The former functions to create a pre-verbal message, namely, a non-linguistic concept of what we would like to say. The latter works to encode a pre-verbal message grammatically, and through this encoding processing a pre-verbal message becomes a verbal message before it is actually articulated. To pursue possible psycholinguistic accounts, this study further narrows down its observation to conceptualizing processes. More specifically, it is presupposed in this study that one possible cause for the positive effects could be attributable to improved conceptual processing from the instruction. It is further presupposed that improved conceptual processing could be reflected in a larger size of the mental lexicon that learners can associate with in their strategic attempts at oral production. Based on these, the following research questions are formulated in this study:

(1) Can lexical CS instruction improve conceptual processing of the English L2 learners at a pre-intermediate level?

(2) Can lexical CS instruction facilitate their procedural lexical knowledge of the target language?

(3) If the answer to the second question is positive, does the learners' proficiency determined by a lexical knowledge test affect their improvement?

Methodology

Participants

In total, 39 Japanese college EFL learners, all non-English majors at a pre-intermediate level, participated in this study. They were enrolled in one of the two general English classes set by the university as different goal-oriented classes. One class aimed mainly to develop their expressive skills and this class was designated as an experimental group. The other was a listening class and it was assigned to a control group. They consisted of 19 and 20 students, respectively, and both of them were taught by the author. Due to the difficulty for the author to give a test to measure the participants' proficiency at a university where he taught them part time, the homogeneity between the two groups with respect to English proficiency was not guaranteed, so the collected data will be compared indirectly in the rest of this study. A rough estimate of these students' English proficiency is, however, about 300-450 in TOEIC scores on the basis of the data released by the university. Furthermore, the learners' proficiency levels between the two classes did not differ substantially since they were classified into the same level, as determined by proficiency criteria determined by the university.

Lexical CS instruction

The control group was set only for comparison purposes, so no specific CS instruction was given to the learners in this class. Instead, they only had ordinary listening exercises based on a textbook. In contrast, the learners in the experimental group were exposed to a 5-week lexical CS training program in the first (or spring) semester, 2007. A pre-test in a picture description format (detailed below) was given prior to the training, and a post-test was also given immediately after the training. Additionally, a delayed-post test was given at the end of the semester – about two months after the program.

The CS instruction was based on the intensive program that the author conducted in the past (Iwai, 2006), with picture description activities and information-gap exercises (e.g., spot the differences and story making) given in class. The following summarizes the main activities each week:

Week 1: Picture description task to let the learners recognize what they can't say

Week 2: Explicit CS instruction and exercise of *approximation*

Week 3: Explicit instruction for *paraphrasing*, and introduction of common expression patterns (e.g., use of relative clause structures)

Week 4: Various description-task activities

Week 5: Interaction with classmates (Review of the entire activities)

Data collection

Data collection (pre-, post-, and delayed post-tests) for this study is based on a referential communication (RC) task (Poulisse, 1990) that has been frequently used in past CS studies. One crucial difference from an ordinary RC task is that usually test takers are requested to describe given task items in sentences or phrases orally; in this study, however, they were requested to jot down only words that could be used to describe the given items. This was to elicit learners' mental lexicon associated with the items while freeing them from the load of grammatical encoding, or *Formulator* processes. For this specific test-giving format, the test in this study is referred to as a CS Word Utterance (WU) test. [The WU test appears similar to a word association test (e.g., Nishiyama, 1996); however, they differ from each other crucially in the point that visual images were mainly used in the WU test and that the test takers were requested to assume that they had to explain (or paraphrase) the given images.]

Two language versions of the WU test, English (an EWU test) and Japanese (a JWU test), were produced to examine the participants' performance in these two languages. In addition, three different tests were made for a counterbalance measure among the three test sessions. Of these three tests, the delayed-post test was not given to the control group. This was because, as reported below, no significant change was observed in this intact group from the pre-test to the post-test so that the delayed post-test was judged unnecessary because no change could be predicted. Figure 1 represents their flow

Each language test consisted of six concrete visual images (photos or illustrations; e.g., plumber and cicada), and three abstract nouns (e.g., sympathy and jealousy) that were presented along with a Japanese translation in the EWU tests. Each task item was displayed on a computer monitor for 40 seconds (in reference to Iwai, 2006), and

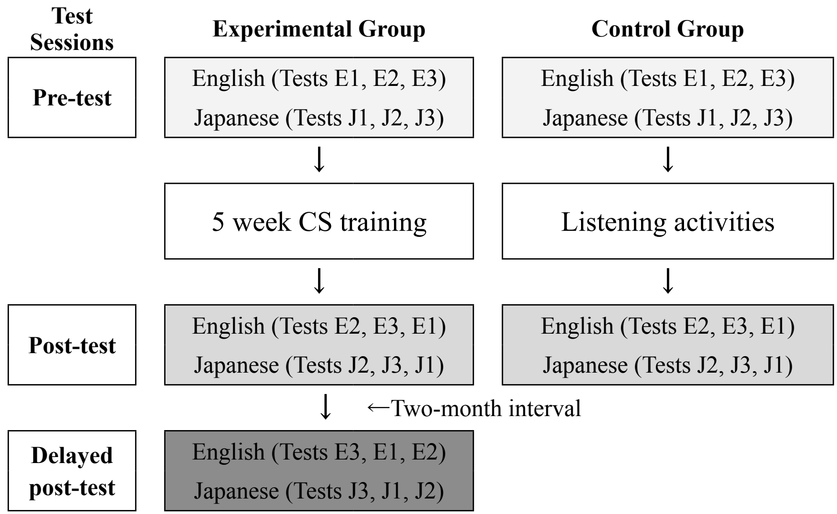

Figure 1. Flow chart of the EWU tests and the JWU tests

the participants were asked to write down all words that "sprang" to their mind to describe the item (see Table 1, Table 2, and Table 3 in Appendix 1 for some sample test items and actual responses to them.) They were also directed, especially in the Japanese test, not to write an exact word even if they knew it since their strategic attempts to describe the items were being tested.

In addition to the WU test, two vocabulary tests in a multiple-choice format were given at the pre-test session to measure the learners' English vocabulary size and depth, both of which were produced on the basis of Mochizuki, Aizawa, and Tono (2003). The former contained 90 questions, and the latter, 25 questions. The correlation coefficient between the results of these tests (by a Pearson's correlation) were significantly high (r = 534, $p < .001$), so only the results of the lexical size test were used in this study.

Analysis method

The lexical data obtained by the WU test can be examined quantitatively and qualitatively; however, due to space restrictions, only the results of quantitative analyses are presented in this study. In analyzing the data, the concrete task items and the abstract ones were treated separately since the learners' test performance differed noticeably between these two categories. Furthermore, longitudinal within-subject comparisons, rather than between-subject comparisons, are made throughout the rest of this study because the two-group homogeneity could not be tested, as mentioned already.

Results

Descriptive statistics of the pre-test

First, the descriptive statistics of the EWU and JWU tests at the pre-test session was obtained to grasp the entire performance tendencies of all participants (*N* = 39) and examine the homogeneity among the three test versions. The following graphs depict them (also see Table 4 and Table 5 in Appendix 2):

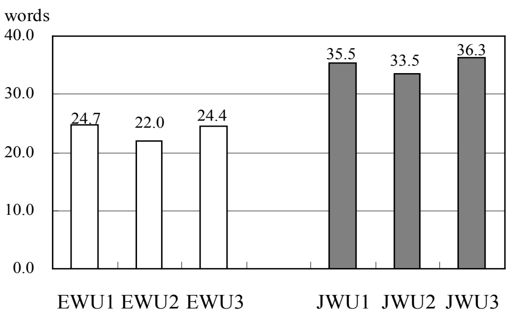

Figure 2. Means of 3 pre-tests (Concrete items)

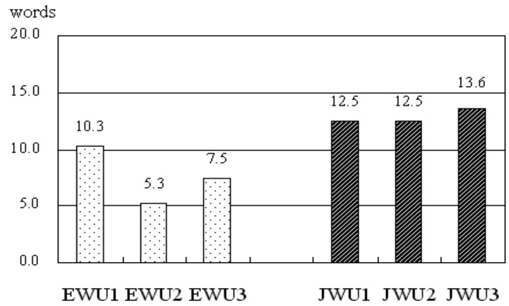

Figure 3. Means of 3 pre-tests (Abstract nouns)

According to Figure 2 for the concrete items, mean word counts for the three EWU test versions ranged from 22.0 to 24.7 (about 3.7 to 4.1 words per task item or w/i), and those for the JWU test, from 33.5 to 36.3 (about 5.6 to 6.1 w/i). The results of a one-way ANOVA revealed no statistical significance either in the EWU [$F(2,36) = .550$, $p = .582$, n.s.] or the JWU [$F(2,36) = .371$, $p = .693$, n.s.]. Figure 3 presents mean word counts of the abstract noun items. Of the three English test versions, the mean word count of EWU Test 2 was 5.3 (or 1.8 w/i), which was significantly lower [$F(2,36) = 5.953$, $p < .01$] than the other two tests, 10.3 (3.4 w/i) and 7.5 (2.5 w/i). In contrast, no such significant difference was yielded among the three JWU tests [means ranged from 13.6(4.5 w/i) to 12.5 (4.2 w/i), and $F(2,36) = .237$, $p = .790$, n.s.]. These results indicate that the WU tests can be regarded as being homogeneous among the three test versions, except for the abstract noun items in the English WU tests. In spite of this non-homogeneity, the three WU tests in both languages are treated without making any distinction from here since, as stated already, they were counterbalanced in the three test sessions so that the influence of within-group differences can be considered minor.

Longitudinal within-group comparison

Next, longitudinal changes were examined in each group, and the results are displayed in Figure 4 (concrete items) and Figure 5 (abstract items) (also see Table 6 and Table 7 in Appendix 2).

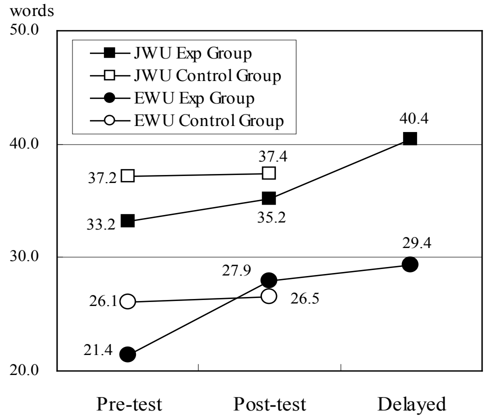

Figure 4. Means of Pre-, Post-, & Delayed post-test (Concrete items)

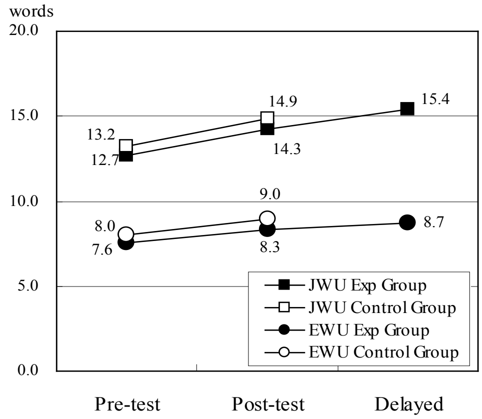

Figure 5. Means of Pre-, Post-, & Delayed post-test (Abstract nouns)

Several prominent features can be identified in the results for the experimental group. First, regarding the concrete items, this group marked, both in English and in Japanese, a steady increase in the mean word counts from the pre-test stage to the post-test stage, and also from the post-test stage to the delayed-post test stage. The significance of this was confirmed by a one-way ANOVA repeated measure (tests repeated at three stages): $F(2, 36) = 11.488$, $p < .001$ for the EWU test, and $F(2, 36) = 6.346$, $p < .01$ for the JWU test. In addition, a post-hoc test (Bonferroni) showed that the development in English was mainly from the pre-test stage to the post-test stage, and that in Japanese was mainly from the post-test stage to the delayed post-test stage. Second, contrary to the concrete items, no substantial changes were found in the abstract noun items in either language, and this was verified by the same statistical test: $F(2, 36) = .784$, $p = .464$, n.s. for the EWU test, and $F(2, 36) = 1.752$, $p = .188$, n.s. for the JWU test. Overall, these results indicate that, as far as the concrete items are concerned, the CS training was highly influential in the learners' English performance and the positive effects were long-lasting; and moreover, the training also brought about positive effects in their Japanese performance, although they were not expected.

These test results were saliently contrastive to those of the control group. That is, the latter group marked no substantial change either in the concrete items or in the abstract noun items, regardless of the language condition, and this was confirmed by a *t*-test (paired-samples from the pre-test to the post test). As a result, the following were obtained: $t = 0.111$, $p = .913$, n.s. in the EWU tests and $t = 0.067$, $p = .948$, n.s. in the JWU tests, both in the concrete items; and $t = 1.160$, $p = .260$, n.s. in the EWU tests and $t = 1.423$, $p = .171$, n.s. in the JWU tests, both in the abstract noun items. No change in the control group indicates that the positive effects of the experimental group were not due to their familiarity with the repeated test measure.

Participants' lexical knowledge and their test performance

In this analysis, the 19 participants in the experimental group were divided into three groups, according to their scores on the lexical size test. Their group means (out of a maximum 90 points) were 75.67 for the highest group (High: $N = 6$), 62.33 for the middle group (Mid: $N = 6$), and 52.9 for the lowest group (Low: $N = 7$), and the significant group difference among them was assured by a one-way ANOVA, $F(2, 18) = 58.745$, $p < .001$ and by a post-hoc multiple-comparison (Bonferroni).

Once these three groups were formed, their performance on the EWU tests was compared in each test session (also see Table 8 and Table 9 in Appendix 2). The statistical outcomes from this analysis tested by a one-way ANOVA were all non-significant either in the concrete items or the abstract noun items at any one of the three test sessions (their details are omitted due to space restrictions). With these combined results, a plausible conclusion is that the learners' lexical knowledge at this pre-intermediate level was not a decisive factor to differentiate their EWU performance, and that the training was equally effective regardless of their possession of English lexical knowledge.

Discussion

The results presented in the preceding section are now interpreted and discussed in terms of Levelt's speed production theory. To test one of the two psycholinguistic processing components theorized by him, namely, the Conceptualizer and Formulator, the data collection of the present study was carried out by liberating the learners from grammatical encoding (Formulator). Consequently, it was found that: (1) the CS training developed the trainees' performance on the concrete items remarkably in both English and Japanese, and these positive effects were not temporary; (2) however, such positive effects were not noticed in their performance on the abstract noun items; and (3) these results in 1 and 2 were unrelated to the participants' relative lexical knowledge of English measured by the test.

A possible account for these outcomes can only be attributed to the processing enhancement in the Conceptualizer component since the processing in Formulator was controlled in this study. This could be further explained as follows. Through the CS training, the participants could expand their flexibility in creating pre-verbal messages, i.e., the mental views or images of what they wanted to say, and the flexibility was reflected in the increase of the number of words associated with the given task items. One unexpected but interesting outcome was that the effects were positively transferred to their Japanese performance although the training was given only for English oral production. This could be explained by the theoretical conceptualization of Cook's (2008) multi-competence model, which predicts the advantage of bilingual minds in language use since such minds allow L2 learners to apply two linguistic capacities to their language processing.

Stating these results and interpretations in different ways, we can now answer the three research questions of this study. Regarding the first one (i.e., whether the lexical CS instruction can improve conceptual processing of the English L2 learners at a pre-intermediate level), the answer is affirmative. Furthermore, the training could enhance their processing in their L1 as well. The answer for the second research question (i.e., whether the lexical CS instruction facilitates the participants' procedural lexical knowledge of the target language) is assumptive since it was not possible to test procedural knowledge directly. However, judging from the fact that no new lexical knowledge was taught intentionally, the answer would be highly probable. In other words, the CS training was effective to change the participants' declarative English lexical knowledge (i.e., knowledge of *what*) into procedural knowledge (i.e., knowledge of *how*). And finally, the answer for the third research question (i.e., whether the learners' differing amounts of lexical knowledge affects the results of the CS training), the answer was absolutely negative. In other words, their relative lexical knowledge at the pre-intermediate level was not a decisive factor for their conceptualizing processes, and these learners apparently could all equally benefit from the CS training.

One implication these results and interpretations is the importance of changing L2 learners' declarative L2 knowledge to procedural knowledge. In many of the author's previous studies and classroom activities in the past, he frequently wondered why the EFL learners could not explain in English even some simple items. There seem to be two reasons for this: one is that they cannot formulate (or encode) what they want to say; and the other is that their mind to create pre-verbal concepts is empty so that they can say very little since "nothing comes from nothing". The findings in this study appear to show that the CS training can facilitate the learners' conceptual processing as reflected on the number of words for conceptual processing, although the study has no intention of arguing that the CS training is the best means to do so.

Finally, attention should be paid to the fact that the positive results from the CS training were limited only to the concrete task items. This is especially important for a pedagogical reason. That is, the results indicate that the concrete items are preferable to use in the CS training with the lower-level learners since the abstract nouns do not appear to guarantee a high return to the learners.

Conclusion

The psycholinguistic mechanisms for the effects of CS instruction were investigated empirically in this study by limiting its scope to a small portion of learners' strategic oral production. The paramount importance of the findings is that CS instruction is highly promising to improve their conceptual processing in the target language as well as in their native language. However, this conclusion is still tentative due to the following three reasons. First, the data obtained from the study were analyzed only quantitatively, and qualitative analyses are unarguably required to clarify through what conceptual processing each respondent derived words from each task item. Second, the test condition of this study (i.e., elimination of formulating processing) was loosely defined; so that, in a strict sense, its validity needs to be tested. Third, the influence of one more speech production component (i.e., Formulator) has been left in this study and, thus, has to be tested in the future study. With all these limitations in mind, it should be noticed that the study does not claim that teaching CSs is the best means to improve the learners' flexibility in conceptual processing, and unarguably, this needs to be tested by comparing it with other teaching methods. In spite of these limitations, the author believes that the present study contributes to rationalizing the CS instruction theoretically.

Acknowledgements

This study was supported by the academic research grant offered by the Japan Society for the Promotion of Science [Grant Code: Scientific Research (C) 19520499]. I am deeply grateful to my colleague, Carol Rinnert, for her valuable suggestions to improve the quality of this study. I also want to thank the two anonymous reviewers for their insightful comments.

IWAI, Chiaki is a professor at Hiroshima City University. For more than 15 years, he has been examining how L2 learners use and develop strategic competence. He is also interested in the issues of interlanguage pragmatics and of English as an international lingua franca. <iwai@intl.hiroshima-cu.ac.jp>

References

Bialystok, E. (1990). *Communication strategies: A psychological analysis of second-language use.* Oxford: Basil Blackwell.

Cook, V. (1993). *Linguistics and second language acquisition*. New York: St. Martin's Press.

Cook, V. (2008). Multi-competence: Black hole or wormhole for second language acquisition research. In Z-H. Han (Ed.), *Understanding second language* (pp. 16-26). Clevedon: Multilingual Matters.

Dörnyei, Z. (1995). On the teachability of communication strategies. *TESOL Quarterly*, *29*, 55-85.

Iwai, C. (2006). *Linguistic and pedagogical values of teaching communication strategies: Integrating the notion of communication strategies with studies of second language acquisition*. Unpublished doctoral dissertation, Hiroshima City University.

Iwai, C., & Gobel, P. (2004). Instructional effects of communication strategies. In M. Swanson & K. Hill (Eds.), *JALT2003 Conference Proceedings* (pp. 162-174). Tokyo: JALT.

Kasper, G., & Kellerman, E. (1997). Introduction: Approaches to communication strategies. In G. Kasper & E. Kellerman (Eds.), *Communication strategies: Psycholinguistic and sociolinguistic perspectives* (pp. 1-13). London: Longman.

Kellerman, E. (1991). Compensatory strategies in second language research: A critique, a revision, and some (non-)implications for the classroom. In R. Phillipson, E. Kellerman, L. Selinker, M. Sharwood Smith, & M. Swain (Eds.), *Foreign/second language pedagogy research* (pp. 142-161). Clevedon: Multilingual Matters.

Kellerman, E. (1998). When words fail: From communication strategies to strategies for communication. In K. Malmkjær & J. Williams (Eds.), *Context in language learning and language understanding* (pp. 91-112). Cambridge: Cambridge University Press.

Kitajima, R. (1997). Influence of learning context on learners' use of communication strategies. *JALT Journal*, *19*(1), 6-23.

Konishi, K., & Tarone, E. (2004). English constructions used in compensatory strategies:Baseline data for communicative EFL instruction. In D. Boxer & A. D. Cohen (Eds.), *Studying speaking to inform second language learning* (pp. 174-198). Clevedon: Multilingual Matters.

Levelt, W. J. M. (1989). *Speaking*. Cambridge, Mass.: M.I.T. Press.

Nakatani, Y. (2005). The effects of awareness-raising training on oral communication strategy use. *Modern Language Journal*, *89*, 76-91.

Nishiyama, M. (1996). A study of non-fluent bilinguals' mental lexicon by means of word association test. *Research Report of Kobe City College of Technology*, *34*, 117-124.

Mochizuki, M., Aizawa, K., & Tono, Y. (2003). Eigo goi-no shido manual [Instruction manual for teaching English vocabulary]. Tokyo: Taishukan.

Poulisse, N. (1990). *The use of compensatory strategies by Dutch learners of English*. Dordrent, Holland: Foris Publications.

Salomone, A. M., & Marsal, F. (1997). How to avoid language breakdown? Circumlocution! *Foreign Language Annals*, *30*(4), 473-484.

Skehan, P. (1998). *Cognition and approach to language learning*. Oxford: Oxford University Press.

Yule, G., & Tarone, E. (1997). Investigating communication strategies in L2 reference: Pros and cons. In G. Kasper & E. Kellerman (Eds.), *Communication strategies: Psycholinguistic and sociolinguistic perspectives* (pp. 17-30). London: Longman.

Appendix 1

Sample test items and responses [see Iwai (2006) for actual visual images.]

Table 1. Test items in English

Semantic Category	Test 1	Test 2	Test 3
Persons	plumber	lumberer	sanitation worker
Insects	green caterpillar	ladybug	cicada
Animals	peacock	ostrich	lizard
Tools (cleaning)	rag	duster	broom
Vehicles	crane	cement track	street cleaner
Clothes (hats)	mortarboard	sun visor	sombrero
Personality	consideration （思いやり）	sympathy （なさけ）	concern （配慮）
Emotion	irritation （いらだち）	impatience （短気）	anxiety （不安）
Fortune	prospect （予想）	prediction （予知）	assumption （憶測）

Table 2. Test items in Japanese

Semantic Category	Test 1	Test 2	Test 3
Persons	mechanic	welder	ceramic artist
Insects	firefly	cricket	mantis
Animals	peacock	rhinoceros	anteater
Tools (cleaning)	rake	dustpan	mop
Vehicles	tow rack	fire truck	garbage truck
Clothes (hats)	top hat	beret	ski mask
Personality	鈍感 (insensitivity)	繊細 (delicacy)	純情 (purity)
Emotion	ねたみ (jealousy)	にくしみ (hatred)	あこがれ (adoration)
Fortune	願い (desire)	願望 (yearning)	期待 (expectation)

N.B.: These items were randomly ordered in the actual tests

Table 3. Sample responses (Semantic category of <Persons>) Student (ID# 01 Male: Experimental Group)

Test	Test #	Target item	Responses
English Pre-test	Test 1	plumber	fix, water, pipe, man
English Post-test	Test 2	lumberer	person, man, working, mountain, cut, tree
English DPost-test	Test 3	sanitation worker	clean, dust, garbage, gather, people, man, worker
JPN Pre-test	Test 1	mechanic	車, 直す, 修理, 男性 (car, fix, repair, man)
JPN Post-test	Test 2	welder	溶接, 強い, 光, 電圧, 熱, 火傷, 目 (welding, strong, light, voltage, heat, burn, eyes)
JPN DPost-test	Test 3	ceramic artist	芸術, 男性, 創る, 陶器, 粘土, 焼く,かま (art, man, create, ceramics, clay, burn, furnace)

N.B. DPost test = delayed post-test. English translations in parentheses are added by the author.

Appendix 2

Table 4. Descriptive statistics of 3 pre-tests (concrete items)

Pre-test	N	M	SD	words/ item
EWU1	13	24.7	8.750	4.1
EWU2	11	22.0	3.406	3.7
EWU3	15	24.4	6.791	4.1
JWU1	13	35.5	6.691	5.9
JWU2	11	33.5	8.419	5.6
JWU3	15	36.3	9.846	6.1

N.B.: EWU = English word utterance test

Table 5. Descriptive statistics of 3 pre-tests (Abstract nouns)

Pre-test	N	M	SD	words/ item
EWU1	13	10.3	4.191	3.4
EWU2	11	5.3	3.289	1.8
EWU3	15	7.5	3.226	2.5
JWU1	13	12.5	4.352	4.2
JWU2	11	12.5	5.126	4.2
JWU3	15	13.6	4.595	4.5

JWU = Japanese word utterance test

Table 6. Descriptive statistics of pre-, post-, and delayed post-tests (concrete items)

EWU tests	M/SD	Pre-test	Post-test	Delayed
Exp Group	M	21.4	27.9	29.4
(N = 19)	SD	5.430	8.891	6.693
Control Group	M	26.1	26.5	
(N = 20)	SD	7.210	13.900	

N.B. Exp group = Experimental group

JWU tests	M/SD	Pre-test	Post-test	Delayed
Exp Group	M	33.2	35.2	40.4
(N = 19)	SD	7.904	10.809	12.451
Control Group	M	37.2	37.4	
(N = 20)	SD	8.501	12.326	

Table 7. Descriptive statistics of pre-, post-, and delayed post-tests (Abstract nouns)

EWU tests	M/SD	Pre-test	Post-test	Delayed
Exp Group	M	7.6	8.3	8.7
(N = 19)	SD	4.194	3.497	3.885
Control Group	M	8.0	9.0	
(N = 20)	SD	3.974	4.839	
JWU tests	**M/SD**	**Pre-test**	**Post-test**	**Delayed**
Exp Group	M	12.7	14.3	15.4
(N = 19)	SD	5.239	7.745	6.760
Control Group	M	13.2	14.9	
(N = 20)	SD	3.968	5.241	

Table 8. Means and SDs of pre-, post-, delayed post tests (by 3 vocabulary groups: Concrete items)

EWU tests (N)		M	SD
Pre-test	High (6)	23.5	6.380
	Mid (6)	21.5	5.857
	Low (7)	19.6	4.198
Post-test	High	29.5	8.550
	Mid	26.3	9.288
	Low	28.0	9.967
Delayed	High	33.3	7.448
	Mid	27.5	6.716
	Low	27.6	5.255

Table 9. Means and SDs of pre-, post-, delayed post tests (by 3 vocabulary groups: Abstract nouns)

JWU tests (N)		M	SD
Pre-test	High (6)	32.7	7.090
	Mid (6)	30.7	6.121
	Low (7)	35.7	10.012
Post-test	High	36.5	12.243
	Mid	30.3	7.737
	Low	38.3	11.800
Delayed	High	42.7	9.026
	Mid	34.3	8.189
	Low	43.7	16.928

N.B. The numbers in parentheses represent the number of group members.