BUILDING AN EMPIRICALLY-BASED MODEL
OF EFL LEARNERS' WRITING PROCESSES

MIYUKI SASAKI

Nagoya Gakuin University, Japan

Abstract. This chapter investigated Japanese learners' processes of English expository writing using multiple data sources including their written texts, videotaped writing behaviors, and stimulated recall protocols. Two groups of Japanese EFL writers (12 experts and 22 novices) were compared both cross-sectionally and longitudinally. The study tested the following eight hypotheses formulated as a result of a pilot study (Sasaki, 2000): (1) EFL writing experts write longer texts at greater speed than EFL writing novices; (2) after two semesters of process writing instruction, neither the quantity nor the speed of the novices' writing improves; (3) the experts spend a longer time before starting to write than the novices; (4) after the instruction, the novices spend a longer time before starting to write; (5) while writing, the experts stop to reread or refine their expressions more often than the novices, whereas the novices stop to make local plans or translate their ideas into L2 more often than the experts; (6) after the instruction, the novices stop to reread more often while making fewer local plans; however, they still have to stop to translate as often as before; (7) the experts tend to plan a detailed overall organization, whereas the novices tend to make a less detailed plan; (8) after the instruction, the novices learn to do global planning, but it is qualitatively different from the experts' global planning. The obtained results are presented as flow-chart diagrams that represent the writing processes of the different groups of EFL learners.

Keywords: Japanese EFL learners, empirical model of writing processes, expert writers, fluency, global planning, local planning, novice writers, pausing behaviors, plan monitoring, stimulated recall protocols, writing expertise, writing strategies.

1   INTRODUCTION

Composing process has been a major focus of L2 writing research for the past several decades (Cumming, 1998; Krapels, 1990; Silva, 1993). Basically following the designs of first language (L1) composition studies, researchers have investigated various aspects of L2 writing processes for different groups of participants. Of particular interest to the present study were those studies that examined writers' micro-level cognitive processes while writing. Zamel (1983), in one of the earliest studies,

analyzed detailed observational data collected while six 'advanced' ESL (English as a second language, *i.e.*, learning English in an English-speaking environment) students were completing 'a course-related writing task' (Zamel, 1983: 169). Among these six students, four were identified as 'skilled' and two as 'unskilled' based on experienced readers' 'holistic assessments' (p. 172) of their writings, and these students spent four to eighteen hours writing several drafts. Zamel did not use the participants' 'think-aloud' data while writing in spite of the fact that they were 'used in most process studies' (Zamel, 1983: 169), because 'there is some doubt about the extent to which verbalizing aloud one's thoughts while writing simulates the real composing situation' (Zamel, 1983: 169). In contrast, Raimes (1985), adopting the methods employed in L1 process writing studies (*e.g.*, Hayes & Flower, 1983; Swarts, Flower, & Hayes, 1984), examined concurrent think-aloud data collected from eight unskilled (determined by holistic measures of their essays) ESL students who wrote a narrative within a 65-minute class period. A much larger scale study was Cumming's (1989) investigation of 23 French-speaking college students' English writing processes using their written texts and think-aloud data. The study was notable because (1) it applied multivariate statistical analyses, which was made possible by the relatively large sample size, (2) it compared students' writing processes for three different tasks (letter writing, summary, and argumentation, one to three hours each), and (3) it introduced controlled variables of L1 writing expertise and L2 writing proficiency. Using the participants' decision statements in the think-aloud data, Cumming focused on four aspects of writing the students attended to while writing (language use, discourse organization, gist, and procedure for writing) and five categories of problem-solving behaviors (heuristic searches with and without resolution, problem resolution, problem identification, and knowledge telling). More recently, Bosher (1998), using a modified version of Cumming's (1989) coding systems, examined the L2 writing processes of three Southeast Asian ESL college students with different educational backgrounds. Bosher's study was unique in that she used, as alternative data to think-aloud protocols, stimulated retrospective protocols collected from the participants who recalled their writing processes while watching their own videotaped writing behaviors. Finally, Roca de Larios, Murphy, and Manchon (1999) again using the participants' think-aloud protocol data while writing, conducted two studies investigating the 'restructuring' process where the writers seek 'an alternative syntactic plan once the writer predicts, anticipates, or realizes that the original plan is not going to be satisfactory for a variety of linguistic, ideational or textual reasons' (Roca de Larios *et al.*, 1999: 16). Unlike the other studies described above, Roca de Larios *et al.*'s study concentrated on the particular writing strategy of restructuring that had 'received very little attention in research on composing' (Roca de Larios *et al.*, 1999: 16).

These previous studies that examined part of or the entire process of L2 writing commonly found that (1) skilled L2 writers were similar to their L1 counterparts in that they tended to plan more, revise more at the discourse level, and spend more time exploring the most appropriate ways to solve the given task (*e.g.*, Cumming, 1989; Raimes, 1987; Roca de Larios, Murphy, & Manchon, 1999; Zamel, 1982, 1983); (2) unskilled L2 writers were similar to their L1 counterparts in that they tended to plan less and revise more at the word and phrase level (*e.g.*, Raimes, 1985,

1987; Roca de Larios, Murphy, & Manchon, 1999; Zamel, 1983), but they were different from their L1 counterparts in that they were relatively less concerned about surface level revisions (*e.g.*, Raimes, 1985, 1987) and in that they showed more commitment to the given assignment (*e.g.*, Raimes, 1985, 1987); (3) there appears to be a 'writing expertise' which is independent of L2 proficiency, affecting L2 writing (*e.g.*, Bosher, 1998; Cumming, 1989; Raimes, 1985, 1987); (4) students' attention patterns and problem-solving behaviors while writing differed according to their L1 writing expertise and the type of tasks they were involved in (*e.g.*, Cumming, 1989).

Because researchers have realized that L2 writers' strategies are similar to those used for L1 writing, many studies from the late 1980's on have also compared the same participants' L1 and L2 writing processes. It should be noted that here again think-aloud protocol data were the main sources of analysis for most studies. In these studies, the participants' L2 is mostly English (but see Cumming, Rebuffot, & Ledwell, 1989 and Whalen & Ménard, 1995 for exceptions), but their L1s greatly vary. For example, Jones and Tetroe (1987) compared six college-level Spanish-speaking ESL students' planning behaviors while these students wrote two English and one Spanish descriptive expositions. Arndt (1987) compared six Chinese postgraduate EFL (English as a foreign language, *i.e.*, learning English in a non English-speaking environment) students' processes of writing expositions in L1 and L2 (completed within one hour each). Similarly, Skibniewski (1988) compared three college-level Polish EFL students' processes of writing expository essays in L1 and L2. In contrast to Jones and Tetroe's or Arndt's study, Skibniewski could compare the differential effects of writing expertise on the three students' L1 and L2 writing processes because they had distinctly different writing skills both in L1 and L2 (*i.e.*, skilled, average, and unskilled). Similarly, Cumming, Rebuffot, and Ledwell (1989) compared the summary writing processes in English and French of 14 English-speaking college students with different writing expertise. Using Cumming's (1989) coding scheme, Cumming *et al.* specifically focused on the participants' problem-solving behaviors. Finally, two more recent and larger scale studies were completed by Whalen and Ménard (1995) and Uzawa (1996). Whalen and Ménard analyzed 12 English speaking participants' planning, evaluation, and revision strategies at three different levels of discourse (pragmatic, textual, and linguistic) while writing argumentative texts in their L1 and L2 (French) within a maximum of two hours for each. On the other hand, Uzawa compared 22 Japanese ESL students' processes of writing first drafts of descriptive expositions in Japanese (30 minutes) and in English (one hour), as well as their processes of translating a magazine article from Japanese into English (one hour). In addition to comparing overall characteristics of each writing process, Uzawa compared attention patterns employed for the three types of writing.

Admitting the noticeable individual differences among the participants reported by some of these comparative studies (*e.g.*, Arndt, 1987), we can also conclude that they have commonly found that (1) L1 and L2 writing strategies, whether the writers were skilled or unskilled, were basically similar, which indicates that L1 writing strategies can be transferred to L2 writing (*e.g.*, Arndt, 1987; Cumming, Rebuffot, & Ledwell, 1989; Jones & Tetroe, 1987; Moragne e Silva, 1988; Skibniewski, 1988; Uzawa, 1996; Whalen & Ménard, 1995); (2) compared with their L1 writing proc-

esses, students' L2 writing processes, especially the higher-order cognitive operations, were negatively affected by their limited L2 proficiency (e.g., Moragne e Silva, 1988; Whalen & Ménard, 1995); and (3) the quality of written L2 texts is more strongly associated with the quality of the students' L1/L2 writing strategies rather than with their L2 proficiency (e.g., Cumming, Rebuffot, & Ledwell, 1989; Jones & Tetroe, 1987).

Although these previous studies provided insight into L2 learners' writing processes, their designs were not without limitations. First, they investigated mainly ESL learners whose educational backgrounds were typically heterogeneous, and whose L2 proficiency was high enough so that they could receive their education in their L2. Even when EFL learners were examined, their L2 proficiency tended to be high (e.g., Arndt, 1987; Skibniewski, 1988). Second, even though some studies included 'skilled' versus 'unskilled' contrasts (mainly among student writers), virtually no studies have included a 'novice' versus 'expert' contrast where 'experts' were those who used L2 writing for professional purposes.[1] Furthermore, many previous studies have employed cross-sectional designs only, and thus lacked developmental perspectives. Including multiple perspectives where novice writers are compared with expert writers as representatives of their ultimate possible goals of achievement, or where the novice writers are compared before and after a certain period of writing instruction with other intervening variables controlled, is crucial for building a more comprehensive and dynamic model of L2 writing processes.

Another limitation of the previous studies of L2 writing processes is their almost exclusive use of think-aloud protocols as the main data source (but see the above description of Zamel, 1983 and Bosher, 1998 as exceptions). Although collecting concurrent verbal reports is an effective way to obtain real-time data on the participants' writing processes (Ericsson & Simon, 1993), it entails various inherent problems (Smagorinsky, 1994), some of which are especially relevant to the present study. First, it is very difficult for some potential participants to produce 'think-aloud' data while writing in L2. It appears even more difficult when they are asked to speak in their L2 (e.g., Raimes, 1985, 1987) because many L2 writers often think in their L1 while writing (e.g., Cumming, 1989; Cumming, Rebuffot, & Ledwell, 1989; Uzawa, 1996). Moreover, even when participants were allowed to speak in any language they wished, some expressed difficulty with the task. For example, Whalen and Ménard (1995), who seem to have allowed the participants to choose the language they spoke in, admitted that ten potential participants (compared to the 12 who actually produced the data for the study) could not perform this difficult task, and thus were excluded from the study. Finally, even if researchers can manage to obtain analyzable data from participants (see Hayes & Flower, 1980, characterizing the nature of analyzing protocol data as 'following the tracks of a porpoise,' p.

9), there is always the danger of 'reactivity'. Previous empirical studies (e.g., Janssen, van Waes, & van den Bergh, 1996; Stratman & Hamp-Lyons, 1994) have reported that the think-aloud condition appeared to have significantly affected the quality and content of the participants' cognitive activities while writing.

With these methodological limitations in mind, I conducted a precursor of the present study as a pilot study (Sasaki, 2000). It investigated the writing processes of three types of L2 writers (professional, and more- and less-skilled) with similar cultural and educational backgrounds, both cross-sectionally and longitudinally (i.e., developmentally),[2] using multiple data sources collected through a less disruptive method than the think-aloud technique.[3] The method was similar to the one used in Bosher (1998) in that the participants produced recall protocols while watching their video-taped writing behaviors, but it was different from Bosher's method in that the participants could choose the language(s) in which they produced the protocols, and in that the data were coded by a coding scheme specifically developed for this type of data (Anzai & Uchida, 1981).

The pilot study was also motivated by the results of two preceding product-oriented studies (Hirose & Sasaki, 2000; Sasaki & Hirose, 1996). Sasaki and Hirose cross-sectionally investigated factors that could explain Japanese EFL students' English writing ability. We found that the participants' L2 proficiency, L1 writing ability, and metaknowledge of L2 expository writing (e.g., how to achieve unity and coherence in a paragraph) significantly explained the students' L2 writing ability variance (52% of the variance was accounted for by L2 proficiency, 18% by L1 writing ability, and 11% by metaknowledge). We also found that good writers were significantly different from weak writers in terms of their attention to overall organization while writing in L1 and L2, their writing fluency in L1 and L2, their confidence in L2 writing for academic purposes, and their experiences of regularly writing more than one paragraph in L2 in high school.   Based on these results, Hirose and Sasaki further examined the teachability of two of these explanatory factors, metaknowledge of L2 writing and regular L2 writing experience. The results indicated that teaching the metaknowledge to the students over 12 weeks significantly improved their metaknowledge, but not their L2 writing ability in general. In contrast, the instruction of metaknowledge combined with regular journal writing sig-

---

1 Although several studies such as Cumming (1989) included writers with professional experience, they were experts in L1 writing rather than L2 writing. I believe that research into the differences between the writing processes of novice and expert L2 writers is necessary to build a comprehensive model of L2 writing processes because experts' writing ability represent an ultimate goal (and also an ultimate achievement limit) that any L2 learners with similar backgrounds can accomplish (Grabe & Kaplan, 1996).

2 In the present study, I used the term 'longitudinal' as synonymous with 'developmental.' When I classified studies, I followed Isaac and Michael's (1981:42) definition of 'developmental': To investigate patterns and sequences of growth and/or change as a function of time. In the pilot study, Sasaki (2000), I investigated the changes in eight student writers' writing process during six months of process writing instruction, and thus I called the study 'longitudinal.'

3 Although the stimulated recall protocol method employed both for Sasaki (2000) and the present study is obviously less disruptive than the think-aloud method (all the participants in both studies could successfully complete the task), it might arguably have entailed some reactivity problems such as the possibility that the participants had been affected by the existence of the video-camera(s). Moreover, it is also true that the recall protocol method can only induce what the participants can recall, or what they think they were thinking about at the point of time in question. Unlike the think-aloud data, what the participants recall may not be a faithful reproduction of what they were thinking about at that particular moment.

nificantly improved mechanical aspects (e.g., spelling, punctuation, capitalization), but not the overall quality of their L2 writing.

When I conducted the pilot study (Sasaki, 2000), very few studies had investigated Japanese EFL learners' L2 writing processes before, and thus, the study inevitably became exploratory in nature: I tried to select as the targets of analysis as many aspects as possible of the participants' writing behavior (e.g., time spent before starting to write, pausing behaviors, writing strategies) that seemed to be important for building an empirical model of their writing processes. At that point, which aspects of those writing behavior would characterize the writing processes of the three different groups of participants was not yet clear. Furthermore, because the pilot study used relatively small samples (four experts, four more-skilled students, and four less-skilled students), the significance of the findings could not be tested by statistical procedures. In order to claim generalizability of the results obtained in the pilot study, hypotheses had to be formed based on the results and then tested with samples that would be large enough for statistical procedures to be legitimately applied (the relationship between the pilot study and present study is in a sense similar to that between exploratory and confirmatory factor analyses, where confirmatory factor analysis tests the relationships among variables that were constructed as a result of exploratory factor analysis; see Bollen, 1989).

The present study thus replicates the general design of the pilot study by investigating the writing processes of Japanese EFL learners both cross-sectionally and longitudinally. The data analyzed in the present study included the participants' written texts, their pausing behaviors while writing, stimulated recall protocols, and analytic scores given to the written texts. The present study tests the following eight hypotheses formulated as a result of the pilot study. These hypotheses reflect the differences found among the participant groups in the pilot study:[4]

1) EFL writing experts write longer texts at greater speed than EFL writing novices.

2) After two semesters of process writing instruction, neither the quantity nor the speed of the novices' writing improves.

3) The experts spend a longer time before starting to write than the novices.

4) After the instruction, the novices spend a longer time before starting to write.

5) While writing, the experts stop to reread or refine their expressions more often than the novices, whereas the novices stop to make local plans or translate from L1 to L2 more often than the experts (see Appendix for the definitions of Rereading, Rhetorical Refining, Local Planning, and Translating from L1 to L2).

6) After the instruction, the novices stop to reread more often while making fewer local plans. However, they still have to stop to translate from L1 to L2 as often as before.

7) The experts tend to plan a detailed overall organization (i.e., Global Planning), whereas the novices tend to make a less detailed plan (i.e., Thematic Planning).

---

4 Unlike the pilot study, the novices were not divided into the more- and less-skilled writers in the present study in order to make the research design less complex for the application of statistical procedures.

8) After two semesters of process writing instruction, the novices learn to do global planning, but it is qualitatively different from the experts' global planning.

Hypotheses 1, 3, 5, and 7 were concerned with the cross-sectional differences between EFL writing experts' writing processes and those of EFL writing novices whereas Hypotheses 2, 4, 6, and 8 were concerned with the effects of a process-writing instruction on the novices' writing processes. Hypotheses 1 and 2 were related to the variable of 'writing fluency'. This variable had been chosen as a target of investigation in the pilot study (Sasaki, 2000) because the good and weak writers in Sasaki and Hirose (1996), one of the product-oriented studies that motivated the pilot study, were significantly different in terms of writing fluency. Hypotheses 3 and 4 were related to writers' planning behavior, which has been investigated in many previous studies of both L1 and L2 writing processes (e.g., Cumming, 1989; Raimes, 1987; Roca de Larios, Murphy, & Manchon, 1999; Zamel, 1982, 1983). The four strategies referred to in Hypotheses 5 and 6 (Rereading, Rhetorical Refining, Local Planning and Translating from L1 to L2) were the ones the experts used differently from the novices among the 21 strategies investigated in the pilot study. Finally, Hypotheses 7 and 8 were concerned with the issue of writing expertise whose existence has been hypothesized in some previous studies (e.g., Bosher, 1998; Cumming, 1989; Raimes, 1985, 1987). In the pilot study, I speculated that particular characteristics of the experts' global planning were similar to those of what Flower, Schriver, Carey, Haas, and Hayes (1992) called 'constructive planning,' which was exclusively used by L1 writing experts after long years of training and experiences.

2   METHOD

2.1   Participants

Two groups of Japanese EFL learners (a total of 34) participated in the present study: an expert writer group (n = 12) and a novice writer group (n = 22). The experts (10 men and 2 women) were operationally defined as Japanese native speakers who had learned English mainly through formal education in Japan until they graduated from universities, and who now regularly write academic papers in English at least once a year. They had studied English (both inside and outside the classroom) for an average of 23.5 years (SD = 6.0 years). Although I didn't specifically measure their English proficiency for the present study because of time constraints,[5] I assumed that their English proficiency was high (six of them reported that their best TOEFL [Test of English as a Foreign Language produced by Educational Testing Service, Princeton, NJ] scores were over 600). They were four applied linguists, two linguists, one communication researcher, three economists, one engineer, and one chemist, with a mean age of 36.8 years. Although they had spent an average of 4 years in English-speaking countries (seven of them had M.A.s, and three of them

---

5 The time taken for the composition writing session and stimulated recall (2.5 to 3 hours) was the maximum I could ask the busy experts to spare for the present study.

had Ph.D.s from universities in English speaking countries), most of their current life is anchored in Japan, a non-English speaking environment.

The novices were all 18 year-old college freshmen (4 men and 18 women), majoring in British and American studies.[6] Candidates for those novices were randomly selected from a sample of 75 students (three classes combined), and asked to participate in the present study. Those who agreed participated in the present study. They were paid for their work. They had studied English for six years through formal education in Japan. None of them had been abroad longer than one month. They were judged to be 'novice L2 writers' because the results of a background questionnaire (see Appendix A of Sasaki & Hirose, 1996) indicated that they had received little L2 writing instruction, including instruction on matters such as 'organizing a paragraph centered on one main idea' or 'developing a paragraph so that the readers can follow it easily.' The mean total score on the Secondary Level English Proficiency (SLEP) Test (Educational Testing Service, 1993) was 40.68 (SD = 4.98), which indicates that these students' English proficiency ranged from low- to mid-intermediate.

The writing processes of the 22 novice writers were also compared before and after two semesters (a total of six months interrupted midway by a two-month summer vacation) of process writing instruction. Although the instruction was basically intended to employ a 'process approach' (Silva, 1990: 15), it also incorporated some 'current-traditional' (Silva, 1990: 13) aspects (i.e., teaching students representative prescriptive writing patterns) because metaknowledge (e.g., the meaning of topic sentence, unity, coherence, and the organization of English expository writing) that could be acquired through such a current-traditional approach had been found to be significant in explaining L2 writing quality of a similar group of students (Sasaki & Hirose, 1996). The 22 novice students in the present study received instruction on paragraph writing in English with 53 other students once a week for 90 minutes. These students did not take any other English writing classes during these two semesters.

The instructor (the researcher) taught them process writing strategies such as planning and revising, based on Bereiter and Scardamalia's (1987) ideas of 'Promoting the development of mature composing strategies' (Bereiter & Scardamalia: 245), using Hashiuchi's (1995) *Paragraph Raitingu Nyuumon* [Introduction to Paragraph Writing], a composition textbook with special emphasis on process writing. Thus, in the first class, the students were told that writing is an interactive process between what they write and what they want to write, and that such a process is cyclical, starting with planning and followed by writing and revising. Furthermore, in each chapter of the textbook (the class covered nine chapters in total), the students first learned rhetorical patterns such as comparison, classification, and expressing opinions, and then were instructed to write a similar paragraph themselves. Before they started to write, they discussed the following points with their instructor and peers:

---

6 *The novices' mean age was significantly different from that of the experts' (t = 11.26, df = 32, p = .000). Therefore, it is possible that these two groups may differ not only in writing expertise but also in cognitive abilities. The relationship among age, cognitive abilities, and writing expertise, however, remains to be studied in future studies.*

1) What is the purpose of this writing?
2) Who are the readers of this writing?
3) What content should I include to make it effective?
4) How should I express the content to make it effective?

After they wrote the first drafts, most drafts were read by their peers or the instructor. When time allowed, these readers gave written comments on which points of the writing were interesting or appealing and which were difficult to understand. After receiving these comments, the students were asked to revise their writing with the ideas of unity and cohesion in mind. A total of nine texts were written during the class, but none of them were graded (the students were graded on the basis of their participation in the class and the scores of the mid-term and final examinations, which included writing compositions).

2.2 Data collection procedure

For the present study I followed a slightly revised version of the technique employed in the pilot study (Sasaki, 2000) as explained below. The technique was originally developed by Anzai & Uchida (1981) for investigating Japanese children's L1 writing process as an alternative data collection method to the concurrent think-aloud procedure,[7] but I found it can also be successfully used for investigating Japanese EFL students' writing processes.

The L2 experts and the 22 novice students wrote an argumentative composition according to Prompt 1:

> There has been a heated discussion about the issue of school uniforms in the readers' column in an English newspaper. Some people think that wearing high school uniforms is a good custom, whereas others believe high school students should be given an opportunity to choose what they wear. Now the editor of the newspaper is calling for readers' opinions. Suppose you are writing for the readers' opinion column. Take one of the positions described above, and write your opinion within 30 minutes. (Original in Japanese, translated by the author)

The 22 novice writers then wrote according to Prompt 2 after the two-semester instruction period ended:

> There has been a heated discussion about the custom of non-Christian Japanese celebrating Christmas in the readers' column in an English newspaper. Some people think that it is a good custom, whereas others believe we should abandon such a custom. Suppose you are writing for the readers' opinion column. Take one of the positions described above, and write your opinion within 30 minutes. (Original in Japanese, translated by the author)

---

7 *Researchers in the field of L1 Japanese writing have also found that it was difficult to collect think-aloud data. For example, Uchida (1989) reported that only two out of ten potential participants produced analyzable think-aloud data in a study investigating children's revising processes in L1 Japanese (see also Uchida, 1986 for discussion of this issue).*

One week after the novices wrote for Prompt 2, I had to ask two of them to write for Prompt 3 because the video camera broke when they were writing for Prompt 2.

> There has been a heated discussion in an English newspaper about introducing English as an elementary school subject in Japan. Some people think that English should be taught at the elementary school level, whereas others believe that it is too early. Take one of the positions described above, and write your opinion within 30 minutes. (Original in Japanese, translated by the author)

For these two students, the data for Prompt 3 were used for subsequent analyses. Prompts 1 to 3 were similar in the sense that they dealt with controversial issues that the students were familiar with. Prompts 1 and 2 were also used in the pilot study (see Sasaki, 2000 for the reasons why I selected these argumentative topics).[8]

All compositions were scored by two EFL writing specialists independent of the present study, according to Jacobs et al.'s (1981) English Composition Profile. In order not to make the raters aware of which topic was written first, all compositions were rated on one occasion after having been completely mingled with no indication of when they were written, or which participant groups they belonged to. The inter-rater correlation (Pearson correlation coefficient) for the content subscore was .93;

---

8 *If these two prompts had been alternated with half of the eight students before and after the instruction (i.e., half of them receiving Prompt 1 before the instruction and Prompt 2 after the instruction with the other half receiving Prompt 2 before the instruction and Prompt 1 after the instruction), I could have avoided possible topic effects on the students' composition scores and their use of writing strategies. However, if I had alternated the prompts for the novice I group, I would also have had to alternate the prompts for the expert group for a fairer comparison. In such a case, I would have had to consider the possible effects of these two different topics on the participants' use of writing strategies, especially when the experts' writing strategies were compared with those of the novices for one composition written on only one occasion. Previous studies (e.g., Carter, 1990; Cumming, 1989; Flower, Schriver, Carey, Haas, & Hayes, 1992) suggest that writers may change their writing strategy use according to different topics. Thus, I decided not to alternate the prompts for the novices. On the other hand, I used similar but different prompts for novices I and II because I was afraid that maturation effects caused by giving the same prompt before and after the instruction (e.g., the novices might have thought about the topic over the two semesters, and thus produced much better quality compositions than they might have if they had not written for the same prompt before the instruction, or they would not plan for the second time because they had already thought about the issue for the first time) might be stronger than possible topic effects (especially when Prompts 1 to 3 were intended to induce very similar argumentative writing, see Sasaki, 2000; Sasaki & Hirose, 1996). I had to make this compromise that might have introduced topic effects for the novices I and II comparison because I wanted to compare the experts and novices on the most equal basis possible while still comparing the same novices before and after the instruction. In other words, I wanted to incorporate into a single study both a cross-sectional design and a longitudinal design using the same participants. This was a real dilemma. But I concluded that topic effects, if they existed at all, would be larger for the expert-novice (inter-group) comparison rather than for the novices I and II (intra-group) comparison. Of course, however, I was aware of the fact that comparing novices I and II could be problematic because of possible topic effects (see the Results and Discussion Section).*

---

the organization subscore, .86; the vocabulary subscore, .93; the language use subscore, .93; the mechanics subscore, .59; and the total score, .97. I judged that these correlations were acceptable for this sample population size.

To collect the data, I asked the participants to come individually to a room and write the compositions in a quiet atmosphere. I asked them to finish writing within about 30 minutes but told them they could take more time if they wanted to. Consequently, some of them took longer than 30 minutes to finish, but all of them finished writing within one hour. The relatively short time allocation was chosen for the purpose of making the following video-watching session manageably short (approximately two hours). As in the pilot study, the participants were not allowed to use dictionaries.

Before the participants started to write, I obtained permission to videotape them while they were writing. I then began to videotape them with one camera focused on their hand/pencil movement, and another focused on their entire writing behavior including their eye/head movements. Unlike the pilot study, I used two cameras to obtain clearer and more holistic views of which part of the text the participants were working on and how they were writing. As in Anzai and Uchida (1981) and in the pilot study, I asked the participants how much planning and what kind of plan they had made before starting to write just after they started to write the first word, when their memory was still fresh. The participants were not told that they would be questioned right after they started writing. When they answered the questions, they were told that their answering time would not be included in the 30 minutes allocated for writing. After the first question session, I let them continue writing without interruption until they finished.

Immediately after they finished writing, each participant and I together watched the participant's writing process on videotape on two monitors, which simultaneously were showing the participant's hand/pencil movement and physical writing behavior. Every time the participants stopped writing for a period longer than three seconds on the videotape,[9] I asked them to explain, either in Japanese or English, what they had been thinking about. No leading question (e.g., 'Did you think such-and-such?') were used to avoid biasing the students' answers (Ericsson & Simon, 1993). This continued until they had finished the entire process of writing on the tape. All participants gave their accounts in Japanese except for some English expressions that they used or considered for use in the compositions (see the examples below). The videotaped behaviors projected through two monitors were clear enough to aid the participants' recall. The participants' accounts were all tape-recorded and subsequently transcribed.

From these writing and recall sessions, I obtained and analyzed three types of data: (1) written texts and drafts, (2) time the participants spent before starting to write and time they spent writing the whole composition, and (3) participants' retrospective accounts, while watching their video-taped writing performance, of what they were thinking about when they stopped writing.

---

9 *Anzai and Uchida (1981) used pauses longer than two seconds instead of three. Longer pauses were used in the present study because it was concluded that three seconds were the shortest possible pauses that could be handled based on the pilot study results.*

2.3 Coding of the protocol data

In the pilot study, I used a revised version of Anzai & Uchida's (1981) coding system. Their coding system was developed in a careful and well-designed empirical study, and successfully used to investigate the participants' Japanese L1 writing strategies. Following Anzai and Uchida, I defined the term 'strategy' in the present study as a writer's mental behavior employed to achieve a certain goal in the 'ill-structured problem-solving' (Anzai & Uchida: 46) activity of writing. Based on Anzai and Uchida's coding system I encoded the pilot study participants' retrospective protocol data into 21 strategies (Global Planning, Thematic Planning, Local Planning, Organizing, Conclusion Planning, Plan Retrieving, Information Retrieving, Generating Naturally Generated Ideas, Generating Description-Generated Ideas, Verbalizing a Proposition, Rhetorical Refining, Mechanical Refining, Sense of Readers, Translating L1 to L2, Rereading, L2 Proficiency Evaluation, Local Text Evaluation, General Text Evaluation, Resting, Questioning, and Impossible to categorize; see Appendix C of Sasaki, 2000 for more details). For the present study, I also encoded the data using the same system, but mainly focused on the six strategies of Global Planning, Local Planning, Thematic Planning, Rereading, Rhetorical Refining, and Translating from L1 to L2, for the purpose of testing Hypotheses 5 to 8, which emerged from the pilot study. As exemplified in the Appendix, the participants' protocol accounts were encoded according to the definition of each writing strategy. One chunk of those accounts that consisted of several propositions was judged to be one category if it as a whole matched the definition of a single category.[10]

Before I coded the data, a subset of the data randomly selected from six out of the 56 protocols (11% of the total sample population) was coded by another trained applied linguist using the same categories. Cohen's Kappa value, which excluded the chance value from the two coders' agreement rate, was calculated for the 348 decisions on strategy classifications (Takeuchi, 1989). The Kappa value was 0.84 with a 95% confidence interval of .79 to .88. Having judged that this agreement rate was acceptable, I coded the rest of the data myself.

2.4 Determining the participants' writing styles

I also classified the participants' writing styles according to a revised version of Anzai and Uchida's (1981) writing style classification formulae based on the results of their analysis of 40 Japanese children's use of L1 writing strategies described above. I revised the formulae so that they could properly identify the qualitative differences found in the pilot study between the experts and novices (this part of the analysis was related to Hypotheses 7 and 8). In the pilot study, the experts' global planning appeared to include not only a detailed planning of overall organization, but also the

---

10 For example, in the Appendix, the example of Rhetorical Refining consisted of one phrase and four clauses, but it was treated as one example (or a chunk) of Rhetorical Refining in the analysis. In contrast, the example of Rereading consisted of two clauses, and it was also treated as one example of Rereading.

behavior of seeking the most effective solution based on the assessment of the given working environment (i.e., writing an argumentative composition within 30 minutes to one hour) whereas the novices' global planning included only planning of overall organization. Any previous classification of planning could not be used here because both the experts' and the novices' behaviors were partially similar to some of the planning categories proposed in previous studies, but not completely the same (e.g., the experts' and the novices' behaviors appeared to include both what Hayes & Nash, 1996: 44, called 'non-content planning' and 'content planning'). The formulae revised for the present study determined each participant's writing style according to his/her use of the writing strategies of Global Planning, Thematic Planning, and Local Planning as follows (see Appendix for definitions and examples of these strategies).

Type	Description
A	Write according to detailed global planning of both the content of the text and the most effective way to express it based on the assessment of the given task.
B	Write according to detailed global planning of the content of the text.
C	Write according to rough global planning (thematic planning) and consecutive local planning.

Type A is similar to Bereiter and Scardamalia's (1987) 'knowledge transforming model' (Bereiter & Scardamalia, 1987: 10) in that the writers try to express their ideas in the best possible way according to their goal setting, whereas Type C is partially similar to Bereiter and Scardamalia's 'knowledge telling model' (Bereiter & Scardamalia: 10) in that they make many local plans, but different from the model in that the writer makes a global plan, although it is not very detailed. Classifying the participants' behaviors according to these formulae revealed the qualitative differences among the different groups' writing patterns that might have been missed if these groups were compared only in terms of quantifiable variables (e.g., time spent before starting to write; frequencies of use of the writing strategies)

3 RESULTS AND DISCUSSION

I used SPSS Version 6.1 (SPSS Incorporated, 1994) to perform all statistical analyses. An alpha level of .05 was used for the statistical tests unless it was necessary to adjust it for multiple comparisons.

3.1 Composition scores

Although the participants' composition scores were not directly related to the hypotheses to be tested in the present study, I will present them as sources of information that will be helpful for understanding the results related to the hypotheses. Table 1 shows the means and standard deviations for the composition scores (according to Jacobs et al.'s 1981, ESL Composition Profile) for the experts and the novices I and

II (before and after the two semesters of instruction). As I stated in Note 8, there might have been differential topic effects on the quality of novices I and II's compositions because they wrote for different prompts. Thus, correlations between the quality of novices' I and II's compositions were investigated as a possible manifestation of topic effects. The correlations were .53* for the content subscore, .26 for the organization subscore, .53* for the vocabulary subscore, .44* for the language use subscore, .13 for the mechanics subscore, and .51* for the total score (*p<.05). These correlation values mean that those novices who scored high for Prompt 1 tended to score high for Prompt 2 after the instruction in terms of content, vocabulary, and language use in spite of the topic differences. It can be speculated that the organization and mechanics subscores did not correlate highly with each other because these aspects of compositions were deeply related to the content of the instruction (recall the Participants Section). These results suggest that the topic effects, if they existed at all, might not have been strong enough to affect the novice students' writing ability in a drastically different manner. Nonetheless, to make the analysis on a fairer ground, the total and the content scores, which can be considered to have been the most susceptible to topic effects, were excluded from the comparison between these two groups (see also Hirose & Sasaki, 2000 for this issue).

The experts' total score and the five subscores were all significantly higher than those of novices I even after the alpha level was adjusted to .008 by a Bonferroni correction for multiple comparison (Tabachnick & Fidell, 1996). Similarly, novices I's subscores for organization, vocabulary, language use, and mechanics were significantly higher than those of novices II after the alpha level was adjusted to .0125 by a Bonferroni correction. Although there might have been possible topic effects intervening in the novices I and II comparison, it appears that the novices' composition scores did improve as a whole after the two semesters of process instruction.

Table 1: Effects of expertise (experts versus novices I) and writing instruction (novices I and II)[a] on writing scores.

Variable		Experts n = 12		Novices I n = 22		Novices II n = 22		Expertise df = 32	Instruction df = 21
	Max	M	sd	M	sd	M	sd	t	t
Total	200	187.8	7.1	128.0	19.2	148	14.7	13.1	-5.3
Content	60	56.9	1.9	40.3	5.8	46.0	4.2	12.4	-5.2
Organization	40	37.0	1.8	25.9	4.3	30.9	3.0	10.6	-3.8
Vocabulary	40	37.7	1.8	25.8	4.1	28.8	3.7	11.8	-4.7
Lang. Use	50	46.4	2.4	28.1	5.4	33.6	4.9	13.7	-4.6
Mechanics	10	9.8	0.4	7.9	0.7	8.7	0.6	8.9	-5.4

Underlined scores are statistically significant at p<.008 (adjusted by a Bonferroni correction for multiple comparison). [a] These scores are sums of the two raters' scores. [b] Novices I and II are the same students before and after the two semesters of instruction.

The mean scores of the experts and the novices were similar to those of the pilot study, which suggests that the sample populations used in these two studies were similar in terms of their writing ability. The only difference between the present study results and those of the pilot study was that unlike in the pilot study, there were significant differences in the organization, vocabulary, language use, and mechanics subscores between novices I and novices II. In the pilot study, these subscores also appeared to be different between novices I and II, but the differences were not statistically significant. Although the novice group might have improved their writing ability as a result of other factors (e.g., other English classes) because there was no control group with which to compare the true effect of the instruction, the results of the present study at least suggest that process writing instruction like that given in the present study has the potential to improve students' L2 writing (see also previous studies where instruction improved at least some aspects of learners' writing ability; e.g., Hirose & Sasaki, 2000; Shaw & Ting-Kun Liu, 1998).

3.2 Fluency

The participants' writing fluency was measured in terms of the two indices of mean total number of words written in the text (quantity) and mean number of words written per minute (speed; see Table 2). The t-test results indicate that the experts wrote significantly longer texts, and wrote significantly faster than the novices ($t = 5.52, p =.000$ for the quantity, $t = 3.45, p = .002$ for the speed; the alpha level was adjusted to .025 by a Bonferroni correction for multiple comparison). It is not always true that more skilled writers tend to write longer texts than their less skilled counterparts (see Reid, 1990; Silva, 1993), but the experts in the present study did write longer texts as well as faster than the novices. Thus, Hypothesis 1 presented in the Introduction Section was confirmed: The experts were significantly more fluent than the novices. As was noted in the pilot study, and as we will see in the Writing Strategies section below, the fluency difference between the experts and the novices appears to have been partially caused by the fact that the novices had to stop to translate their ideas into English more often than the experts, which was probably caused by the L2 proficiency difference between these two groups.

In contrast with such a difference between the experts and the novices, the differences between novices I and II were not statistically significant ($t = -.93, p =.37$ for the quantity, $t = 1.54, p =.14$ for the speed by matched t-tests; the alpha level was adjusted to .025 by a Bonferroni correction). This might be because the prompts given to novices I and II were different, but considering the fact that the total scores given to their compositions were significantly correlated ($r = .51$; see Section 3.1), we can also suspect that the novices did not improve writing fluency as a result of the two semesters of process writing instruction. It appears that Hypothesis 2 concerning the novice writers' improvement in terms of fluency was also confirmed. This result thus suggests that process writing instruction of two semesters may not help students improve their writing fluency.

Table 2: *Effects of expertise (experts versus novices I) and writing instruction (novices I and II)^a on fluency (number of words, and words per minute: wpm).*

Variable	Experts n = 12		Novices I n = 22		Novices II n = 22		Expertise df = 32	Instruction df = 21
	M	sd	M	sd	M	sd	t	t
Words Written	243.6	88.41	95.7	38.49	103.6	25.76	5.52	-.93
WPM	6.04	2.23	3.61	1.81	3.10	1.03	3.45	1.54

Underlined scores are statistically significant at p<.025 (adjusted by a Bonferroni correction for multiple comparison). ^{a} Novices I and II are the same students before and after the two semesters of instruction.

3.3 Time spent before starting to write

Table 3 presents the mean ratios of the time the participants spent before starting to write to the total time they spent writing the whole composition. Because the total time spent by each participant varied, such a ratio rather than the time itself was used for comparison. For the first composition, the experts on average spent 23% of the total time for the initial planning, whereas the novices spent only 9%. After the instruction, the novices spent 20% of the total writing time for their initial planning. Because ratios of such time values tend to have skewed distributions, and because these values can acquire more normal distributions under logarithmic transformation (Iwahara, 1997), the normality of the distributions of these time ratios and those of their corresponding logarithms were compared by the Shapiro-Wilks normality test (SPSS Incorporated, 1993). Because the values after logarithmic transformation were judged to have more normal distributions,[11] these values were used for subsequent *t*-tests. The *t*-test results show that there was a significant difference between the experts and novices I, and between novices I and II (see Table 3).

Thus, the third hypothesis regarding the difference between the experts' and novices' planning time was confirmed. Although there might be topic effects involved again, the fourth hypothesis regarding the differences between novices I and II appears to have been also confirmed. The experts spent a longer time before starting to write than the novices, and the novices learned to spend a longer time before starting to write after the instruction. Recent research (e.g., Hayes & Nash, 1996) indicates that planning time may not be as strongly related to L1 writing ability as was once believed, but it appears that in this particular case of L2 writing, planning time *is* related to L2 writing expertise as many previous studies have reported (e.g., Bereiter

---

11 *One of these time-ratio values before logarithmic transformation had a distribution that significantly deviated from a normal one, but after the transformation none of the values had distributions that were significantly different from a normal distribution.*

& Scardamalia, 1987; Krapels, 1990). Moreover, the present results suggest that instruction may also influence students' planning time.

Table 3: *Effects of expertise (experts versus novices I) and writing instruction (novices I and II)^a on starting time (proportion).*

Variable	Experts n = 12		Novices I N = 22		Novices II n = 22		Expertise (df = 32)	Instruction (df = 21
	M	sd	M	sd	M	sd	t	t
Starting time	.23	.14	.09	.19	.20	.15	4.13	-3.07

Underlined scores are statistically significant at p<.01. ^{a} The t values were calculated after the ratios were transformed to their logarithms to normalize their distributions.

3.4 Writing strategies

Table 4 shows the mean total numbers of the four strategies (token frequency) used by the participants in the present study (ranging from 63.5 to 87.8). These strategies are related to Hypotheses 5 and 6. The numbers presented in Table 4 were generally larger than those used by the participants in the pilot study (ranging from 32 to 36.25). This might be partly because the experts in the present study spent a longer time (M_{time} = 41 minutes 10 seconds) completing the compositions than their counterparts in the pilot study (M_{time} = 31 minutes 11 seconds). However, this explanation does not apply to the novices because the novices in the present study spent a shorter time (M_{time} = 29 minutes 26 seconds) completing the compositions than their counterparts in the pilot study (M_{time} = 35 minutes 22 seconds). Furthermore, unlike the findings in the pilot study, the experts in the present study appear to have used more strategies than the novices, whereas the novices appear to have used more strategies after the instruction than before (in the pilot study, the experts and novices used similar numbers of strategies whereas novices II used only half as many strategies as novices I).

Because the total numbers of strategies used by each group thus varied, the mean ratios of individual strategy tokens (i.e., frequencies of the strategies used by the participants; if the same strategy was used three times, it was counted as three instead of one) used by the members of each group out of the total number of strategy tokens used by them were employed for comparison.[12] As in the case of the time spent before starting to write, such ratio values tend to have skewed distributions, and these values can acquire more normal distributions by logarithmic transforma-

---

12 *For example, Expert 2 reread 11 times while writing the composition. Because he used a total of 45 tokens of strategies for completing the composition, 11 out of 45, i.e., 0.24 was used as his value for the Rereading strategy. After the individual calculations were completed, those values were averaged for each group for each strategy.*

tion (Iwahara, 1997). Consequently, the normality of the distributions of these time ratios and those of their corresponding logarithms were compared by the Shapiro-Wilks distribution-normality test (SPSS Incorporated, 1993). Based on the judgment that the values after logarithmic transformation were more normally distributed, these values were used for subsequent t-tests.[13] Because t-tests were applied to the same sample population four times, the alpha level was adjusted to .0125 by a Bon-ferroni correction for multiple comparisons. The t-test results show that, unlike the findings in the pilot study, there was no difference between the experts and novices I, or between novices I and II, in terms of Rereading. That is, no differences were found in the ratios of rereading among the three groups. Because the three groups used different total numbers of strategies, this does not mean that they used similar numbers of rereading strategies. Rather, it implies that all groups tended to pay a similar amount of attention to rereading.

Whether 'rereading' is related to the quality of writing has remained an unsolved issue according to the results of previous L1/L2 writing studies. For example, Levy and Ransdell (1995) reported that L1 English writers of better compositions tended to allocate more time for 'reviewing' (Levy & Ransdell: 773) than the other writers. In contrast, Anzai and Uchida (1981) reported in their cross-sectional study of L1 Japanese writers that the participating elementary school children's reading behavior increased in terms of frequency from the second grade to the fifth, but decreased in the sixth grade. In contrast, Breetvelt, Van den Bergh, and Rijlaarsdam (1996) intro-duced the idea of 'time-dependent changes' (Breetvelt et al., 1996: 19) into their investigation of cognitive activities such as rereading while writing in L1 Dutch. Breetvelt et al. reported that a positive correlation between rereading and text quality 'increases during the first half hour, and slowly decreases afterwards' (Breetvelt et al., 1996: 16). As for L2 writers, Zamel (1983) reported that 'All of the students, their writing skill notwithstanding, reread' (Zamel, 1983: 173). However, the more skilled writers tended to reread various aspects of the text ranging from the local text just written to entire paragraphs whereas the least skilled writer reread shorter 'chunks of discourse' (Zamel, 1983: 173). Furthermore, Cumming (1989) found that some of the less skilled L2 writers 'made conscious choices not to monitor their writing, by resolving not to edit or proofread their written compositions' (Cumming, 1989: 114), whereas some of the more skilled writers 'reviewed their previous text, rereading it every few minutes' (Cumming, 1989: 115) when they were concentrat-ing on how to express their ideas. It appears that both more skilled and less skilled writers tend to reread while writing but that they often reread different parts of the text for different purposes at different times. The three groups in the present study might have been differentiated if 'what,' 'why,' and 'when' they reread had been investigated. Such an investigation should be conducted in future studies.

---

13 The values for the experts' Local Planning, novices I's Rhetorical Refining, and novices II's Local Planning and Rhetorical Refining had distributions that significantly deviated from that of a normal distribution. After the logarithmic transformation, none of the values had such abnormal distributions.

Table 4: Effects of expertise (experts versus novices I) and writing instruction (novices I and II)[a] on writing strategies: Mean ratios of the four writing strategies out of the total number of strategies used.

Strategy	Experts n = 12		Novices I n = 22		Novices II n = 22		Expertise (df = 32)	Instruction (df = 21)
	M	sd	M	sd	M	sd	t	t
Rereading	.14	.07	.15	.06	.14	.05	-.09	.021
Rhetorical refining	.19	.13	.03	.03	.03	.04	5.71	-.15
Local planning	.04	.04	.11	.07	.04	.05	-3.67	5.59
Translating L1 to L2	.32	.18	.53	.14	.53	.12	-2.59	.02
Total # of strate-gies	87.8	28.4	63.5	17.2	71.2	14.1		

Underlined scores are statistically significant at p<.0125. [a] The total numbers of strategies means the total number of tokens of the strategies used. Thus, if the same strategy was used three times, it was counted as three instead of one.

In contrast with the participants' rereading behavior, there were significant differ-ences between the experts and novices I in terms of Rhetorical Refining and Local Planning. First of all, the experts used Rhetorical Refining significantly more often than the novices. In Rhetorical Refining, writers try to choose the most appropriate L2 expressions to convey their meanings, as in Example 1 (the participants' ac-counts in all examples below were originally given in Japanese except for the Eng-lish words in quotation marks, and the accounts were translated into English by the author):

Example 1:

I was wondering which expression to choose, 'my feeling was not important' or 'my feeling did not matter.'

(Expert 11)

Because the writers must be proficient enough in English to refine their expressions in such a way, this strategy is probably related to the experts' high English profi-ciency. Roca de Larios et al. (1999) reported a similar phenomenon where the more advanced group of L2 writers spent more time on 'ideational/textual restructuring' (Roca de Larios et al., 1999: 33), including the writers' search for stylistically better expressions, than the less proficient group, who spent more time seeking ways to compensate for their limited L2 proficiency. Furthermore, such a characteristic of advanced writers may be related to the common core of composing competence shared by L1 and L2. For example, reporting that L2 writers with L1 writing exper-tise also demonstrated a similar behavior of paying special attention to word/phrase

choice when they wrote in L2, Cumming (1989) speculated that such a behavior may be a manifestation of what Gardner (1983) called 'linguistic intelligence' (Cumming, 1989: 73) that transcends L1 and L2 differences.

As for Local Planning, the novices stopped for local planning significantly more often than the experts. This behavior is related to the writing style differences between these two groups (see the Writing Style section below). Furthermore, the novices stopped to translate their ideas into L2 more often than the experts ($t = -2.59$, $p = .024$, $ns.$): About half of the novices' strategies were devoted to translating generated ideas into English whereas only one-third of the experts' strategies involved translation. Although the difference between the two groups was not statistically significant, we may at least say that the novices had a tendency to stop for translation more often than the experts because the obtained p value was close to the determined alpha level (.0125) for statistical significance, and because applying a Bonferroni correction tends to become too conservative as the number of comparisons increases (Nagata & Yoshida, 1997). As was speculated in the pilot study, this result is probably related to the difference in L2 proficiency between the experts and the novices. Furthermore, it can be speculated that the novices' spending so much time on translation would be a factor limiting their fluency. A similar relationship among fluency, L2 proficiency, and mental translation was also reported in Sasaki and Hirose (1996) based on a sample population of comparable EFL backgrounds.

In contrast, after the two semesters of instruction, novices II were still not able to rhetorically refine their expressions very often, and they still had to stop to translate often. However, they made significantly fewer local plans while writing. Although we have to take possible topic effects into consideration here again, we could hypothesize that the novice students would use similar strategies for the similar argumentative topics given the significantly high correlations between the quality of novices I and II's compositions (recall Section 3.1). Thus, Hypotheses 5 and 6 were largely confirmed, except for the case of rereading and translating from L1 to L2. The experts stopped to refine their expressions more often than the novices, whereas the novices stopped to make more local plans. Furthermore, although the difference was not significant, the novices tended to stop to translate from L1 to L2 more often than the experts. After two semesters of instruction, the novices made fewer local plans. However, they still had to stop to translate as often as before. Here again, it is suggested that students can change their use of writing strategies as a result of instruction, which accords with the results of previous L1 and L2 studies (e.g., Anzai & Uchida, 1981; Bereiter & Scardamalia, 1987; Kobayashi & Rinnert, 1998; Cohen, Weaver, & Li, 1998).

3.5 Writing styles

Table 5 presents the relative distribution of different writing styles among the three groups. As explained in the Method section, these three writing styles are related to the three strategies of Global Planning, Local Planning, and Thematic Planning. Most of the experts were classified as Type A (Write according to detailed global planning of both the content of the text and the most effective way to express it

based on the assessment of the given task) or Type B (Write according to detailed global planning of the content of the text), and most of the novices were classified as Type C (Write according to rough global planning and consecutive local plans). After the two semesters of instruction, most novices changed to Type B. Although here again we must consider possible topic effects for the novices I and II comparison, we can assume that the novices would be likely to use similar writing styles for the similar argumentative topics (Prompts 1 to 3), especially when novices I and II's composition quality was significantly correlated (recall Section 3.1).

The results of a chi-square test and McNemar's test indicated that there was a significant difference between the experts and the novices, and novices I and II ($\chi^2 = 22.33$, $p < .001$ for experts/novices; z for McNemar's test = 3.90, $p < .001$ for novices I/II). Thus, Hypotheses 7 and 8 were both confirmed: The experts tended to plan detailed overall organization whereas the novices tended to make a less detailed plan; after the two semesters of instruction, the novices learned to do global planning of the content of the text, but their global planning was different from that of the experts in that the novices did not plan how to implement the content in the most effective way. It is interesting that the results related to Hypothesis 8 concerning the effects of instruction concur with those of previous developmental product-oriented studies in that instruction can direct learners' attention to more global aspects of writing, although the aspect studied varied (e.g., organization skills for Berman, 1994; revision skills for Kobayashi & Rinnert, 1998). On the other hand, the present results also suggest that such shift of the learners' attention patterns over a relatively short period of instruction time is probably related to Type B writing style, but not to Type A writing style, which seems not to have been acquired by the present novice participants after two semesters.

Table 5: Relative distribution of different writing styles (percentages).

Style[14]	Experts $n = 12$	Novices I $n = 22$	Novices II $n = 22$
Type A	58.3	0	0
Type B	33.3	13.6	90.9
Type C	8.3	86.4	9.1

$\chi^2 = 22.33$ ($p < .001$) for Experts/Novices I; z for McNemar's test = 3.90 ($p < .001$) for Novices I/II.

14 Styles: A: Write according to detailed global planning of both the content of the text and the most effective way to express it (cf. Hayes & Nash's, 1996 content and non-content planning); B: Write according to detailed global planning of the content of the text; C: Write according to rough global planning (i.e., thematic planning) and consecutive local planning (see Appendix for the descriptions of the different types of planning).

As in the pilot study, the experts' global planning can be characterized by two types of behaviors that were absent from the novices' global planning. First, as can be seen in Example 2, the expert writers assessed the characteristics of the given task at a relatively early stage of planning.

Example 2:

Because this is a letter to a readers' column in a newspaper, I was thinking about using my own experiences [to support the idea that students should not be forced to wear uniforms][15].

(Expert 11)

Secondly, the expert writers sought the most appealing way of presenting the text to the probable readers while searching though available resources in long term memory as in Example 3:

Example 3:

Just presenting several reasons to support my opinion here *would not be so interesting to the readers*. But my directly saying that school uniforms should be abolished *would not be appealing, either*. So, I wondered if there are any other angles to look at the same thing. Then, although it is not entirely a new perspective, it occurred to me that *the voices of the students who are actually forced to wear uniforms would be the most convincing to the readers*.

(Expert 9)

In contrast to these behaviors observed in the experts' global planning, none of the novices showed such characteristics, even after the two semesters of instruction. They simply planned what they were going to write for overall organization as in Example 4 (see Flower *et al.*, 1992 for a similar finding):

Example 4 (Novice 6):

Researcher	Writer
What were you going to write before you started to write?	Because I agree with the opinion [introducing English as an elementary school subject], I was going to write it.
Have you decided what you are going to write in the end?	Yes.
What are you going to write?	I am planning to conclude the composition by writing that we should use more and more English these days.
I see. Have you decided what you are going to write in the middle?	Yes[16] (showing the researcher the draft he had written).

---

15 The insertions in brackets are supplementary explanations given by the author.
16 The draft contained the following seven Japanese sentences [translated by the researcher]: English should be taught in elementary schools. We should make children get used to English from the time they are small. We should make children get familiarized with English. We should teach the joy of learning English --> They can continue to study English when

Have you? Oh, this is your draft?	Yes.
Have you decided everything you are going to write in the middle?	Yes.
Can you briefly explain what you are planning to write?	I thought of two reasons, so I was planning to elaborate each reason by added one or two sentences. I was planning to write these two reasons so that they would lead to the concluding sentence.
I see. Did you think about any other thing?	No.

Regarding the experts' global planning, it is also noteworthy that, unlike the findings in the pilot study, five out of the 12 experts (41.7%) were not concerned about 'the most effective way to express' the content they had planned in the beginning, and thus did not belong to Type A. One of them did not even do any global planning before starting to write. He simply planned he was going to write in the beginning of the composition (*i.e.*, local planning), and started to write. However, a closer look at their accounts of what they did *after* starting to write reveals that they also made great efforts to assess the characteristics of the given task and shape the text to meet the probable readers' needs and purposes. The experts' frequent use of rhetorical refinement strategies (see Table 4) is one such example. In addition, in their local planning, these experts tried to search for the best way to express their ideas according to their assessment of the characteristics of the given task as in Example 5:

Example 5:

[Here] I was thinking very hard about what I should write next (Researcher: You mean what you were going to write?) I wondered what kind of people would support the idea 'that the uniformity and discipline should be maintained in school.' (Uh, huh.) That means, I thought very hard here, well, that this idea will be convenient for the people who are on the side of educational administration. (Uh, huh.) So, I wondered if I should say something like the Ministry of Education or the Board of Education, but I don't know how I can express such ideas...

(Expert 6 who was categorized as belonging to Type B)

None of the novices exhibited such behaviors while writing even after the two semesters of instruction.

Cumming (1989) reported similar phenomena where the L1 writing experts organized their ideas 'through advance planning or *emergent planning* (emphasis added)' (Cumming, 1989: 114). That is, some experts in Cumming's study planned the overall organization of the composition before starting to write just like many of the experts in the present study, whereas the others 'planned each composition as it was emergent on the page' (Cumming, 1989: 115) while paying attention to the overall organization that was emerging. In contrast with the novice writers, however, they never lost control over their writing or got stuck. They kept writing by 'amending, adding to, and reordering their phrases frequently, reconsidering how best to

---

they become high school students --> English will be needed in the future international society. Therefore we should study English more.

state what they were most recently deciding they wanted to state' (Cumming, 1989: 116). We can find some similarity between these emergent planners and those experts in the present study who did not do elaborate pre-writing planning. It can thus be speculated that although efforts to find the most effective way to express their ideas were often observed in the experts' pre-writing planning in the present study, such effort or behavior itself, rather than the pre-writing global planning, might be at the innermost core of writing expertise. Moreover, if I am allowed to speculate further, such behavioral characteristics of the expert writers might be a manifestation of a more general cognitive ability, or what Bachman (1990) called 'strategic competence' (Bachman, 1990: 98), which enables a language user whatever the skill they may be using, to assess the context, plan an appropriate action, and actually execute it in the most effective way. Here writing expertise might in fact be a part of strategic competence rather than an independent ability.

Finally, one last qualitative finding related to the participants' writing style was that, as in the pilot study, both global and local planning appeared to guide or monitor the participants' writing process, whether they were the experts or the novices. I did not use '(Global or Local) Plan Monitoring' as an independent writing strategy category in the present study because the pilot study results revealed that the phenomenon of plan monitoring was observed in many identified strategies as a kind of 'secondary' activity engaged in by the participants. See Examples 6 through 8 (see also Appendix C of Sasaki, 2000 for the details of these strategies):

> Example 6 ('Local planning' monitoring of Novice 7's writing, which was manifested in her 'Translating from L1 to L2')
>
> Here I wanted to write [in English] 'you can wear uniforms only when you are students'
>
> Researcher: Yes.
>
> But I couldn't think of an English word to express the meaning of 'to wear.'
>
> Researcher: Is that why you stopped writing here?
>
> Yes, I was trying to find the right word.

In Example 6, what Novice 7 was actually doing is trying to translate her idea 'you can wear uniforms only when you are students,' but by doing so, she was simultaneously implementing the local plan she had previously made. In other words, that local plan was guiding her translating activity here.

In contrast, in Example 7, Novice 17 originally planned to say 'Students can wear whatever they want to at school,' and retrieved that idea from his rough draft in Japanese. When it came to the time he actually put that idea on the paper, however, he wanted to revise it by adding the clause 'Some people think that' before expressing it in English. Here, we can see that Novice 17's local plan was giving a general direction to his writing by having been retrieved although it was implemented in a more elaborated way.

> Example 7 ('Local Planning' monitoring of Novice 17's writing, which was manifested in his 'Plan Retrieving')
>
> Researcher: Were you reading the top part of your rough draft written in Japanese?
>
> Yes, I looked at the original sentence (in the draft) 'Students can wear whatever they want to at school [originally said in Japanese]' and I was trying to change the sentence to 'Some people think that students can wear whatever they want to wear (originally said in Japanese).'

Finally, in Example 8, Expert 6 was wondering if he should make the first paragraph longer, but realized that doing so might ruin the overall coherence that had been planned in the beginning. Here we can see that Expert 6 was monitoring his writing process by referring to his global plan.

> Example 8 ('Global Planning' monitoring of Expert 6's writing, which was manifested in his 'Rereading')
>
> I was reading the end of this paragraph, and I was wondering if I should end this paragraph like this.
>
> Researcher: I see.
>
> I felt I should write more [according to my plan], but I thought, 'Wait a second, it would be disastrous [the whole composition would lose balance] if the first paragraph gets longer than this!'
>
> Researcher: Hmmm. I see.
>
> So, I was thinking about concluding this paragraph with this last sentence.

A similar phenomenon of using plans to monitor or guide the writing process has been reported in studies of both English and Japanese L1 writing (Anzai & Uchida, 1981; Faigley, Cherry, Jolliffe, & Skinner, 1985; Uchida, 1990).

3.6 Summary of the characteristics of the participants' writing processes

The findings of the present study can be summarized as follows.
1) The experts wrote longer texts at greater speed than the novices.
2) After two semesters of process writing instruction, the novices did not improve their writing fluency. However, their writing ability seems to have improved.
3) While writing, the experts stopped to refine their expression more often than the novices, whereas the novices stopped to make local plans more often. The novices also tended to stop to translate from L1 to L2 more often than the experts.
4) After two semesters of instruction, the novices made fewer local plans than before, but they stopped to translate from L1 to L2 as often as before.
5) The experts spent a longer time before starting to write, planning both detailed overall content of the text and the most effective way to express the content, whereas the novices spent a shorter time making a less detailed plan.

6) After two semesters of instruction, the novices learned to spend a longer time before starting to write, planning detailed overall content of the text.

7) The experts' global planning and subsequent writing was based on their assessment of the characteristics of the given task for successfully achieving the task. Such behavior appears to be a manifestation of a core aspect of writing expertise that cannot be acquired over a short period of time.

8) Both 'Global Planning' and 'Local Planning' guided the participants' writing processes, whether they were experts or novices.

Figure 1 presents summary illustrations of these results. These flow-chart illustrations for the three types of EFL writers (experts, novices I, and novices II) are not presented as faithful records of these writers' micro-level writing processes (cf. Zimmerman's 2000 'formulating model'), but as illustrations of the findings and speculations that resulted from the present study. In other words, these models (if I can call them 'models') are crude, but representative of the differences found among the three different groups in the present study. The illustrations presented in Figure 1 do not include the 15 other writing strategies (e.g., Conclusion Planning, Generating Naturally Generated Ideas, Mechanical Refining, Local Text Evaluation) that were included in the pilot study because the differences among the participants in terms of these 15 strategies were not tested in the present study. Nevertheless, I believe that the Figure 1 illustration will help the readers to better understand the differences among the different groups of writers, which were supported (if not confirmed) by the results of the present study.

In Figure 1, experts spend a significantly longer time for pre-writing planning, where they plan both the overall content of the text and the way to express it in the most effective way based on assessment of the characteristics of the given task. This behavior is a manifestation of writing expertise, which may transcend both L1 and L2 writing, and which may be part of what Bachman (1990) called 'strategic competence.' After they complete the global planning, they usually make several local plans to achieve the global plan. When they make these local plans, they also tend to search for the most effective way to express their ideas. This is also considered to be a manifestation of writing expertise. After making local plans, experts spend time refining their expressions before producing the following text output, which subsequently leads to retrieval of the local plan for the following content. Such a procedure may be interrupted by occasional translation of the generated ideas from L1 to L2. This process continues until the writers feel that they have finished writing (whether or not they are satisfied with the whole written product), with the original global planning guiding the entire writing process (i.e., they use their global plans to monitor their writing process).
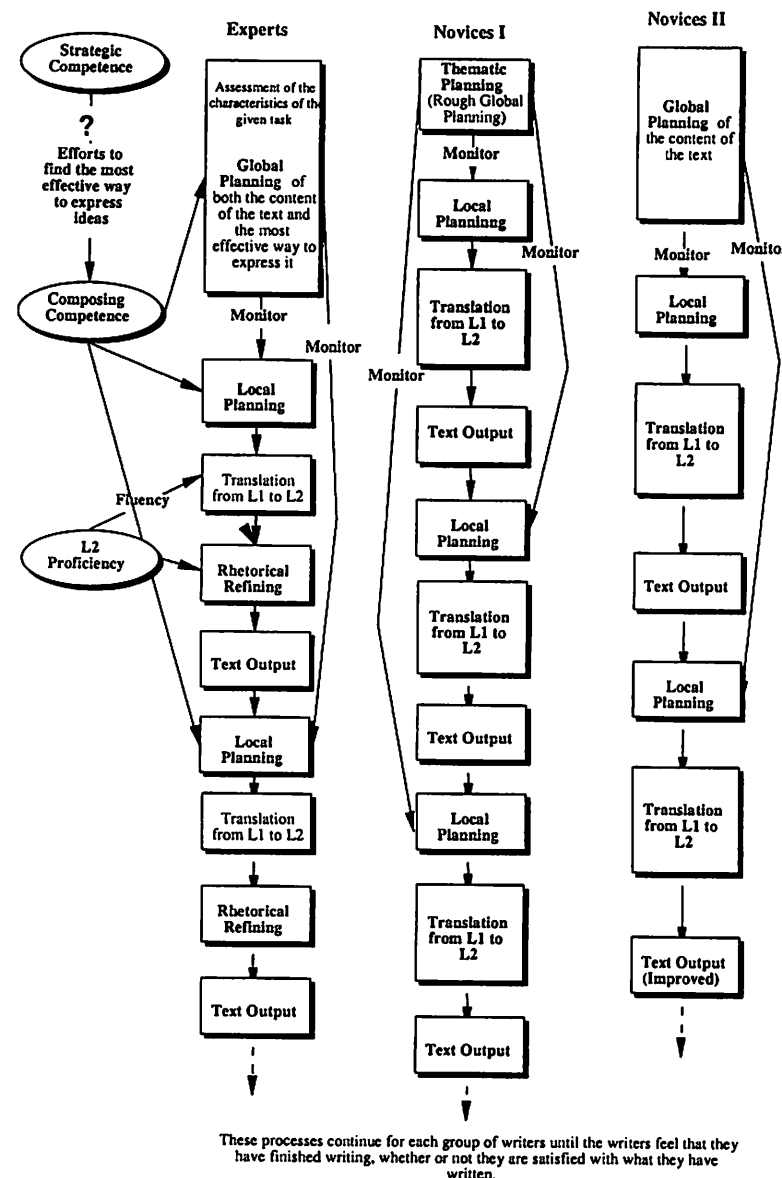

These processes continue for each group of writers until the writers feel that they have finished writing, whether or not they are satisfied with what they have written.

Figure 1. Flow-chart summary illustration of the present study results. See note next page.

Note. In Figure 1, each writing strategy is enclosed in a square. The unidirectional arrows indicate one-way flow from one writing process to the next. The length of the straight lines with arrows representatively suggest the relative length of time which the process enclosed in the upper square might take: the longer the arrows, the longer it might take to complete the process. The strategies written in boldface are the ones used often, and the strategies written in standard typeface are the ones used less often.

In contrast, novices I without any process writing training, make a thematic plan or a rough global plan followed by a local plan. In order to realize this local plan, the novices have to spend time translating the generated ideas into L2 because of their low L2 proficiency. After managing to produce one piece of L2 text, they make a second local plan to be translated into L2. Probably because so much of their attention is paid to the translating activity, they usually cannot spend much time refining their expressions. They are satisfied as long as they can translate what they want to express. Although the original thematic plan guides their writing processes, they have to make a relatively detailed local plan every time they run out of related ideas about the topic, which prevents them from writing longer and faster. This continues until they feel that they have nothing more to say. Finally, novices II after two semesters of process instruction have learned to make a detailed global plan of what they should say about the given topic. Thus, they no longer have to stop to make so many local plans every time they finish expressing an idea about the topic. Probably because the original global plan orchestrates their writing process more effectively than before the instruction, they can now produce better-quality writing with improved coherence and organization. However, the period of two semesters is not long enough to improve their L2 proficiency, and they still have to spend much time on translation. They still cannot refine their expressions very often, and they still cannot write faster or produce longer texts. In order to attain the writing expertise exemplified in the experts' flow-chart, the novices will need 'consistent practice in a variety of similar contexts to the point of proceduralization or automaticity' (Grabe & Kaplan, 1996: 129) through many years of experience. They may also lack some strategic competence or 'linguistic intelligence' (Gardner, 1983: 73), which most likely cannot be acquired through instruction. The extent to which such writing expertise can be taught is an issue that definitely should be investigated in future studies.

4   CONCLUSION

Using multiple data sources, this chapter investigated the cognitive processes of Japanese EFL writing experts and novices while writing an argumentative exposition in English. Although the novices' within-subject comparison might have been affected by possible topic effects, the study is still significant in that it incorporated both cross-sectional and longitudinal investigation of sample sizes that were large enough to allow the hypotheses to be tested by statistical procedures. However, the results of the present study cover only part of the complex mechanism of L2 writing

processes (see, for example, Grabe & Kaplan's, 1996: 226, model of writing as communicative language use as a more comprehensive model). Moreover, the obtained representative patterns of different types of EFL writers' writing processes are still preliminary findings, and thus are subject to revision and elaboration in light of the results of future studies.

First, the present study should be complemented by qualitative studies that examine the details of EFL writers' individual writing processes, which were not captured in this empirical confirmatory study. For example, how each participant within a particular group in the present study used different types of strategies (a total of 21 types) should be individually followed, and compared with the ways the other participants used these strategies in terms of content and order. Adopting the notion of 'writing signatures,' suggested by Levy and Ransdell (Levy & Ransdell, 1996: 158) to refer to the writing processes that are characteristic of individual L1 writers, should be useful in such investigation. Furthermore, when such individual comparisons are made, it might also be possible to investigate the changes in the writers' psychological state in terms of new knowledge formation (*e.g.*, whether the writers feel that they have created new ideas as a result of their writing) as advocated by Galbraith (1999). Detailed qualitative studies exemplified by Levy and Ransdell and by Galbraith would be useful for filling the gaps in the models presented in Figure 1.

Secondly, future studies should investigate affective and emotional factors that were not treated in the present study. Affective factors such as motivation or attitude have not been given much attention in previous writing models, but have begun to be recognized as crucial for understanding writing behavior (*e.g.*, Hayes, 1996). In Sasaki and Hirose (1996), a product-oriented study that targeted a similar sample population of Japanese EFL writers, we also found that confidence in L2 writing for academic purposes was one of the factors that distinguished the good and weak writers. Adding the results of future studies that examine *how* such affective factors influence the actual writing processes would make the target writing process models more comprehensible.

Finally, the model should also go beyond the cognitive domains over more contextualized domains in order to make it more realistic. If we assume that writing cannot be conducted in a social vacuum and that writing is 'social construction' (Cumming, 1998: 61), we cannot ignore the social/cultural contexts where the writing takes place. If we hope to extend the scope of our research in that direction, an ethnographic approach such as the one advocated by Grabe and Kaplan (1996) might be a good place to start. Thus, we should ask 'the basic question: *Who writes what to whom, for what purpose, why, when, where, and how?*' (Grabe & Kaplan, 1996: 203). For example, the present study can be replicated with different types of writing under different conditions. Using other topics and types of tasks, such as writing a complaint letter or a project report[17] to be completed within a longer span of time with the help of dictionaries, may reveal somewhat different processing phases. The study can also be replicated with other sample populations that have different L1 or L2 and varied educational/cultural backgrounds. When we conduct

---

17 These examples were identified by Tannenbaum, Rosenfeld, and Breyer (1996) as writing tasks that might be required in the field of English for international communication.

such studies, we should be careful to 'situate writing socially as part of their explanations' (Grabe & Kaplan, 1996: 215) while also paying attention to the writer's own (*i.e.*, *emic*) perspective (Ramanathan & Atkinson, 1999). Incorporating the results of these studies will surely contribute to enriching the content and usefulness of the comprehensive L2 writing process model we can eventually hope to build.

ACKNOWLEDGEMENTS

APPENDIX

SIX ENCODED STRATEGIES FOR PARTICIPANTS' RETROSPECTIVE ACCOUNTS OF THEIR WRITING PROCESSES IN THE PRESENT STUDY

These six strategies are among the 21 strategies used for the pilot study. Only six of these 21 strategies were investigated in the present study in order to test the eight hypotheses that were formulated as a result of the pilot study. The participants' accounts in all examples were originally given in Japanese except for the English words in quotation marks.

1   GLOBAL PLANNING

Definition: Detailed planning of overall organization
Example: I am planning to write that there are several merits for students to have uniforms, and that the rule of wearing uniforms reflects the fact that people should obey rules of the society they are in. I am for the opinion that students should wear uniforms in the sense that by wearing them students will learn that it is natural for the school to have some rules because it is the epitome of a real society. (Researcher: Have you decided what you are going to write in the end?) In the end, I was thinking about summarizing the second half of what I said in the beginning. (Have you decided what you are going to write in the middle?) Oh yes, in the middle, I will talk about the purposes of schools. There are usually two purposes for schools. (Yes.) Well, to study and to learn how one should act in the real society in the epitome of a society. A school is a little society, you know. It is a commonplace people share. So, by writing these purposes, from the viewpoint of these purposes, to study and to learn how to act in a common place, for the second purpose, in a common place, there are always some rules, and wearing uniforms is a typical example of such rules. So, I start with the purposes of schools, and according to the purposes, it is OK to have a rule of wearing uniforms. That will be my supporting sentences. Then next, I thought how this issue of whether or not students wear uniforms should be discussed, there must be various ways to look at this issue, but I will try to present my own opinion about the merits and demerits [of wearing uniforms], ah, for example, [wearing uniforms is] economical, and the students don't have to think about their own fashion. I was thinking about presenting some merits. As for the demerits, [wearing uniforms] is not individualistic, emphasizing the standardized uniformity. (Yes.) These things have been said about uniforms, so I can discuss such demerits. And after that, although [wearing uniforms] is actually non-individualistic, but if the students want to emphasize their own individuality, they can do it by modifying their uniforms in spite of the existence of the uniforms. (Yes.) So, the demerits are not so strong. (Yes.) I was trying to write something like that. That will be short, but after that, I was thinking about summarizing the content before the end. (Expert 7).

2   THEMATIC PLANNING

Definition: Less detailed planning of overall organization.
Example: I was going to write that senior and junior high schools should have uniforms. (I see. Have you decided what you are going to write at the end?) Not particularly. (Not yet?) No. (Have you decided what you are going to write in the middle?) Not yet. (Novice 7).

3   LOCAL PLANNING

Definition: Planning what to write next.
Example: In the beginning, I wondered what I should write here. (Novice 2).

4   RHETORICAL REFINING

Definition: Refining the rhetorical aspect(s) of an expression
Example: ... the idea of 'both sides of the coin,' (Hmmm.) I first wondered if I should make it 'either side' (of the coin), then I thought it would be OK to make it 'both sides of the coin' after all, or something like that ... (Expert 5)

5   TRANSLATING FROM L1 TO L2

Definition: Translating the generated idea into English.
Example: I was wondering how I could express in English the idea that students should wear uniforms. (Novice 5)

6   REREADING

Definition: Rereading the already produced sentence.
Example: I was reading the sentence I just wrote. (Novice 19)