Chapter 1 : Introduction

Where Did the Topic Come From and How Did the Research Develop?

The recast was the topic of my thesis research. It is a popular type of oral feedback that the native speaker (NS) and the English as a second language (ESL) teacher provide when they engage in a conversation with a non-native speaker (NNS). An anecdote about the NS recast relates to my own experience. Once I had a conversation with a NS friend. I talked about my day: "I was *thinking to* explore Old City today." My friend reacted to me and said something. I continued, "I was thinking to go to the City Hall, and I did. There was an interesting tour." Then, my friend once again reacted to me; I heard her repeating "thinking." I wondered why she repeated a verb; she could have acknowledged what I said by saying "uh-huh" or "yeah." A few more turns continued; then, finally, she looked at me and said, "Toshiyo, you've been saying 'thinking to', but it's wrong! It's 'thinking OF'. 'I was thinking OF doing so-and-so!'" "Oh!" In fact, she had been recasting, "thinking of," but I did not notice that her utterances were recasts until she gave me the overt colloquial correction. That was my first encounter with the recast. Since then, I have heard many NS friends giving me countless recasts in various situations, and I have found them very useful.

The recast was first introduced to me as an academic topic in a second language acquisition (SLA) course in graduate school. The course, taught by an interaction SLA researcher, provided me with an extensive theoretical account for the recast being a "promising" form of feedback. However, I soon found that the topic was controversial linguistically, psycholinguistically, and pedagogically. The most notable debate was over the efficacy of the recast in relation to its pedagogical validity: i.e., studies by Lyster (1998b) and Mackey and Philp (1998). Lyster questioned the

effectiveness of the recast as corrective feedback based on his immersion classroom data, whereas Mackey and Philp (1998) insisted, based on their experimental data, that the recast contributed to their ESL learners' learning. Both arguments made sense to me in reference to my own experience of recasts: Lyster's argument was consistent with my being unaware of my friend's recast and her corrective intention, whereas Mackey and Philp's points were consistent with my later experience noticing NS recasts in the non-instructive contexts. The issues around the recast were confusing. Thus, this simple but complex verbal behaviour became my topic for investigation.

In the process of reading various studies on the recast, I found that the crucial shortcoming in the recast studies was the lack of data regarding consciousness. The claim for the recast's efficacy was solely based on the theoretical assumption that this type of feedback could attract the students' attention. There were not enough data to reveal why teachers make recasts and how learners actually pay attention to and process them. I thought the discussion about the effect of recasts in the literature was not very informative. Therefore, I decided to conduct a study which would incorporate the teacher's intention and adult students' attention in understanding recasts occurring in second language classroom discourse.

The question motivating me to prepare my research proposal was simple: "What is the recast? Is it really as effective as many SLA theorists say?" I wondered if teachers always gave recasts with the intention of correcting the student's language. I questioned if all recasts equally commanded the student's attention. I wanted to know how students would attend to them. In conducting my study, I aimed at three goals. First, I wanted to collect recasts in natural L2 classroom discourse and compile descriptions of them. Second, I wanted to examine the relationships among types of recasts, the teacher's intentions, and learners' attention to them. Finally, I attempted to examine whether any types of recast contribute to learners' second language development.

At the analysis stage where I attempted to incorporate teacher's intention and students' attention in relation to recasts, I found myself in trouble. The recasts I found were categorized according to different descriptions. The teacher and the students provided me with insightful introspective information. However, I was not able to give a good account of the relationship between the recasts and the participants' thoughts. I was not equipped with a theoretically sound tool to explain, for instance, *why* some items were better attended to than others, or *why* the teacher gave recasts in different ways. It took me some time to determine this problem was rooted in the theoretical framework I had been using in formulating my research.

The SLA research that I was familiar with took the cognitive-interactionist approach. Cognitive-interactionist SLA researchers are concerned with the representation of linguistic knowledge in the mind and the mechanisms for learning a second language. Drawing on research in linguistics and adopting discourse analysis as a research method, these researchers have investigated the process of SLA (e.g., Hatch, 1978; Pica, 1994). The researchers, focusing exclusively on language exchanged by interlocutors during interaction, developed a theory to show that "input" becomes integrated in the brain of the learner so that the learner becomes able to make "output." Even though the researchers discuss the second language learning "process," their descriptions are often limited to the relationship between "processed language" and learner performance. A diagram of their model (e.g., Gass, 1997, p. 3), for example, illustrates 12 different stages of processing linguistic input. The entries for the stages and sub-functions, which are connected by arrows, are either converted forms of language (e.g., "comprehended input") or functions working upon the language (e.g., "hypothesis formation"). The mechanical, almost computer-program-like flow-chart diagram illustrates exactly how cognitive-interactionist SLA researchers view the SLA process: as a second language processing machine. In the diagram, there is no room for the "intention" and "attention" data I collected from the teacher and the

students.

I, therefore, needed a different discipline in order to give a coherent account of what I found in my data. My alternative was the sociocultural theory of mind and SLA research based on this theory. I learned that the theory values interaction between individuals as the basis for development of human cognition. At first I thought the emphasis on interaction in sociocultural theory was similar to that in cognitive-interactionist SLA theory. This is not the case. The crucial differences are that human cognition in sociocultural theory is conceived of as "action" (Wertsch, 1998) that takes place primarily in the social sphere. This human action is mediated through semiotic-tools. One of the most powerful mediational tools, theorists claim, is language. Linguistic utterances, such as reasoning, questioning, recasting, are not simply chains of linguistic code or performance, separable from human mental functions. They are manifestations of "human-agent-acting-with-mediational-tools," inseparable parts of human agency. My initial desire to understand the recast in relation to the participants' cognitive functions was, re-framed in sociocultural theory, to understand the significance of recasts in relation to the teacher's and the students' agency.

This is the background of my thesis research. Because all the twists and turns are important components of my study, they have inevitably become parts of the thesis. In particular, the discovery of sociocultural SLA and the re-discovery of cognitive-interactionist SLA cannot be ignored. Although the illustrations and descriptions of cognitive-interactionist SLA were always in the textbooks and literature, I did not realize its distinctive differences from sociocultural SLA theory until I realized that I needed to frame my study from a different perspective. All the twists and turns and discoveries needed to be woven into the writing of this study.

The Organization of the Book

This book is composed of seven chapters. The current chapter, Chapter 1, introduced a short and rather personal story about the emergence and development of the thesis study and revising the study for this book. Reflecting my twists and turns in terms of a theoretical framework, the report of my thesis work is composed of two phases of study. Phase 1, Chapters 2 to 4, is based on the cognitive approach to SLA research and the recast. Phase 2 of the study, the sociocultural SLA approach, is presented in Chapters 5 and 6.

Chapter 2 presents a review of the literature relevant to recast studies. It first focuses on theoretical issues in cognitive-interactionist SLA: the linguistic aspect (i.e., input), the psycholinguistic aspect (i.e., attention and noticing), and the pedagogical aspect (i.e., form-focused instruction). The latter part of the chapter focuses on empirical studies investigating recasts exclusively or inclusively. The empirical findings are illustrated under two categories: description of recasts and evaluation of the effectiveness of recasts.

Chapter 3 presents the research design. After stating the three operationalized research questions, the research context and procedures for data collection and analysis are illustrated. It is important to note that a unit of analysis for the recast in this study was a short episode including it, rather than a turn consecutive to an erroneous turn as in other previous studies. The two data collections adopted to accommodate two of my research goals are worth mentioning. In order to collect data to investigate the teacher's and the learners' cognitive activities, a "stimulated recall interview," an introspective data collection method, was adopted. For assessing the effectiveness of recasts in relation to the learner's attention, tailor-made grammaticality judgment tests were developed.

Chapter 4 presents the findings of the three research questions regarding the description of recasts in the classroom, the students' attention to recasts, and the effectiveness of the recast. The recasts found

in the classroom are described according to detailed characteristics. The students' attention to recasts is also discussed according to the recast characteristics. The recasts and student attention are analyzed both quantitatively and qualitatively. The grammaticality judgment tests results are presented numerically in relation to the characteristics of the recasts and to student attention.

Chapters 5 and 6 compose the second phase of the study. Chapter 5 presents a review of the literature in sociocultural theory and sociocultural SLA research. The first part of the literature review focuses on the important and relevant concepts in the theory: the zone of proximal development and, in particular, its scaffolding functions. The latter part of the literature review introduces feedback research within the sociocultural framework. The differences in feedback research between cognitive and sociocultural SLA are also noted.

Chapter 6 presents four case studies within the sociocultural SLA framework of the recast regarding feedback and the cognitive activities of the teacher and the students. The four cases are comprised of two episodes that were successful in linguistic instruction, one episode that was successful in content instruction, and one episode that caused much linguistic debate among the students.

Finally, Chapter 7 concludes the thesis by providing a summary and implications of the research. Regarding the theoretical concerns, problems in the current definitions of recasts and the application of sociocultural theory in future SLA research are presented. The recasts, group work, and learners' agency are also considered in respect of their pedagogical implications. Finally, the limitations of this study are also discussed.

Chapter 2 : Phase 1 - Cognitive-interactionist SLA Research

Introduction

Research interest in the recast stems from mainstream cognitive-interactionist second language acquisition (SLA) research into the mechanisms of second language (L2) learning. Cognitive-interactionist SLA researchers see L2 learning as the process of building up and restructuring the knowledge system of the target L2 so that it can eventually be utilized automatically for speaking and understanding (Lightbown & Spada, 1993, p.25). A currently dominant view of the L2 learning mechanism among cognitive-interactionist SLA researchers is that the language to which the L2 learners are exposed becomes "comprehensible input" through interactive verbal negotiation and learners' cognitive processing of the language; the comprehensible input is eventually integrated into the learners' interlanguage grammar, enabling the learners to produce the target language (see Ellis, 1998a; Gass, 1997; Long, 1996). Through empirical research it is possible to test this theoretical model by investigating the effects of different verbal feedback moves, including the recast. Investigation into feedback has pedagogical value because identification of effective feedback is clearly useful for L2 teachers and learners. The recast is, in fact, advocated by some cognitive-interactionist SLA researchers as an effective means for drawing learners' attention to language form while maintaining the flow of communication in the communicative L2 classroom (Long, 1996).

My review of the literature first focuses on understanding the cognitive-interactionist framework and its focus on the recast within the broader area of SLA research in general. Because cognitive-interactionist SLA researchers are concerned with both external (i.e., the language to which L2 learners are exposed in a given environment) and internal (i.e.,

learners' cognitive functions) factors contributing to L2 learning (Ellis, 1994, p.16), studies of these two factors are first reviewed. I then turn to research on form-focused instruction and feedback and conclude with a review of empirical recast research.

External Factors: Linguistic Input and Interaction

In SLA research, the language to which language learners are exposed is "input." This is presumed to be the fundamental and essential source for language acquisition. The exact characteristics of input and its function, however, vary according to researchers' theoretical emphases within the SLA framework. There are those who emphasize universal grammar (UG): Krashen (1982), for instance, advocated the view that input is necessary and sufficient for L2 acquisition. He argued that learners' L2 learning should target the acquisition of intuitive, rather than conscious, linguistic competence achieved through exposure to comprehensible input. His vague and ambiguous notion of "comprehensible input" was later expanded and refined by other researchers (see Gass, 1997; Long, 1983; 1996; Pica, 1994). Cognitive-interactionist SLA researchers contend that linguistic input becomes comprehensible through verbal interaction between a native and a non-native speaker[1)] because the interaction process involves modifications of utterances. It is hypothesized that such interaction, particularly that in which negotiation of meaning is involved, enables learners to comprehend, intake, and learn the L2 input.

---

1) Obviously, the verbal interaction focused on in SLA research does not occur exclusively between a native and a non-native speaker: a learner's interactions with another peer (NS or NNS) or a non-native L2 teacher are also important interaction contexts in L2 classroom and SLA research. Therefore, labelling the participants of interaction as a native and a non-native speaker can be misleading because this labelling suggests that interaction in the classroom is by default native speaker/non-native speaker interaction. In this study, however, I used these terms because the participating teacher was a native speaker of English and the students, non-native, and this study is about the teacher/student interaction.

The theoretical shift from the input hypothesis to the interaction hypothesis parallels the theoretical argument within first language (L1) acquisition research (FLA). In FLA research, the Chomskyan innatist L1 researchers view a child's first language acquisition as a process whereby the language acquisition device (LAD) containing the UG is activated through exposure to the language around the child. This UG-based view has been challenged by other L1 researchers who focused on caretakers' subtle modifications in their verbal interactions with children (Snow & Ferguson, 1977).

For these cognitive-interactionist researchers in both L1 and L2 acquisition research, input is not a single entity. Expanding the linguistic view of input as "linguistic evidence," they view some input for a language learner as "positive evidence" to show the ideal usage and other input as being "negative evidence" to show inappropriate usage (Schachter, 1991). Much negative evidence is theoretically available for learners through the caretaker or the native speaker's reaction to their inappropriate use. The issue of whether negative evidence is available in reactive input (e.g., recasts) is controversial. It has been debated within the L1 research literature (e.g., Bohannon, & Stanowicz, 1988; Grimshaw & Pinker, 1989), and has recently begun to attract attention in the SLA literature (e.g., Leeman, 2000; Long, 1996). Identification of the specific roles of various types of input, particularly reactive types of input, is important for cognitive-interactionist researchers because it will better inform language teaching practitioners. Recast studies in both L1 and L2 research relate to these cognitive-interactionist researchers' interest in understanding and evaluating this type of feedback.

Internal Factors: Learners' Cognitive Mechanism

Although cognitive-interactionist L2 researchers share with the L1 cognitive-interactionists the theoretical view of the role of interactive input, their focus on reactive input has caused them to raise questions about the influence of L2 learners' internal factors in their learning,

namely, L2 learners' cognitive mechanisms. Unlike the participants in L1 studies, who are young children learning both language and world concepts more or less simultaneously, many participants in L2 studies are adult L2 learners who have acquired and established their own world-view through their respective L1s.

In theorizing about the internal mechanism of L2 learning, Schmidt (1990, 1994, 1995; Schmidt & Frota, 1986) argued that learners' attention to and noticing of a target feature was necessary for learning. This view is largely accepted among many cognitive and interactionist SLA researchers (see, for example, Ellis, 1998a; Gass, 1997; Long, 1996; Skehan, 1998; VanPatten, 1996). According to Schmidt (1990, 1994), two aspects of human consciousness are particularly relevant in SLA research: the state of consciousness and cognitive activities. The former is, in Schmidt's terms, awareness indexed on a level of sensitivity continuum, ranging from unintended attention to the highest level of overt understanding.[2)] The dichotomy between unintentional and intentional levels of awareness corresponds to the cognitive-interactionist researchers' view of native-like performance and the beginner L2 learners' performance. The native-like performance is considered as automatic performance in a language operating under subconscious, implicit linguistic knowledge; in contrast, L2 learners' performance is intentional, deliberate, and effortful, and is influenced by explicit understanding of target language grammar (e.g., Ellis, 1998a; Gass, 1997; Long, 1996).

The cognitive activities in Schmidt's description of human consciousness are paying attention, noticing and understanding. Attention is a "mechanism" (Jackendoff, 1987 as cited in Schmidt, 1995, p.18) or a

2) Another facet of consciousness Schmidt refers to is "intention." The notion is often narrowly applied by cognitive-interactionist SLA researchers, namely in contrastive studies between incidental learning and intentional learning. However, intention is an important variable in SLA research, applicable in other aspects of the learning process. For example, as Harley (1994) pointed out, being enrolled in an L2 course is also an intentional activity by the learner.

sense that triggers noticing, and paying attention is the activity with a range of intention to direct such sensitivity toward an object. Noticing is a consequence of paying attention; it is the "conscious registration of the occurrence of some event" (Schmidt, 1995, p.29). Understanding is a more overt realization than noticing, or "recognition of a general principle, rule or pattern" (Schmidt, 1995, p.29) that may be integrated into one's knowledge.

An important question for cognitive-interactionist SLA researchers is what makes L2 learners pay attention and notice during interaction. Learners' attention is considered to be attracted to salience of form. What makes a form salient is unclear (Gass, 1997); however, the features thought to increase perceptual salience discussed in SLA research are frequency of input (either highly frequent or highly infrequent), intonation contour, position (e.g., juxtaposition), and form-focused instruction devices (e.g., input enhancement).

Form-Focused Instruction, Feedback, and Recasts

In a pedagogical context, research into the external and internal factors which influence L2 learning connects to the question of the role of instruction in L2 learning. Formal instruction was once considered to have little impact on learners. Krashen (1976, 1982), for example, denounced the role of instruction, referring to empirical studies which found that L2 learners from different backgrounds, whether receiving instruction or not, showed the same developmental sequence of certain linguistic aspects (e.g., morphological development). However, this view has been challenged because L2 learners who received much comprehensible input were not as successful as anticipated. For example, despite meaningful input, young learners of French as an L2 in immersion programs in Canada failed to achieve native-like competence in productive L2 use (Allen, Swain, Harley, & Cummins, 1990; Harley & Swain, 1984; Swain, 1985). These researchers now argue that classroom L2 instruction is necessary.

The L2 instruction recently advocated is form-focused instruction (FFI). FFI is defined as "pedagogical effort which is used to draw the learners' attention to language form either implicitly or explicitly" (Spada, 1997).[3] FFI is conceived as an effective approach to L2 instruction because it incorporates both the interaction hypothesis and the noticing hypothesis (Doughty & Williams, 1998; Long, 1983, 1988, 1991; Spada, 1997). Studies into FFI involve identification and evaluation of instruction in overt and/or covert formats (e.g., Spada, 1997, p.74), for example, a teacher's various reactive verbal behaviours (i.e., feedback) and implicit communicative negotiation devices (e.g., clarification, confirmation). Feedback is presumed to be an effective form of FFI because it either implicitly or explicitly signals or highlights the linguistic problem in the learners' performance.

Implicit corrective feedback within the cognitive-interactionist framework is called a "recast," i.e., a reformulation of the learner's erroneous utterance.[4] The content of the recast is contingent upon the learner's original meaning; thus, the utterances by the student and the teacher contrast at the level of linguistic form (Leeman, 2000). Because the contrasting form is provided in juxtaposition to the learner's erroneous utterance, the linguistic difference is intended to become salient to the learner (Schmidt & Frota, 1986); the recast is, therefore, presumed to be perceptually salient (Leeman, 2000; Long, 1996). Because the recast

3) FFI that Spada (1997) defines includes what Long (1991) refers to as *focus on form* (FonF) and *focus on formS*. The former is an instructive technique which "overtly draws students' attention to linguistic elements as they arise incidentally in lessons whose overriding focus is on meaning or communication" (Long, 1991, p.46), and the latter refers to traditional grammar instruction of isolated linguistic features. Although Long (e.g., 1991) denounces *focus on formS*, many cognitive-interactionist SLA researchers consider it also a useful means of instruction as long as it is provided within communicative contexts (see Doughty & Williams, 1998).

4) Among cognitive-interactionist SLA researchers, recasts are exclusively defined as "implicit corrective feedback" perhaps because they adopt the innatists' positive and negative evidence dichotomy. This definition, however, contains a logical contradiction: feedback that is "salient" should more likely be "explicit."

provided immediately subsequent to the learner's erroneous utterance is salient, it is thought that the learner will pay attention to the recast and compare the two forms. Thus, the recast functions as corrective feedback. This hypothesis still requires support based on empirical data on learner attention.

Recast Study Findings to Date

Although the term recast only came into the SLA literature fairly recently (Doughty, 1994), recasting has been frequently observed in various interaction contexts; in NS-NNS interaction studies, for instance, recasting has been called "modification" and "completion or elaboration" (Pica, Holliday, Lewis, & Morgenthaler, 1989). In classroom interaction studies, terms such as "paraphrasing" (Spada & Fröhlich, 1995), "repetition with change" and "expansion" (Chaudron, 1977) have been used. There are a number of studies in L2 research which have directly and indirectly investigated the effect of recasts. The findings in these studies can be outlined under two headings:(1) description of recasts (in particular, of their frequency), and (2) evaluation of the effectiveness of recasts in (a) observational (mainly classroom) studies and in (b) experimental laboratory studies.

Description of Recasts

Recasts have been found to be a highly frequently occurring type of feedback in L2 classrooms and in NS-NNS interaction. In four content-based French immersion classrooms, Lyster and Ranta (1997) found that 55% of teacher feedback moves were provided in the form of recasts. Panova and Lyster (2002) found 55% of an ESL teacher's feedback was in the form of recasts. Doughty (1994) also found that approximately half the teacher feedback provided in a French as a foreign language (FL) classroom was in the form of recasts (see also Chaudron, 1988, p.145). Sheen (2004) reported that recast feedback occupied 68% of teacher feedback moves in two ESL classrooms in New Zealand and 83% of two

EFL teachers' feedback moves in Korea. Recent studies also revealed that L2 learners performing pair- or group-work in communicative ESL classes also give each other recast feedback. Morris (2002), for instance, reported 68% of peer feedback moves during a collaborative task were in the form of recasts. In an experimental study examining implicit NS feedback (i.e., clarification and confirmation requests) and the resulting modified NNS output, Pica, Holliday, Lewis, and Morgenthaler (1989) found that confirmation requests in the form of recasts and repetitions occurred across the three task types more than two-thirds of the time. In Braidi's (2002) study, 25.56% of Japanese ESL learners' erroneous utterances received NS implicit negative feedback in the form of recasts and confirmation checks. Iwashita (2003) also found that 30 to 40% of learner errors in forming the target structure received Japanese NS interlocutors' recast feedback during communicative task-based interaction.

Recasts frequently occur as responses to grammatical and phonological errors.[5] In the French immersion classrooms (Lyster, 1998a), teachers provided recasts for 72% of grammatical errors and 64% of phonological errors, while only 38% of lexical errors were treated by recasts. The dataset in Mackey, Gass, and McDonough (2000) also showed that during a communication task, NS interlocutors reacted with recasts[6] to ESL learners' morpho-syntactic errors 44% of the time.[7] Oliver (1995), who employed a more detailed categorization of error

5) Although some researchers distinguish "errors" from "mistakes," I do not make this distinction in my study. Both "errors" and "mistakes" are used interchangeably to indicate "erroneous utterances."

6) Mackey et al. (2000) identified three patterns of conversational interaction: that involving (1) recast, (2) negotiation, and (3) combination of negotiation moves and recast. The NS feedback called "recast" here is NS feedback containing only the recast.

7) This percentage is based on the number of recast feedback episodes in their Tables 7 and 8. In their study, ESL learners' phonological errors received recast feedback 9% of the time and negotiation-recast combination feedback 16% of the time.

types, found that her NS child participants least frequently provided recasts to NNS children's lexical errors (e.g., word choice or order, word or subject omission), and gave recasts most frequently to grammatical errors (i.e., singularity or plurality, subject-verb agreement). Oliver attributed these differences to comprehension difficulties: When the message's meaning was comprehensible despite the error, NS children responded to NNS with recasts. The lexical errors, such as word choice, caused more comprehension difficulties than the singularity or plurality mistakes.

Although recasts seem to be a useful resource for learning, some researchers are sceptical about their corrective effect since the remedial intention in recasts is often ambiguous in specific discourse contexts. Interlocutors, especially teachers in communicative L2 classrooms, are found to repeat and rephrase both well- and ill-formed learner utterances. Lyster (1998b) found that teachers' non-corrective repetition of L2 learners' well-formed utterances occurred almost as frequently as teacher recasts of ill-formed utterances in the content-based immersion classrooms he studied. A similar distribution of teacher recast and non-corrective repetition was observed in Doughty's study (1994). Doughty's dataset (p.105) show that the FL teacher in her study provided 40% of non-corrective repetition to well-formed learner utterances and 57% of recasts to one-error-utterances. Further, Lyster (1998b) found in the immersion classrooms that the recasts often accompanied affirmative comments of the truth value of the learner's reply, which made the recasts almost identical to non-corrective repetition.

Evaluation of the Effectiveness of Recasts

Based on observation

The effectiveness of recasts as corrective feedback is inconclusive in current L2 research. In observational studies, learners' repairs of their errors immediately after recast feedback have been considered an indication that recasts are an effective form of feedback. In general,

however, recasts appear less likely than other types of feedback to trigger learner uptake (i.e., repair incorporating the modified form). In classroom-based studies, Lyster and Ranta (1997) found immersion students' uptake after recast feedback occurred only 31% of the time in contrast to, for example, 88% in response to clarification requests. In Panova and Lyster's (2002) study, adult learners' uptake after recast occurred 40% while uptake after clarification request or elicitation occurred 100%. Learners' infrequent uptake was also found in experimental studies using communicative tasks. The dataset in Mackey and Philp (1998) showed that learners responded to recast feedback only about 30% of the time during a communicative task (p.349).[8] Similarly, Mackey et al. (2000) showed NNS learners did not respond in 44 out of 66 (68%) episodes where NSs provided recasts during a communicative task. In Pica et al. (1989), the most common reaction (44%) of Japanese ESL learners to NS recasts was to acknowledge the NS with a yes/no confirmation (pp.80-81).

However, researchers who conducted studies in FL contexts emphasize that learners are likely to react to teacher recasts. Doughty (1994), for example, reported that the college FL students in her study reacted to teacher recasts most frequently among the four feedback types under investigation, though 87% of teacher feedback received no response (p.106). Sheen (2004) found that Korean adult EFL learners reacted to teacher recasts frequently: 82% of teacher recasts received learner reaction, with 70% of these being correctly repaired. A high rate of learner uptake was found even in an ESL context: in a dataset from New Zealand ESL classrooms composed of mainly Asian students preparing for admission to colleges, 73% of teacher recasts received learner reaction, 66% of them being successfully repaired. Ohta (2000) also found that Japanese as FL (JFL) students reacted to teacher recasts provided in the class. Indeed, she found that the students were most likely to react to

8) This percentage is based on the number of turns excluding "no-opportunity" instances.

teacher recasts when they were not the addressees of the recast feedback.

Although uptake is an important and observable source for understanding the impact of the feedback (e.g., the feedback is noticed as correction), it is not in itself sufficient evidence to evaluate the effect of the feedback. Learners' successfully repaired uptake does not necessarily mean that long-term learning has occurred. It is also possible that learners' uptake does not fully represent their cognitive processing of the feedback: the lack of uptake does not necessarily mean that the learner was not thinking. Occurrence of learner uptake is also influenced by the discourse context. Oliver (1995) argued that the lack of learner response was attributed to conversational constraints in that the NS interlocutor often did not provide opportunities for the learners to repair their utterances. Oliver analyzed her NS-NNS interaction data taking this lack of opportunity into consideration; when turns in which a NNS's incorporation of a NS's recast was either impossible or inappropriate were excluded, the proportion of NNS incorporations of NS recasts increased from 10% to 35%.

Based on experimental test performance

Another consideration regarding the effectiveness of recasts is learners' test performance. Employing various pre- and post-treatment performance tests in experimental laboratory studies, recast treatment has been compared to different types of input and feedback treatment. For example, Ortega and Long (1997) compared the effect of recasts as a reactive model to the provision of a pre-emptive model on two grammatical aspects in Spanish (i.e., direct object topicalization and adverb placement). They measured development in the selected grammatical rules using pre- and post-treatment oral picture-description tests. The post-treatment test included items used in the treatment, as well as new items. They found that learners in the recast condition performed significantly better with one of the two target grammatical features (adverb placement) than in the model condition.

Ayoun (2001) compared the effect of recasts, pre-emptive models, and traditional grammar instruction (i.e., combination of explicit rule explanation and feedback on exercise answers) with respect to passé composé and imparfait in French. The researcher provided the participants with group-specific treatments of the written-format computer-based instruction and tasks for 4 sessions. After the 5th session, the participants wrote a composition as a post-test and the researcher measured the learners' correct usage of the target structure in their writing. The post-test analyses showed that those in the recast group performed significantly better than the traditional grammar instruction group, though the difference between the recast and model groups was not significant.

Carroll and Swain (1993) also employed pre- and post-treatment tests in a study of the effects of different feedback strategies on adult ESL learners' ability to recognize verbs which do or do not alternate in dative sentences. They measured the effect of different treatments using the same items as in the treatment in the post-treatment recall test, and found that the group that received recast feedback (i.e., "modeling and implicit negative feedback" in their terms) performed second best among the four treatment groups in the short-term recall session next to the group who received metalinguistic explanation feedback. However, only the latter maintained a long-term advantage.

Through pre- and post-treatment production, Mackey and Philp (1998) showed that intensive recast treatment had a positive delayed effect on learners' use of target question forms. The researchers compared the accuracy and complexity of students' production of questions before and after the treatments. The advanced learners, who were exposed to intensive recast feedback during the treatment period, produced developmentally higher question forms 78% of the time in the post-test while the other advanced learners with interactive negotiation treatment produced such forms only 17% of the time. No differences were found among the less advanced groups.

Using a time-series design, Ishida (2004) also studied the effect of intensive recast feedback. In her small-scale study, Ishida measured changes of target structure in the oral production of 4 JFL learners during communicative activities before and after intensive recast treatment. The learners' target-like use of the focused linguistic structure (i.e., Japanese aspectual form) increased after recasting treatment began and was maintained during the post-test sessions.

Han (2002) and Doughty and Varela (1998) focused on tense marking. Two groups of students in Han's (2002) study produced picture-based narratives. The first narrative and the last two narratives served as the pre-, post-, and delayed post-test respectively. Han compared the learners' tense-marking in their narratives after the instruction session in which the recast group received recast feedback on their production, and the other group received no feedback. She found that the recast group increased consistent use of appropriate tense marking, and the consistency remained in the delayed post-test. On the other hand, the tense marking by no-feedback group members continued to fluctuate.

Doughty and Varela (1998) implemented a "corrective recast" treatment in their experimental classroom study of a theme-based ESL class. They implemented feedback composed of repetition of the learner's ill-formed sentence (with emphasis on the incorrect verb form), followed by a recast of the target linguistic form (i.e., tense marking), and compared the learners' performance in class assignments before and after the feedback treatment, as well as with learner performance in a comparison class. The ESL students who received the treatment showed greater improvement in marking past tense accurately in their oral production than the students in the comparison group. The treatment students also increased their attempt in the post and delayed post oral tests to mark past tense, but the form was not accurate (i.e., they used interlanguage forms).

The studies examining the effect of recasts through learner production reveal positive findings; unfortunately, they are not

comprehensive. For instance, the positive findings could be attributed to the narrow discourse contexts of the experimental environments. Recasts in these experimental studies were often isolated from other feedback strategies and were provided intensively.

Variables influencing the effect of recasts

Recent studies have indicated that some of the positive findings regarding the effectiveness of recasts are conditional: The target language forms (e.g., Ortega & Long, 1997; Ayoun, 2001; Iwashita, 2003), the levels of learners' proficiency (e.g., Mackey & Philp, 1998; Sheen, 2004), and the levels of learner attention are influential variables. Recasts seem most useful in helping learners to acquire a linguistic form which they know but cannot properly use yet. Ortega and Long (1997), for instance, found object topicalization may have been too difficult. Ayoun (2001) also found that learning the imparfait from recasts seemed difficult for her participants. Positive learning results related to relatively basic structures such as tense marking (Doughty & Varela, 1998; Han, 2002) and questions (Mackey & Philp, 1998; Philp, 2002).

Recast feedback seems likely to be more effective for advanced learners. Those who learned developmentally higher question forms in Mackey and Philp (1998) were advanced learners. In the study by Sheen (2004), the learners who had higher educational backgrounds tended to make reactive moves in response to recast feedback. The role of learner's attention cannot be ignored. Philp (2002) found that learners' abilities to notice and recall recasts varied with their proficiency level. Williams (1999) reports a similar finding with respect to learner attention and proficiency. In her study into the learners' production of Language Related Episode (LRE) (Swain, 1998; Swain & Lapkin, 1995), Williams found advanced learners are more likely than less advanced learners to produce LREs.

Reflecting the importance of learner attention, a new approach for evaluating recasts has been adopted. As a preliminary step in evaluating

recasts based on learners' cognitive processes when they receive feedback,[9] Mackey et al. (2000) studied the relationship between feedback and learner perception. They found that implicit negative feedback, especially in the form of recasts, to ESL learners' morpho-syntactic errors was less accurately perceived as such than feedback to lexical and phonological errors, despite the fact that recasts of morpho-syntactic errors were one of the most frequent types of feedback provided (47% of all NS implicit negative feedback). Further analyses regarding recast feedback to ESL learners indicated that feedback episodes composed of recasts were less likely to be perceived accurately than feedback episodes composed of negotiation or a combination of negotiation and recasts. Their dataset showed that recasts were inaccurately perceived 75% of the time while the negotiation and negotiation-recast combination feedback were inaccurately perceived only 13% and 12% of the time respectively.

Mackey et al. (2000) indicated that recast types (i.e., recast-alone or recast and negotiation combined) influence learner perception. Lack of learner repetition and modification (uptake) also appears related to their inaccurate perception. Their study showed that learners' verbal reports also provide researchers with complementary data for understanding the effect of recast feedback.

Implications for This Study

According to recast research carried out in controlled settings to date, recasts appear to be useful for learners. Findings from classroom studies are inconclusive, though the strong indication is that recasts, particularly in communicative classrooms, are difficult for learners to identify as feedback on form (Nicholas, Lightbown, & Spada, 2001). The theoretical assumption for recasts being effective is their potential salience

9) Ohta (2000), for example, examined "private speech" described in sociocultural theory to trace learners' cognitive activities.

for drawing learners' attention. Empirical studies into the relationship between recast feedback and learners' attention are few. The approach taken by Mackey et al. (2000) to gather data about learners' attention through stimulated recall seems promising. Similar studies in different contexts, especially in classroom contexts (e.g., FL classrooms and content-based L2 classrooms), would provide rich empirical data. The present study was conducted to provide such data.

Chapter 3 : Research Design

Introduction to Present Research

In the present study, I gathered empirical data to understand the relationship between recasts and learners' attention in a communicative L2 classroom. Further, I evaluated the effectiveness of recasts for L2 learning. In this chapter, I first present the research questions and then describe the research design, including the research participants, context, data collection procedures, and data analysis procedures.

Research Questions

The principal research questions guiding this phase of the study were:

What are the roles of recasts in communicative classroom interaction? Do recasts have a role as corrective feedback and facilitate adult L2 learners' learning by attracting their attention to language as cognitive-interactionist SLA theory presumes? If so, how effective are they?

The principal questions were translated into the following operational research questions:

1. What recasts were the students exposed to in a communicative theme-based EFL classroom?

2. In what ways did the students attend to the teacher's utterances and recasts during interaction?

3. Which recasts related to the students' L2 learning as measured by grammaticality judgment tests?

The Research Context

The data were collected in a theme-based English Discussion class in a private Japanese women's college specializing in English education in an urban area in western Japan. The research was conducted in the second semester (i.e., ten weeks) of the two-semester English Discussion class.

Participants

The English Discussion class was composed of 28 female, first-year students, between the ages of 18 and 20, and an American EFL teacher. The students had passed entrance examinations in order to be admitted to the college. They were placed in this upper-intermediate class[10] according to a placement test administered at the beginning of the academic year.

Of the 28 students, eight students, Aiko, Eiko, Fumiko, Hisako, Keiko, Shoko, Tokiko, and Yasuko,[11] volunteered to participate in this study. They agreed to be video- and audiotaped during the lessons and to be interviewed outside of the class. Although they were at the same "proficiency level" based on the school placement test, individual differences were observed. According to their teacher Ms. Johnson's evaluation, Shoko and Keiko were stronger students in the class, and Tokiko was the weakest of the eight. These students were divided into two groups of four during the semester. As shown in **Table 1**, Groups 1A and 1B worked together in the first five weeks of the semester. In the sixth

10) This level designation was relative to the students in this college. The student language proficiency level at an "upper-intermediate class" in this college is considered roughly equivalent to TOEFL 400-450 (Katayama, 2001, p.172). Because the full reference of Katayama (2001) includes the name of the college in which this study was conducted, the reference is not included in this thesis in order to keep the anonymity of the participants.

11) Names are pseudonyms. The eight volunteer students were compensated financially for their out-of-class interview time. Each of the eight students was paid six thousand yen.

Table 1: Student Groups

Session 1 Week 1-5 (theme: human rights)	Group 1A	Shoko, Yasuko, Hisako, Tokiko
	Group 1B	Keiko, Fumiko, Aiko, Eiko
Session 2 Week 6-10 (theme: environment)	Group 2A	Shoko, Yasuko, Aiko, Eiko
	Group 2B	Keiko, Fumiko, Hisako, Tokiko

week, two members from each group changed groups, forming Groups 2A and 2B.[12)]

The teacher, Ms. Johnson, had taught English for more than 15 years in Japan and for 10 years at this college at the time of this study. She encouraged students to speak in English during class. In order to maximize students' opportunities to use English as well as to make her interaction with students more feasible, Ms. Johnson adopted a group-work strategy in teaching (Teacher Recall 1, September 10). She grouped her students into seven groups of four, and most tasks were assigned as group work. She monitored the students' use of English in group activities, and a group which used Japanese extensively was sometimes penalized by having classroom participation points deducted.

The Classroom

The English Discussion course in this college curriculum was one of five compulsory English courses for first year students: (1) English Grammar, (2) Phonetics, (3) Reading, (4) Academic Writing, and (5) English Discussion. The latter three courses were designed to increase students' ability to use the language skilfully by immersing the students in a language learning environment with academic themes. The themes (e.g., "Human Rights" and "Environmental Issues") were adopted across the three courses so that the students could read, discuss, and write about the

12) In the sixth week, all students were separated into different groups. Although Ms. Johnson directed students to move to different tables and form groups with new members, the volunteer students formed different groups among the eight at my request.

same themes for five weeks, recycling both the language and the content. The English Discussion class met twice a week for 70 minutes each time throughout each 10-week semester.

The Discussion course emphasized developing fluency in oral English communication. Living in an EFL context (i.e., having few opportunities or little need to communicate in English on a daily basis) and having learned English in traditional teacher-fronted classrooms in high school, the students were, in general, hesitant to express themselves in English. Thus, they were expected to improve their communication skills through discussing themes such as "human rights" and "environmental issues" during the two semesters the course was offered. The students were also expected to "understand and use the vocabulary of the topics with a fair degree of accuracy" as well as "to expand their horizons" by becoming aware of current events in their society and in the world (College Catalogue, p.16). The course textbook contained many communicative tasks, such as information gap and ranking tasks, interviews and surveys, to facilitate students' discussion. In contrast to the FL classroom in previous recast studies (e.g., Ohta, 2000), the teacher in this theme-based class placed little emphasis on the instruction of linguistic components. This was probably because these students were concurrently taking courses such as English Grammar and Academic Writing, which had a more linguistic orientation.

Data Collection Procedure[13)]

The data collection involved cycles of a set procedure. A basic cycle was composed of (1) classroom observation and videotaping of a 70-minute class period, (2) administration of a Grammaticality Judgment test (a GJ Mini-test), and (3) a stimulated recall interview with the students within a week of the videotaping. This cycle was conducted

13) A glossary of abbreviated terms appears in Appendix A.

Table 2: Data Collection Schedule

Weeks in the Semester	Data Collection Week	Video Taping	Discourse Dataset	GJ Test*	Student Stimulated Recall	Teacher Stimulated Recall
1	–	–	–	–	–	–
2	–	–	–	–	–	–
3	1	Lesson 1	Dataset 1	Mini-test 1 (Part of Test 1)	SR 1	–
4	2	Lesson 2	Dataset 2	Mini-test 2 (Part of Test 1)	SR 2	–
5	3	Lesson 3	Dataset 3	Mini-test 3 (Part of Test 1)	SR 3	–
6	4	Lesson 4	Dataset 4	Mini-test 4 (Part of Test 1)	SR 4	Teacher SR 1
7**	5	Lesson 5	Dataset 5	Mini-test 5 (Part of Test 1)	SR 5	–
8	6	Lesson 6	Dataset 6	Mini-test 6 (Part of Test 1)	SR 6	–
9	7	–	–	–	–	–
10	8	–	–	–	–	–
1 week after	9	–	–	Test 2 (Mini-tests 1-4 and 6)	–	–
2 weeks after	10	–	–	–	–	Teacher SR 2

* Each Mini-test was administered to individual eight students before the Stimulated Recall interviews were conducted.

** Because two students were absent from Lesson 5, the data from this week were excluded from analyses.

approximately each week and was repeated six times[14] from the third to eighth week within the 10-week semester. Additionally, I administered a

14) Due to the absence of two volunteer students in one videotaped lesson (Lesson 5), data from five sessions were used for analyses in this study.

composite GJ Test (Test 2) after the term was completed (approximately three weeks after the last cycle). I also conducted two stimulated recall interview sessions with the teacher in the middle of, and after, the semester. **Table 2** illustrates the overview of data collection schedule.

Classroom Observation and Videotaping

During classroom observation, I was seated at the back of the room.[15] A video-camera and a multi-directional microphone were set up for each of the two groups of four students. In order to capture the students' dialogues and activities, the camera was placed near the group table and the microphone at the centre of the group table. Two small cassette recorders with built-in microphones were also on the group table as audio back-up.[16]

The three major activities in the class were newspaper article reports, discussion activities in groups, and summing up of discussion outcomes. These activities took place either in a teacher-fronted whole-class format or in groups. Teacher-fronted interaction in this class often showed a typical IRF sequence (e.g., Cazden, 1988; Mehan, 1979): teacher-initiation (by calling upon a student), students' reply (or reporting), and teacher's acknowledgement, evaluation, or follow-up comments. Ms. Johnson did not impose any rules for conducting group discussion; the students were equal participants. However, perhaps because of their relative higher proficiency in oral English, Keiko and Shoko tended to lead their group discussion by reading aloud task instructions and questions and asking their members for their opinions.

The dialogues on the tapes, including both teacher-fronted whole

15) Although I was a silent researcher-observer, I was occasionally included in the discussion activities. Ms. Johnson referred to me during discussions as an example of someone who had lived in North America. She also encouraged her students to ask me for help when she was not readily available to answer their questions.

16) Back-up audiotapes ended up being used in place of five videotapes that had sound problems.

class discussions and individual group discussions, were transcribed. The transcripts were analyzed for error treatment episodes (ETEs) and recast episodes (REs). An ETE was defined as a sequence of feedback turns to deal with one aspect of non-target-like language use found in a learner's utterance. An episode started when a learner made an erroneous utterance which was reacted to by the teacher and ended when the focus shifted away from the error.[17] The error treatment types were taken from Lyster and Ranta (1997) and Lyster (1998b). These constituted (1) Recasts (R); (2) Explicit Correction (EC), through which a target form was explicitly provided; (3) Clarification requests (CR), through which difficulty in comprehension of the non-target-like utterance was indicated and a repetition or a reformulation was required; (4) Metalinguistic feedback (MF), through which comments, explanations, or questions about the divergence of the learner's utterance were provided without explicitly providing a more target-like form; and (5) Elicitation (El), through which the interlocutor attempted to elicit the more target-like form from the previous speaker (as in Lyster & Ranta, 1997). An RE was a subcategory of ETE whose feedback move involved at least one recast within the episode. (A detailed explanation of REs is provided in the "Analysis of classroom discourse" section, pp.28-30.) The REs identified during the class period served as the database (i.e., Discourse Dataset in Table 2) for the Mini-test and the "stimuli" for the student stimulated recall interview for that cycle.

Uptake Claim Survey

After each of the six videotaped lessons, the eight students were

17) This definition of an episode as a unit of analysis may allow "embedded" (i.e., an episode dealing with one problem is embedded in another episode which deals with another problem) and "expanded" (i.e., a solution in one episode develops into a new episode dealing with an overlapped but slightly different problem) episodes (see Fortune & Thorp (2002) for discussion about natures of episodes). Because the teacher tended to focus on only one language problem when providing feedback, no teacher ETEs were embedded or expanded.

asked to complete an uptake claim survey. An uptake claim is defined as a self-report of having learned something in a lesson (Allwright, 1984), a type of learner verbal report data used to identify learners' allocation of attention and noticing in their language learning during the lesson. The survey questionnaire, taken from Slimani (1987, 1992), asked the students to relate in as much detail as possible the points they recalled from the preceding lesson. The questionnaire contained subcategories such as "words and phrases," "spelling," "pronunciation," "grammar," "language usage," and "other" (Appendix B). The questionnaire was distributed at the end of the class. Students spent approximately five minutes to complete the survey each time.[18)]

Grammaticality Judgment Tests (GJ tests)

Rationale for using GJ tests in this study

Grammaticality Judgment (GJ) tests were used in this study to assess learning outcomes. Among the testing methods in the psycholinguistics of language abilities (e.g., elicited imitation, cloze test), I chose GJ tests for measuring learning from recasts, focusing on the L2 learners' receptive linguistic ability. Unlike an elicited imitation test, which assesses L2 learners' language abilities manifested in actual performance, GJ tests evaluate the state of knowing. Tapping the state of knowing was the purpose of assessment in this study.

The use of GJ tests in SLA research is controversial. Their validity has been questioned because L2 learners may rely on pure guessing or prescribed grammatical knowledge to judge given item sentences; therefore, the test may evaluate their explicit grammatical knowledge, rather than abstract, intuitive grammatical knowledge in the L2 (e.g., Cowan & Hatasa, 1994; Gass, 1994). In fact, Goss, Zhang, and Lantolf

18) Because some students needed to leave the class quickly for their next appointments, they filled out and submitted the forms later the same day. This arrangement resulted in missing data; four uptake claim questionnaires were not retrieved (Hisako (2), Keiko (1), and Shoko (1)).

(1994) found that, in taking GJ tests, advanced L2 learners used a variety of strategies, including referring to explicit grammatical knowledge. The concern that the GJ test assesses explicit grammatical knowledge is problematic for some cognitive-interactionist SLA researchers who define L2 learning as the ability to internalize explicit knowledge into native-like implicit knowledge. The GJ test is an inefficient assessment of such ultimate L2 learning because it is unclear whether the test assesses L2 learners' use of their explicit or implicit grammatical knowledge. However, the distinction between explicit and implicit L2 knowledge was not a concern in my study. Rather, the focus was on learners' competence in distinguishing a grammatically incorrect sentence from the correct one that they were exposed to in the teacher recast.

Construction and administration of GJ tests

The GJ tests in this study were group-specific, tailor-made tests, modeled on a format suggested by Gass (1994). The GJ tests contained between 8 to 15 test items.[19] My construction of the GJ tests began with the compilation of error treatment episodes (ETEs), particularly recast episodes (REs), from the classroom discourse transcripts, both teacher-fronted and group interactions (see the "Classroom Observation and Videotaping" section, pp.21-22). The test items were comprised of sets of ill-formed utterances (by a student) and the well-formed utterances (by the teacher) identified in the ETE and RE dialogues, which became the test items. Two NSs proofread the sentences for the tests to verify the grammaticality or non-grammaticality of each utterance.

The ill-/well-formed sentences were often modified to make them into GJ test items. That is to say, a student's ill-formed sentence, which tended to contain more than one mistake, was turned into a test item containing one mistake which contrasted with the teacher's model. For

19) I intended to construct GJ tests throughout the data collection sessions containing the same number of items; however, the final two weeks' GJ test contained only eight items due to the infrequent occurrence of feedback in the lessons.

example, a student's ill-formed utterance, "*If man and many … every people understand woman's thinking, this problem is … we can good … good situation*" was made into an ill-formed test item, "*If men understand women's thinking, we can good situation*." This ill-formed item was contrasted with a well-formed item, "*If men understand women's thinking, the situation will become better*." This well-formed sentence item incorporated the teacher's isolated recast, "*Become better*?" As seen in this teacher recast, some of the recasts were fragments, and needed modification. Because of the nature of the GJ test, no test items were derived from phonological REs or from an RE that dealt with the L1 use (Japanese).

A test based on each videotaped lesson (i.e., Mini-tests 1 to 6) was administered to the eight students within a week after the six individual videotaped lessons. These mini tests are referred to as Test 1 later in the thesis. The test was prepared in two versions: Group A and Group B. Each version contained the same test items derived from teacher-fronted episodes and different group-specific items derived from group-based episodes. In other words, the students were tested on the specific sentences to which they had been exposed. In order to make the length of each version the same, items from non-recast ETEs (drawn mostly from the group discussion) were included in the shorter tests.[20] **Table 3** summarizes the number of test items for each GJ Mini-test.

Each student took the tests individually at her own pace. The test instructions were in Japanese. Adapted from Gass (1994) and Swain and Lapkin (1998), the test asked the student to judge the sentences using the following choices: "Absolutely Correct," "Probably Correct," "Probably Incorrect," "Absolutely Incorrect" or "Not Sure" (see Appendix C for sample items). Each student took about 5 to 10 minutes to complete each test.

In the ninth week of data collection (i.e., three weeks after Mini-test

20) Group B tended to receive less feedback than Group A.

Table 3: GJ Test Item Numbers

GJ Test	Total items	T-fronted items	Group specific items
Mini-test 1	15	2	13
Mini-test 2	15	4	11
Mini-test 3	15	6	9
Mini-test 4	15	5	10
Mini-test 5*	8	4	4
Mini-test 6	8	4	4

*Mini-test 5 was excluded from analyses.

6), the eight students took GJ Test 2. Test 2 was composed of the same test items that individual students took as part of Test 1. The items were arranged randomly. Each student took the test at her own pace. It took them about 25 to 40 minutes to complete the test.

Students' Stimulated Recall Interviews

Rationale for conducting stimulated recall interviews in this study

Stimulated recall is a research method used for collecting introspective verbal data in psychology (e.g., Ericsson & Simon, 1984) and recent L2 learning research (Cohen, 1994; Færch & Kasper, 1987; Gass & Mackey, 2000) as a means of exploring the cognitive processes people use for solving problems. The rationale behind this method is based on an information processing approach "whereby the use of and access to memory structures is enhanced, if not guaranteed, by a prompt that aids in the recall of information" (Gass & Mackey, 2000, p.17). These verbal reports are one type of data usable for understanding learners' cognitive involvement such as paying attention.

Limitations of the introspective data collection method have been identified. One major concern is reliability. Claims have been made, for example, that results vary according to the instructions given to the learners as well as to the characteristics of the participants in the research (Olson, Duffy, & Mack, 1984 as cited in Cohen, 1994). Differences in

human cognitive activity, however, are to be expected, as are the learners' idiosyncratic claims of uptake from the same lessons, as shown in Slimani (1987, 1992).

Stimulated recall interview procedure

Stimulated recall interview sessions were conducted with the eight students individually within a week of videotaping.[21] The length of each session was between 30 and 50 minutes. For the stimulated recall session I selected the target REs,[22] and a few episodes of other types of feedback and other activity moments (as distracters so that the students would not become over-sensitized to recasts and correction).[23] Each student was asked to recall and recount what she was thinking at that moment. The baseline question during viewing was "what were you thinking then?" A set of follow-up questions was asked of all the participants in the third and sixth interviews in order to elicit the students' beliefs about learning. (See Appendix D for detailed examples of questions.)

The stimulated recall sessions were conducted in Japanese because of the complex nature of the information students were asked to provide. The students' verbal reports were audiotaped and later transcribed for analysis.

---

21) The interval time between the event and the stimulated recall is a methodological consideration. Although Gass and Mackey (2000) suggest an ideal stimulated recall is one given immediately after the event, it was not feasible in this research design due to constraints such as the time required for constructing tailor-made GJ tests and students' schedules.

22) The number of target REs differed among the students because the teacher provided a different number of recasts in each group interaction.

23) In other studies, the students were allowed to identify error correction episodes (Roberts, 1995) or to stop the video and comment on the episodes of their choice (Mackey et al., 2000). I did not choose these methodologies due to time constraints imposed by the schedules of the students. There simply was not enough time to go through the videos made of classroom interaction with each student.

Teacher's Stimulated Recall Interview

Teacher's stimulated recall interview procedure

In that feedback is a teacher's reaction to learners' language use, feedback provision is a dialectic process. Capturing the teacher's thoughts and intentions at the moment of feedback provision was, therefore, necessary to complete the picture of feedback episodes.

Stimulated recall interviews with the teacher were conducted twice: in the fourth week of data collection and four weeks after the final videotaped class (see Table 2). Each recall session was approximately 1.5 hours long. The stimulated recall interview procedure was similar to that used with the students; many of the same REs used in the student sessions were shown to the teacher as stimuli. Additional ETEs were shown to elicit the teacher's pedagogical beliefs. There were two baseline questions for the teacher: "What were you thinking then?" and "Why did you do this?" (see Appendix E for detailed examples of questions). The recall sessions were conducted in English and were audiotaped for transcription and analysis.

Data Analysis

The data consisted of the recast episodes (REs) identified in the classroom discourse, the students' written reports from the uptake claim survey, the stimulated recall interviews, and the GJ test results. Although the data were collected over six cyclical sessions, those from the five sessions in which all participants were present were used for analysis in this study. One session, when two students were absent, was not included.

Analysis of Classroom Discourse Data

A recast in this study is operationally defined as either an isolated or expanded rephrasing of a learner's non-target-like utterance provided by the teacher immediately after the non-target-like utterance. Non-target-like utterances, or "error types," are those with linguistic problems

(i.e., grammatical, lexical, phonological) (e.g., Lyster, 1998a), L1 use (e.g., Lyster & Ranta, 1997),[24] incomplete and fragmented utterances (e.g., Farrar, 1990), or combinations of these problems. I focused on these problem areas for which recasts could be provided, reflecting actual occurrences of recasts in the natural classroom interactions.

A recast episode (RE) started with a non-target-like language use that was reacted to by an interlocutor, involved a sequence of one or more feedback turns including at least one recast, and ended when the focus shifted away from the error.[25] An RE could include the student's reaction to the feedback occurring before the topic shift.[26] As presented earlier, an RE is a sub-category of "Error Treatment Episode" (ETE) in which a non-target-like utterance was reacted to by an interlocutor in any form of feedback (Lyster & Ranta, 1997). An RE may contain a single recasting treatment (i.e., Single Recast Episode (SRE)) as in Example 1. In this episode, the teacher gave a recast in response to Shoko's non-target-like utterance (turn 5B-018) and then moved on to the next topic, thus ending the episode.

Example 1:[27]

5B-016	Ms. Johnson:	Ok. Everything was on sale. Why?
→5B-017	Shoko:	Because … baseball winner.
→5B-018	**Ms. Johnson:**	**Ok. Because they won the Japan series.** Do you like baseball?

24) Panova and Lyster (2000) considered teacher's rephrasing feedback to L1 use as translation rather than recast.

25) It was often difficult to tell what the linguistic focus of a recast was.

26) The students did not always react to the teacher's recasts. An RE in which no student reactions occurred before the topic shift and after the teacher feedback was considered as a "no-uptake" or "no-chance" RE as described later.

27) Transcript line numbers are composed of the lesson number (e.g., fifth lesson), the group (e.g., Group B), and the turn (e.g., 016). The arrows indicate turns in an RE. Teacher recasts are bolded. The basic classroom discourse transcript conventions were: a hyphen for a discontinued utterance, three dots for a pause, and xx for unclear utterance.

An RE may contain more than one recast as in Example 2, which is a compound recast episode (CPRE). In this example, the teacher provided a recast (turn 3C-471) to complete the student's utterance. Even so, the student's utterance was not correct in the context and so the teacher provided another recast which was incorporated into her expanded turn (turn 3C-473).

Example 2:[28]

3C-469	Ms. Johnson:	Ok. Masako?
→3C-470	Masako:	Studying hard to en- … ent…
→3C-471	**Ms. Johnson:**	**Enter**
→3C-472	Masako:	enter examination
→3C-473	**Ms. Johnson:**	**Ok. Studying hard for examinations. Ok. That's a problem, yeah.**

REs may also contain other error treatments[29] as in Example 3, which consists of a complex recast episode (CXRE). In this example, the teacher put Asako's problematic sentence into a more target-like form and added explicit corrections.[30]

Example 3:

→6C-004	Asako:	… My opinion is … cats are more dangerous animal than dog because

28) The C in the transcript line number stands for "Class," indicating that dialogue took place during the teacher-fronted interaction.

29) Other types of error treatments were adopted from Lyster and Ranta (1997): Explicit Correction, Clarification Requests, Metalinguistic Feedback, Elicitation, and Repetition.

30) Explicit correction is indicated in italics in Example 3. My interpretation of explicit correction is broader than Lyster and Ranta's (1997) which contained a clear indication of a student utterance as an incorrect utterance. In my data, Ms. Johnson infrequently rejected her students' utterances as inaccurate utterances. Her explicit corrections were colloquial grammatical explanations; that is, she gave grammatical rules without using grammatical terms. Because she did not use metalinguistic terms, such grammatical corrections were not coded for "metalinguistic feedback" but for "explicit correction."

they … they keep going when they met a car. They never change their way, and they … run over.

→6C-005 **Ms. Johnson:** **Ok. Yeah, … cats are … Cats are at more danger.**
Ok. *So something dangerous is going to hurt something else. At danger is they can be hurt. Ok.*

REs varied with respect to students' reactions. The reaction types considered were: (1) student uptake, in which the student who received a teacher recast in response to her erroneous utterance either repeated or incorporated the teacher's recast in the following turn; (2) no student uptake ("no-uptake"), in which the student did not repeat or incorporate the teacher's recast even though there was a chance to do so (e.g., a long enough pause after the teacher's turn); and (3) no uptake opportunity ("no-chance"), in which the student did not repeat or incorporate the teacher's recast because there was no chance to do so. The no-uptake category included the student's semantic, rather than linguistic, reactions such as acknowledging the teacher's reaction to her utterance by nodding or saying "yeah" or "ah." The original student uptake devised by Lyster and Ranta (1997) in teacher-student classroom interaction represented the successful students' reaction (i.e., correct repetitions or incorporations of the teacher's recast), and was distinguished from unsuccessful students' reaction (i.e., repetitions or incorporations of the teacher's recast that are still erroneous). However, I did not make a distinction between the two because the students' linguistic attention was more important than their linguistic behaviour in this study.

Reliability of the coding for REs and reliability of codes applied to each RE was calculated with assistance from a second coder (see detailed coding systems in Appendix F).[31] A doctoral student in a graduate

31) Prior to coding the classroom discourse data for REs, I coded them for ETEs.

program related to second language education in Canada, whose L1 was English, was the second coder. I first explained the coding system and she practiced it on a set of classroom discourse data from this study. Then, we individually coded 20% of the data (i.e., two lesson transcripts). After individual coding, we compared our coding results and resolved discrepancies. The level of inter-coder agreement was 90%.

Analysis of Stimulated Recall Data[32)]

Different coding schemes were developed and used for the analysis of the stimulated recalls from the students and from the teacher.

Coding schemes for student stimulated recall interviews

Table 4 below describes the list of codes and their definitions that I developed and used for analysis of the student recalls. The students' recalls referred to their receptive (listening) and/or productive (speaking) activities; therefore, I differentiated the recalls according to the activities to which they referred. The two speaking-related attention-levels were: (1) "Speaking for Content" (Sp-C), where the students gave content-related accounts for saying something in the conversation, and (2) "Speaking for Language" (Sp-L), in which they gave language-related accounts for saying something in the conversation.

The students' listening-related recalls differed regarding the direction of their attention while listening; they sometimes made recalls specifically in relation to the teacher's utterances. Therefore, I established

---

Another doctoral student in the graduate program related to second language education in Canada, whose L1 was English, was the second coder of this part. I first explained the coding system and he practiced it on a set of data. Then, we individually coded 20% of the data (i.e., two lesson transcripts). After individual coding, we compared our coding results and resolved discrepancies. The level of inter-coder agreement was 87%.

32) The basic transcript conventions for stimulated recall interview data were: three dots means a pause, three dots in square brackets means I cut some unclear or redundant utterances.

Table 4: Codes for Student Stimulated Recall Interview Protocols

Levels of attention	Definition	Example Quotes*
Not Listening **No-Lis**	・Confession or explanation of doing something else. ・Thoughts or reflection about the student's emotional response or personal matter, unrelated to the given activity.	"I was not listening then. I was playing with my cell phone. I wanted to return her mail early." "I was thinking at the time, 'I'm tired, I don't like these'."
Listening for Gist **Gist**	・Comments about the student's listening behaviour, often about superficial matters such as presentation styles and listening conditions.	"I was listening, but it was difficult to listen to her." "Her group presented in a different way from ours. They said everything in a smooth flow."
Listening for Content **Lis-C**	・Comments relating to the concrete content meaning of what the student heard. They include the student's own reflective comments on the content of the lesson or discussion.	"Tokiko was explaining that we should search for rights … relating to the Declaration of Human Rights." "I didn't think this right was so important so when I listened to Shoko, I did not agree with her."
Listening for Language ** **Lis-L**	・Comments relating to aspect(s) of language (e.g., lexicon, grammar, or phonology) and/or language use of what the student heard. This includes identifying language problems.	"I did not understand the meaning of 'electric fan'." "She said 'I have' and I thought it was wrong."
Hearing the Teacher **Hear-T**	・Observational comments about the teacher's utterance. This includes the student's appreciations of the teacher's erudition.	"I was thinking 'oh, she gives a follow-up question today'." "I was thinking 'Ms. Johnson knows everything'."
Listening to the Teacher **Lis-T**	・Concrete comments relating to the meaning of the teacher's utterance. This includes the student's reflection, and/or (dis)agreement with the teacher.	"The teacher said that adults don't need love, but I thought if they don't have love, they can't love others, so I was not sure."
Noticing teacher Feedback **NFB**	・Comments indicating that the student noticed the teacher's feedback as feedback.	"I was wondering how to say this and the teacher said [a sentence] so I thought "oh, that's how to say it."
Speaking for Content **Sp-C**	・Explanation about the student's own utterance in relation to expressing ideas.	"Because I did my homework, I thought I'd give my answers."
Speaking for Language **Sp-L**	・Explanation about the student's own utterance in relation to clarifying language use.	"I thought [saying 'I have'] was wrong so I said 'you have?' "

*Examples are translated from Japanese into English.
**Listening for Language was further categorized in reference to Language Related Episodes (Swain & Lapkin, 1998). The subcategories included lexicon, form, pronunciation, and discourse.

categories exclusively for the students' comments about the teacher. Further, the students' recalls in relation to their listening in general indicated different degrees of attention. I categorized the four degrees of attention they paid while listening in general: (1) "Not Listening" (No-Lis), (2) "Listening for Gist" (Gist), (3) "Listening for Content" (Lis-C) and (4) "Listening for Language" (Lis-L). No-Lis included the students' comments that they were not listening or their emotional comments such as they were feeling tired. Gist included their comments about presentation styles or listening conditions. Such comments were superficial and did not clearly indicate their comprehension of the event and conversation content. The third category, Lis-C, was for the students' recalls indicating they listened to and understood the conversation content. Their understanding included their awareness of being uncertain or confused about the conversation, too. The final Lis-L category included the students' comments reflecting their thoughts and attention to the use of language.

The receptive attention exclusively paid to the teacher was divided into three categories: (1) "Hearing the Teacher" (Hear-T) (2) "Listening to the Teacher" (Lis-T), and (3) "Noticing teacher Feedback" (NFB). The Hear-T category was for the students' superficial, observational comments about the teacher's verbal behaviour. This included, for example, their appreciation of the teacher's erudition. The Lis-T category included the students' recalls which were reflective of and/or specific about the content and significance of the teacher's utterance. The final category, NFB, contained students' comments indicating they noticed the teacher had provided feedback.

The student stimulated recall data were coded in the original language, which was mainly Japanese with the occasional mixing of English words or phrases. I coded all the interview transcripts. A Japanese doctoral student in a graduate program related to second language education in Canada assisted me as a second coder of the student stimulated recall interview data. I first explained the coding schemes to

him, and he then practiced them on a small set of recall data from this study. After the practice, the second coder and I individually coded the same 20% of the stimulated recall data, which was composed of one recall session transcript from each of the eight students. After individual coding, we compared our coding results and resolved discrepancies. The inter-coder agreement was 89%.

The coded data were further clustered according to four levels of attention (see **Table 5**). "No Attention" (NoA) was the lowest attention category, and included recalls coded for No-Lis, Gist, and Hear-T. The recalls did not show clear or explicit understanding of the event and discussion. The second category, "Attention to Content" (AC), included stimulated recalls coded for Lis-C, Sp-C, and Lis-T. These comments showed the students' involvement in the conversation and understanding (or lack of understanding) of the meanings of events and discussion content. The third category, "Attention to Language" (AL), included Lis-L and Sp-L. These recalls reflected the students' attention to language (e.g., lexicon, form, phonology, or discourse) in relation to the conversation they heard. The final attention category was "Noticing Feedback" (NFB), which was the same as "Noticing teacher Feedback."

Table 5: Attention Categories

Levels of Attention		Corresponding Recall Protocols
No Attention	(NoA)	No-Lis, Gist, Hear-T
Attention to Content	(AC)	Lis-C, Sp-C, Lis-T
Attention to Language	(AL)	Lis-L, Sp-L
Noticing Feedback	(NFB)	NFB

The codes were applied to all recalls the students made in response to the ETEs they watched during the stimulated recall interview sessions, and the entire set of recalls was used for qualitative analysis. For the quantitative analysis, the "highest" level of attention code occurring in the target RE was tallied, assuming that AC was more important than NoA, AL was more important than AC, and NFB was more important than AL. In

other words, if a student's recall about a specific RE contained comments coded for "AC," "AL," and, "NFB," the RE was counted as an NFB RE, whereas if comments on another RE contained "AC" and "AL" codes, the RE was counted as an AL RE. The target REs focused on for the quantitative analysis were the same 14 REs from which GJ test items had been constructed.[33] Depending on the group the individuals belonged to, the students were exposed to four to seven group REs along with the same five teacher-fronted REs. Individual students actually viewed and made between 6 and 12 recalls on REs as shown in **Table 6**. Fumiko's recall of teacher-fronted REs is short by 1 because of the stimulated recall interview time constraints; in one stimulated recall session, she could spare only 30 minutes due to her next appointment schedule.

Table 6: Students' Recalls with Respect to REs

	T-fronted context	Group context	Total
Fumiko	4	2	6
Keiko	5	2	7
Eiko	5	4	9
Aiko	5	4	9
Hisako	5	4	9
Tokiko	5	5	10
Yasuko	5	6	11
Shoko	5	7	12
Total	39	34	73

Coding schemes for teacher's stimulated recall interview

The codes for the teacher's recall interview were developed based on the current data, with reference to other schemes used in teacher belief-

33) Among the 25 REs, 11 REs were not used for the qualitative analysis. Eight of these 11 REs were not applicable to GJ test items, 1 set of GJ Test items from 1 RE turned out to be problematic (RE1204), and 2 REs were not included in the stimulated recall sessions due to interview time constraints (REs108, 109). See Appendix G for detailed profile of episodes.

system research (Borg, 1998). Much of the teacher's recall consisted of reporting the beliefs which influenced her teaching strategies. This is likely attributed to a long interval between the events and the stimulated recall interviews, and also because of the base-line question "Why did you do this?" The coding scheme contained components of her beliefs about teaching and learning.

As shown in **Table 7**, five major belief systems were identified. The first category, "Provider," consisted of Ms. Johnson's belief in her role as an L2 input provider, a linguistic correction provider, and a provider of opportunities for language practice. The second category, "A teacher in a theme-based class," was for her comments regarding theme teaching in the theme-based discussion class. The third category, "A teacher in the class/school," involved Ms. Johnson's concerns for her students in general, for the well-being of her class and for the school curriculum. The "students' language" category contained her comments regarding students' linguistic accuracy and performance (namely audibility of their utterances), and her expectations of her students. The expectations could be as high as to think the student "should know it because it's easy stuff," or could be as low as to think the students "should not know it yet." The "Students' performance" category was for the teacher's comments regarding their task performance, classroom participation, and readiness to learn.

The teacher's stimulated recall data were coded with assistance from a second coder. A doctoral student in a graduate program related to second language education in Canada, whose L1 was English, was the second coder. I first explained the coding system. Then, we individually coded the entire teacher interview data independently. After individual coding, we compared our coding results and resolved discrepancies. The inter-coder agreement was 96%.

Analysis of Uptake Claim Form

All entries the students made on their uptake claim forms were listed

Table 7: Codes for Teacher Stimulated Recall Interview Protocols

Categories	References	Example Quotes
Provider of	Input	"I think that they are linguistically ready to handle what I told them."
	Correction/ Intervention	"I never plan 'Oh, today I'm going to do grammatical correction."
	Practice opportunities	"When I call on some person … it's usually because I feel that person can do it but needs a little bit more speaking practice."
A teacher in a theme-based class	Theme instruction	"I find [the students] are very quick on [claiming their rights], but uh, recognizing that rights also bring with them responsibilities is something not that clear" (talking about an activity under "human rights" theme.)
A teacher in the class/ school	General instruction	"I want them to feel that I'm paying attention to them."
Students' language	Accuracy	"If it's too inaccurate to understand what they want to say, that's a big problem. … but if their grammar doesn't get in their way of meaning, then it's a small problem."
	Audibility	"It was hard to hear. Her voice will carry, but only when she wants it to."
	Positive expectation	"She should know that, that's easy stuff to say."
	No expectations	"I think they didn't know the word 'restroom' for 'bathroom'."
Students' performance	Task	"It's a class. Are you here? Are we all together? So … they looked like they were not paying attention."
	Participation	"She seemed like … she wanted to do it [=speaking about a language school advertisement]" (Speaking)
	Readiness	"Some students don't respond well to correction, and some students really want it."

on spreadsheets according to each group in each lesson. Claimed items were then checked against the written classroom materials, namely, textbook and activity prompts and the classroom interaction transcript to identify when the students were exposed to and/or used the claimed item.

Analysis of Grammaticality Judgment Tests (GJ tests)

The total number of items administered in the five Mini GJ tests (Test 1) was 68. However, the number of test response entries actually used in the analysis varied according to individual students. This was because the number of group-specific test items varied according to the groups the individual students belonged to. The total of 68 also included items from non-recast ETEs, and these were excluded from the analysis as well.[34] As a result, the maximum total number of the individual student's responses to GJ test items derived exclusively from REs ranged from 16 to 26 (see **Table 8**).

Table 8: Students' GJ Test Responses with Respect to REs

	T-fronted context	Group context	Total
Fumiko	12	4	16
Keiko	12	4	16
Eiko	12	8	20
Aiko	12	8	20
Hisako	12	8	20
Tokiko	12	10	22
Yasuko	12	14	26
Shoko	12	14	26
Total	96	70	166

As described earlier (in the "Coding schemes for student stimulated recall interviews" section, pp.39-43), the students did not view and

34) Despite the proofreading, some test items were later considered inappropriate.

Table 9: Students' GJ Test Responses with Respect to REs with Recall Data

	T-fronted	Group	Total
Fumiko	7 [35]	4	11
Keiko	9	4	13
Eiko	9	8	17
Aiko	9	8	17
Hisako	9	8	17
Tokiko	9	10	19
Yasuko	9	12	21
Shoko	9	14	23
Total	70	68	138

provide stimulated recall comments to all the REs from which GJ test items were constructed (see Table 6). Thus, the number of individual student's responses to GJ test items corresponding to the REs about which the students provided their stimulated recalls (i.e., students' attention data) ranged between 11 to 23 as shown in **Table 9**.

GJ test outcomes were analyzed for two purposes: (1) to obtain the total number of correct judgments by individual students as a means of measuring their learning, and (2) to understand students' correct and incorrect judgments in relation to their attention during the REs. Although using one set of data for both purposes would be ideal, the size of the GJ test dataset corresponding to the attention data was very small. I, therefore, decided on two approaches to the GJ test data. For the first purpose of analysis, I used the maximum number of responses to RE-based GJ test items (i.e., a total of 166 items, ranging between 16 to 26 individually). This allowed me to analyze the maximum number of responses to commonly answered items: 12 teacher-fronted items and 4 to 14 group-specific items (see Table 8). For the second purpose, I

35) Fumiko did not have enough time to view and comment on all the target REs during one stimulated recall session. The session needed to finish in 30 minutes due to her schedule that day.

focused on the responses to test items from the REs corresponding to stimulated recall data (i.e., the REs to which the students made reference in the recalls). In this data-pool were 70 teacher-fronted items (ranging between 7 and 9 individually) and 68 group-specific items (ranging from 4 to 14 individually) (see Table 9).

Each student's correct judgment, that is either absolutely or probably correct choices on grammatical sentence items and absolutely or probably incorrect choices on ungrammatical sentence items, was counted for both Test 1 and Test 2.[36)]

36) Data were originally analyzed using the four categories; however, the results were similar to those found when the absolute and probable categories were combined. Therefore, the analysis reported here is based on combining the absolute and probably categories.

Summary of Datasets

Table 10 summarizes the datasets collected and analyzed in this study.

Table 10: Summary of Datasets

Data Analyses	Data for discussion		
Error Treatment Moves	Recast		
Recast Episodes	Delivery SRE CPRE CXRE	Error Grammar Lexicon Incomplete sentence Pronunciation L1	Reaction Uptake No－uptake No－opportunity
Grammaticality Judgment Test	Test 1 Correct judgments	Test 2 Correct judgments	
Students' Stimulated Recall (levels of attention)	No-attention Attention to content Attention to language Noticing feedback		
Uptake Claim Form	Words and phrases Spelling Pronunciation Grammar Ways of using the language		
Teacher's Beliefs Interviews	Provider of L2 learning opportunities Theme learning Human education Students' language use Students' performance Challenges		

Chapter 4 : Phase 1 - Findings and Discussion

This chapter presents the findings for each of the research questions:

1. What recasts were the students exposed to in a communicative theme-based EFL classroom?
2. In what ways did the students attend to the teacher's utterances and recasts during interaction?
3. Which recasts related to the students' L2 learning as measured by GJ tests?

Presentation of the findings will be followed by discussion.

Findings for Research Question 1:

What Recasts Were the Students Exposed to in a Communicative Theme-Based EFL Classroom?

It was evident in the classroom and from the audiotaped discourse transcripts that Ms. Johnson gave feedback less frequently than expected. I present the quantitative findings from my first research question focusing on (1) the occurrence and (2) the characteristics of REs in Ms. Johnson's classroom. I also provide some comments from Ms. Johnson's recalls that add to the quantitative findings.

The total number of ETEs where Ms. Johnson provided any kind of feedback was only 27 throughout the 350-minute recording time. ETEs in each lesson ranged from three to eight, and the average was 7.7 per lesson. Among these ETEs, 25 (93%) were REs.[37] Other ETEs without recasts included one instance of elicitation and one instance of metalinguistic feedback.

37) Because the number of REs in this study is low, any inferences drawn must be considered tentative. See Appendix H for all REs.

Recast Episodes in Classroom Interaction

As shown in **Table 11**, Ms. Johnson provided a total of 25 REs in her class. The majority (68% or 17) consisted of SREs, followed by approximately 20% (5) of CXREs and 10% (3) of CPREs. Ms. Johnson's RE delivery pattern was similar in the two different interaction contexts in the class: teacher-fronted and group. In both contexts, SREs accounted for approximately 70% of the sub-total of REs, CPREs for about 10%, and CXREs for about 20%.

Table 11: REs and Delivery Type

Delivery Type	T-fronted context		Group context		Total in class	
	Number	%	Number	%	Number	%
SRE	9	64.3	8	73	17	68
CPRE	2	14.3	1	9	3	12
CXRE	3	21.4	2	18	5	20
Total	14	100	11	100	25	100

When classified according to error type, 15 REs in total were found to focus on grammatical REs (60%) (see **Table 12**). The next most frequent error type treated in the class was incomplete sentence problems (16% or 4). Lexical, L1 use, and phonological problems were least treated by recasting. Ms. Johnson gave recasts in a similar manner to each language problem in the two interaction contexts of the total REs; approximately 60% of these were in response to grammatical problems in both the teacher-fronted and the group contexts; approximately 15% were in response to incomplete sentences. Although REs to treat lexical problems occurred in both the teacher-fronted and group contexts, they were less than 10%. REs to phonological problems occurred only in the teacher-fronted activities.[38)]

38) One instance of L1 use, which was treated by elicitation, occurred in a teacher-fronted ETE.

Table 12: REs and Error Type

Error Type	T-fronted context		Group context		Total in class	
	Number	%	Number	%	Number	%
Grammar	8	57.2	7	64	15	60
Lexicon	1	7.1	1	9	2	8
L1 use	0	0	1	9	1	4
Phonology	3	21.4	0	0	3	12
Incomplete	2	00.3	2	18	4	16
Total	14	100	11	100	25	100

As shown in **Table 13**, uptake REs accounted for 28% of the total REs, no-uptake REs for 20%, and no-chance REs for 52%. Given that no-uptake REs included the teacher's provision of a pause for uptake to which the student did not react, Ms. Johnson was found to have provided chances for uptake in almost half of her REs (i.e., 28% + 20%). Among the uptake opportunities provided by Ms. Johnson (7 uptake + 5 no-uptake REs), the students actually responded about half the time. Ms. Johnson appeared slightly more likely to provide uptake opportunities in the group context (54% of the time) than in the teacher-fronted context (43% of the time).

Table 13: REs and Learner Reaction Type

Reaction Type	T-fronted context		Group context		Total in class	
	Number	%	Number	%	Number	%
Uptake	4	29	3	27	7	28
No-uptake	2	14	3	27	5	20
No-chance	8	57	5	46	13	52
Total	14	100	11	100	25	100

Ms. Johnson appeared to have provided recasts in a similar manner in both teacher-fronted and group contexts; however, closer examination of the data shows that individual students' exposure to teacher recasts varied. Tables 14 to 16 show the sub-distributions of REs by group. As

seen in **Table 14**, students in Group 1A were exposed to more REs (45% of group REs or 5) than other group members: most REs were SREs. **Table 15** shows that the REs to Group 1A concentrated on grammatical REs whereas Group 2B did not receive any grammatical REs. **Table 16** shows that Group 1A, in addition to two uptake REs, received three no-chance REs, but Group 2A did not receive any no-chance REs. Groups 1B and 2B each received one no-chance RE.

Table 14: Group REs and Delivery Type

Delivery Type	1A	1B	2A	2B	Group Total	
	Number	Number	Number	Number	Number	%
SRE	4	2	1	1	8	73
CPRE	1	0	0	0	1	9
CXRE	0	0	1	1	2	18
Total	5	2	2	2	11	100

Table 15: Group REs and Error Type

Error Type	1A	1B	2A	2B	Group Total	
	Number	Number	Number	Number	Number	%
Grammar	5	1	1	0	7	64
Lexicon	0	1	0	0	1	9
L1 use	0	0	0	1	1	9
Phonology	0	0	0	0	0	0
Incomplete	0	0	1	1	2	18
Total	5	2	2	2	11	100

Table 16: Group REs and Learner Reaction Type

Reaction Type	1A	1B	2A	2B	Group Total	
	Number	Number	Number	Number	Number	%
Uptake	2	0	1	0	3	27
No-uptake	0	1	1	1	3	27
No-chance	3	1	0	1	5	46
Total	5	2	2	2	11	100

In class, Shoko and Yasuko in Group 1A and 2A had the greatest exposure to teacher recasts, and Keiko and Fumiko in Group 1B and 2B had the least.

Ms. Johnson's Recasts

In order to further understand Ms. Johnson's recasts, I compared them across the RE types. As shown in **Table 17**, which summarizes the occurrence of REs with respect to delivery type of recasts and learners' reaction, the majority (59%) of SREs were no-chance REs. That is to say, Ms. Johnson tended to provide recasts in a single turn, and without waiting for a student's uptake move, she expanded her utterance in the same turn. Thus, her students could have thought responding to her recast inappropriate, even if they had noticed the correction. They could also have shifted their attention to the next topic raised in Ms. Johnson's expanded turn.

Table 17: Distribution of RE Delivery Type and Learner Reaction Type

	SRE		CPRE		CXRE		Total	
	Number	%	Number	%	Number	%	Number	%
Uptake	2	12	2	67	3	60	7	28
No-uptake	5	29	0	0	0	0	5	20
No-chance	10	59	1	33	2	40	13	52
Total	17	100	3	100	5	100	25	100

Table 18 summarizes error types in relation to delivery types and learner reaction types. The majority (80%) of grammatical problems were treated by SREs (68%). In addition, 75% of incomplete sentence problems were treated by SREs.[39] The REs treating incomplete sentence problems provided no-uptake chances. Many grammatical REs (67%) were also no-

39) This finding was rather unexpected because incomplete sentences require the listener (i.e., the teacher in this case) to infer the intended message meaning. The teacher's CXREs were more anticipated than SREs for incomplete sentence problems (see Mackey et al. 2000).

uptake chance REs, yet approximately 27% of grammatical REs received students' uptake reaction.

Table 18: Distribution of Error Type Across Delivery Type and Reaction Type

	Grammar (15 REs)		Lexicon (2 REs)		L1 use (1 RE)		Phonology (3 REs)		Incomplete (4 REs)		Total (25 REs)	
	Number	%	Number	%	Number	%	Number	%	Number	%	Number	%
Delivery												
SRE	12	80	1	50	0	0	1	33	3	75	17	68
CPRE	1	6.7	0	0	0	0	1	33	1	25	3	12
CXRE	2	13.3	1	50	1	100	1	33	0	0	5	20
Reaction												
Uptake	4	27	1	50	0	0	2	67	0	0	7	28
No-uptake	1	6	1	50	0	0	1	33	2	50	5	20
No-chance	10	67	0	0	1	100	0	0	2	50	13	52

A numerical summary of RE findings shows that Ms. Johnson provided recasts most frequently on grammatical errors within her extended turns and provided her students with few chances for uptake. Her recasting corresponds to findings from previous studies (Lyster, 1998a; Mackey et al., 2000; Oliver, 1995). However, the frequent SRE with few uptake opportunities raises a question about Ms. Johnson's intention in "error correction" through recasting: Did Ms. Johnson intend to instruct her students in linguistic accuracy? Oliver (1995) explained the differences between recasting and negotiation in relation to transparency of original meaning. She found her young NS participants gave recasts to erroneous NNS utterances whose meaning was transparent, whereas negotiation feedback was given to NNS utterances whose content was opaque. Assuming young children involved in peer activities would not have a clear intention for language instruction, I question whether Ms. Johnson had a clear intention to provide error correction in her class.

Her intention seemed to be directed to smooth communication rather than linguistic intervention. This appeared in Ms. Johnson's way of treating incomplete sentence problems. An incomplete sentence requires

the listener to infer the message meaning. The meaning might be negotiated through clarification requests. For instance, the NS and near native interlocutors engaged in a communicative task in Mackey et al. (2000) provided a recast and negotiation combination when they found it necessary to clarify or confirm the message's meanings. Ms. Johnson, however, frequently treated incomplete sentence REs with SREs instead of CXREs. This means she inferred incomplete message meanings and modeled a sentence for the students.[40)]

Ms. Johnson's recall interviews confirm that emphasis in class was on conveying meaning rather than on linguistic instruction. She emphasized the value of her students' efforts to communicate their ideas rather than their accurate production of the English language. She also indicated that she inferred and understood her students' intended meaning in their utterances. She accepted her students' ungrammatical utterances as long as "their grammar does not get in their way of meaning," and if it did, she would "check" the meaning.

Excerpt 1: Teacher's recall 1

> If it's too inaccurate to understand what they want to say, that's a big problem. [...] but if their grammar doesn't get in their way of meaning, then it's a small problem. It's something that's kind of annoying and that you'll hope eventually they'll get the correct grammar to go with their meaning but as a second language speaker myself, I know when I speak Japanese, my grammar sometimes is just horrendous, but I'm trying to get my idea across, [...] and I think that if my students get to a point where they can get up and talk to people ... in grammatical or ungrammatical language ... and communicate their ideas, I would be extremely satisfied. [...] If there are a little bit of pronunciation differences from native speakers, who cares? If there're some grammatical problems, well those will take care of themselves

40) Ms. Johnson's expanded SREs after her student's incomplete utterance resembles a mother's expansion (e.g., "non-corrective recasts" in Farrar, 1992) of her child's incomplete utterances, as observed in L1 research (e.g., Brown & Hanlon, 1970; Farrar, 1992).

> down the line if it's necessary, and if it's not necessary, why worry about it? But if they can't communicate their ideas, that's a problem. So I think if I'm getting their ideas, it's gonna be Ok. If I'm not sure, then I'm gonna check. (Teacher's recall 1: Nov. 2)

Ms. Johnson's overt instructional interest in communication needs to be acknowledged, particularly when the findings are compared and contrasted with previous experimental studies. The nature of controlled research design is such that recasting in experimental studies is a means of intentional error correction. Doughty and Varela (1998), for instance, called their intentional recasting "corrective recasts." Their corrective recasting involved two teacher turns; in the first turn, the teacher repeated the student's erroneous utterance with a rising intonation, and in the second turn, after a pause for the student's repair, she provided the correct form (i.e., a recast) with falling intonation. This is similar to my CXRE (i.e., repetition + recast) category.

Summary of Findings for Research Question 1

The students had little exposure to linguistic feedback in Ms. Johnson's class yet, among the limited linguistic feedback, they heard the recasts most frequently. Although recasts occurred in a similar manner in both the teacher-fronted and the group contexts, the sub-distribution of group recasts was different for each group. Thus, an individual student's experience of recasts in the class was different from her peers'.

Ms. Johnson was more likely to recast grammatical errors in SRE. Lexical and incomplete sentence problems were likely to receive CPREs or CXREs; the teacher sometimes inferred meaning of incomplete sentences and provided SREs. Many SREs did not provide the students with uptake opportunities.

The teacher's primary instructional interest was in facilitating students' communication in English. For instance, she inferred her student's intended meaning in the student's incomplete sentences, and responded in SREs rather than through negotiation. Her recall comments

support her recasting in the classroom: SREs with expanded turns were likely to maintain the flow of conversation. Although Ms. Johnson rarely focused on language in her class, she used CXREs when she encountered lexical or incomplete sentence problems and linguistic attention was required.

Although cognitive-interactionist SLA researchers seldom question the recast provider's intention, Ms. Johnson's intention seems to determine the types of recasts used. Recasts similar to CXREs were employed in previous studies intentionally investigating the corrective function of recasts. Ms. Johnson's CXREs were also associated with her attention to language. CXREs, therefore, may be a means to realize the teacher's linguistic intention.

Findings for Research Question 2:

In What Ways Did the Students Attend to the Teacher's Utterances and Recasts During Interaction?

In this section, I present findings from the students' uptake claim entries and stimulated recall data. I examined the students' stimulated recall data both quantitatively and qualitatively. The quantitative findings were exclusively based on the dataset composed of a total of 73 REs (see Chapter 3, p.43). The qualitative discussion was on the basis of the students' entire recall data of ETEs and REs that they viewed during the stimulated recall sessions.

The data show the students paid attention differently according to the characteristics of the recast but more so according to the interaction contexts. The qualitative data also add insight into the students' autonomous cognitive activities. The outline of the findings for Research Question 2 are: (1) the students' immediate recalls of their "learning," (2) the students' attention patterns during the REs in relation to (a) their interaction contexts, (b) delivery types, (c) error types, and (d) student reaction types, (3) the students' attention to the teacher and her utterance, and (4) the students' detailed accounts of paying or not paying attention

in (a) the teacher-fronted activities, and (b) the group activities.

Students' Immediate Recall of Their "Learning" [41]

The students made a total of 156 claim entries about their learning from the five lessons, ranging from 22 to 43 per lesson. The students' claims are concentrated in the "words and phrases" category, which was one of six categories included in the claim form (i.e., "words and phrases," "spelling," "pronunciation," "grammar," "language usage," and "other"). Some items were repeated across categories on the same claim form. For instance, a student recorded "constitution" in both the "words and phrases" and "pronunciation" categories. In a few instances, some students repeatedly recorded the same items on the different claim forms. For example, one student recorded "property" in her claim form for the first and second sessions.

I traced back the items claimed by the students to (1) written classroom materials, namely textbook and activity prompts; (2) teacher's utterances addressed to the class; (3) teacher's utterances addressed to the group; (4) utterances by a peer (s) in the group and/or in the class; and (5) the student's own utterance produced in the group or in the class. The students' claims were idiosyncratic, as Slimani (1987) found. Among the 156 item entries, nearly 42% were single report entries. There was a total of 17 entries for which more than three students made claims. I found the sources of 16 of these claims in the classroom transcripts or in the written materials. Among these 16 entries, two clearly originated in two instances of group discussions exclusively among students and six related to group discussions among the teacher and the students. Three entries related to incidences of teacher-fronted interaction. Five entries were traced back to every aspect of the classroom activity: teacher talk in

41) Learning in this section is independent from the learning I define based on the grammaticality judgment test results. The students' uptake claim for learning is defined as "what individual learners claim to have learned from the interactive classroom events" (Allwright, 1984 as cited in Slimani, 1992, p.198).

teacher-fronted discussion, group discussion, and materials.

Among the 156 items, only five were related to REs: "right," "responsibility," "force," "citizens don't," and "citizen doesn't." "Right" was claimed by one student only once. The other items were claimed by more than three students. "Force," "citizens don't," and "citizen doesn't" were from two group REs. "Responsibility" was claimed by six students on the Lesson 2 claim form, and it was identified in two REs in the same lesson. However, neither RE focused on this word as an error treatment. In fact, the concept of responsibility was the theme of the second lesson, and the students had intensive exposure to the term and used it during the lesson. Attributing the claim for "responsibility" to either RE would, therefore, be inappropriate.

In the immediate recalls of their learning from a lesson, the students identified various lexical-level items. The infrequent report for the grammar category indicated that the students did not perceive language problems at a morpho-syntactic level (see Mackey et al., 2000). Although few claims related to REs, the students reported items relating to the teacher's utterances during the group discussions. The teacher was, therefore, the source of much of their language learning.

Students' Attention in Relation to REs

Students' attention and interaction contexts

As shown in **Table 19**, the students paid active attention to about 80% of the teacher's recasts in the class, in general. Among a total of 73 recalls, students' "No Attention" (NoA) recalls (i.e., "not listening," "gist," and "hearing the teacher") numbered 15 (21%). The majority of attention recalls (34 or 47%) consisted of "Attention to Content" (AC). Although "Attention to Language" (AL) recalls and "Noticing Feedback" (NFB) recalls occurred individually only 16% of the time (12), a total of 32% of students' attention (16% of AL and 16% of NFB) was oriented toward language.

As shown in the "T-fronted" and "Group" columns in Table 19, the

Table 19: Attention Recalls

	T-fronted context		Group context		Total in class	
	Number	%	Number	%	Number	%
NoA	13	33	2	6	15	21
AC	17	44	17	50	34	47
AL	5	13	7	20	12	16
NFB	4	10	8	24	12	16
Total	39	100	34	100	73	100

students paid attention differently in the two forms of interaction. The students were less attentive in the teacher-fronted interaction (13 or 33% NoA) than in the group contexts (2 or 6% NoA). Students' content-oriented attention (AC) was relatively low (17 or 44%), and their language-oriented attention (AL and NFB) was only 23% (13% +10%) in the teacher-fronted contexts. Their AC recall was 50% and their AL and NFB recalls totalled 44% in the group setting.

That these students tuned out the conversation during the teacher-fronted discussion was usually obvious in the video, as evidenced by whispering with their group mates, playing with stationery or cell phones, looking at dictionaries or reading different written material. Students' stimulated recalls revealed that they found interaction in teacher-fronted episodes distant and less involving. For instance, Shoko tuned out during the teacher-fronted interaction because she felt bored by not getting opportunities to give her direct verbal reaction to the speakers.

Excerpt 2: Shoko's recall session 4, November 1.

> I was not listening then. [...] When I saw myself in the video, I thought I looked very bored. I don't like to listen to others. Just listening to them is not fun. I always want to talk. So I tune out when others give opinions, and since I'm bored, I find other things to do. For instance, I look in the dictionary because reading random pages in the dictionary is fun.

In contrast to the remoteness of teacher-fronted interaction, group

discussion was direct and immediate for the students; therefore, the teacher's utterances to the group were usually meaningful and relevant to their talk. For instance, Fumiko recalled her reflection on Ms. Johnson's comments in her group while the members discussed recycling.

Excerpt 3: Fumiko's recall session 4, November 2.[42]

> Yeah… the teacher said there was not enough space in her apartment even though she wanted to recycle something, so I thought there was a situation for some people that they could not recycle things even though they wanted to.

The students also asked the teacher for help in facilitating their group discussion, and in preparing their answers.

Excerpt 4: Tokiko's recall session 1, October 5.

> That was when we had different opinions and we became confused, and we needed the teacher's help.

Excerpt 5: Keiko's recall session 3, October 19.

> We asked the teacher if this [item] was similar to '[the concept of peace]', then she told us it was. So I thought we were doing the right thing and had the right answer.

Because the teacher talk in the group discussion was immediate to their interest and needs, the students' attention to Ms. Johnson and her utterances was higher than it was in the teacher-fronted discussions.

Students' attention and delivery types

Table 20 summarizes the distribution of students' attention recalls across delivery types of REs in the classroom. SREs were more likely to be associated with AC recalls (52%), followed by NoA recalls (26%). Student's language-oriented attention (AL and NFB) in relation to SREs

42) What I coded as reflecting the students' attention to content is highlighted by a single underline. The omitted words contextually evident in their Japanese recalls are supplied in square brackets in English translation.

was 22%. Student's attention to CPREs concentrated on NFB. Students' attention in CXREs was more likely to be AC (40%); their language-oriented attention (AL and NFB) added up to 50%. Thus, in general, CPREs and CXREs were more likely than SREs to be associated with language-oriented attention (AL and NFB).

Table 20: Attention Recalls and Delivery Type in Class

	SRE		CPRE		CXRE		Total	
Attention	Number	%	Number	%	Number	%	Number	%
NoA	13	26	0	0	2	10	15	21
AC	26	52	0	0	8	40	34	47
AL	8	16	0	0	4	20	12	16
NFB	3	6	3	100	6	30	12	16
Total	50	100	3	100	20	100	73	100

As in the previous section that found contextual differences in students' attention, a similar contextual difference is observed in the distribution of their attention across delivery types. As shown in **Table 21**, SREs in the teacher-fronted contexts were most likely to be associated with NoA recalls (48%). The students paid language-oriented attention (AL and NFB) in only 13% of teacher-fronted SREs. Most teacher-fronted CXREs were related to AC (50%). A total of 37.5% of teacher-fronted CXREs involved the students' paying language-oriented attention (AL and NFB). As an independent category, however, NFB recalls accounted for even fewer of the total CXREs. Students' NFB recalls in the teacher-fronted SREs and CXREs were 9% and 12.5% respectively.

As shown in **Table 22**, the group SREs were more likely to be related to AC recalls (63%), followed by language-oriented recalls (AL and NFB) (30%). Seven percent of the student attention in SREs was NoA. Students' attention in the group CPREs and CXREs was associated with NFB recalls.

In sum, SREs were less likely than other delivery types to relate to language-oriented attention. However, the students were more likely to

Table 21: Attention Recalls and Delivery Type in the Teacher-fronted Setting

	SRE		CPRE		CXRE		Total	
Attention	Number	%	Number	%	Number	%	Number	%
NoA	11	48	—		2	12.5	13	33
AC	9	39	—		8	50	17	44
AL	1	4	—		4	25	5	13
NFB*	2	9	—		2	12.5	4	10
Total	23	100	—		16	100	39	100

*All recalls coded for NFB concurrently contained AL recalls.

Table 22: Attention Recalls and Delivery Type in the Group Setting

	SRE		CPRE		CXRE		Total	
Attention	Number	%	Number	%	Number	%	Number	%
NoA	2	7	0	0	0	0	2	6
AC	17	63	0	0	0	0	17	50
AL	7	26	0	0	0	0	7	20.5
NFB*	1	4	3	100	4	100	8	23.5
Total	27	100	3	100	4	100	34	100

*Seven of these eight recalls coded for NFB (87.5%) concurrently contained AL recalls.

be attentive to language in SREs if the episode occurred in the group interaction than in the teacher-fronted interaction. CXREs were likely to relate to students' AL and NFB recalls, and the students were more likely to make NFB recalls in the group than in the teacher-fronted CXREs.

Students' attention and error types

As shown in **Table 23**, students' attention expressed in their recalls in relation to grammatical REs was most likely to be AC (31 recalls or 54%), followed by NoA (13 or 23%). Their language oriented attention (AL and NFB) during grammatical REs was 23% (13). Students' attention during lexical REs were more likely to be associated with AL recalls (58%),

followed by AC (17%) and NoA (17%). One student made a NFB recall (8%) in relation to a lexical RE. The students were more likely to make AL recalls in relation to incomplete sentence REs (75%). The students' AL recalls in grammatical REs were fewer than those in respect to lexical or incomplete sentence REs (i.e., 4% as opposed to 58% and 75%). However, their NFB recalls were the highest in grammatical REs: 90% of all NFB recalls (i.e., 11 of 12 NFB recalls).

Table 23: Attention Recalls and Error Type in Class

	Grammar		Lexicon		Incomplete		Total	
Attention	Number	%	Number	%	Number	%	Number	%
NoA	13	23	2	17	0	0	15	21
AC	31	54	2	17	1	25	34	47
AL	2	4	7	58	3	75	12	16
NFB	11	19	1	8	0	0	12	16
Total	57	100	12	100	4	100	73	100

As shown in **Table 24**, in the teacher-fronted context, students' attention in relation to grammatical REs was most likely to be AC (51.6%), followed by NoA (35.5%). Only 30% of students' recalls in respect to grammatical REs were attention to language (AL). However, the students made AL recalls along with NFB recalls in relation to three other

Table 24: Attention Recalls and Error Type in the Teacher-fronted Setting

	Grammar		Lexicon		Incomplete		Total	
Attention	Number	%	Number	%	Number	%	Number	%
NoA	11	35.5	2	25	—		13	33
AC	16	51.6	1	12.5	—		17	44
AL	1	3.2	4	50	—		5	13
NFB*	3	9.7	1	12.5	—		4	10
Total	31	100	8	100	—		39	100

*All recalls coded for NFB concurrently contained AL recalls.

grammatical REs.[43] Therefore, a total of 13% of students' attention during the teacher-fronted grammatical REs was paid to language. In contrast, the students oriented a total of 63% of attention to language during lexical REs: 50% of AL recalls and 12.5% of NFB recalls. The students made AC recalls only 12.5% of time in relation to teacher-fronted lexical REs.

As shown in **Table 25**, students' attention expressed in their recalls in relation to group-based grammatical REs was most likely to be AC, followed by NFB (30.8%). Although students' AL recall in relation to the group-based grammatical REs was only one instance (3.8%), they actually made AL recalls along with NFB recalls in seven grammatical REs. In other words, approximately 27% of AL recalls were made along with NFB recalls. In relation to group-based lexical REs, the students made more AL recalls than AC recalls; however, there were no NFB recalls. In respect to incomplete sentence REs in the group context, the students made AL recalls (75%) most as well. However, there were no NFB recalls regarding incomplete REs, either.

Table 25: Attention Recalls and Error Type in the Group Setting

	Grammar		Lexicon		Incomplete		Total	
Attention	Number	%	Number	%	Number	%	Number	%
NoA	2	7.7	0	0	0	0	2	6
AC	15	57.7	1	25	1	25	17	50
AL	1	3.8	3	75	3	75	7	20.5
NFB*	8	30.8	0	0	0	0	8	23.5
Total	26	100	4	100	4	100	34	100

*Seven of these eight recalls coded for NFB (87.5%) concurrently contained AL recalls.

43) As illustrated in the "Analysis of stimulated recall data" section (pp.42-43), the students' attention codes were counted only once for numerical analysis.

In sum, the students appeared more likely to pay attention to content (AC) in the grammatical REs, and to language (AL) in the lexical and incomplete sentence REs. However, they were more likely to notice feedback in grammatical REs than in lexical or incomplete sentence REs. In fact, the students made most NFB recalls along with AL recalls with respect to grammatical REs. There seemed a close connection between the students' attention to language and noticing feedback during grammatical REs. The interaction contexts seemed to have influenced students' language-oriented attention (AL and NFB), particularly with respect to grammatical REs; their language-oriented attention was higher in group-based than teacher-fronted grammatical REs.

Students' attention and learner reaction in REs

Table 26 summarizes the distribution of recalls for different levels of attention across the three reaction-type REs: uptake, no-uptake, and no-chance. The connection between students' attention and the reaction in REs should not be interpreted as causal. The occurrence of uptake in the data was not attributed to the students' noticing feedback. For instance, the uptake in the teacher-fronted uptake REs were reactions of non-participant students (i.e., the students in the classroom but not participating in this study), and the attention recalls from the participant students were based on their indirect experience of the REs as "side-

Table 26: Attention Recalls and Learner Reaction in Class

	Uptake		No-uptake		No-chance		Total	
Attention	Number	%	Number	%	Number	%	Number	%
NoA	6	22	0	0	9	23.7	15	21
AC	6	22	2	25	26	68.4	34	47
AL	5	19	6	75	1	2.6	12	16
NFB	10	37	0	0	2	5.3	12	16
Total	27	100	8	100	38	100	73	100

listeners." In the group contexts, three students in one group were "side-listeners."[44)]

In general, the students were likely to pay language-oriented attention (AL and NFB) in relation to uptake REs. Ten NFB recalls were associated with uptake REs; this was approximately 83% of all NFB recalls (i.e., 10 of 12 NFB recalls). On the other hand, the students were less likely to recall paying language-oriented attention in relation to no-chance REs. Nine NoA recalls, which was 60% of all NoA recalls, related to no-chance REs. Students' lack of reaction did not mean no attention was paid, however. The students were actually paying attention to language 75% of the time in no-uptake REs and about 8% in no-chance REs.

As shown in **Table 27**, the students were likely to make language-oriented recalls (AL and NFB) in relation to uptake REs in the teacher-fronted context: a total of 44% language-oriented recalls. The three instances of these recalls (18.8%) were composed of both AL and NFB recalls. On the other hand, the majority (61%) of recalls expressed in relation to no-chance REs in the teacher-fronted context were AC recalls; there were less than 10% language-oriented recalls (AL and NFB) regarding no-chance REs. In general, however, the students made many

Table 27: Attention Recalls and Learner Reaction in the Teacher-fronted Setting

	Uptake		No-uptake		No-chance		Total	
Attention	Number	%	Number	%	Number	%	Number	%
NoA	6	37.5	—		7	30.4	13	33
AC	3	18.75	—		14	60.9	17	44
AL	4	25	—		1	4.35	5	13
NFB*	3	18.75	—		1	4.35	4	10
Total	16	100	—		23	100	39	100

*All recalls coded for NFB concurrently contained AL recalls.

44) The students made uptake moves in three of the group REs. In two incidences, the addressee students made NFB recalls, and in the other incidence, the addressee student made an AC recall.

NoA recalls regarding both uptake and no-chance REs: approximately 38% and 30% respectively.

As shown in **Table 28**, uptake REs were likely to relate with students' language-oriented recalls (AL and NFB) in the group context as well. The majority of student recalls (63.6%) regarding group uptake REs were NFB, followed by AC (27.4%). Students' AL recalls in relation to group uptake RE were only 9%; however, their language-oriented attention (AL and NFB) totalled approximately 73%. In addition, six out of the seven recalls that were coded for NFB in relation to uptake REs concurrently contained AL recalls. The students paid attention to language (AL) for 75% in relation to no-uptake REs, and to message content (AC) for 80% in relation to no-chance REs.

Table 28: Attention Recalls and Learner Reaction in the Group Setting

	Uptake		No-uptake		No-chance		Total	
Attention	Number	%	Number	%	Number	%	Number	%
NoA	0	0	0	0	2	13	2	6
AC	3	27.4	2	25	12	80	17	50
AL	1	9	6	75	0	0	7	20.5
NFB*	7	63.6	0	0	1	7	8	23.5
Total	11	100	8	100	15	100	34	100

*Seven of these eight recalls coded for NFB (87.5%) concurrently contained AL recalls.

In sum, the students were more likely to make NFB recalls in relation to uptake REs than other reaction type REs, and AC recalls in relation to no-chance REs than other reaction REs. The students' AL recalls made in relation to no-uptake REs suggested that the students were still paying attention to language even though they were not verbally reacting. The interaction contexts seemed to influence the students' attention to the REs of different learner reaction types, too. The students were more likely to make NFB recalls in relation to group rather than teacher-fronted uptake REs. The tendency of the students to make relatively frequent NFB recalls

with respect to uptake REs may be related to the fact that the student who originally made the mistake responded to the teacher's recast. Hearing the student reacting to the teacher might have helped the side-listener students to pay attention to the language and realize the corrective function of teacher utterances (e.g., Ohta, 2000).

Summary of the students' attention in relation to REs

It was found that the students paid attention differently according to the characteristics of recasts and to the interaction context. Overall, the students were more attentive in the group context than in the teacher-fronted context. This was perhaps because the discussion was more interesting and relevant for the students in the group than in the teacher-fronted context. In general, SREs were less likely to be associated with students' language-oriented attention yet those in the group contexts were associated with moderate instances of NFB recalls. On the other hand, CPREs and CXREs were relatively successful in attracting students' attention to language. This finding was similar to Mackey et al.'s study (2000) in which the students were more likely to interpret "recast and negotiation," a feedback category similar to CXREs, as feedback. In general, grammatical REs were less likely than lexical and incomplete sentence REs to be related to students' AL recalls; however, they were connected to the highest instances of NFB recalls in both the teacher-fronted and group contexts. The uptake REs were more likely to be related to students' language-oriented attention than no-chance REs; however, the instances of NFB recalls in relation to uptake REs were higher in the group than in the teacher-fronted context.

Some tendencies concerning the students' attention within different types of REs were observed. Their attention in grammatical and lexical REs contrasted. Although lexical REs were connected highly to AL recalls, few NFB recalls occurred in association with lexical REs. In contrast, although students' AL recalls tended to be low in relation to the grammatical REs, the students were more likely to make NFB recalls

regarding grammatical REs; in fact, many NFB-coded grammatical REs concurrently contained AL recalls. This may be attributed to the students' knowledge of prescribed grammar learned through high school English classes. They may have had the explicit grammatical knowledge to validate the teacher's grammatical recast as feedback. In contrast, lexical problems (i.e., mostly word choice) were semantic violations, and the students might not have had enough lexical resources to deduce the teacher's correction by her recast. The different attention allocations in uptake, no-uptake or no-chance REs were also interesting. As I noted earlier, the students might have been aware of the teacher's recasts because of the response by the student to whom the recast was directed. The students' attention to language in no-uptake REs was evidence to support the argument that evaluating the efficacy of recast on the basis of learner's immediate reaction is not sufficient (Mackey & Philp, 1998).

Students' Attention to the Teacher

Because NFB occurred infrequently, and because it was felt during the stimulated recall interview that the students were not paying much attention to the teacher's talk, an additional analysis was conducted to explore the students' general attention to their teacher. For this analysis, all references made to the teacher, i.e., "Hearing the teacher," "Listening to the teacher," and "Noticing teacher feedback" (see the "Coding schemes for student stimulated recall interview" section, p.41), were tallied for all ETEs in the recalls.[45] There was a total of 103 ETE-related recalls in relation to 20 ETEs on which the eight students commented during their stimulated recall sessions (i.e., 61 teacher-fronted and 42 group contexts). Among these recalls, 52 (i.e., 22 in teacher-fronted and 30 in group contexts) were teacher-related comments. This was

45) Students' stimulated recall coded for Lis-L was not included in this analysis because comments coded for Lis-L usually included comments on other students' language use. It was difficult to distinguish the comments on the teacher's language use from those on other students'.

approximately 50% of the recalls. In other words, the students paid attention to their teacher about 50% of the time. Their meaning-oriented attention to teacher talk in those ETEs averaged about 41% of the time, and their noticing teacher feedback was found as little as 14% of the time.

Students' Accounts of Their Attentiveness in the Teacher-Fronted Interaction

Some general accounts made by the students explaining their reduced attentiveness during the teacher-fronted interaction were presented earlier (the "Students' attention and interaction contexts" section, pp.61-65). There were two aspects in the classroom in particular that contributed to their reduced attentiveness in the teacher-fronted interaction: inaudibility and established classroom-specific discourse patterns.

Incomprehensible utterances

As illustrated earlier, the students were less attentive to the teacher-fronted discussion because they felt it was distant and irrelevant. In addition to these affective reasons, inaudibility and incomprehensibility of classmates' utterances were attributed to the physical distance that prevented the students from listening carefully. Students frequently spoke in soft voices when they were called upon.[46] When they gave opinions spontaneously, their utterances were fragmented. If they prepared for a public talk (e.g., newspaper article reporting activity), their language approximated written language and was therefore difficult to process.

Excerpt 6: Keiko's recall session 1, October 9.

> (Talking about the newspaper article report activity,) I don't usually listen to them. Even when I listen to them, they usually read aloud things written

46) Ms. Johnson also acknowledged her students' inaudibility in the whole-class discussion (Teacher recall 1).

in the newspaper. There are many difficult words that I cannot understand.[47]

Yasuko expressed her frustration with her classmate's unclear pronunciation.

Excerpt 7: Yasuko's recall session 1, October 6.

(Talking about the newspaper article report activity given by Masako,) Isn't there something? … I mean … pronunciation is … maybe different? She may be pronouncing well, but I thought it was hard to listen to her… besides she spoke as she looked down …

Discourse patterns

Discourse patterns particular to a classroom setting, such as "Initiation-Response-Feedback (IRF)" (Mehan, 1979), also influenced students' attention to classroom talk. For instance, the students paid little attention to the teacher when she summed up a student's reply to her initial question in teacher-fronted activities. Checking group opinions in class after a small-group discussion activity was also repetitive for the students, and they paid little attention to the teacher. Hisako explained that she listened to her classmates' opinions more carefully than to the teacher's summary.

Excerpt 8: Hisako's recall session 3, October 20.

I listen to the answers … and do not listen to the follow-up explanations. … For example, in this case, I listened to [the answers] 'A, C, E', and thought those answers were the same as our group answers, but did not pay much attention to the teacher's explanation.

On other occasions, the students tuned out entirely because their classmates' answers were the same as their own.

47) What I coded as reflecting the students' attention is highlighted by a single underline (i.e., attention to content), a double underline (i.e., attention to language), and bold face (i.e., noticing feedback).

Excerpt 9: Eiko's recall session 4, November 1.

> It's boring… I cannot help thinking other things [during the answer checking]. If there are different opinions, [I could pay attention] but the opinions are usually the same.

However, students' personal needs and interests sometimes overrode the discourse pattern. Occasionally, they paid attention to the teacher during teacher-fronted activities because they could not understand the previous classmate's utterance.

Excerpt 10: Shoko's recall session 3, October 19.

> Yeah, I think I was paying attention to the teacher when she spoke. [When classmates gave opinions in class] sometimes I could not understand, but if the teacher rephrased [what the classmate said] I could understand.

Communication breakdowns between the teacher and students were unusual discourse pattern which drew their attention to language. The teacher, being unable to hear and understand a student's quiet presentation in one lesson, walked up to the presenter and read aloud what the presenter had prepared to say. This unusual teacher-fronted interaction drew other students' attention; they turned their faces to the teacher and the presenter. Some students in the recall sessions reported that their attention was attracted because "Ms. Johnson said she did not know [the word]" (Hisako's recall session 2) and they thought the presenter had pronunciation problems (Tokiko's recall session 2).

Students' Accounts of Their Attentiveness in the Group Interaction

Students' recalls elicited in relation to the group interaction REs illustrate their comprehensible language use. Their AL recalls were associated with language-related difficulties such as unknown words, mismatches between what they heard and what they knew, and "the gap" (Swain, 1985, 1995) between what they wanted to say and what their L2 knowledge allowed them to say.

Unknown words for comprehension[48)]

Most language-related difficulties were vocabulary-related, and the students found many unknown words in the input, either in the written materials, in the teacher's utterances, or in their peers' utterances. For example, Shoko's language problem in Lesson 3 was "declaration" in the activity worksheet dealing with the Universal Declaration of Human Rights. She did not know either the meaning or pronunciation of this word (session 3, October 19). Her group members became involved in an extensive discussion to clarify the meaning, and they practiced the pronunciation of the word. Keiko reported her appreciation of information gap tasks as opportunities for learning "new words and phrases that [her] classmates used" during the group activities (session 1, October 9).

Mismatches

The students sometimes felt confused when the language they heard did not match the expected meaning. For instance, Hisako reported:

Excerpt 11: Hisako's recall session 4, November 2.

> (Talking about her reaction after Tokiko's opinion that "young [men] care [about recycling things in order to protect the environment],") in my mind, young men were more careless [than older people], but Tokiko said young man cares more, so I thought she made a mistake with *comparatives*."

Tokiko's different opinion made Hisako pay attention to the linguistic structure. Instead of understanding and accepting Tokiko's utterance, she attributed the difference to Tokiko's linguistic problem.

On several occasions, the students made comparisons between their intended content and the linguistic form provided by the teacher. Fumiko recalled that Ms. Johnson "perhaps mistook" her group member's intended message content and supplied inappropriate words (session 4,

48) Some vocabulary-related problems addressed in the stimulated recalls related to the students' attempts at speaking English will be discussed under the "gap" subtitle.

November 2).

Gaps for production

Students' language-related attention was sometimes related to their endeavour to produce comprehensible output (Swain, 1985). The students were often aware that they did not have enough vocabulary to express their ideas. For instance, Tokiko tried to confirm the meaning of the unfamiliar word "biodegradable" by referring to an idiomatic translation from Japanese, namely that something biodegradable can be resolved into "soil."[49] In the stimulated recall about this episode, Tokiko reported:

Excerpt 12: Tokiko's recall session 4, November 2.

> I was thinking ... being able to get resolved, and I wanted to say *tsuchi* [soil], but I didn't know what *tsuchi* is in English, so I couldn't. So I said "return ..." and gestured to mean the ground, the earth. ...

Tokiko's awareness of the lack of the necessary word (i.e., "soil") in her English vocabulary compelled her to attend to Ms. Johnson when she supplied her with a different word, "rust." Tokiko expected, in the negotiation process, that the teacher would provide her with the appropriate English expression, although the teacher did not.

Excerpt 13: Tokiko's recall session 4, November 2.

> The teacher said ... "rush"? ... "Rust"? "Rust." ... Did it mean *tsuchi* [soil] ? It didn't, did it? I wanted her to use the word for tsuchi. But she didn't.

The language-related attention that the students paid during their attempt at comprehensible output often contributed to their noticing teacher feedback. Shoko also gave an account of explicitly noticing the teacher's feedback to her linguistic problem.

---

49) What Tokiko actually said was "it returns to ...," which was almost a direct, literal translation from the Japanese idiomatic expression, "[biodegradable things] go back to soil."

Excerpt 14: Shoko's recall session 2, October 16.

> I said 'freedom of thinking.' I was not certain if it should be 'thinking' or 'thought.' I didn't come up with 'thought' then, so I said 'freedom of thinking' then I felt it might be wrong. **Then the teacher said 'freedom of thought.' So I thought, 'Oh, oh. I was wrong - just as I thought.'**

As in Tokiko's case, Shoko heightened her language-oriented attention on the basis of her awareness of her linguistic shortcomings in expressing herself.

The students were sometimes aware of linguistic problems in other students' performance and this heightened their own language-oriented attention. For instance, Fumiko recalled both noticing a linguistic problem in a classmate's utterance and noticing Ms. Johnson's feedback.

Excerpt 15: Fumiko's recall session 2, October 12.

> I thought [Mayumi] couldn't say what she wanted to say in English, she probably wanted to say 'better.' **So [when the teacher said 'better?',] I thought that's it.**

Shoko's recall in a group RE was more concrete and precise.

Excerpt 16: Shoko's recall session 1, October 5.

> Tokiko's sentence was strange, but we could guess what she wanted to say. The teacher understood her. She said, "Ah," and **rephrased it in a correct or maybe in a better sentence. So I thought, "that's how to say it."** When I heard Tokiko, I thought it sounded wrong, but I didn't know how to express the idea better either. I was wondering "how to say this?" and **the teacher said a sentence, and I thought that's the right expression.**

Being involved in the conversation, Shoko seems to have shared not only message conent but also the linguistic problem with her peer. She compared and contrasted the non-target-like and target-like forms, although the non-target-like form in the question was not her own utterance. The students independently and autonomously identified their

difficulties during the classroom interaction, which contributed to their voluntary attention to the language.

Summary of Findings for Research Question 2

Students' attention to the teacher's recasts during interaction was multifaceted. The linguistic characteristics of the REs, interaction contexts, and conditions at the moment of RE influenced in what ways the students paid attention to the REs. The characteristics of REs, i.e., the delivery types in which recasts are provided, students' responses to the teacher recasts, and the error types to which recasts are provided, had been focused on in previous cognitive-interactionist SLA research. In my data, the students were found more likely to attend to language in CPREs and CXREs than in SRE, and more likely to attend to language in uptake REs than in no-uptake or no-chance REs. They were likely to make AL recalls in lexical REs, and their noticing of feedback was more likely to occur in grammatical REs. The students also made AL recalls in relation to a total of 10 NFB-coded grammatical REs at the same time. Thus, the students' attention to language (AL) in grammatical REs was often related to their noticing feedback.

These overall tendencies in students' attentiveness according to the linguistic characteristics of REs seem, however, to be influenced by the interaction contexts. The students were more likely to attend to the teacher in the group than in the teacher-fronted context. The students paid more attention to language in group SREs than in teacher-fronted SREs, for instance. They paid language-oriented attention (AL and NFB) in relation to grammatical REs more in the group contexts than in the teacher-fronted contexts. Their noticing feedback in relation to uptake REs was better in the group than in the teacher-fronted context.

The students were attentive in the group because the discussion there was more meaningful, immediate, and relevant. Students' recalls regarding the group REs revealed their reflections on the dialogue. Their reflection could be on the unfamiliar vocabulary they heard in another's

utterances or mismatches between the meaning they anticipated and the language they heard. The students were also found to have realized linguistic problems before they noticed feedback. In contrast, the students found many challenges in paying attention in the teacher-fronted discussion. The discussion between the teacher and other students in different groups was often felt to be remote and irrelevant. Students' attention in the whole-class discussion, especially during the debriefing of group discussion outcomes, was selective. They often focused on getting the content information they needed (i.e., correct answer) either in other students' or the teacher's utterances, and did not pay much attention to language.

Some of the findings seem to support previous results and arguments. For instance, CPREs and CXREs overall related to students' language-oriented attention. This finding supports Doughty and Varela's (1998) claim for the efficacy of the "corrective recast." The finding also complements Mackey et al.'s (2000) finding that students more accurately perceived error correction if recasts were provided in combination with negotiation. The students were also found less likely to make AL recalls than AC recalls in relation to grammatical REs. This finding supports the finding in Mackey et al. (2000) that morpho-syntactic recasts were least accurately perceived as feedback on grammar. However, the students in my study were most likely to notice feedback in grammatical REs, and most NFB recalls in respect to grammatical REs were concurrently made with AL recalls. The finding that uptake REs were better attended to than no-uptake or no-chance REs supports Lyster's (1998b) concern that the teacher's recasts might not be effective because the students seldom respond. However, the students were also found to have paid attention to language even during no-uptake or no-chance REs; therefore, as Mackey and Philp (1998) argued, evaluating the efficacy of recasts exclusively on the basis of students' uptake is dangerous. In their uptake claims for learning, the students showed their independent and idiosyncratic learning, just as the students in Slimani's (1987, 1992) study did. Many

claimed items were related to the teacher's utterances; therefore, the teacher was an important resource for language learning for them, yet the students chose whether or not to pay attention to her, particularly during the teacher-fronted interaction.

Findings for Research Question 3:

Which Recasts Related to the Students' L2 Learning as Measured by GJ Tests?

As illustrated on Tables 8 and 9, the number of GJ test items varied from participant to participant. Thus, the results are discussed on the basis of correct answer percentages. I discuss learning in relation to GJ tests. The GJ tests measured the students' receptive ability to judge the grammaticality of given sentences. The students' productive ability was not measured. Because no pre-tests were administered, whether the students already knew the linguistic aspects tested in GJ tests is not clear. However, the very fact that the students produced erroneous utterances to which the teacher gave recasts (i.e., group recasts) would seem to indicate their limited understanding of those linguistic aspects. The knowledge of the students who did not produce the target erroneous utterance (i.e., in the teacher-fronted REs and the other three students in the group REs) was unclear. However, their lack of understanding of those linguistic problems was suggested in their stimulated recall comments. Therefore, the GJ test items represented, in general, unknown linguistic aspects for the students. It is unclear in this study whether the students' learning as measured was isolated item learning or systemic development as the test item sentences were similar to those to which the students were actually exposed. Finally, because the total of all item numbers, particularly those from group-based REs, is small, interpretation of the data requires caution.

First, I present the overall GJ Test results, then I present them in relation to characteristics of REs (i.e., delivery types, error types, and learner reaction types), and finally in relation to learners' attention (i.e.,

NoA, AC, AL, and NFB).

Correct Judgment and Interaction Contexts

Table 29 displays the eight students' Test 1 results. The average correct judgment of the total items was approximately 59% (SD = 11.1), and the median was 54%. Their correct answers ranged from 50% to 80%. Regarding the teacher-fronted items, the median of correct judgments was approximately 54% and the average was approximately 56% (SD = 15.9) with a range of 41.7% to 83.3%. The median and average with group items was better than that with the teacher-fronted items: approximately 61% and 63% (SD = 10.8) respectively with a range of 50% to 75%. Although their median scores in class total and teacher-fronted items were slightly above 50%, the students achieved an average of approximately 60% on Test 1. This indicated in general that recasts contributed moderately to

Table 29: Correct Judgments in GJ Test 1

	T-fronted RE items		Group RE items		Total	
	Correct/Total	%	Correct/Total	%	Correct/Total	%
Performance						
Tokiko	5/12	41.7	6/10	60	11/22	50
Shoko	5/12	41.7	8/14	57.1	13/26	50
Fumiko	6/12	50	2/4	50	8/16	50
Yasuko	7/12	58.3*	7/14	50	14/26	53.8
Eiko	5/12	41.7	6/8	75*	11/20	55*
Keiko	7/12	58.3*	3/4	75*	10/16	62.5*
Aiko	9/12	75*	5/8	62.5*	14/20	70*
Hisako	10/12	83.3*	6/8	75*	16/20	80*
Description						
Median	54.15		61.25		54.4	
Mean	56.25		63.08		58.91	
SD	15.88		10.78		11.09	

Note: Group median and means are percentages within the category. Asterisks (*) indicate the performance in the upper half of the median.

immediate learning.

Table 30 displays the eight students' Test 2 results. The median correct judgment of total items was approximately 64% and the average was about 60% (SD = 11.58). Correct answers ranged from 45% to 75%. The median correct judgment with teacher-fronted items was 62.5% and the average was approximately 54% (SD = 16.69), with a range of 25% to 66.7%. The median of correct judgements with group items was 67% whereas the average was approximately 62.7% (SD = 23.76), with a range of 25% to 92.9%. In general, Test 2 results were slightly better than those in Test 1, although the scores have a wider spread than in Test 1. The medians for total, teacher-fronted, and group items were 63.75%, 62.5%, and 67%, respectively, higher than for Test 1 (54.4%, 54.15%, 61.25%).

Table 30: Correct Judgments in GJ Test 2

	T-fronted RE items		Group RE items		Total	
	Correct/Total	%	Correct/Total	%	Correct/Total	%
Performance						
Tokiko	8/12	66.7*	7/10	70*	15/22	68.2*
Shoko	4/12	33.3	13/14	92.9*	17/26	65*
Fumiko	7/12	58.3	1/4	25	8/16	50
Yasuko	3/12	25	9/14	64	12/26	46.2
Eiko	8/12	66.7*	7/8	87.5*	15/20	75*
Keiko	8/12	66.7*	2/4	50	10/16	62.5
Aiko	6/12	50	3/8	37.5	9/20	45
Hisako	8/12	66.7*	6/8	75*	14/20	70*
Description						
Median	62.5		67		63.75	
Mean	54.18		62.74		60.24	
SD	16.69		23.76		11.58	

Note: The students' names are in the same order as in Table 29. Group median and means are percentages within the category. Asterisks (*) indicate the performance in the upper half of the median.

Test 2 results were maintained or improved from Test 1 despite the 3 to 9 week interval between them.[50] However, the improvement should not be simply interpreted as maintenance of language learning through recasts provided in the class. My research design provided the students with opportunities to re-experience the REs through video watching prior to Test 2. Thus, the Test 2 results may reflect the students' additional learning through the stimulated recall sessions.

Although performance in general improved on Test 2, the changes were unique to each student. The scores were spread over a wider range than in Test 1. Tokiko and Shoko, for instance, performed better on Test 2 than on Test 1; on the other hand, Aiko performed poorly on Test 2. Shoko performed better on group-based items than teacher-fronted items, but Aiko did better on teacher-fronted than group items. These differences can be attributed to internal variables. One important variable that I did not control was the students' proficiency. According to Ms. Johnson's evaluation, Tokiko was the weakest student among the eight, and this explains her low achievement in Test 1. However, the proficiency alone does not explain the results, because Shoko, whom Ms. Johnson evaluated as the strongest student in the class, also performed relatively poorly on Test 1. Aiko was a middle level student among the eight. Another stronger student, Keiko, performed well in Test 1, but not in Test 2. Therefore, there must have been other variables influencing the individual students' learning.

Correct Judgment and RE Types

Tables 31 to 33 illustrate the distribution of correct judgments in GJ Test 1 across the characteristics of the original REs. The numbers given in the tables are percentages of correct judgments over the total GJ items that belong to the REs of the given characteristics.

50) Test 2 was administered three weeks after the final weekly set of Test 1 (i.e., Mini-test 6). The interval between the first part of Test 1 (i.e., Mini-test 1) and Test 2 was 9 weeks (see Table 2).

As shown in **Table 31**, the distribution of students' correct answers for delivery types and reaction types was moderately even. They were correct more than 50% for each sub-type in these two characteristics. However, their answers according to error type were unevenly distributed. The students were more likely to answer lexical items incorrectly than grammatical or incomplete sentence items.

Table 31: Correct Test 1 Responses (%) with Respect to REs in Class

	Delivery			Error			Reaction		
	SRE	CPRE	CXRE	Gram.	Lex.	Inc.	Uptake	No-uptake	No-chance
All REs	58.9 (66/112)*	63.6 (14/22)	53.1 (17/32)	63.6 (75/118)	33.3 (8/24)	58.3 (14/24)	61.1 (33/54)	56.3 (9/16)	57.3 (55/96)

*Numbers in parentheses are raw scores (i.e., correct responses over total item numbers in the category). The same rule applies in the following tables.

The students' Test 1 answers on teacher-fronted items were spread out among the sub-types in each category. As shown in **Table 32**, their answers to CXRE-items were less likely to be correct than to SRE- or CPRE-items. Under error type, lexical items were less likely to be correctly answered than grammatical or incomplete sentence items. The students' correct answers were almost evenly distributed between uptake and no-chance RE items.

Table 32: Correct Test 1 Responses (%) with Respect to Teacher-fronted REs

	Delivery			Error			Reaction		
	SRE	CPRE	CXRE	Gram.	Lex.	Inc.	Uptake	No-uptake	No-chance
T-fronted	64.3 (36/56)*	56.3 (9/16)	37.5 (9/24)	64.1 (41/64)	25.0 (4/16)	56.3 (9/16)	53.1 (17/32)	—	57.8 (37/64)

With regard to GJ Test 1 responses to the group RE items, although the students' correct judgments were more than 50% across the categories, they were relatively unevenly distributed (see **Table 33**). For instance, SREs were less likely to be answered correctly than those from CPREs

and CXREs. Lexical REs were less likely to be judged correctly than grammatical or incomplete sentence items. Items from uptake REs were more likely to be answered accurately than no-uptake and no-chance REs.

Table 33: Correct Test 1 Responses (%) with Respect to Group REs

	Delivery			Error			Reaction		
	SRE	CPRE	CXRE	Gram.	Lex.	Inc.	Uptake	No-uptake	No-chance
Group	53.6 (30/56)*	83.3 (5/6)	100 (8/8)	63.0 (34/54)	50.0 (4/8)	62.5 (5/8)	72.7 (16/22)	56.3 (9/16)	56.3 (18/32)

As shown in **Table 34**, the students' answers in Test 2 across delivery types and reaction types were distributed relatively evenly. They answered nearly 60% correct for each sub-type in the delivery category. Their correct answers were spread among the categories under error type. The students were less likely to answer the lexical items correctly than grammatical or incomplete sentence items. The students' correct judgements were unevenly distributed among the reaction types. The students were more likely to correctly judge items from uptake REs than from no-uptake and no-chance REs.

Table 34: Correct Test 2 Responses (%) with Respect to REs in Class

	Delivery			Error			Reaction		
	SRE	CPRE	CXRE	Gram.	Lex.	Inc.	Uptake	No-uptake	No-chance
All REs	59.8 (67/112)*	63.6 (14/22)	59.4 (19/32)	64.4 (76/118)	41.7 (10/24)	58.3 (14/24)	70.4 (38/54)	56.3 (9/16)	55.2 (53/96)

Regarding the Test 2 responses to the teacher-fronted RE items, the students' correct answers were evenly distributed across delivery categories. As shown in **Table 35**, although lexical items remained more difficult than grammatical or incomplete sentence items, nearly 44% of the students' answers were correct. Their answers in relation to learner

reaction varied more than they did in Test 1. The students were more likely to correctly answer the items from uptake REs than those in no-chance REs.

Table 35: Correct Test 2 Responses (%) with Respect to Teacher-fronted REs

	Delivery			Error			Reaction		
	SRE	CPRE	CXRE	Gram.	Lex.	Inc.	Uptake	No-uptake	No-chance
T-fronted	57.1 (32/56)*	50.0 (8/16)	50.0 (12/24)	57.8 (37/64)	43.8 (7/16)	50.0 (8/16)	59.4 (19/32)	—	51.6 (33/64)

As shown in **Table 36**, the students' correct answers were unevenly distributed in relation to group-based Test 2 items. Their answers to SRE-items were least likely to be correct among the three delivery types, as were the answers to lexical items among the error types. Uptake REs were most likely to be answered right among the reaction types.

Table 36: Correct Test 2 Responses (%) with Respect to Group REs

	Delivery			Error			Reaction		
	SRE	CPRE	CXRE	Gram.	Lex.	Inc.	Uptake	No-uptake	No-chance
Group	62.5 (35/56)*	100 (6/6)	87.5 (7/8)	72.2 (39/54)	37.5 (3/8)	75.0 (6/8)	86.4 (19/22)	56.3 (9/16)	62.5 (20/32)

In sum, the students' answers on SRE-items were more likely to be correct than on CPRE- and CXRE-items in the teacher-fronted contexts, but in the group-contexts, CPRE- and CXRE-items were more likely to be answered correctly. Regarding the error type, the students were least likely to answer lexical items correctly. They were correct on lexical items at most only 50% of the time (i.e., group-based Test 1). This result is in contrast to their achievement with grammatical and incomplete sentence items in Test 2 (group), in which over 70% of their judgments were correct judgments. Finally, the students performed successfully on items from group uptake REs on Test 1, and on uptake REs in both contexts on

Test 2.

Correct Judgment and Students' Attention

As shown in **Table 37**, the students were found to have judged 43% of the items correctly from NoA REs on Test 1. The students' correct judgment was most highly related to their NFB recalls (74%), followed by AC recalls (61%). Only 46% of the students' answers on the items derived from AL REs were correct. However, for the purpose of evaluating the students' language oriented attention for their language learning, this category seemed to under-represent the actual effect of AL because many AL recalls, as discussed earlier (the "Summary of the students' attention in relation to REs" section, pp.71-72), occurred along with NFB recalls (see also Excerpts 14, 15, and 16, p.78). Therefore, I considered an additional category (i.e., the "AL + (NFB)" category on Table 38) in which all AL recalls are related to grammaticality judgment results.[51] In other words, this category included the sum of the AL entry (i.e., 11 correct judgments over 24 AL-coded RE items) and the correct judgments in relation to NFB REs in which AL recalls co-occurred (i.e., 15 correct judgments over 21 NFB-coded REs which also includes AL recalls). The students achieved an average of almost 60% on the AL + (NFB) items.

Table 37: Correct Test 1 Responses (%) with Respect to Attention in All REs

	Attention in REs				
	NoA	AC	AL	AL+(NFB)	NFB
Total in class	43.3 (13/30)*	60.7 (37/61)	45.8 (11/24)	57.8 (26/45)	73.9 (17/23)

51) Not all NFB recalls corresponded to AL recalls. NFB recalls which accompanied AL are put in parentheses in order to distinguish it from the other general NFB category.

Table 38 shows the students' correct answers on the teacher-fronted Test 1 items in relation to their attention. The students were found to have judged 39% of the items correctly from NoA REs in this form of interaction. They were more likely to correctly judge items from the AC REs than items from NFB REs. The students were less likely to correctly judge items from AL REs than those from AL + (NFB) REs: their judgments were 50% and 53% orrect respectively.

Table 38: Correct Test 1 Responses (%) with Respect to Attention in Teacher-fronted REs

	Attention in REs				
	NoA	AC	AL	AL+(NFB)	NFB
T-fronted	38.5 (10/26)*	66.7 (18/27)	50 (5/10)	52.9 (9/17)	57.1 (4/7)

As shown in **Table 39**, the students' highest correct judgments (81%) were on NFB-related items on group-based Test 1. In this context, the students were more likely to correctly answer items from AL + (NFB) REs (61%) than those from AC REs (56%). Their answers on items from AL REs were only 43% orrect. The students judged three items out of four correctly (75%) even though they did not pay attention (NoA).

Table 39: Correct Test 1 Responses (%) with Respect to Attention in Group REs

	Attention in REs				
	NoA	AC	AL	AL+(NFB)	NFB
Group	75.0 (3/4)*	55.9 (19/34)	42.9 (6/14)	60.7 (17/28)	81.3 (13/16)

The students' performance on Test 2 shows a similar pattern to that of Test 1: higher percentages of correct answers related to NFB-items. As shown in **Tables 40 to 42**, more than 80% of the students' judgments were correct on all NFB-related items. Test items from AL + (NFB) REs were also associated with correct judgments. The students had an overall

average of approximately 70%; in particular, 75% of their judgments for the group context were correct. Approximately 60% of items derived from AC and AL REs were correct in any interaction context. The students were also found to achieve an average of 50% on their answers to NoA items.[52)]

Table 40: Correct Test 2 Responses (%) with Respect to Attention in All REs

	Attention in REs				
	NoA	AC	AL	AL+(NFB)	NFB
Total in class	53.3 (16/30)*	62.3 (38/61)	58.3 (14/24)	73.3 (33/45)	87.0 (20/23)

Table 41: Correct Test 2 Responses (%) with Respect to Attention in Teacher-fronted REs

	Attention in REs				
	NoA	AC	AL	AL+(NFB)	NFB
T-fronted	50.0 (13/26)*	59.3 (16/27)	60.0 (6/10)	70.6 (12/17)	85.7 (6/7)

Table 42: Correct Test 2 Responses (%) with Respect to Attention in Group REs

	Attention in REs				
	NoA	AC	AL	AL+(NFB)	NFB
Group	75.0 (3/4)*	64.7 (22/34)	57.1 (8/14)	75 (21/28)	87.5 (14/16)

The distribution of correct answers according to the students' attention levels showed important connections between noticing teacher feedback and learning. If the students noticed that the teacher provided feedback in response to a language problem, they were able to judge the

52) Interpretation of Test 2 results requires caution. Due to my research design, in which the students answered Test 1 before watching the video stimuli, there is a possibility that the test items made the students more attentive to the feedback, thereby facilitating learning from the video, even though both the tests and video scenes included distracters.

grammatical and ungrammatical sentences accurately. However, their attention to language alone (AL) did not connect to as much learning as their attention to language along with noticing feedback (AL + (NFB)) did. This was perhaps because identifying linguistic problems was not enough for the students to achieve "learning." Rather, they needed to find a solution to the issue they uncovered.

Summary of Findings for Question 3

The GJ test results indicate that the students learned the language, though moderately, averaging approximately 60%[53], through classroom recasts. The relationship between the linguistic characteristics of REs and students' test performance appeared strong: the lexical items were constantly more difficult for the students than grammatical items were. However, the delivery type and reaction type of REs did not seem to differentiate much based on the students' performance. Rather, the attention the students paid seemed to influence their performance. In general, items on both Test 1 and Test 2 that were derived from NFB REs were more likely to be judged correctly. The students' attention to language (much of which co-occurred with NFB) was also associated with their correct judgments. Because CPREs and CXREs provided in response to grammatical errors during the group interaction were better related than other REs to students' NFB recalls, they also connected well to students' accurate judgments. In the teacher-fronted REs, SREs provided in response to grammatical and incomplete sentence problems were more likely to relate to students' NFB recalls; these REs were better

53) The students could get 50% correct answers by chance. However, I do not think that was the explanation for their performance. As long as I observed, the students engaged in the GJ tests seriously. Although the students were able to achieve nearly 50% correct judgments on NoA items (i.e., 43% in Test 1 total and 53% in Test 2 total), their poor performance on NoA items in the teacher-fronted context (i.e., the low attention context) on Test 1 supports the interpretation that the test results and the students' attention were related.

associated with students' correct judgment on the tests.

Discussion: Recasts, Learning, and Cognitive-interactionist SLA

The findings for the three research questions reveal the complexity of recasts, learners' attention, and learning in a communicative L2 classroom. In this section, I discuss recasts and L2 learning based on the current research findings and on current cognitive-interactionist SLA theories. My discussion will be categorized by the relationships between (1) attention and learning, (2) interaction and learning, and (3) "linguistic environment" and learning.

Attention and Learning

The findings in this study support the noticing hypothesis (Schmidt, 1990, 1994, 1995; Schmidt & Frota, 1985). The students were found to have performed consistently well on the NFB-related items on GJ tests. On the group-based NFB items, in particular, they correctly judged 81% of the Test 1 items and 87.5% of the Test 2 items. As Schmidt suggested, noticed items are learned well.

However, the relationship between levels of attention and learning appear complex. Because the noticing hypothesis claims that items to which the learner pays attention can be learned, it is assumed that for language learning, language-oriented attention is more facilitative than attention to content. The students' GJ test results, however, revealed that attention to language (AL) alone did not relate better to GJ test performance than did their attention to content (AC).

Although the students' attention to language, when treated as an exclusively independent category (i.e., AL), appeared not as facilitative for learning, many AL recalls coincided with the students' NFB recalls. When these concurrent AL recalls (i.e., AL + (NFB)) were considered in relation to the students' correct judgment, they were found to relate to nearly 60% correct judgments in Test 1 and 80% correct judgments in Test 2. Therefore, AL was possibly an important prerequisite for NFB. Because

AL recalls were often recognition of language-related problems (i.e., "noticing the gap" in Swain, 1985, 1995), such noticing might have facilitated noticing feedback. Noticing the gap, therefore, may be an important sub-category in language-related attention contributing to language learning because it enhances students' autonomous noticing of feedback. Paying attention to language "to notice a gap" and solving the problem are important aspects that facilitate learning. However, not all students' AL recalls were associated with NFB. When NFB did not coincide with AL, the students were unable to resolve their noticed problem. The REs that left the students with unresolved language-related problems (i.e., AL) may not be facilitative for L2 learning at that moment.

The students' recalls evidenced, to some extent, that their attention to content could play a role in learning certain aspects of language. Neither the comprehensible input hypothesis nor the interaction hypothesis has clearly developed the role of learners' attention to content. Although negotiation of meaning, for instance, has been claimed to be facilitative in L2 learning (e.g., Gass, 1997; Long, 1996; Pica, 1994), interaction SLA researchers have not yet shown how attention to content meaning during the negotiation could contribute to L2 learning other than triggering linguistic manipulation between the interlocutors. As shown in students' excerpts 11, 15, and 16 (pp.63-65), AC recalls sometimes coincided with AL and/or NFB recalls. Contrary to the view that attention to content is tantamount to passively "decoding" meaning, the students related their attention to content to their attention to language in a spiral fashion. Hisako, for example, thought Tokiko's utterance contained a mistake because the meaning of Tokiko's utterance did not match that which she, Hisako, would have expressed. Similarly, comprehending what her peer wanted to say, Fumiko noticed a language problem and contrasted it with what she would have expressed on the basis of her, Fumiko's, comprehension. The students' active attempts to make sense of what they heard (or read), in fact, involved unvoiced language production processes. Comprehension of input that facilitates L2 learning, therefore, involves

learners' unvoiced output, which mediates their attention to language, and if possible, solves language problems.

Interaction and Learning

The interaction context played an important role in this study. The REs in the group context were more facilitative than those in the teacher-fronted context because the students found the group interaction more meaningful and interesting, and the teacher feedback provided in the group more immediate and relevant than in teacher-fronted REs. These findings are applicable to other recast studies (Carroll & Swain, 1993; Han, 2002; Ishida, 2004; Mackey et al., 2000; Mackey & Philp, 1998; Oliver, 1995; Ortega & Long, 1997). For instance, many empirical studies employed communicative dyadic tasks. A NS interacting with a NNS in a dyadic activity is presumably similar to Ms. Johnson's interacting with four students during the small group activity. As Ms. Johnson devoted her attention to facilitating her students' task performance, the NS participant could have worked enthusiastically with the NNS participant. In the studies that employed communicative tasks (Ishida, 2004; Mackey et al., 2000; Mackey & Philp, 1998; Oliver, 1995), the NNS must have encountered many linguistic problems, and the NS's recast should have been meaningful, immediate, and relevant.

Because experimental studies tend to focus on certain aspects of language learning, the NNS participants were intensively exposed to the same grammatical aspects during the experimental task (Carroll & Swain, 1993; Han, 2002; Ishida, 2004; Mackey & Philp, 1998; Ortega & Long, 1997). In addition, the NS utterances could be delivered in such a way that the linguistic aspect could be efficiently noticed. The transcripts available from the experimental studies indicate that recasts are often given in isolated SREs (e.g., Mackey & Philp, 1998). That the delivery type of RE is influenced by the contexts is also observed in classroom-based recast studies. An L2 classroom where learning linguistic aspects of language was the focus, as opposed to communicating in the language, also

contained isolated recasts. The transcripts from Ohta (2000) illustrate that the teacher's recasts were either in SREs or CPREs. In the interaction where the participants share the understanding that language is the focus, short and isolated recasts may be facilitative. In fact, the students in my study learned from grammatical REs, many of which were provided in SREs.

The previous study (Doughty & Varela, 1998) in which recasts were found to have an impact on a large number of people employed a means to make the recast relatively explicit. In the content-based ESL class in Doughty and Varela's (1998) study, the effective recast was a complex recast (i.e., similar to a CXRE and a CPRE). The teacher employed the "corrective recasts" so that her correction was relevant to the mistake and immediately available for noticing and correction.

The feedback provision employed by the teacher in Doughty and Varela's (1998) study was intentional for experimental purposes, and her corrective purpose in recasting was quickly understood by the students. Her recasts were non-metalinguistic, but rather explicit and salient. Ms. Johnson employed CXREs and CPREs in her theme-based L2 classroom. Her inclination to provide corrective feedback, although it occurred infrequently, coincided with her CPREs and CXREs. In her class, where theme teaching often overrode L2 teaching, these RE types were effective because they could deliver her corrective intention to the students and draw their attention without her giving metalinguistic explanations.

Taking account of interaction contexts involves taking into account the participants. The cognitive-interactionist SLA research interaction framework has neglected the teacher's and learners' involvement beyond their utterances. This issue will be discussed in the following section, in relation to "linguistic environment."

"Linguistic Environment" and Learning

My findings for the relationship between linguistic characteristics of REs and students' attention may support the previous studies claiming

that recasts are effective in facilitating L2 learning. For instance, CPREs and CXREs overall related to students' language-oriented attention. This finding supports the study by Doughty and Varela (1998) in which they claim the efficacy of "corrective recasts." The finding also supports Mackey et al.'s (2000) findings on the basis of Schmidt's noticing hypothesis that recasts to morpho-syntactic mistakes may not be effective: in their study, the morpho-syntactic recasts were least accurately perceived as feedback on grammar. In my study, the grammatical REs in general were found more likely to relate to students' attention to content than attention to language.

The entire set of findings suggests that the relationships among the characteristics of REs, students' attention, and their learning are not simple. For instance, despite the tendency of the students to pay attention to content during grammatical REs, the students in my study tended to make more NFB recalls in relation to grammatical REs than other error type REs. In addition, the students performed constantly well on grammatical GJ test items. On the other hand, the students' performance on lexical GJ test items was poor, even though they tended to make AL recalls in relation to lexical REs. When these general outcomes were compared between the two forms of interaction, the results differed; for example, more NFB recalls in relation to lexical than grammatical REs in the teacher-fronted, and relatively better Test 1 performance on the lexical items in the group rather than teacher-fronted context. As discussed in the previous section, the interaction contexts seemed to have an important connection to the students' attention and their GJ test performance rather than the linguistic characteristics of REs.

Delivery types, learners' reaction types, and error types can be facets of the linguistic environment created through interaction. The cognitive-interactionist SLA researchers share the assumption that interaction that modifies input into comprehensible input plays an important role in L2 learning. The evidence in this study suggested that the error type dealt with in the RE interaction has important connections to L2 learning. More

importantly, the findings regarding the impact of interaction contexts suggested the importance of individual participants' involvement in the interaction. In other words, L2 teaching/learning is not a process triggered by the linguistic environment alone or by the linguistic exchange that takes place between the teacher and the learner. The linguistic input made available for the learner is in fact the reflection and the product of the teacher's (and other learners') spontaneous thoughts in reaction to a learner's utterances. The teacher and the students interact, using language, in order to achieve their (multiple) goals of "learning English as a FL" in the classroom. Their verbal interaction is the evidence of human agency in action. This view of interaction, substantially different from and broader than that of interaction as linguistic environment perceived by cognitive-interactionist SLA researchers, has been adopted and applied by sociocultural SLA researchers. In the latter half of this thesis, I discuss recasts, interaction, and L2 learning in the sociocultural SLA framework.