

神奈川大学大学院

言語と文化論集

特別号 2014年8月

英語英文学専攻 博士後期課程修了 (2014年3月)

博士学位論文

Factors Affecting the Effect of Error Corrections

大場衣織

神奈川大学大学院 外国語学研究科

大場衣織氏の博士論文刊行に寄せて

神奈川大学大学院教授

石 黒 敏 明

神奈川大学大学院刊行の『言語と文化論集』に外国語学研究科の推薦で、大場衣織氏の博士論文が掲載されることになり、博士論文審査主査として研究科に感謝を述べるとともに、長年この研究に従事してきた大場氏に対してお祝いを申し上げたい。以下、大場氏の博士論文完成に至る経緯と博士論文の一部を紹介したい。

修士論文「インプット貧困環境におけるキャストとプロンプトの有効性」を2008年に完成した後、一貫して「学習者のエラーと教師の訂正」について研究を続けてきた。2009年と2010年には『神奈川大学大学院 言語と文化論集』にそれぞれ論文を投稿し査読を受け掲載されている。それ以外に学会活動として「神奈川大学英語英文学会」にて2008年、2010年、2011年に論文を口頭発表した。最近の研究業績として2013年に出版された『英語学・英文学の理論と応用』（共著）御茶ノ水書房出版をあげることができる。本書は、日本の大学で英語教育に携わっている著名な研究者による英語教育一般に関する論文集である。

大場氏は神奈川大学に2008年に提出した日本語による修士論文完成後、上記の学会で数々の論文を発表し、2013年9月神奈川大学大学院に英文による博士論文を提出した。その論文は外国語研究科設立以来、欧米言語文化専攻に提出された三人目の博士論文である。論文副査の東京大学名誉教授ジョン・ボチャラリ教授、神奈川大学デビッド・アリン教授、高橋一幸教授、久保野雅史准教授には、論文の作成過程において温かいご助言と指導を受け、主査として深く感謝している。次に、大場氏による博士論文の一部を紹介する。

本研究は、日本人英語学習者（高校生）に対する英語母語話者教員の訂

正の頻度、そのタイプや効果について調査したものである。リサーチ・クエスションとして、1) 教師による訂正の方法に多様性があるか、2) 教師2名の訂正方法に異なる傾向があるか、3) 教師は学習者のいかなる種類の誤答に注意が引かれるか、4) 訂正に対する学習者の反応や修正の頻度は、学習者の習熟度によって異なるか、5) 日本人学習者にも訂正は言語習得に有効であるか。これらを調べるためのデータを収集し、結果を分析した。

結果分析後の考察では、「教師間の要因」、「教師Aの要因」（教師Aがレベルの異なるクラスにおいて、異なる訂正をしたかどうか）、「学習者要因」が誤答訂正にどの程度作用するか等を考察した。

初めに「教師間の要因」、すなわち両教師の間に訂正の頻度において、さらに訂正の種類の一例プロンプトにおいて差が見られた。これらの個人差は、どの教師を被験者にするかで研究結果に違いがでてくる可能性を示唆している。しかしながら、教師間のリキャストにおいては有意差が見られなかった原因は、極めて似た環境で英語を担当した教師2名を対象としたことが挙げられる。

次に「教師Aの要因」について考察した。過去の第二言語習得研究では母語話者教師は学習者の能力によって言語の使用を変えるという報告が多い。この知見から教師Aもレベルの異なるクラス間において訂正の使用を変えることが予想できた。しかし本研究では教師Aはレベルの異なるクラス間において、訂正の頻度を変えたものの、その訂正の種類、すなわちリキャストやプロンプトにおいては調整することはなかった。つまり、教師Aは学習者の習熟度によって訂正方法を変えていないということになる。このことから、訂正に関してはこの教師はクラス別に使い分けることが困難だったのではないかと解釈した。

最後に「学習者要因」に関して、すなわち学習者の習熟度と訂正の有効性の効果（「アップテイク」と「リペア」）について考察した。学習者によるこれらの反応の差は、予測より少ない結果であった。先行研究では、学習者の習熟度と教師による訂正の有効性には関係があると分析しているが、本研究では進学クラスと普通クラスを比較した結果、訂正の有効性に差は

見られなかった。別の解釈をすると進学クラスと普通クラスとも教師による訂正が同等に有効であったとも言える。この結果を大場氏は、次のように解釈した。教師による訂正の有効性は、学習者の習熟度より、むしろ授業のコンテキスト（第二言語学習環境と外国語学習環境）や学習項目に依存するのではないか。授業形態や学習項目のみならず、学習者の年齢、使用教科書、授業の進行度に渡って同等な二つのクラスを観察したことが、このような結果に至ったと考えられる。

本論文の優れている点は以下の通りである。第一点として、過去に発表された誤答訂正に関する膨大な論文を英文で読み込み、理解を深めたことは、外国語として英語を学習した研究者にとっては、大きな功績と評価される。第二点は、その結果として、本研究で解明しようとしたリサーチクエッションへの着眼点が高く評価できる。第三点として、リキャスト、プロンプト、アップテイク、リペアなどの分類に二名の判定員が関わり、その判定の信頼性を高めた点は、通常の手続きとは言え評価できる。第四点として、過去の研究結果と比較し、類似点や相違点を深く考察している点が高く評価できる。特に、異なる結果について、その原因を単一の変数（学習者の熟達度）と解釈しないで、学習者を囲む複雑な環境を総合的に考慮し、授業形態や学習項目にも目を向けて考察している点を挙げることができる。第五点として、本論文の構成は伝統的な科学的論文の形式をとり、さらに英語自体の推敲を何度も繰り返し、かなり立派な表現となっている点を評価したい。

最後に、本研究のリミテーションをあげると、本研究は教師による訂正の短期的な効果を測ったもので、長期的な効果への推論ができない点である。教師による訂正の長期的効果を調べる研究が将来実施されることが待たれる。

以上の論文評価を持って、本論文が博士（文学）の学位を授与するのに十分値すると認められ、今回の博士論文刊行に至った。最後に、この博士論文の出版を機会に、大場氏が英語教育の専門家として大いに飛躍すること、また後に続く後輩の目標となることを願いながら、大場氏の博士論文刊行祝いの言葉とする。

Acknowledgements

It is with deep gratitude that I wish to acknowledge the contributions made to this dissertation by many people, only a few of whom can be mentioned here.

I am grateful to Professor Toshiaki Ishiguro who is my adviser in Kanagawa University. Without his intellectual advice and encouragement throughout my years of master's and doctor's courses, this research would have been impossible. Particularly, his comments on the earlier version of this dissertation were helpful for improving it.

I am also thankful to the second readers: Professors John Boccellari, Kazuyuki Takahashi, Masashi Kubono, and David Aline.

I wish to express my gratitude to Emeritus Professor of Kanagawa University, Mikiko Ishii for her enduring encouragement and patient support.

Finally, I express my deepest gratitude to the students and the teaching staff of Yokohama Sougakkan high school, who allowed my wishes for collecting data on English learners in Japan. Undoubtedly, their cooperation was necessary for the purpose of this research.

Factors Affecting the Effect of Error Corrections

A Dissertation

Submitted to the Faculty of

Kanagawa University

In Partial Fulfillment of the Requirements for

the Degree of Doctor of Philosophy

By Iori Oba

September 2013

Copyright 2013

By

Iori Oba

Contents

List of Tables and Figures

Chapter

I. Introduction	1
II. Literature Review	3
A. Second Language Acquisition Studies Concerning Classrooms	3
B. Research Methods Used in Previous Studies	11
C. Instruments Used in Previous Studies	16
D. Reliability and Validity in Previous Studies	26
E. Summary	32
III. Method of Study	35
A. Research Questions	35
B. Research Procedures	36
C. Data Collection	38
D. Instruments	40
E. Reliability and Validity	41

IV. Results	43
A. Results from Present Data	43
B. Summary	51
V. Discussions and Conclusions	54
A. Teacher Factors	54
B. One-Teacher Factors	58
C. Learner Factors	61
D. Efficacy of Error Corrections for English Learners in Japan	64
E. Conclusions	68
Bibliography	72
Appendix A	81
Appendix B	82

List of Tables

Table

1	Frequency of teacher feedback	43
2	Distribution of feedback types	44
3	Frequency of teacher feedback	46
4	Distribution of feedback types	47
5	Distribution of error types (teacher A)	48
6	Distribution of error types (teacher B)	48
7	Uptake, repair, and no uptake following teacher feedback	49

List of Figures

Figure

1	Frequency of error corrections in two different-level classes	44
2	Percentage distribution of feedback types in the two different-level classes	45
3	Frequency of error correction in two same-level classes	46
4	Distribution of feedback types in the two same-level classes	47
5	Distribution of error types	49
6	Distribution of uptake and no uptake	50
7	Distribution of uptake and repair	50

Chapter I

Introduction

In second language research, many studies have investigated procedures in actual language classrooms. Research on teachers' use of corrective feedback has been ongoing for several decades. Corrective feedback refers to teachers' corrections of language learners' erroneous utterances, including a teacher's correct paraphrasing of a student's incorrect utterance or a teacher's request for a student's reformulation of a previous utterance. Research on the effectiveness of corrective feedback has been conducted in both second and foreign language contexts.

However, there is little research on the effectiveness of corrective feedback among Japanese language learners of English. In addition, there has been even less research on younger learners (students in junior high and high schools). Furthermore, the results of some previous studies are inconsistent with those of other studies. Some of the causes for such a discrepancy may be: 1) different second or foreign language settings, 2) different classroom content-structure settings, or 3) different proficiency levels of the learners. The present study will be conducted in order to shed further light on these topics, and to see what aspects affect the results of classroom-based research.

The first research question in this study is whether a particular teacher varies the use of feedback. To answer this research question, one teacher who teaches advanced and average-level classes will be observed.

The second research question is whether there are any differences between

the two teachers in the use of feedback when teaching learners at the same level. To investigate this question, two teachers who teach the same-level classes will be observed.

The third research question is what kind of learners' errors tend to attract the most attention of the teachers. The two teachers of the same-level classes will be observed and the learners' four error types will be aligned in the order of frequency.

The fourth research question is whether the frequencies of learners' uptake (i.e., learners' immediate reactions to the teacher corrections) and repairs (i.e., the correct reformulations of errors) following teacher corrections may be different, according to the learners' proficiency levels. The advanced class and the average one will be observed.

The fifth research question is whether the teacher corrections are immediately beneficial to the learners who study English in Japan. To see the effectiveness of the teacher corrections, the ratios of the learners' uptake and repairs will be calculated. In order to examine these aspects of teachers' corrections, the literature related to teacher's feedback will be reviewed in the following chapter.

Chapter II

Literature Review

In this chapter, the previous studies related to classroom research and teacher feedback are introduced. The discussion in this chapter covers 1) second language acquisition studies concerning classrooms, 2) research methods used in previous studies, 3) instruments used in previous studies, and 4) reliability and validity in previous studies.

A. Second Language Acquisition Studies Concerning Classrooms

According to Ellis (1994), second language acquisition studies concerning actual classrooms can be divided into three perspectives. The following quote describes the first perspective: “The first perspective is that found in *comparative method studies*. These seek to compare the effect of different language teaching methods on L2 learning” (565). Ellis (1994) explained the second perspective as follows:

A second perspective involves going inside the ‘black box’ of the classroom itself. It views the classroom as a place where interactions of various kinds take place, affording learners opportunities to acquire the L2. (565)

Ellis (1994) presents the third perspective as follows:

The third perspective involves investigating the effects of formal instruction. In this case, instruction is viewed as an attempt to intervene directly in the language learning process by teaching specific properties of the L2. . . . Researchers have been particularly interested in whether instruction directed at specific grammatical

items and rules has any effect on interlanguage development. (565)

Furthermore, different aspects of classroom interactions can be investigated: teachers' talk, teachers' questions, error corrections, learners' participation, and activities done in small groups. However, these aspects are strongly connected to each other and it is difficult to investigate individual effects.

Research on teacher feedback in second language acquisition was inspired by the results of first language acquisition research conducted by the first language researchers (e.g., Farrar; Bohannon and Stanowicz). In particular, Bohannon and Stanowicz (1988) found that more than 70% of parents' recasts and their expanded repetitions followed ill-formed sentences uttered by children:

Results indicated that adults were more likely to repeat verbatim a well-formed sentence than an ill-formed sentence. In contrast, adults were more likely to repeat with changes, or request clarification of, a sentence containing syntactic or phonological errors than well-formed sentences. (684)

Kail and Nelson (1993) discussed the phenomenon of maternal speech:

Parents, as well as other adults, do provide feedback by repeating well-formed statements and by recasting poorly formed ones. This does not occur for every utterance; in fact, a majority of children's errors—about two thirds—go uncorrected. However, the amount of feedback is sufficient for children to reject incorrect hypotheses about syntactic rules and retain correct hypotheses. (167)

Kail and Nelson (1993) also referred to a specific feedback type produced by parents: "Several investigators have shown that use of recasts is associated with more rapid language acquisition" (168). Given the results of first language acquisition research, some second language researchers assumed that corrective feedback is beneficial for second language acquisition. Thus, they started to investigate the effectiveness of corrective feedback in the field of second language learning.

One of the earlier studies for corrective feedback in second language research was conducted by Fanselow (1977). He observed eleven teachers who taught in ESL classrooms and found that there were sixteen types of treatments following the learners' errors, including verbal and non-verbal behaviors. He commented on the results of his study in the following way: "The similarity of behavior among the teachers did not provide as much insight into the treatment of errors as was hoped for" (583).

Chaudron (1977) conducted research using a discourse analysis framework. He observed three teachers and categorized their corrective feedback into four major types: "repetition with change," "repetition with change and emphasis," "repetition with no change," and "repetition with no change and emphasis." Chaudron (1977) concluded his findings as follows:

The differences between teachers can readily be observed, and variations for any one teacher will also be recognized. This is to say that any use of the present model in actual observation will have to take into account individual teachers' "basal" features or types of corrections, so that special intonation and uses of stress would be evaluated relative to each other and to the teacher's "unmarked" reactions. (44)

Doughty and Varela (1998) emphasized the potential effectiveness of negative feedback:

Taken together, these findings suggest that not only do adults provide negative evidence to children but that children notice this information and make use of it in acquisition. Furthermore, examination of examples of the child-directed discourse reported in the above-mentioned studies shows clearly that the provision of negative evidence via recasting does not halt communication between parent and child but, rather, is relatively incidental to the primary goal of mutual understanding. These findings provided a basis for

predicting that recasting would be the ideal FonF¹ procedure to be implemented in our study, provided that such recasting could successfully be accomplished in a classroom setting. (117)

Long (1996) offered an explanation for the terminology “recast,” which was noted in the citations above:

Recasts are utterances which rephrase a child’s utterance by changing one or more sentence components (subject, verb, or object) while still referring to its central meanings. (484)

Long (1996) also emphasized the effectiveness of teacher feedback:

Negative feedback obtained during negotiation work or elsewhere may be facilitative of L2 development, at least for vocabulary, morphology, and language-specific syntax, and essential for learning certain specifiable L1-L2 contrasts. (414)

Ammar and Spada (2006) compared a group of students who received teachers’ feedback and a group who did not receive teachers’ feedback. The target feature of the study was the possessive case of the third-person singular “her/his.” The group receiving teachers’ feedback acquired the ability to use the possessive case correctly. The results of the study showed a significant difference for the group that received the teacher feedback. The results of previous studies indicate that corrective feedback may promote second language ability.

In psychology, there was an argument that the provision of negative feedback was not sufficient to establish its usefulness in language acquisition. For example, Farrar (1990) indicated that the provision of input has a selective effect on language acquisition:

These results suggests that in addition to the general effects often demonstrated in previous work, very specific links between input and acquisition can be established also. The findings also support claims

¹ The abbreviation for “Focus on Form” (this note is mine).

by several investigators (e.g. Newport et al. 1977, Goldin-Meadow, 1982) who argued that adult input will be particularly effective in assisting the child in acquiring language-specific characteristics, such as grammatical morphemes, compared to language-universal characteristics. (621)

Similar to the argument above, this issue permeated the field of second language acquisition. VanPatten (1990) argued the following point:

How do learners get intake from input? It is generally acknowledged that not all of input is available for language processing, that much of input is “noise.” Intake is thus defined as a subset of the input that the learner actually perceives and processes. (287)

Swain (1998) also proposed that learners develop their second language ability not only by receiving input but also by perceiving the target features and using them:

The collection of studies discussed in this chapter (Kowal & Swain, 1994, 1997; LaPierre, 1994; Swain & Lapkin, 1995, 1996, in progress) shows evidence of learners noticing the “gap” in their interlanguage, that is, noticing the difference between what they want to say and what they are able to say. As proposed by the output hypothesis, this happens as the students try to produce the target language. (79)

However, the statements quoted above from Swain and VanPatten are basically different. Swain emphasized that the production of language is an important part of acquiring a language, while VanPatten emphasized that production of language only develops the accuracy and fluency of oral production. There are some interesting studies that compared input practice with output practice (e.g., VanPatten 287-301; Dekeyser and Sokalski). In the study by Dekeyser and Sokalski (1996), 82 learners of Spanish participated and were divided into three groups (the control group, the input practice group, and the output practice group).

The target features of their study were two Spanish morphosyntactic features. Pretests, posttests, and delayed posttests were conducted among the groups. Dekeyser and Sokalski (1996) concluded their study as follows:

The first hypothesis (for the direct object clitic², input practice would be better than output practice for comprehension and no worse for production) was confirmed only in part. The immediate posttest shows the interaction effect that one would expect on the basis of the previously mentioned research on the specificity of skill acquisition: Input practice is significantly better for comprehension tasks, and output practice significantly better for production tasks. — The second hypothesis (for the conditional³, output practice would be better than input practice for production and no worse for comprehension) was also confirmed only in part. For the conditional, output practice was better than input practice for both production and comprehension tasks in the immediate posttest. (613)

Although there is an argument for comparing the effectiveness of output practices with that of input practices, many researchers still consider recasts as input practices and the ideal corrective feedback.

Lyster and Ranta (1997) identified six types of teacher corrections: 1) recast, 2) clarification request, 3) meta-linguistic feedback, 4) elicitation feedback, 5) repetition, and 6) explicit corrections. The most observable type was recast; the next was clarification request, followed by meta-linguistic feedback, elicitation, and repetition. Lyster revised the term meta-linguistic feedback as “meta-linguistic clue” in a later study and categorized those four types (clarification request, meta-linguistic clue, elicitation, and repetition) into prompts. In their research, the least observable type among these categories was

2 The direct object clitics usually precede the verb and the subject follows the verb in many cases in Spanish (this note is mine).

3 The conditional form of the verb. Verbal agreement morphology (this note is mine).

explicit correction. The correction type prompts have been described in various ways in previous studies. For example, some studies have called this type “negotiation of meanings,” “negotiation of form,” and “prompts.”

According to Ellis (2006), corrective feedback can be distinguished into two categories: 1) input-providing corrective feedback and 2) output-pushing corrective feedback. Input-providing corrective feedback provides correct linguistic forms through recasts or explicit corrections. Output-pushing feedback does not provide correct forms, but instead forces learners to correct the utterances themselves through prompts.

Recasts are often described as implicit feedback because they can correct students’ erroneous utterances without disturbing classroom contexts. However, some scholars have questioned the effect of recasts, such as Mackey and Philp (1998) who argue the following: “It is difficult to identify whether learners who repeated the recast were actually perceiving the recast as feedback or simply another way of saying the same thing” (351). Lyster and Ranta (1997) also made a similar point. Furthermore, Mackey and Philp indicated that recasts may be effective only for developmentally ready learners as “It was the developmentally ‘ready’ learners who made the greater gains in terms of sustained increase in higher-level structures” (352). There are other studies that have referred to the limitation of the effectiveness of recasts. For example, Nicholas, Lightbown, and Spada (2001) proposed that the learners were not aware of recasts in content-based or communication-based classrooms but they were aware of recasts in focus-on-forms classrooms:

The results of the classroom studies indicate that the classroom context (particularly the communicative and/or content-based classroom) may make it difficult for learners to identify recasts as feedback on form and hence difficult for them to benefit from the reformulation that recasts offer. (744)

They interpreted the finding as “The effectiveness of recasts may depend in part

on the overall developmental level of proficiency or inter-language variety of the learner” (752). VanPatten (1990) also emphasized the relationship between learners’ awareness of linguistic forms and the inputs provided by teachers:

The results of the present study offer evidence that conscious attention to form in the input competes with conscious attention to meaning, and, by extension, that only when input is easily understood can learners attend to form as part of the intake process. (296)

On the other hand, many studies have found that teacher feedback prompts are effective to learners in any situation (e.g., Havranec and Cesnik; McDonough; Ammar and Spada). Ammar and Spada (2006) concluded that:

Overall, prompts were more effective than recasts and that the effectiveness of recasts depended on the learners’ proficiency. In particular, high proficiency learners benefited equally from both prompts and recasts, whereas low proficiency learners benefited significantly more from prompts than recasts. (543)

Nobuyoshi and Ellis (1993) conducted a small-scale study to observe low-level adult learners, and they concluded the following:

This study provides some support for the claim that ‘pushing’ learners to improve the accuracy of their production results not only in immediate improved performance but also in gains in accuracy over time. Two of the learners in the experimental group showed significant gains in accuracy, whereas none of the learners in the control group did so. (208)

Many scholars have studied prompts not only for their effectiveness but also for their role in language learning, which can lead learners to perform “self-repair.” Self-repair refers to cases in which learners correct their erroneous utterances without teachers’ correct models. Only prompts can lead learners to self-repair, so scholars have assumed these to be the most useful corrections. According to Havranec and Cesnik (2001), self-repaired utterances produced by

learners work effectively for developing second or foreign language learning.

B. Research Methods Used in Previous Studies

Researchers who engage in classroom research use a variety of research methods. Chaudron (1988) illustrated four types of research methods: psychometric, interaction analysis, discourse analysis, and ethnographic. He explained the psychometric approach as follows:

The most traditional approach to the study of L2 classrooms involves comparison of the effects of specific instructional programs or methods on student learning outcomes, as measured by standardized proficiency tests or instruction-related achievement tests. This program (context)-product approach was taken by most language education researchers in the post-World War II period until the 1970s. (28)

Ellis (1994) interpreted the psychometric method as an experimental method that includes pre- and post-tests with experimental and control groups. The psychometric studies conducted by other scholars are discussed below.

For the first example, Ammar and Spada (2006) compared a group of students that received teacher feedback with a group that did not receive teacher feedback in order to examine learners' acquisition of the possessive case of the third-person singular "her/ his."

For the second example, Dekeyser and Sokalski (1996) conducted an experimental study. In their study, three groups were compared: a control group, an input practice group, and an output practice group. To examine the effectiveness of input and output, practices, pre-tests, post-tests, and delayed post-tests were conducted among them.

As for the interaction analysis approach, researchers used instruments and coded the classroom conversation in real-time or after the observations. Furthermore, the instruments included some categories and researchers checked

the items written in the categories. Chaudron (1988) explained interaction analysis as follows:

Of the earlier instruments for observation of classroom interaction, the most well known was that of Moskowitz (1968, 1970, 1971). . . Moskowitz's adaptation for L2 classrooms involved the separate simultaneous coding for language of each behavior (1970), and later, categories for drill and feedback behaviors (1976) and the like. (31)

The studies using interaction analysis are introduced below.

First, Moskowitz (1976) conducted research to answer the question “what makes a good teacher good?” He compared outstanding teachers' classes to normal teachers' classes using the Foreign Language Interaction System. The Foreign Language Interaction System classifies not only teachers' verbal and non-verbal behavior, but also those of students. In this case, a trained observer wrote down a category number every time a different behavior was used in a real time.

Second, Fanselow (1977) conducted research using the instruments called FOCUS (Foci for Observing Communications Used in Settings). He classified the behavior of people sending and receiving messages in both teaching and non-teaching settings. FOCUS is able to distinguish five characteristics of communication: the source (the person who communicates), the medium (linguistic medium or non-linguistic medium), the use (how the mediums were used to communicate), the content (what areas were communicated), and the pedagogical purpose.

Third, Spada and Frohlich (1995) described the COLT (Communicative Orientation of Language Teaching) framework. According to Spada and Frohlich, COLT was divided into two parts: part A and part B. Part A represented teaching practices in terms of content, focus, and organization of activity types. Part B represented specific aspects of the language produced by teachers and students. The COLT framework is generally used to characterize the nature of interaction

between teachers and students, and between students and other students.

The discourse analysis approach is research method that uses the framework of discourse analysis. Chaudron (1988) explained this approach as follows:

A major step was made in the discourse analysis of L1 classrooms with the research of Sinclair and Coulthard (1975) on British elementary school classrooms with minority children. They incorporated both the linguistic and sociolinguistic traditions in their conception of classroom interaction as a hierarchically structured system of “ranks,” analogous to the rank scale approach to sentential linguistic description (see Halliday 1961). (40)

Sinclair and Coulthard (1975) noted a crucial step in developing classroom research. They emphasized that the basic unit of interaction is not a single utterance, but at least two utterances produced by different speakers. According to Coulthard (1978), individual utterances were called “moves.” Furthermore, Sinclair and Coulthard (1975) noted that each unit begins with an “initiating” move and a “responding” move that follows the “initiating” move. A third “follow-up” move is necessary in classroom interaction. Similar to the terms defined by Sinclair and Coulthard above, Sacks (1974) introduced the different terminologies of “turn” and “pair.” Coulthard (1978) explained the difference between move and turn as follows:

Despite the different labels, the units proposed by Sacks and Sinclair et al. are remarkably similar and one can often read ‘move’ for ‘turn,’ ‘exchange’ for ‘pair,’ and so on. (23)

The following are representative studies that have applied the discourse analysis procedure.

Sinclair and Coulthard (1975) conducted research in a British elementary school. They tried to discover how much of conversations were pedagogical and how much were linguistic. They categorized the teachers’ “opening moves,” the

learners' "answering moves," and the teachers' "follow-up moves." These categories were devised to observe classroom conversations. They also identified an "act," which denotes the functions of utterances (e.g., request a response from a listener).

Chaudron (1977) adopted the method described by Sinclair and Coulthard (1975). He employed discourse analysis to analyze classroom conversations between the learners and the teachers. He used the terms (moves and act) demonstrated by Sinclair and Coulthard (1975). Second language classroom researchers rarely employed the discourse analysis approach before Chaudron (1977), so it may be argued that Chaudron contributed a new approach for the field of second language research.

Lyster and Ranta (1997) also adopted the terms (e.g., turns) derived from discourse analysis, though these might have been inspired by the study of Schegloff, Jefferson, and Sacks (1977).

The final research technique, the ethnographic method, was illustrated by Chaudron (1988) as follows:

The procedures for conducting ethnographic research involve considerable training, continuous record keeping, extensive participatory involvement of the researcher in the classroom, and careful interpretation of the usually multifaceted data (for description of such procedures see Wilson 1977; Mehan 1979; and Levine et al. 1983). The result of such an investigation is usually a detailed description of the research site, and an account of the principles or rules of interaction that guide the participants to produce their actions and meanings and to interpret the actions and utterances of others.

(46)

The representative studies using the ethnographic approach are presented below.

Schmidt and Frota's (1986) study is well-known as the diary study. Schmidt used a diary to record his language-learning experiences in classes of Portuguese

as a second language. Afterwards, Schmidt and Frota proposed the well-known “noticing hypothesis.” Schmidt consciously noticed linguistic forms in the second language input before he acquired the target linguistic forms. He reached the conclusion that learners needed to focus on the target linguistic features consciously before they can acquire the target linguistic forms.

As mentioned above, Chaudron (1988) presented four basic approaches to investigate actual classrooms. On the other hand, Long (1980) described two basic approaches for classroom research (each approach has specific subcategories) as follows:

Two basic approaches characterize classroom research on second language learning: interaction analysis and anthropological observation. Within each approach several methods are available for data collection and analysis. Procedures described are the use of behavioral observation systems, discourse analysis, ethnography, constitutive ethnography, and diary studies. (1)

The definition of the research approaches presented by Long and Chaudron shared many aspects, although Long’s categorization included more specific subcategories.

Fraenkel and Wallen (1996) also explained that various research methodologies have been applied to educational research. In contrast to the definitions provided by Chaudron or Long, they presented three general research methods: descriptive, associational, or intervention type studies. They described descriptive research as follows:

In educational research, the most common descriptive methodology is the survey⁴, as when researchers summarize the characteristics (abilities, preferences, behaviors, and so on) of individuals or groups, or (sometimes) physical environments (such as schools). Content

⁴ Fraenkel and Wallen (1996) define “survey” as follows: Information is collected to describe the characteristics of a group.

analysis⁵, qualitative, and historical⁶ research methodologies are also primarily descriptive in nature. (13)

Fraenkel and Wallen (1996) explained associational research as follows:

Research that investigates relationships often referred to as associational research. Correlational⁷ and causal-comparative⁸ methodologies are the principle examples of associational research. Examples of associational studies include studying relationships (a) between achievement and attitude, between childhood experiences and adult characteristics, or between teacher characteristics and student achievement. (14)

They described intervention studies as follows:

In intervention studies, a particular method or treatment is expected to influence one or more outcomes. Such studies enable researchers to assess, for example, the effectiveness of various teaching methods, curriculum models, classroom arrangements, and other efforts at influencing the characteristics of individuals or groups. . . . The primary methodology used in intervention research is the experiment. (14)

According to Fraenkel and Wallen (1996), some studies combined more than two types of these three approaches.

C. Instruments Used in Previous Studies

According to Long (1980), there are three basic types of instruments

5 Fraenkel and Wallen (1996) define “content analysis” as follows: The contents of a communication are analyzed to look for patterns or relationships that may exist.

6 Fraenkel and Wallen (1996) define “historical” as follows: Information from the past is analyzed to better understand what took place.

7 Fraenkel and Wallen (1996) define “correlational” as follows: Within one group, individual scores on one attribute are compared with scores on another attribute.

8 Fraenkel and Wallen (1996) define “causal-comparative” as follows: Two or more existing groups of people are compared to study possible causes or consequences of differences between them.

according to the recording procedures: category system, sign system, and rating scale. Long defined the category system as follows: “When each event is coded each time it occurs we are dealing with a true category system” (6). Many previous studies have used category systems (e.g., Chaudron 29-46; Moskowitz; Spada and Frohlich; Lyster and Mori). In the present study, the category system was adopted because this investigation draws on discourse analysis as part of its methodology. The advantage in using the category system is that it can describe classroom conversations more precisely. Moreover, Long (1980) explained that instruments may be different in terms of the items that the instruments contain. He divided items into three types: low-inferences, high-inferences, and mixed-inferences. Mixed-inferences were adopted in this study because specific and less-specific categories can be contained in the mixed-inferences. Less specific items are easy to determine by researchers, although the specific phenomenon that appears in classroom interactions cannot be found.

Instruments can be selected according to the units that researchers employ. Long (1980) classified the units into arbitrary units and analytic units as follows:

Arbitrary units involve the selection of a (usually short) time period, e.g., three seconds, at the end of which the most prominent event of the period is coded. . . . Analytic units, e.g., the move, episodes, or speech act, inevitably mean irregular coding; they do not adhere to time schedules, and they involve more abstraction from the data and far greater variety in their realization. (10)

Following the study by Sinclair and Coulthard (1975), the analytic units derived from discourse analysis have been used in research on second language learning. For example, Oliver (2000) coded the NS-NNS⁹ conversations into three parts: 1) initial turn, 2) response, and 3) reaction. He proposed that this represents a cyclical pattern, and that the NNS reaction becomes the initial turn in

⁹ The abbreviation for “Native Speakers vs. Non-Native Speakers.”

the next three-part exchange. In other words, the focal points in previous classroom observations were: first utterances by learners, second utterances by teachers, and third utterances by learners.

For the initial turn, a categorization of learners' errors was needed because initial turn contains the learners' errors. Chaudron (1977) categorized learners' errors according to different types: linguistic, content, lexical items, and interaction and discourse. Chaudron explained the interaction and discourse type as follows:

... insofar as these last ones are orally manifested, as in speaking out of turn, taking up the wrong question in the lesson, using English in the immersion context, on occasion failing to speak, and not speaking in complete sentences (regarding this last category, cf. Mehan 1974). (32)

Fanselow (1977) also referred to the study by Mehan (1974), and categorized learners' errors into more specific categories: incorrect function words, incorrect pronunciation, incorrect agreement, incorrect word order, incorrect content words, full form, and different task. Fanselow explained that some teachers asked students to change full forms to contractions, and he called the students' production of the full form "full-form" type. "Different task" was another new category presented by Fanselow. When a student's response was linguistically correct but different from what the teacher had asked the student to do, it was called "different task."

Recent work in classroom research has applied simpler classifications compared to those used by Chaudron (1977), Fanselow (1977), or Mehan (1974). For example, Lyster and Ranta (1997) classified learners' errors into five types: L1 (the learner's first language), gender, grammatical, lexical, phonological, and multiple (learner errors including more than one error). Oliver (2000) used more concise definitions of learner errors:

They were rated as either correct, nontargetlike (e.g., ungrammatical

utterance or one containing an obvious pronunciation error), or *incomplete* (e.g., ellipsis or an interrupted attempt). (130)

The classification demonstrated by Oliver did not contain the lexical category of the learners' errors. Thus, it seems to be impossible to grasp the whole range of learner errors.

As for the error classification explained above, some errors could be interpreted as both a phonological and a lexical error. For example, if a learner said "fifty" instead of "fifteen," it could be interpreted as either the learner being confused by the lexical items or as the learner failing to pronounce the last phoneme /n/. When more than two interpretations are possible, a specific classification for learner errors would be needed. In reference to studies conducted by other scholars, the specific classification is summarized by the present author below. For errors involving the level of phonemes, the three directions presented below were followed.

1. When a learner makes a phonemic error, which can alter the meaning, it will be classified as a lexical error. "Fifty" /fifti/ for "fifteen" /fifti:n/ is classified as a lexical error.

2. When a learner makes a phonemic error that can alter the meaning and affect grammatical aspects, it will be categorized as a grammatical error. For example, when a learner produced "named"/neimd/ for "name" /neim/, it is classified as a grammatical error.

3. When a learner makes a phonemic error that does not alter the meaning, it will be classified into the category of lexical errors. For example, "*freshermen" for "freshmen" is classified as a lexical error.

The directions below were adopted to classify errors concerning bound morphemes.

1. Inflectional morphemes (the plural "-s," "-ed," "-en," "-ing," "-s," the third person singular "-s," "-er," and "-est") are categorized as grammatical errors.

2. Derivational morphemes are categorized as lexical errors. For example, “They have their own *cultural/ culture.”

As for the level of words, the two main classifications shown below were used.

1. When a learner’s uttered words have problems relating to rhythm and stress, it will be categorized as a phonological error.

2. When a problem occurs in content words and applies to the three conditions explained below, it will be categorized as a lexical error.

2.1. Confusion with sense relations; for example, “foot”/ “leg.”

2.2. Collocational errors; for example, “high building”/ “*high pencil.”

2.3. Misformations; for example, “*baby car”/ “stroller.”

When the learner errors were at the level of phrases and clauses, the following procedures were taken.

1. When a learner’s uttered phrases or clauses have problems with intonation or rhythm, it will be categorized as a phonological error.

2. When a problem occurs in content words or function words that are a part of a learner’s uttered phrases or clauses and applies to the three conditions below, it will be categorized as a grammatical error.

2.1. Omissions: “Omission errors are characterized by the absence of an item that must appear in a well-formed utterance” (Dulay et al. 154). For example, “*She sleeping” has omitted “is” before “sleeping.”

2.2. Additions: “Addition errors are the opposite of omissions. They are characterized by the presence of an item which must not appear in a well-formed utterance” (Dulay et al. 156). For example, “is” is redundant in “*She is sleeps.”

2.3. Misorderings refer to:

As the label suggests, misordering errors are characterized by the incorrect placement of a morpheme or group of morphemes in an utterance. For example, in the utterance “He is all the time late.” all the time is misordered. (Dulay et al. 158)

For example, “is” is located at an incorrect position in “*She sleeping is in the bed.”

Response turn also needed classification for the teacher corrections as well as the initial turn by the learners. Lyster and Ranta (1997) and Lyster and Mori (2006) confirmed six types of teacher feedback: explicit correction, recast, clarification request, meta-linguistic clue, elicitation, and repetition. The study by Lyster and Ranta (1997) defined “explicit correction” as follows:

Explicit correction refers to the explicit provision of the correct form. As the teacher provides the correct form, he or she clearly indicates that what the student had said was incorrect (e.g., “Oh you mean,” “You should say”). (46)

An example of explicit corrections in the present data is as follows:

(1)

S¹⁰: The Yokohama station.

T¹¹: You don’t have to say “the Yokohama station,” just Yokohama station. [Explicit correction]

S: Ah, yeah? No?

Recasts were explained by Lyster and Ranta as “recasts involve the teachers’ reformulation of all or part of a students’ utterance, minus the error” (46). An example of recasts in the present data is as follows:

(2)

S: *Fudousan* (Real estate agent).

T: Real estate agent. [Recast]

S: Real estate.

T: Agent. [Recast]

In the first turn of the interaction, the learner could not say “real estate agent” in English. Therefore, the learner said “real estate agent” in his first language

¹⁰ The abbreviation for “Student.”

¹¹ The abbreviation for “Teacher.”

(Japanese). In the second turn, the teacher responded to the learners' ill-formed utterance by giving the correct model: "real estate agent." In the third turn, the learner repeated the correct model provided by the teacher but he could correct only part of his ill-formed utterance. Therefore, in the fourth turn, the teacher completed the utterance by adding "agent." The second turn uttered by the teacher is a type of "repetition with change," as defined by Chaudron (1977). Moreover, this turn corresponds to the four properties of recasts noted by Long (1996):

- (a) they are a reformulation of the ill formed utterance, (b) they expand the utterance in some way, (c) the central meaning of the utterance is retained, and (d) the recast follows the ill formed utterance. (434)

Lyster and Ranta (1997) and Lyster and Mori (2006) classified the latter four types of feedback (clarification request, meta-linguistic clue, elicitation, and repetition) into prompts. Clarification requests tell the learner that there is a misunderstanding in the learners' utterances. When teachers use prompts, they do not give a correct model to learners. For example, "What?," "I don't understand you," and "I'm sorry?" are the typical phrases used for clarification requests. Lyster and Mori (2006) explained a clarification request as follows:

- The teacher uses phrases such as "Pardon?" and "I don't understand" after learner errors to indicate to students that their utterance is ill-formed in some way and that a reformulation is required. (271)

Lyster and Ranta (1997) defined a meta-linguistic clue as follows:

- Meta-linguistic feedback contains either comments, information, or questions related to the well-formedness of the students' utterances, without explicitly providing the correct form. (47)

An example of a meta-linguistic clue found in the present data is as follows:

(3)

S: *Demo jyuukuji han no houga iino?* (Well, in English do you say

nineteen thirty?)

T: We don't usually say nineteen thirty. [Meta-linguistic clue]

S: *Suuji de iino?* (Is it ok to say the number?)

In the first turn of this interaction, the learner formulated his question in his first language. The content of his question asked how to express 7:30 p.m. in English because it is possible to say either seven o'clock or nineteen o'clock in Japanese. In the second turn, the teacher gave the learner the following information: "We don't usually say nineteen thirty." The teacher narrowed the learners' choices, but the learner ultimately did not find the answer to the question.

For the category of elicitation, Lyster and Ranta (1997) demonstrated that there were at least three techniques of elicitation:

First, teachers elicit completion of their own utterance by strategically pausing to allow students to "fill in the blank" as it were (e.g., "C'est un..."). Second, teachers use questions to elicit correct forms (e.g., "Comment on dit ca?," "Comment ca s'appelle?," "How do we say X in French?"). Such questions exclude the use of yes/no questions: A question such as "Do we say that in French?" is metalinguistic feedback, not elicitation. Third, teachers occasionally ask students to reformulate their utterance. (48)

An example of the first type of elicitation found in this study is as follows:

(4)

T: What are you practicing?

S: (pointing on a handout) *Kore* (This).

T: Why?

S: *Yarette iwareta* (Because the teacher told me to do it).

T: Because...? [Elicitation]

The classification for reaction was considered as well. Lyster and Ranta (1997) classified learners' reactions into two divisions: "uptake" and "no uptake." They explained the phenomenon of uptake and no uptake as follows: "If

corrective feedback is provided by the teacher, then it is either followed by uptake on the part of the student or not (no uptake entails topic continuation)” (45). They defined uptake as follows:

Uptake in our model refers to a students’ utterance that immediately follows the teachers’ feedback and that constitutes a reaction in some way to the teachers’ intention to draw attention to some aspect of the students’ initial utterance (this overall intention is clear to the student although the teachers’ specific linguistic focus may not be). (49)

Specifically, Lyster and Ranta classified uptake into two types, denoted as repair and needs-repair:

There are two types of student uptake: (a) uptake that results in “repair” of the error on which the feedback focused and (b) uptake that results in an utterance that still needs repair (coded as “needs-repair”). (49)

They classified repair uptakes into four types: 1) repetitions, 2) incorporations, 3) self-repairs, and 4) peer-repairs. In this study, I refer to the categorization of repetitions and self-repairs and do not refer to the incorporations and peer-repairs because peer-repairs were not recorded in this observation and the definition of incorporations given by Lyster and Ranta was ambiguous. The definitions of repetitions and self-repairs from Lyster and Ranta (1997) are as follows:

Repetition refers to a student’s repetition of the teacher’s feedback when the latter includes the correct form. . . . Self-repair refers to a self-correction, produced by the student who made the initial error, in response to the teacher’s feedback when the latter does not already provide the correct form. (50)

An example of repetition observed in the present data is as follows:

(5)

S: Bring back.

T: Bring it back.

S: Bring it back. [Repair-repetition]

In the first turn of this interaction, the student omitted an object. In the second turn, the teacher provided the correct model of the utterance. Finally, the student repeated the correct form provided by the teacher. An example of self-repair is also shown below:

(6)

S: Taro and Ichiro put out with my house.

T: What does it mean? Taro and Ichiro put out with my house?

S: Stay? [Repair-self]

In the first turn of this interaction, the student uttered a lexically incorrect utterance. In the second turn, the teacher indicated that he could not understand what the student said. The student corrected his utterance in another way in the third turn. The student corrected his ill-formed utterance without the teacher's correct model.

Lyster and Ranta (1997) classified needs-repair into six types: acknowledgement, same error, different error, off-target, hesitation, and partial repair. First, "acknowledgement" is as follows:

Acknowledgement generally refers to a simple "yes" on the part of the student in response to the teacher's feedback, as if to say, "Yes, that is indeed what I meant to say (but you've just said it much better)" (see Calve, 1992). Acknowledgement may also include a "yes" or "no" on the part of the student in response to the teacher's metalinguistic feedback. (50)

Second, they illustrated "same error" as signifying "uptake that includes a repetition of the student's initial error" (50).

Third, "different error" is defined as follows:

Different error refers to a student's uptake that is in response to the teacher's feedback but that neither corrects nor repeats the initial

error: instead, a different error is made. (50)

Fourth, “off target” is described as follows:

Off target refers to uptake that is clearly in response to the teacher’s feedback turn but that circumvents the teacher’s linguistic focus altogether, without including any further errors. (50)

Fifth, “hesitation” is defined as follows: “Hesitation refers to a student’s hesitation in response to the teacher’s feedback” (Lyster and Ranta 50).

Finally, “partial repair” refers to “uptake that includes a correction of only part of the initial error” (Lyster and Ranta 50). Learners’ needs-repair can lead to teachers’ further feedback until learners can produce the correct forms.

D. Reliability and Validity in Previous Studies

Whichever approach is taken by a researcher, research has to be meaningful for others. From this perspective, classroom researchers need to be aware of reliability and validity. Validity is the measure that determines whether a method or an instrument is appropriate for the purpose of research. Mackey and Gass (2005) explained the notion of validity as follows:

After spending a great deal of time and effort designing a study, we want to make sure that the results of our study are valid. That is, we want them to reflect what we believe they reflect and that they are meaningful in the sense that they have significance not only to the population that we tested, but, at least for most experimental research, to a broader, relevant population. (107)

According to Mackey and Gass (2005), there are two main types of validity: internal validity and external validity. Internal validity refers to the “participant characteristics, participant morality, participant inattention and attitude, participant maturation, data collection, and instrumentation and test effects” (109).

As for participants’ characteristics, one aspect of the characteristics is

language background. In general, experimental studies compare two or more groups, so each group of students must be relatively homogeneous. If one of the two groups has had more experience with a second language, the groups are not homogeneous. Another important characteristic is language learning experience. For example, participants in ESL settings come from different countries. Each country has its own programs for English learning that yield differences in prior English experience among the participants. Researchers have to consider participants' prior language learning experiences in their home country. The other important characteristic can be proficiency level. Even if the participants are in the same-level classes, there are differences among the participants in terms of their individual strengths. For example, some participants are good at speaking and others are good at reading.

Regarding participant mortality, in longitudinal studies, it is difficult to control participants' attendance at all sessions observed by researchers. Researchers must determine how to deal with this situation.

Participant inattention and attitude also need to be taken into account. There is the possibility that the experimentation itself may affect the results through the Hawthorne effect and the halo effect, the boredom and fatigue of participants, and the practice effects of the test material. The Hawthorne effect is described by Mackey and Gass (2005) as follows:

One factor that might affect participant behavior is what is known as the Hawthorne effect, which refers to the positive impact that may occur simply because participants know that they are part of an experiment and are, therefore, "different" from others. (114)

They explained the halo effect as follows: "Participants may also try to please the researcher by giving the answers or responses they think are expected" (114).

Fatigue and boredom among participants are the results of participants' labor when they are asked to perform a test or a task. Inattention is an element that can affect participants' attention. For example, an experiment was conducted

during the week of the participants' examination period. In this circumstance, participants could be inattentive to the experiment.

Participants' maturation is a particularly important concern in longitudinal studies. At the outset of the study, the participants' abilities in the second language could be equivalent, although the progress of each participant's ability might be different. When maturation is considered, a control group that does not participate in the experimental intervention is needed.

Data collection involves a location where the data are collected and a collector who gathers the data. For example, if one of the two groups is given a test in uncomfortable circumstances and the other is not it may skew the data. The environment of participants could affect the results of the studies. This issue of data collection was explained by Mackey and Gass as follows: "One could imagine different results depending on whether or not the interviewer is a member of the native culture or speaks the native language" (115). The last factor of internal validity was not mentioned here because it is the consideration of experimental studies.

Aspects of internal validity were explained above. Now, the aspects of external validity will be explained. External validity refers to sampling, representativeness and generalizability, and collecting information.

First, sampling can be divided into two methods: random sampling and non-random sampling. According to Mackey and Gass (2005), random sampling is as follows:

Random sampling refers to the selection of participants from the general population that the sample will represent. In most second language studies, the population is the group of all language learners, perhaps in a particular context. Quite clearly, second language researchers do not have access to the entire population (e.g., all learners of Spanish at U.S. universities), so they have to select an accessible sample that is representative of the entire population. (119)

The authors also explained non-random sampling as follows:

Nonrandom sampling methods are also common in second language research. Common nonrandom methods include systematic, convenience, and purposive sampling. (122)

Systematic sampling in the citation above focuses on each individual. In this case, researchers have to be sure that participants are truly random. Convenience sampling is the selection of participants who happen to be available for the study. According to Mackey and Gass, convenience sampling is quite common in second language research. Purposive sampling is the method in which researchers knowingly select participants to elicit the phenomenon that is the main focus of the study. Fraenkel and Wallen (1996) explained that random sampling was not a common approach in education research:

In the vast majority of studies that have been done in education, random samples have not been used. There seem to be two reasons for this. First, there may be insufficient awareness on the part of educational researchers of the hazards involved in generalizing when one does not have a random sample. Second, in many studies it is simply not feasible for a researcher to invest the time, money, or other resources necessary to obtain a random sample. (108)

Mackey and Gass (2005) commented on representativeness and generalizability of research as follows:

If researchers want the results of a particular study to be generalizable, it is incumbent upon them to make an argument about the representativeness of the sample. (123)

Fraenkel and Wallen (1996) propose that researchers should adjust the number of subjects when they want their studies to be generalizable:

A recommended minimum number of subjects is 100 for a descriptive study, 50 for a correlational study, and 30 in each group for experimental and causal comparative studies. (111)

Collecting information about participants is one important aspect of research. It allows readers to determine whether the results are acceptable for their research. According to Mackey and Gass, participants' information has to be balanced between two concerns:

The first is the privacy and anonymity of the participants; the second is the need to report sufficient data about the participants to allow future researchers to both evaluate and replicate the study. (124-126)

According to Mackey and Gass, there are some types of validity that influence both internal and external validity, such as validities of content, face, construct, and criterion-related validities. First, Mackey and Gass explained content validity as follows:

Content validity refers to the representativeness of our measurement regarding the phenomenon about which we want information. If we are interested in the acquisition of relative clauses in general and plan to present learners with an acceptability judgment task, we need to make sure that all relative clause types are included. (107)

Second, they described face validity as follows:

Face validity is closely related to the notion of content validity and refers to the familiarity of our instrument and how easy it is to convince others that there is content validity to it. If, for example, learners are presented with reasoning tasks to carry out in an experiment and are already familiar with these sorts of tasks because they have carried them out in their classrooms, we can say that the task has face validity for learners. (107)

Third, construct validity is explained as follows:

Construct validity is an essential topic in second language acquisition research precisely because many of the variables investigated are not easily or directly defined. In second language research, variables such as language proficiency, aptitude, exposure to input, and

linguistic representations are of interest. . . In research, construct validity refers to the degree to which the research adequately captures the construct of interest. (107)

Finally, criterion-related validity is defined as follows:

Criterion-related validity refers to the extent to which tests used in a research study are comparable to other well-established tests of the construct in question. (107)

The validity types have been discussed above. Likewise, researchers need to consider the reliability of their research. Mackey and Gass stated the simple definition of reliability as follows:

Reliability in its simplest definition refers to consistency, often meaning instrument consistency. For example, one could ask whether an individual who takes a particular test would get a similar score on two administrations of the same test. If a person takes a written driving test and receives a high score, it would be expected that the individual would also receive a high score if she or he took the same written test again. (128)

Rater reliability refers to whether a judgment concluded by a researcher is objective and consistent. They argued that there are two types of reliability: rater reliability and instrument reliability. In this section, instrument reliability is not mentioned because the present study does not deal with test instruments.

Rater reliability can generally be divided into inter-rater and intra-rater reliability. Inter-rater reliability involves having two or more judges who judge the same set of data. Intra-rater reliability is a method that uses only one researcher who judges data consistently. For example, one researcher judges the same data at different times (e.g., morning and evening) and ensures his or her judgment is consistent. Frick and Summel (1978) explained that there are a variety of methods for testing reliability of classroom observation instruments. The simplest approach is to have several raters check the same instruments' items,

and calculate the ratio of agreement between one rater and the other rater.

E. Summary.

In the first section, previous studies concerning teacher corrections were discussed in chronological order. Research on teacher corrections was inspired by first language research. Some researchers who investigated first language acquisition demonstrated that parents correct ill-formed sentences produced by their children. According to their observations, the parents used a variety of corrections, so researchers labeled these corrective techniques as “recast” or “request clarifications.” Thus, second language researchers assumed that these corrective techniques would also be beneficial for second language learners. One of the earlier studies on corrective feedback in second language classrooms was conducted by Fanselow (1977). The earlier studies succeeded in reporting what could be expected in actual classrooms; however, the effectiveness of corrections was still unknown. The later studies emphasized the effectiveness of corrections in second language acquisition. For example, Long (1996), Doughty and Varela (1998), and Ammar and Spada (2006) reported that teacher feedback facilitates L2 development.

In psychology, there was an argument that the provision of corrective feedback was not sufficient to establish its usefulness in language acquisition. This issue also spread through the field of second language acquisition and researchers began to investigate favorable corrective techniques. This line of research led to specific categorizations for teacher corrections. For example, Ellis (2006) distinguished between two categories of corrective techniques: input-providing corrective feedback and output-pushing corrective feedback. Lyster and Mori (2006) also made distinctions among corrective feedback types, such as recasts, prompts, and explicit corrections. According to Ellis (2006), recasts and explicit corrections are input-providing types and prompts are output-pushing types. In recent studies in teacher correction research, these input

types and output types have been compared to identify the most effective corrective technique.

In the second section, the research methods used in previous studies were discussed. There are several research methods, and it can be difficult to make clear distinctions among them. However, the distinctions made by some scholars were mentioned in this section. First, Chaudron (1988) illustrated four types of methods: psychometric, interaction analysis, discourse analysis, and ethnographic. Second, Long (1980) described two basic approaches to classroom research: interaction analysis and anthropological observation (each approach has subcategories). Third, Fraenkel and Wallen (1996) explained research methodologies from an educational point of view, so their distinction was more universal than that of Chaudron or Long. Fraenkel and Wallen described three research types: descriptive, associational, and intervention (each approach has subcategories).

In the third section, the classifications used in previous studies were discussed. To analyze teacher-student conversations, the following three parts were examined: 1) initial turn, 2) response, and 3) reaction.

The last section focused on the issues of reliability and validity in the previous studies. Reliability and validity are important aspects to consider when assessing whether appropriate procedures have been taken in a study. Validity has two main types: internal validity and external validity. Internal validity refers to participant characteristics, participant morality, participant inattention and attitude, participant maturation, data collection, and instrumentation and test effects. External validity refers to sampling, representativeness and generalizability, and collecting information.

Reliability refers to the consistency of a study and this measure can be divided into rater reliability and instrument reliability. Rater reliability considers whether a judgment made by a researcher is objective and consistent. Instrument reliability is not mentioned here because instrument reliability relates to test

instruments, which the present study did not use.

The reviews of the above studies present several important research questions to the author who is interested in the corrective feedback among Japanese learners of English in Japan. Those research questions and the study method will be described in the following chapter.

Chapter III

Method of Study

In this chapter, the procedures taken in this study are presented. The procedures adopted in this study referred to the methods used in the previous studies. Some studies that were mentioned in the previous chapter will be explored with more specific details in this chapter. In research design, five important aspects should be taken into account: 1) research questions, 2) research procedures, 3) data collection, 4) instruments, and 5) reliability and validity.

A. Research Questions

Five research questions are presented in this study. The first research question is whether a teacher uses various types of corrective feedback. To investigate this research question, I observed two different proficiency classes: one was an advanced-level class (class A) and the other was an average-level class (class B). Both classes were taught by a native English speaker, called Teacher A. in this study. Teacher A was observed to ascertain if there were any differences in his use of corrections between the higher-proficiency class and the lower-proficiency class. The two research models that are presented below were used to answer this research question. As for the first research model, the independent variables were the learners' English proficiency levels and the dependent variables were the frequencies of teacher feedback. In the second research model, the independent variables were the learners' English proficiency levels and the dependent variables were the types of teacher feedback.

The second research question is whether the two teachers' tendencies for the use of feedback are different even though they are teaching learners at the

same level. Two teachers were observed in order to examine the second research question. Teacher A was an experienced teacher and Teacher B was not as experienced. Both teachers taught the advanced-level classes (class A for Teacher A and class C for Teacher B). Two research models presented below were adopted to assess the second research question. Regarding the first research model, the independent variables were the teachers and the dependent variables were the frequencies of teacher feedback. In the second model, the independent variables were the teachers and the dependent variables were the types of teacher feedback.

The third research question asks which learners' errors are focused on by the teachers. The same procedure was followed as for the second research question. In this research model, the two teachers who taught the learners in the advanced classes were observed and compared. Concerning the first research model, the independent variables were the learners' error types and the dependent variables were the frequencies of Teacher A's feedback. Subsequently, Teacher B was observed as well as Teacher A to determine his priority for the error types.

The fourth research question is whether the frequencies of learners' uptake and repairs following teacher corrections are different according to the learners' proficiency levels. The average-level class (class B) and the advanced-level class (class A) were observed so as to determine the frequencies of uptake and repair in each class. As I mentioned above, both classes were taught by Teacher A and the material progressions of the two classes were also the same. Regarding the research model for this question, the independent variables were the learners' proficiency levels and the dependent variables were the frequencies of uptake and repair following the teacher corrections. The last research question was investigated using the same research model as the fourth research question.

B. Research Procedures

To examine the first research question (whether one teacher has variations in his feedback use) the present study employed the research procedures shown

below.

1. Two different-level classes (class A and B) were observed and recorded.
2. Only subsets of teacher-learner interactions of the recorded materials were transcribed by the first rater.
3. The transcribed data were categorized into the instrument presented in the next section. All the transcribed data were judged by the first rater. Next, 20% of the data were checked by the second rater. The ratio of agreement between the two raters was counted.
4. The frequencies of the independent and the dependent variables were counted.
5. Only 200 samples (100¹² samples from class A and 100 samples from class B) were randomly selected and the frequencies of the teacher feedback were counted.
6. Finally, the data were tested by a chi-square analysis.

The two teachers (Teacher A and Teacher B) were observed to test the second research question, namely whether there are any differences between the two teachers in the use of feedback when teaching learners at the same level. The present study followed the procedures outlined below.

1. The two advanced-level classes (class A and C) were observed and recorded.
2. The second procedure was the same as the first research question.
3. The transcribed data were categorized into the same classification as the first research question.
4. The frequencies of the independent and the dependent variables were counted.
5. Only 200 samples (100 samples from class A and 100 samples from class C) were randomly chosen and the frequencies of teacher feedback were counted.
6. The frequencies were analyzed using a chi-square analysis.

To implement the third research question (what kind of learner error tends to attract the teachers' attention) the same procedures were taken as for research question two.

¹² The present author used the number 100, following Oliver (1998).

1. The two same proficiency classes (class A and class C) were recorded.
2. The second procedure was the same as the first research question.
3. The transcribed data were categorized similar to the first research question.
4. Two hundreds samples (100 samples from class A and 100 samples from class C) were randomly selected and the frequencies of the two teachers' feedback were counted.
5. The samples were tested by Kendall's rank correlation coefficient.

The fourth research question is whether the frequencies of learners' uptake and repairs following teacher corrections are different according to the learners' proficiency levels. The two classes of the different English levels were observed. In addition, the same procedures were taken to investigate the last research question.

C. Data Collection

The data presented in this study was derived from three classrooms in a high school in Kanagawa, Japan. The participants were first-year students in high school and were Japanese native speakers studying English as a foreign language. Three different classes were observed: two of them were the advanced classes (class A and C) and one was an average class (class B). The levels of classes are determined by the results of an entrance examination in March of each year. The observations were made from September to November in 2011 although the data from class C were collected in 2007. Thirty seven students participated in class A, 35 in class B, and 25 in class C. The oral communication classes were observed in order to examine the research questions. In general, oral communication classes in Japan are different from those offered in other countries. They focus not only on content but also on grammar items or idiomatic phrases while content is the main focus in many other countries. The oral communication lessons were held twice a week for 50 minutes each. Consequently, all the students learned English communication for an hour and 40 minutes each week. In addition to the oral

communication classes, they had to take English grammar classes. Classes A and B used the same textbook and the same teaching materials. However, students in the advanced class A usually had a little more vocabulary and phrases than those in class B. All the classes proceeded similarly according to the curriculum of the school. Class C is also an advanced class like class A, but class C was organized in 2007, and a different textbook was used at that time.

Two teachers participated in the present study. Both of them were native English speakers and were males. Teacher A is from England and taught class A and class B. He had been working in the high school for seven years at the time of the study. Native speakers who have been working in a Japanese high school for seven years are considered experienced teachers. Before that, he was an English teacher at an English language school in Japan and he could not speak Japanese. Teacher B is from Australia, taught class C, and he had been teaching in the high school for half a year at the time of the study. He had been an English teacher at an English language school in Japan before working at the high school. He was bilingual in English and Japanese. In general, oral communication classes were team-taught by a Japanese teacher and an English native speaker. Generally speaking, English native speakers speak only in English in their classes. Hence, Japanese teachers have to support the English native speakers when English conversations cannot proceed. In the present study, Teacher A was paired with two Japanese female teachers (one teacher for class A and the other teacher for class B). Teacher B was paired with a Japanese male teacher.

The data were recorded with an audio recorder and transcribed into written materials. Seven lessons were recorded in each class; thus, 21 lessons were recorded in total. Each lesson lasted 50 minutes; the data totaled 17.5 hours in all. The written materials were analyzed using the procedures stated above and the instruments presented below. Furthermore, the validity and reliability of the present study are discussed in the last section of this chapter.

D. Instruments

As I have explained in the previous chapter, each study had its own categorization for learners' errors even though there were some common items, namely grammatical and phonological items. Moreover, lexical items were also common. Therefore in this study, three of these learners' error types were adopted, and one additional error categorization of "interaction and discourse" was added to them. This addition was made because classroom conversations contain not only linguistic deviations but also interactional and discourse-level deviations, as proposed by Chaudron (1977). Furthermore, in my study the students' utterances that are considered inappropriate for a given situation or context were also categorized as an error of interaction and discourse, even though Chaudron did not deal with the phenomenon of inappropriate utterances.

When classifying learner errors, another problem arose, namely that some errors could be interpreted as both a phonological and a lexical error. In that case, the classification criteria for learner errors that are illustrated on pages 19-21 in the previous chapter were used.

To classify teacher corrections, the present study referred to the classification of teacher feedback used by Lyster and Ranta (1997) and Lyster and Mori (2006) because of their clear definitions. In addition to their six types of teacher feedback, a new type of correction, namely "compulsion" Oba (2009) was adopted in this study, and the examples of compulsion were included in the category of "prompt." Consequently, in this study there were seven types of teacher feedback: recast, explicit correction, clarification request, meta-linguistic clue, elicitation, repetition, and compulsion.

The classification for learners' reactions toward teachers' corrective feedback was considered as well. In the present analysis, the terms "uptake" and "no uptake" presented in Lyster and Ranta (1997) were adopted, so their subcategories of "needs-repair" and "repair" were also used. However, some of

their terms were not adopted in this study, namely incorporations and peer-repairs. Also the subcategories of “needs-repair” were not used in the present study because there was no requirement to classify the learners’ needs-repair.

E. Reliability and Validity

In this section, validity and reliability are considered. First, it is said that there are two main types of validity in general: internal validity and external validity. Regarding internal validity, one element of participant characteristics is language learning background. The present study included participants who belonged to already existing classes. All of the participants were in foreign language learning settings, so they were learning English not only in the national education program but also in the local program determined by the high school. The participants in the present study were first-year students in the high school. Thus, they were placed into average-level classes and advanced-level classes according to their scores on the entrance examinations in March. From this point of view, the learners in this foreign language setting possess a similar language background.

Internal validities have been discussed above. Now, external validity has to be considered. External validity refers to sampling, representativeness and generalizability, and collecting information. First, sampling can be random or non-random. In this study, however, random sampling was not realistically possible. Thus, a non-random sampling method was adopted to observe the participants. The total number of participants was 98 (38 learners in group A, 35 in group B, and 25 in group C), plus two teachers. Each group was limited in the number of participants so that non-random sampling was achieved. For data analyses, chi-square analyses were used.

The validity of the present study was discussed above, and will now be followed by a discussion of the inter-rater reliability. In the present study, the inter-rater method was adopted so that the transcribed data could be judged

between the two raters and the ratios of agreement could be calculated. All the transcribed data were categorized by the first rater. Then, 20% of the entire data set was checked by the second rater and the ratio of agreement was more than 90%.

Chapter IV

Results

To answer the five research questions presented in the previous chapter, the results of the data are illustrated in this chapter. The oral communication classes were audio recorded and transcribed into written materials in order to investigate the tendencies of teacher corrective feedback. The results of the study were examined by a chi-square analysis and correlation analysis. A significance of 0.05 was required for all tests.

A. Results from Present Data.

Table 1 shows the total numbers of error corrections provided by the teacher who undertook the two different-level classes. The frequency of teacher feedback in the advanced class is 223 while that of the average class is 626. To see if there was any correlation between the class types and the amount of feedback, the chi-square test was used.

Table 1. Frequency of teacher feedback.

Advanced class	Average class
223	626

The results of the test ($\chi^2 = 191.3$, $p < .001$) show there is a significant relation between the class types and the amount of teacher feedback. The frequencies observed in the two different-level classes are shown in Figure 1 below.

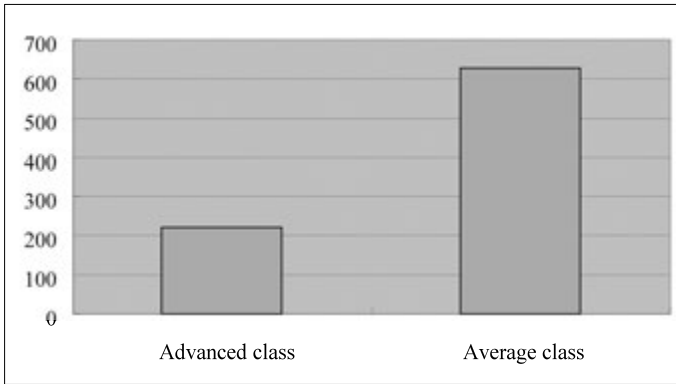


Figure 1. Frequency of error corrections in two different-level classes.

Table 2. Distribution of feedback types.

	Advanced class	Average class
Recasts	79	70
Prompts	4	3
Explicit corrections	17	27

More specific details are illustrated in Table 2. The frequencies of the error corrections and the types of corrections are shown. One hundred samples were selected at random from each group. In both classes, the largest category is recast, which accounts for 79 in the advanced class and 70 in the average class. The other categories are illustrated in decreasing frequency as follows: explicit correction (17 in the advanced class and 27 in the average class) and prompt (four in the advanced class and three in the average class). To examine the correlation between the correction types and class types, the chi-square analysis is adopted;

however, the results show no statistically significant difference between the two groups ($\chi^2 = 2.96$, $p > .05$). Despite this non-significant difference, the tendencies of the frequencies are presented in Figure 2.

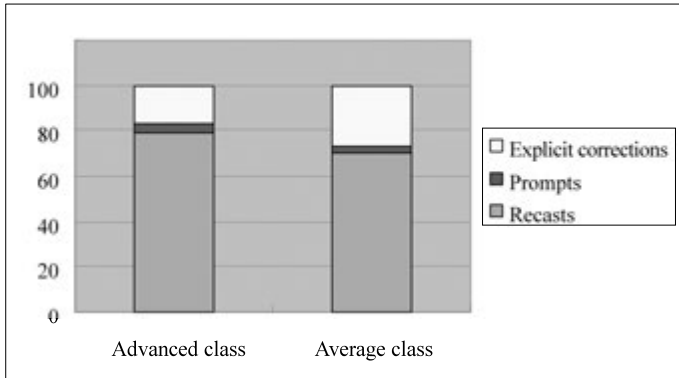


Figure 2. Percentage distribution of feedback types in the two different-level classes.

The total numbers of the two teachers' corrections are displayed in Table 3. The two teachers taught participants of the same age with the same English level. The teacher who taught class A corrected the learners' errors more than the other teacher (223 vs. 168) even though the learners in both classes were at the same English level. The frequencies derived from the two groups were analyzed using the chi-square test ($\chi^2 = 7.73$, $p < .005$) and are shown as significantly different. The frequencies of the teachers' corrective feedback are also displayed in Figure 3 below.

Table 3. Frequency of teacher feedback.

Advanced class A	Advanced class C
223	168

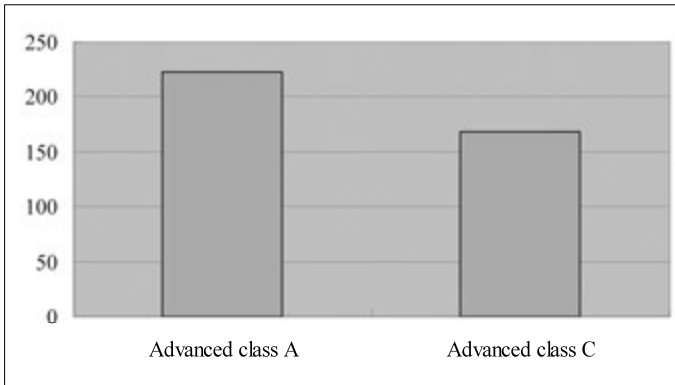


Figure 3. Frequency of error correction in the two same-level classes.

Table 4 indicates the correlation between the correction types and the group types. One hundred samples were selected randomly from class A and class C. For both teachers, the largest category is recast, which is 79 for class A and 70 for class C. The second largest category is explicit correction (17) in class A and prompt (16) in class C. The least frequent type is prompt (4) in class A and explicit correction (14) in class C. As for the chi-square analysis, the correlation between the correction types and the class types are significantly different ($\chi^2 = 8.03$, $p < .05$). Nevertheless, when each feedback category was tested by the chi-square test, recasts and explicit corrections in the two classes did not show a significant difference. Only prompts showed a significant difference ($\chi^2 = 8.00$, $p < .005$).

Table 4. Distribution of feedback types.

	Advanced class A	Advanced class C
Recasts	79	70
Prompts	4	16
Explicit corrections	17	14

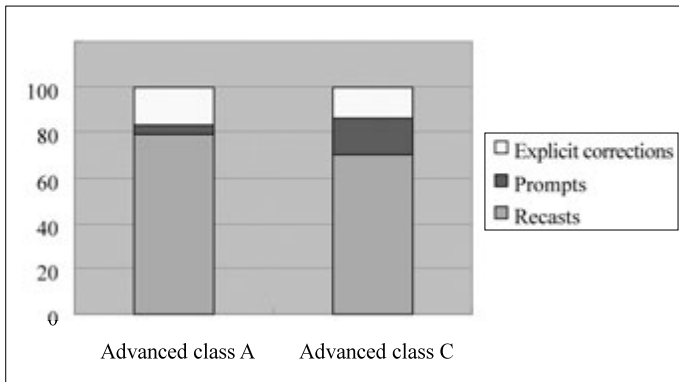


Figure 4. Distribution of feedback types in the two same-level classes.

The two teachers' preferences for different error types are displayed in Tables 5 and 6. The tendency for Teacher A to correct the learners' phonological errors more frequently can be observed in Figure 5. The other types are shown in decreasing frequency as follows: interaction and discourse, lexical, and grammatical. Unlike Teacher A, Teacher B had the tendency to correct the learners' interaction and discourse errors. The other types are displayed in decreasing frequency as follows: lexical, grammatical, and phonological. The two teachers had completely different priorities for the choice of error types (see Figure 5). The two teachers' different priorities were tested by Kendall's rank

correlation coefficient ($\tau = -.167$, $p > .05$) to see if there were any rank order correlations between the teachers; however, the results were not statistically verified.

Table 5. Distribution of error types (Teacher A).

	Phonological	Lexical	Grammatical	Interaction and discourse
The frequency of teacher feedback	70	3	1	26

Table 6. Distribution of error types (Teacher B).

	Phonological	Lexical	Grammatical	Interaction and discourse
The frequency of teacher feedback	14	16	15	55

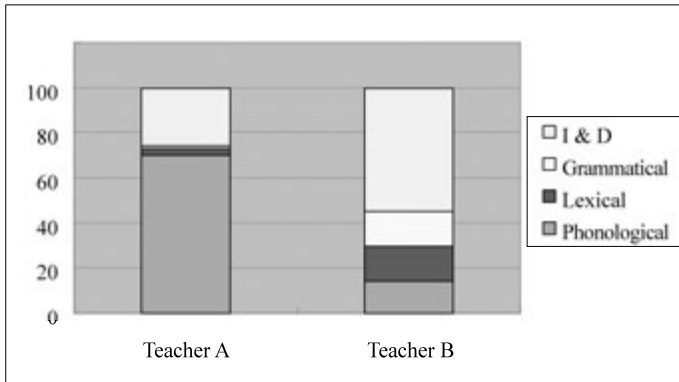


Figure 5. Distribution of error types.

The frequencies of uptake, repair, and no uptake recorded in the advanced and the average classes are displayed in Table 7.

Table 7. Uptake, repair, and no uptake following teacher feedback.

	Advanced class	Average class
Uptake	93	91
Repair	36/ 93	47/ 91
No uptake	7	9
Corrective feedback	100	100

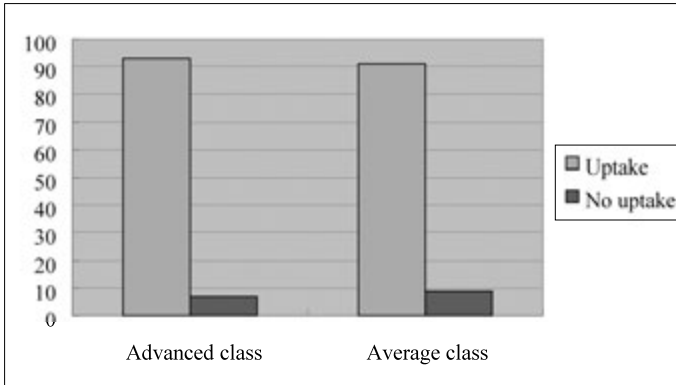


Figure 6. Distribution of uptake and no uptake.

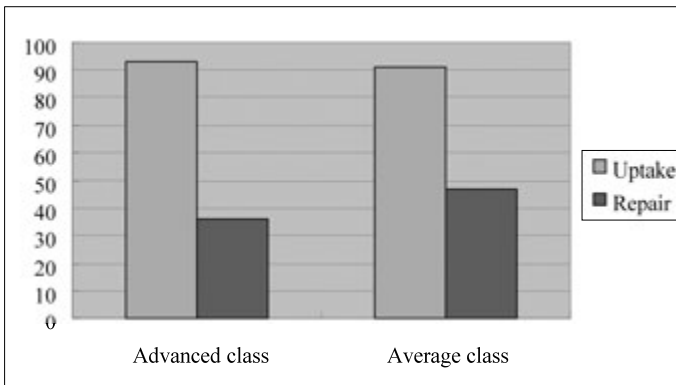


Figure 7. Distribution of uptake and repair.

In the case of the advanced class A, 223 erroneous utterances were corrected by the teacher. Only 100 samples were selected at random from 223 erroneous utterances and are presented in Table 7 and Figures 6 and 7. Figure 6 shows that 93 uptaken utterances were produced by the learners, and Figure 7 indicates 36 repairs derived from these uptaken utterances.

In the average class B, although there were 639 erroneous utterances that

were corrected by the teacher, only 100 utterances were sampled and are shown in Table 7 and Figures 6 and 7. Figure 6 indicates that 91 uptaken utterances are derived from 100 erroneous utterances, and Figure 7 shows that the uptake led the learners to repair 47 utterances. The frequencies of uptake, no uptake, and repair in the two groups were examined by the chi-square analysis ($\chi^2 = 1.30$, $p > .05$). The value of chi shows no significant difference in the two classes.

B. Summary.

The first research question asks whether one teacher uses various types of feedback. To answer this research question, the present author examined the advanced class and the average one that were both taught by the same teacher. Three points were examined: 1) the frequencies of the error corrections in the two classes, 2) the types of corrections provided in the two classes, and 3) the frequencies in use of each feedback type (the third point is investigated in case of any differences found for the second point).

Regarding the first point, Teacher A in the average-level class provided three times as many corrections for errors as he did in the advanced class. Statistically, a significant difference ($\chi^2 = 191.3$, $p < .001$) was found between the two classes. Consequently, the result was interpreted as showing that one teacher varied the frequencies of his corrections, according to the level of the English learners. Concerning the second point, a significant statistical difference in the teacher's correction types could not be supported while he was teaching the different-level groups. Given the non-statistical support, a further investigation for the third point was not conducted.

The second research question addresses whether there are any differences between the two teachers in the use of feedback in teaching learners at the same level. To research the question, two teachers who taught the same-level classes were observed and recorded. Three points were examined for this question: 1) the two teachers' frequencies for corrections, 2) the types of corrective feedback used

by the two different teachers, and 3) the frequencies in the use of each feedback type (the third point is investigated in case of any differences found for the second point).

Statistically, there were significant differences ($\chi^2 = 7.73$, $p < .005$) in the two teachers' frequencies of corrections. From the results of the first point, it was concluded that in statistics, the two teachers' frequencies for error corrections were different. Examination of the second point also statistically verified that there were significant differences between the two teachers in the types of corrective feedback. As for the third point, it could be said that there were some differences in the use of prompts between the two teachers. However, the two teachers' different tendencies to use recasts and explicit corrections could not be statistically supported. In the case of those two feedback types, the results of the present study may not be generalized to the other cases.

The third research question addresses what kinds of learner errors tend to attract attention from the teachers. Two teachers undertaking the same-level classes were observed and the learners' four error types were aligned in the orders of frequency. A correlation between the two teachers' priorities for the four error types was not statistically supported; thus, it was concluded that further investigation will be needed for this issue.

The fourth research question asks whether the frequencies of learners' uptake and repair following teachers' corrections might be different according to the learners' proficiency levels. The advanced class and the average one were observed in order to see if there were any differences in the frequencies of uptake and repair. Statistically, significantly different tendencies in uptaken utterances and repaired utterances produced by the two classes could not be verified, so further considerations will be needed for this topic.

The fifth research question assesses whether the teacher corrections are beneficial to learners who study English in Japan. Statistics were not used for this topic, although in the present study both the higher- and lower-level classes had

high rates of uptake and repair. Uptake in this study refers to a student's utterance that immediately follows a teacher's correction, and repairs mean that a learner can correct his or her previous incorrect utterance. Thus, uptake and repair are the learner's immediate reactions to the teacher's corrections. A large number of the learners' uptake and repairs could be understood as partly beneficial to the learners who participated in this study; although, uptake and repair only indicate a short-term effect.

Chapter V

Discussions and Conclusions

The results obtained from the present analysis address the five research questions presented in the previous chapter. The discussions in this chapter emerge from the findings of these research questions. The discussions address the following four topics: teacher factors, one teacher factors, learner factors, and efficacy of error corrections for English learners in Japan. The former three topics are discussed in order to indicate what may strongly affect the teacher corrections. The final topic relates to the last research question, namely whether teacher corrections are beneficial to learners who study English in Japan. At the end of the chapter, the conclusions derived from these discussions will be presented.

A. Teacher Factors.

This discussion offers an interpretation for the findings from research questions two and three concerning the differences between the two teachers. The first finding from my observation was that the frequencies of feedback provided by the two teachers were different. The same finding was reported in the study conducted by Lyster and Ranta (1997). In their study, among the four teachers, the highest frequency of feedback was 244 and the lowest was 103, and the ratios of feedback among the four teachers were different. The authors state “We do know, however, that teachers provided feedback on 62% of the student turns with errors — this varied roughly from 50% to 70% for the four individuals” (56). According to the present study, the two teachers provided corrections for their learners at different frequencies, which were statistically different. In the case of the feedback frequency, the findings of the previous studies were supported.

In addition to the teachers' feedback frequencies, the relationship of the feedback types between the two teachers was examined. The results indicated that the two teachers' tendencies in the use of recasts were not significantly different nor was their use of explicit corrections. In the case of explicit corrections, the same finding was presented from other studies as well (Lyster and Ranta 37-66; Lyster and Mori 269-300). In my study, a difference was found between the two teachers in their use of prompts. Overall, few differences were found in feedback types between the two teachers, and previous researchers have reported the same phenomenon as the present study. For example, Lyster (1998) revealed that the four teachers in his study had a common tendency in their feedback use as follows:

The findings suggest that the 4 teachers provided corrective feedback somewhat more consistently and less randomly than teachers observed in previous studies (e.g., Allen et al., 1990; Fanselow, 1977). (205)

Unlike the finding noted by Lyster (1998) and the present observation, some researchers have concluded that each teacher has a different characteristic when choosing feedback types. For example, Fanselow (1977) found 16 types of teachers' responses to learners' errors, and the commonality among the 11 teachers was one tactic. Fanselow expressed this finding as "The similarity of behavior among the teachers did not provide as much insight into the treatment of errors as was hoped for" (583). Chaudron (1977) also presented a similar finding:

The differences between teachers can readily be observed, and variations for any one teacher will also be recognized. This is to say that any use of the present model in actual observation will have to take into account individual teachers' "basal" features or types of corrections. (44)

One factor that caused a contradiction between the results of my study and those of Fanselow and Chaudron might be the observational environment. In the present study, the two teachers taught the same level of English learners with the same

age groups, and they conducted their English lessons according to the school curriculum. From that point of view, the two teachers were teaching English in the same context. The results of the present study might be influenced by the fact that the teachers were observed in the same teaching settings.

There is one more finding concerning the relationship of the feedback types between the two teachers. In my study, the use of prompts between the two teachers was statistically different, and a similar result has been reported by previous scholars. For example, Lyster and Mori (2006) investigated whether the occurrences of prompts were different in Japanese immersion and French immersion contexts. They found the following:

The next highest proportion of feedback moves in both settings was attributed to prompts, which comprised 38% and 26% of all feedback moves in FI and JI classrooms, respectively. (284)

In addition, Lyster and Ranta (1997) found that one of the four teachers used more prompts than the other teachers. Given that the present study supported the findings of the previous studies, the occurrences of prompts might be a fluctuating factor in classroom research.

Priorities for the four error types by the two teachers were also examined. The correlation between the two teachers' priorities for the four error types was not verified. Nevertheless, it is clear that in my study, the apparent tendency of Teacher A was to exclusively attend to the learners' pronunciation. The rate of his corrections toward phonological errors was 70% of all the corrections. His attention to phonological errors was unparalleled compared to the other types of errors. A previous study conducted by Nicolas, Lightbown, and Spada (2001) emphasized that pronunciation faults may cause a serious problem in communication. The authors write:

In the ESL contexts, in which phonological errors were relatively frequent and learners came from many different L1 backgrounds, communication was genuinely disrupted when learners'

pronunciation was faulty. (736)

If, as Nicolas et al. claim, pronunciation was the only reason for the phenomenon, Teacher B would also have focused on the learners' pronunciation in the same way as Teacher A. Fanselow (1977) offered another explanation why some teachers focus on particular errors while others do not. The teacher's strong focus on the learners' pronunciation could be interpreted as 1) the learners' pronunciation faults were likely to cause problems in communication, or 2) the teachers had their own priorities for correcting learners' error types.

While Teacher B dealt equally with any type of error, Teacher A strongly focused on the learners' pronunciation. There was no tendency toward any significant prioritizing for Teacher B, although he tended to focus on interaction and discourse types containing L1 items and incomplete utterances. Many researchers indicated that teachers tend to focus on the learners' errors that might cause an interruption in communication. Burt and Kiparsky (1972) classified students' errors into two categories: global errors and local errors. Global errors are those that cause a listener or reader to misunderstand a message or consider a sentence incomprehensible. On the other hand, "local errors affect single elements in a sentence" (Ellis 712). Hendrickson (1978) commented on this classification as follows:

On the basis of how errors affect the comprehensibility of whole sentences, one could build a local-to-global hierarchy of errors that would potentially guide teachers to correct students' mistakes (Burt 1971, Burt and Kiparsky 1972, and Valdman 1975). (391)

In the study by Hughes and Lascaratou (1982), 10 native speakers of English who were not teachers, 10 teachers who were native speakers of English, and 10 non-native-speaker teachers (Greek teachers) were observed. The authors found that the native speakers used intelligibility as a criterion for error evaluations more frequently than the non-native speakers:

Greek teachers made reference to the 'basicness' of the rules

infringed, while the non-teachers depended almost exclusively on the criterion of intelligibility. The English teachers made use of both criteria, but showed some preference for that of intelligibility. (175)

Similar to the previous studies, the present study found that Teacher B tended to focus on the learners' errors that might cause sentences to be incomprehensible. The present author labeled this type of error as "interaction and discourse," following Chaudron (1977), and found that not only Teacher B, but also Teacher A had a high priority for interaction and discourse errors. Thus, the results of this study supported the results from previous studies.

B. One-Teacher Factors.

In this section, research questions one and four, which consider the issue of one teacher's feedback use, will be discussed. To investigate one teacher's variety in feedback use, Teacher A, who participated in the two different-level classes, was observed. The results of the present study revealed two main discoveries: 1) Teacher A changed the frequencies of error corrections depending on the language levels of the learners, and 2) statistically, it was not supported that Teacher A changed his feedback types according to the English proficiency levels of the learners.

Some previous studies have found that a single person can vary his or her use of language according to various situations (e.g., Ishiguro, 1988). For example, Ferguson (1971) noted that native speakers employ a copula "is" in equational clauses in normal communication, although they often omitted it when talking with foreigners. Moreover, a number of previous studies investigated whether native speakers or native-speaker teachers adjust their language use depending on learners' language levels. Hatch (1983) found that native speakers sought the learners' current language stages until they found an appropriate level. My investigation found that the lower-level English class received more corrections than the higher-level English class. In the classrooms, one teacher

adjusted his provision of feedback according to the level of the learners.

Lyster and Ranta (1997) reported that the highest-level class tended to receive more output-pushing feedback types than the other classes:

For example, T3 — whose students have a higher degree of proficiency due to their more intensive and longer exposure to French — uses recasts considerably less than T4, T5, and T6 (39% vs. 66%, 60%, and 68%, respectively). This allows her to draw more on other feedback types and, in particular, on those that are more likely to lead to uptake. Indeed, 70% of T3's feedback turns led to student uptake, whereas 43%, 47%, and 50% of the feedback turns of T4, T5, and T6, respectively, led to uptake. (56)

They interpreted their discovery in the following way:

Thus, given her students' higher level of proficiency, T3 is able to push students more in their output and to rely less on the modeling techniques (i.e., recasts with infrequent uptake) used by the other teachers with less advanced students. (57)

Given the findings of the previous study, it was predicted that one teacher would vary his feedback use, depending on the learners' language levels. According to the present study, the teacher undertaking both the advanced and the average classes changed the frequencies of corrections in the two classes. However, in this case, the teacher's feedback types were not significantly different even though he was teaching different-levels of English learners. In my study, there was no statistical support for this issue, so the result cannot be applied to other cases. However, the factors that resulted in such a discrepancy between the previous study and the present study might be age and the language learning settings. The learners who participated in the previous study were elementary school students and those in my study were high school students. At times, it might be more difficult to make adolescent learners speak a foreign or a second language. For example, Swain (1985) illustrated the case of French immersion classes as

follows:

And my own informal observations indicate that most peer-peer interaction that is not teacher-directed is likely to occur in English rather than in French at this grade level. (246)

Tarone and Swain (1995) reported the following:

One of the most persistent and perplexing problems for researchers as well as for immersion teachers is the tendency of older immersion students not to use the second language (L2) in the classroom, particularly when conversing with each other. (166)

These problems can be seen not only in immersion settings but also in foreign language settings. My questionnaire survey conducted after the observations indicates the learners' non-preference for being forced to answer by their teacher. The output-pushing teaching style is less effective when teachers deal with adolescent learners.

The other explanation may be the different language learning settings. The participants in my observation were studying English in a foreign language setting so both the teacher and learners relied more on language itself. According to Nicolas et al. (2001), recasts, which are input-providing feedback types, were more beneficial in foreign language contexts:

The results of the classroom studies indicate that the classroom context (particularly the communicative and/or content-based classroom) may make it difficult for learners to identify recasts as feedback on form and hence difficult for them to benefit from the reformulation that recasts offer. The exception may be some foreign language classrooms in which students' and teachers' focus is more consistently on the language itself. (744)

Consequently, the current observation in this study and those from previous studies could be affected by the factors of age and the learning setting; although, further discussion will be needed on this topic.

C. Learner Factors.

In this section, research questions one and four, which relate to the learner factors, will be considered. To investigate this aspect, two different-level classes were observed. The results of the present study are as follows: 1) learners' proficiency levels affected the frequencies of teacher feedback; 2) despite non-statistical support, the higher-level learners tended to receive more recasts while the lower-level learners were likely to have more explicit corrections; and 3) despite non-statistical evidence, the number of uptake and repairs were not significantly different when learners at different proficiency levels were compared.

The most significant discovery was that the frequencies of feedback that occurred in the two different-level classes were different. The teacher provided feedback to the average class in approximately three times as many instances as he did in the advanced class. Statistically, there was a significant difference between those two classes.

The second finding was that the teacher used recasts, explicit corrections and prompts in similar ways between the classes of different English proficiency levels. Despite not finding any statistically significant difference, it was observed that the higher-level class tended to receive more recasts than the lower-level class. On the contrary, the lower-level class tended to receive more explicit corrections than the higher-level class. In particular, the phonological errors produced by the lower-level learners were likely to receive more explicit corrections than those produced by the higher-level learners. The study conducted by Ammar and Spada (2006) demonstrated that recasts were less effective for the low-proficient learners in their study. Their study was a quasi-experimental study and the target linguistic structures were grammatical. Therefore, their finding might apply to the results of the present study. In my study, recasts would be less effective than the other corrective techniques for the lower-level learners, so the

teacher might have assumed that explicit corrections would be better for the lower-proficiency students. This finding suggests that an explicit teaching style could be effective for lower-level learners.

The third discovery was that both level classes had a large number of uptake and repairs; thus, significant differences between the two groups could not be found in this study. Despite the non-statistical support, one explanation for this finding might be that the teacher undertaking both classes placed a higher priority on correcting the learners' phonological errors. Some researchers have found that learners' attention to teacher feedback was affected by the learners' error types. Mackey, Gass, and McDonough (2005) explained this as follows:

It may be so because, even when morphosyntactic feedback is provided in interaction, through recasts, learners often do not perceive it as such, whereas when phonological and lexical feedback is provided in interaction, they are more likely to perceive it correctly. (494)

Lyster (1998) also illustrated that recasts can easily lead learners to focus on phonological problems:

Indeed, the unusually high rate of repair following teachers' recasts of phonological errors suggested that students did indeed notice the corrective intentions underlying the teacher's recast, in that they tended to repeat it and get it right. (206)

The fact that the teacher had prioritized the phonological errors and the learners had recognized the teacher's corrections might be one explanation for such frequent uptake.

The second explanation for the frequent uptake produced by the learners was that the research was conducted in an EFL (English as a Foreign Language) context and the focus of the classrooms was on the language itself. The rates of uptake or repair might change according to classroom contexts. Lyster and Mori (2006) reported that the Japanese immersion classrooms in their study had a more

positive effect from the teacher corrections than that of the French immersion context. They write:

The total amount of uptake following feedback was higher in JI classrooms: 76% of all feedback moves were followed by learner uptake in JI classrooms compared to 55% in FI classrooms. (284)

In their study, the participants in Japanese immersion classes had a large number of uptake from recasts. Lyster and Mori interpreted this result as follows:

The JI example is thus more reminiscent of Seedhouse's (2004) description of form-and-accuracy contexts, in which a student is expected to speak accurately and repeat a teacher's recast as a means of discourse practice... which suggests that interaction was more tuned to linguistically accurate behavior in JI classrooms than in FI classrooms. (292)

Slimani (1992) observed grammar lesson classrooms for foreign language learners and concluded his study as follows:

Bringing particular linguistic features to the class's attention appears to be a rather valuable characteristic of uptake as most of the uptaken items were focused upon during instruction. (215)

Most notably, in the case of recasts, it has been claimed by some scholars that the effectiveness of recasts depends on the classroom context (a communication class vs. a grammar class or a foreign language context vs. a second language context). For example, Nicolas et al. (2001) summarized that the classroom context (the communicative and/or content-based classroom) may make it difficult for learners to identify recasts as feedback. Exceptions might be some foreign language classrooms in which students' and teachers' foci are more consistently on the language itself. In this study, the foci of the classes were on language itself; thus, high rates of uptake were recorded in both level classes. As Nicolas et al. have suggested above, the high rates of uptake or repair in my study could be understood as the learners maintaining their focus on the targeted linguistic

features, in which case the learners might be aware of the aim of the teacher's corrections.

D. Efficacy of Error Corrections for English Learners in Japan.

According to Ellis (2006), feedback can be distinguished into input-providing corrective feedback and output-pushing corrective feedback. Input-providing corrective feedback provides correct linguistic forms through recasts or explicit corrections. Output-pushing feedback does not provide correct forms, but forces learners to correct the utterances themselves through prompts.

With regard to the input-providing type, several researchers have found that recasts often do not lead to learners' uptake. This finding suggests that the learners in these studies had no conscious awareness of the corrected linguistic features. Lyster and Ranta (1997) discussed this phenomenon: "Our results indicated that almost 70% of all recasts did not lead to uptake" (56). They also added the following:

There is little evidence that they can actually notice the gap (see Schmidt & Frota, 1986) between their initial use of nontarget forms and the teacher's reformulation, given the ambiguity of recasts from the classroom learner's perspective. (57)

Lyster and Ranta also questioned the repaired utterances produced by the learners: "A repair in which the student simply repeats what the teacher has said does not necessarily imply that the feedback has been understood as such" (54). Their discussion raised questions about the basic role of recasts.

In second language acquisition theory, the benefits of receiving input have been widely debated. For example, "the input hypothesis" defined by Krashen (1985) is as follows:

The Input Hypothesis claims that humans acquire language in only one way — by understanding messages, or by receiving 'comprehensible input.' We progress along the natural order

(hypothesis 2) by understanding input that contains structures at our next ‘stage’ — structures that are a bit beyond our current level of competence. (2)

Later, Schmidt and Frota (1986) presented the “noticing hypothesis,” which claims that learners must notice the new linguistic features in the input in order to acquire them. Given these theories concerning inputs, recasts that provide learners with input are assumed to be effective for second or foreign language learning. Indeed, in the present study, a large number of the recasts led the learners to uptake: 92% of the recasts led the learners to uptake in the advanced class and 87% of them in the average class. The learners in both classes displayed high rates of uptake through recasts. Moreover, it seemed that the participants in my study were aware of the aim of the teacher’s recasts. In the example below, one learner in the advanced class was telling the teacher that she had difficulty with the pronunciation of the word “text,” which includes a consonant cluster at the end of the word. This type of consonant cluster does not appear in Japanese, so the student had difficulty with it.

(7)

S: “Text” difficult. Oh alright did you get my ... [Phonological]

T: Did you get my text? [Recast]

S: Text. [Needs-repair]

T: Text text text. [Recast]

S: Text text text. [Needs-repair]

T: Ok, so this here. Did you get my and then text. [Explicit]

S: Did you get my text? [Needs-repair]

The learner’s attention was focused on the phonological detail because she was trying to improve her pronunciation. The learners were trained to focus on linguistic features and were directed to follow the teacher’s feedback. In certain educational environments, recasts could be more useful, as Nicolas et al. (2001) suggested.

The “output hypothesis” offered by Swain suggested a new perspective for second language acquisition. Prompts are categorized as output-pushing feedback because they can provide learners opportunities to produce more accurate output. Swain (1985) claimed that learners develop their second language ability not only by receiving input but also by perceiving the target features and using them:

I would like to suggest that what is missing is output. Krashen (1981b) suggests that the only role of output is that of generating comprehensible input. But I think there are roles for output in second language acquisition that are independent of comprehensible input. (248)

The study by McDonough (2005) emphasized the importance of learners’ modified output. The author writes:

Additionally, negative feedback in the form of clarification requests may indirectly contribute to question development by creating opportunities for learners to modify their output. Thus, this study provides empirical support for the output hypothesis (Swain, 1985, 1993, 1995) and strengthens claims for an association between modified output and ESL question development (Mackey, 1997). (94)

As McDonough (2005) stated above, learners’ modified output is indirectly effective for second language learners. It would be effective for second or foreign language learners if they create modified output more frequently and more accurately. Schmidt (1983, 1990) and Gass and Varonis (1994) indicated the important role of teachers’ prompts. One of the significant roles provided by prompts is to lead learners to self-repair. Ammar and Spada (2006) also highlighted the role of prompts:

First, they unambiguously indicated the presence of an error and, therefore, encouraged and directed students to think about alternative forms. Second, once the learners were aware of the fact that there was a problem in the form that they had used to express their

meaning, they were given metalinguistic clues to help them identify the nature and locus of the error. (563)

It has been argued that learners' self-repaired utterances are only generated from prompts. In the case of the present research, few prompts were observed (see Chapter Four). For Teacher A, only 4% of all feedback was made up of prompts in the advanced class and 3% in the average one. As for Teacher B, 16% of all his feedback was made up of prompts. All prompts were followed by some kind of learner response; however, there were only a few cases in which they repaired their utterances. This result implied that the learners could not correct their previous incorrect utterances even though the teacher directed them to do so. An example of a conversation between the teacher and a learner is shown below:

(8) The situation: The teacher told the student to finish exercise one. The student finished the practice and was doing exercise two.

T: What are you practicing?

S: (pointing on a handout) *Kore* (This).

T: Why?

S: *Yareteiwareta* (Because the teacher told me to do it). [I & D]

T: Because...? [Elicitation]

S: *Owattakara* (I finished the first one). Finish. [Needs-repair]

T: You finished the first one? Okay.

In Example 8, the teacher encouraged the student to correct the previous erroneous utterances with the phrase, "Because...?" The student had to compose sentence following "because." However, the student was not able to express her thought in a correct sentence. So, she expressed her ideas by using the word "finish" because it was too difficult for her to construct an English sentence. The learners who participated in this observation often had difficulty constructing sentences including subjects and verbs. The learners tended to utter a single noun or verb, such as in the following example. In the first line, the learner asked "May I go to a bathroom?" in Japanese. In the second line, the teacher pushed the

learner to speak English. Finally, the learner said a single noun “toilet.”

(9)

S: *Toireni itte iidesuka?* (May I go to a bathroom?) [I & D]

T: What’s that in English? [Prompt]

S: Toilet. [Needs-repair]

Generally speaking, learners who study English in Japan continuously practice targeted sentences or grammar rules through English programs, but they often fail to construct sentences because they cannot use appropriate grammatical rules. One explanation for this phenomenon is provided by Gass and Torres (2005). In their study, vocabulary items were defined as “less complex” and grammatical structures were categorized as “more complex.” Regarding the degree of complexity, I agree with them and observed that the learners in my study were likely to have difficulty in the more complex language areas. Gass and Torres (2005) suggested a solution for this problem:

If a particular area of language is simple (i.e., noncomplex), learners can gather appropriate information on their own. When the language form or rule is more complex, external intervention may save learners time (although Hulstijn & de Graaff do not state that learners are unable to figure out complex phenomena on their own). (7)

In general, foreign language learners in Japan rarely receive interventions from others because: 1) there is limited time for English lessons, 2) English lessons do not always have interactive teaching styles, and 3) English classes are not always composed of a small student population. These three aspects might be the main reasons for preventing learners in Japan from receiving external interventions from others.

E. Conclusions.

In conclusion, one of the main topics, namely teachers’ factors, should be taken into account. Overall, there were statistical differences between the two

teachers in their frequencies of correcting learners' errors even though the teachers were observed under similar circumstances and during the same period. Specifically, there were significant differences between the two teachers in their use of prompts. On the contrary, a significant difference was not found in the use of their recasts or explicit corrections. The two teachers were teaching English under similar circumstances, so they might not be significantly different in the use of the latter two feedback types. These discoveries confirm the results from previous studies conducted by other scholars. As for the priorities in the four error types, there was no correlation between the two teachers. While Teacher B dealt equally with all types of errors, Teacher A strongly focused on the learners' pronunciation. Teacher A's strong focus on the learners' pronunciation could be interpreted as follows: 1) the learners' pronunciation faults were likely to cause problems in communication, and 2) the teachers had their own priorities for correcting the learners' error types.

The second factor examined in this study was the learner factor. Many scholars have conducted research with the assumption that there are relationships between learners' second or foreign language levels and the effectiveness of teacher corrections. Concerning the frequencies of the teacher corrections, there was a significant difference between the higher-level group and the lower-level group; however, significant differences could not be supported in the other aspects (e.g., the types of corrective feedback and the ratios of uptaken and repaired utterances). Despite non-statistical support, different tendencies could be observed in the two groups. First, the higher-level class tended to receive more recasts. Second, the lower-level class was likely to receive the teacher's explicit corrections. Recasts could be less effective than the other corrective techniques for the lower-level learners, so the teacher in my study might have assumed that explicit corrections would be better for the lower-proficiency students. This finding suggests that an explicit teaching style may be more effective for lower-level learners.

One teacher's variation in feedback use was also investigated, together with the learner factor. In the study of second language acquisition, researchers have reported that native speakers and/ or native-speaker teachers adjust their language use in consideration of learners' language levels. Similar to previous studies, the teacher observed in this study adjusted the frequencies of the corrections he provided for the two different-level classes. Nevertheless, that the teacher adjusted his correction types according to the level of learners could not be inferred from this study.

The efficacy of teacher corrections was illustrated in the discussions, and some conclusions for this topic can be drawn. The participants in my study produced a large number of uptake and repairs when the teacher provided them with corrective feedback. From this point of view, the teacher corrections yielded partly beneficial outcomes for the learners. The reason for this phenomenon could be the high rates of uptake derived from the teacher's recasts. In contrast to results of the recasts, the teacher could not lead the learners to repaired utterances when he used prompts.

Finally, some limitations of the present study should be acknowledged. First, although observational classes should be sampled randomly, it was difficult to conduct random sampling in the present study. Hence, the teachers and learners who participated in this study were not selected by random sampling. Second, the learners were tested by an entrance examination conducted by the high school; thus, the levels of the learners were not evaluated for this observation. Third, generally in classroom observations, the concepts of uptake and repair are used to determine the effect of teacher corrections; however, these concepts only indicate short-term effects. Moreover, uptake following recasts does not always relate to learners' learning. Like many other studies, long-term effects could not be inferred from the present study. There are some experimental or quasi-experimental studies that have examined the long-term effects of teacher corrections. However, in general, the targeted structures in those studies relate to

grammatical aspects. Thus, future research needs to determine the long-term effects for learners' acquisition of lexical or phonological features.

Bibliography

- Allen, Patrick, et al. "Aspects of Classroom Treatment: Toward a More Comprehensive View of Second Language Education." The Development of Second Language Proficiency. Ed. B. Harley, P. Allen, J. Cummins and M. Swain. New York: Cambridge University Press, 1990.
- Allwright, Richard L. "Turns, Topics, and Tasks: Patterns of Participation in Language Learning and Teaching." Discourse Analysis in Second Language Acquisition Research. Rowley: Newbury House Publishers, Inc, 1980.
- Ammar, Ahlem, and Nina Spada. "One Size Fits All? Recasts, Prompts, and L2 Learning." Studies in Second Language Acquisition 28 (2006): 543-574.
- Block, Ellen L. "See How They Read: Comprehension Monitoring of L1 and L2 Readers." TESOL Quarterly 26 (1992): 319-343.
- Bohannon, John N., and Laura Stanowicz. "The Issue of Negative Evidence: Adult Responses to Childrens' Language Errors." Developmental Psychology 24 (1988): 684-689.
- Brown, James D. Understanding Research in Second Language Learning. Ed. H. Altman and P. Stevens. New York: Cambridge University Press, 1988.
- Burt, Marina. From Deep to Surface Structure. New York: Harper & Row, 1971.
- Burt, Marina, and Carol Kiparsky. "Global and Local Mistakes." New Frontiers in Second Language Learning. Ed. J. Schumann and N. Stenson. Rowley: Newbury House, 1975.
- Cadierno, Teresa. "Formal Instruction from a Processing Perspective: An Investigation into the Spanish Past Tense." The Modern Language Journal 79 (1995): 179-193.
- Calve, Pierre. "Corriger ou ne pas corriger, la n'est pas la question." The Canadian Modern Language Review 48 (1992): 458-471.

- Carl, James. Errors in Language Learning and Use. Ed. C. N. Candlin. London: Longman, 1998.
- Cathcart, Ruth L., and Judy Olsen. "Teachers' and Students' Preferences for Error Correction of Classroom Conversation Errors." On TESOL '76 (1976): 41-53.
- Chaudron, Craig. "A Descriptive Model of Discourse in the Corrective Treatment of Learners' Errors." Language Learning 27 (1977): 29-46.
- . Second Language Classrooms: Research on Teaching and Learning. Ed. M. Long and J. Richards. New York: Cambridge University Press, 1988.
- Cicourel, Aaron V., et al. Language Use and School Performance. New York: Academic Press, 1974.
- Coulthard, Malcolm. "Discourse Analysis in English – A Short Review of the Literature." Language Teaching and Linguistics: Surveys. Ed. V. Kinsella. New York: Cambridge University Press, 1978.
- Davies, Eirlys E. "Error Evaluation: The Importance of Viewpoint." ELT Journal 37 (1983): 304-311.
- Dekeyser, Robert M., and Karl J. Sokalski. "The Differential Role of Comprehension and Production Practice." Language Learning 46 (1996): 613-642.
- Donato, Richard. "Collective Scaffolding in Second Language Research." Vygotskian Approaches to Second Language Research. Ed. J. Lantolf and G. Appel. New Jersey: Ablex, 1994.
- Doughty, Catherine, and Elizabeth Varela. "Communicative Focus on Form." Focus on Form in Classroom Second Language Acquisition. Ed. C. Doughty and J. Williams. New York: Cambridge University Press, 1998.
- Dulay, Heidi, and Marina Burt. "Remarks on Creativity in Language Acquisition." Viewpoints on English as a Second Language. Ed. M. Burt et al. New York: Regents, 1977.
- Dulay, Heidi, Marina Burt, and Stephen D. Krashen. Language Two. New York:

Oxford University Press, 1982.

Ellis, Rod. The Study of Second Language Acquisition. New York: Oxford University Press, 1994.

Ellis, Rod, and Gary Barkhuizen. Analysing Learner Language. New York: Oxford University Press, 2005.

Ellis, Rod, and Younghee Sheen. "Recasts in Second Language Acquisition." Studies in Second Language Acquisition 28 (2006): 543-574.

Fanselow, John F. "Beyond RASHOMON-Conceptualizing and Describing the Teaching Act." TESOL Quarterly 11 (1977): 17-38.

---. "The Treatment of Error in Oral Work." Foreign Language Annals 10 (1977): 583-593.

Farrar, Michael J. "Discourse and the Acquisition of Grammatical Morphemes." Journal of Child Language 17 (1990): 607-624.

Ferguson, Charles A. "Absence of Copula and the Notion of Simplicity: A Study of Normal Speech, Baby Talk, Foreigner Talk and Pidgins." Pidginization and Creolization of Languages. Ed. D. Hymes. Cambridge: Cambridge University Press, 1971.

Fraenkel, Jack R., and Norman E. Wallen. How to Design and Evaluate Research in Education. New York: McGraw Inc, 1996.

Frick, Ted, and Melvyn I. Summel. "Observer Agreement and Reliabilities of Classroom Observational Measures." Review of Educational Research (1978): 157-184.

Gass, Susan. "Second Language Vocabulary Acquisition." Annual Review of Applied Linguistics 9 (1988): 92-106.

---. "Integrating Research Area: A Framework for Second Language Studies." Applied Linguistic 9 (1988): 198-217.

Gass, Susan, and Maria J. A. Torres. "Attention When?." Studies in Second Language Learning 27 (2005): 1-31.

Goldin-Meadow, Susan. "The Resilience of Recursion: A Study of a

- Communication System Developed without a Conventional Language Model.” Language Acquisition: State of the Art. Ed. E. Wanner and L. Gleitman. Cambridge: C.U.P., 1982.
- Halliday, Michael A. K. “Categories of the Theory of Grammar.” Word 17 (1961): 241-291.
- Hama, Mika, and Ronald P. Leow. “Learning without Awareness Revisited.” Studies in Second Language Acquisition 32 (2010): 465-491.
- Hatch, Evelyn M. “Second Language Learning – Universals?” Bilingualism/Travaux de Recherches sur le Bilinguisme 3 (1974): 1-17.
- Havranek, Getaurd, and Hermann Cesnik. “Factors Affecting the Success of Corrective Feedback.” EUROSLA yearbook 1 (2002): 99-122.
- Hendrickson, James M. “Error Correction in Foreign Language Teaching: Recent Theory, Research, and Practice.” The Modern Language Journal 62 (1978): 387-398.
- Hughes, Arthur. and Chryssoula Lascaratou. “Competing Criteria for Error Gravity.” ELTJournal 36 (1982): 175-182.
- Ishiguro, Toshiaki. “Simplification and Elaboration in Foreign Language Teacher Talk and Its Source.” Diss. Stanford University, 1986.
- Izumi, Shinichi. “Output, Input Enhancement, and the Noticing Hypothesis: An Experimental Study on ESL Relativization.” Studies in Second Language Acquisition 24 (2002): 541-577.
- Kail, Robert V., and Rita W. Nelson. Developmental Psychology. New Jersey: Prentice Hall Inc, 1993.
- Kowal, Maria, and Merrill Swain. “Using Collaborative Language Production Tasks to Promote Students’ Language Awareness.” Language awareness 3(2) (1994): 73-98.
- . “From Semantic to Syntactic Processing: How Can We Promote Metalinguistic Awareness in the French Immersion Classroom?” Immersion Education: International Perspectives. Ed. K. Johnson and M.

- Swain. New York: Cambridge University Press, 1997.
- Krashen, Stephen D. Second Language Acquisition and Second Language Learning. Oxford: Pergamon, 1981.
- . The Input Hypothesis: Issues and Implications. London: Longman, 1985.
- . Second Language Acquisition and Second Language Learning. Ed. C. N. Candlin. New York: Prentice Hall International Ltd, 1988.
- . "The Input Hypothesis and Its Rivals." Implicit and Explicit Learning of Languages. Ed. N. C. Ellis. London: Academic Press, 1994.
- LaPierre, Donna. "Language Output in a Cooperative Learning Setting: Determining Its Effects on Second Language Learning." M.A. thesis, University of Toronto, 1994.
- Lavine Harold G., et al. "Teaching Participant-Observation Methods: A Skills-Building Approach." Anthropology and Education Quarterly 11 (1983): 38-54.
- Long, Michael H. "Teacher Feedback and Learner Error: Mapping Cognitions." TESOL '77. Ed. H. D. Brown et al. (1977): 278-293.
- . "Inside the 'Black Box': Methodological Issues in Classroom Research on Language Learning." Language Learning 30 (1980): 1-42.
- . "Question in Forigner Talk Discourse." Language Learning 31 (1981): 135-156.
- . "The Role of the Linguistic Environment in Second Language Acquisition." Handbook of Second Language Acquisition. Ed. W. Ritchie and T. Bhatia. New York: Academic Press, 1996.
- Long, Michael H., and Charlene J. Sato. "Classroom Foreigner Talk Discourse: Forms and Functions of Teachers' Questions." Classroom Oriented Research. Ed. H. Seliger and M. Long. Rowley: Newbury House Publishers, Inc, 1983.
- Long, Michael H., and Peter Robinson. "Focus on Form: Theory, Research, and Practice." Focus on Form in Classroom Second Language Acquisition. Ed.

- C. Doughty and J. Williams. New York: Cambridge University Press, 1998.
- Lyster, Roy. "Negotiation of Form, Recasts, and Explicit Correction in Relation to Error Types and Learner Repair in Immersion Classrooms." Language Learning 48 (1998): 183-218.
- Lyster, Roy, and Leila Ranta. "Corrective Feedback and Learner Uptake." Studies in Second Language Acquisition 20 (1997): 37-66.
- Lyster, Roy, and Hidenori Mori. "Interactional Feedback and Instructional Counterbalance." Studies in Second Language Acquisition 28 (2006): 269-300.
- Lyster, Roy, and Kazuya Saito. "Oral Feedback in Classroom SLA." Studies in Second Language Acquisition 32 (2010): 265-302.
- Mackey, Alison, and Jenefer Philp. "Conversational Interaction and Second Language Development: Recasts, Responses, and Red Herrings?" Modern Language Journal 82 (1997): 338-356.
- Mackey, Alison, and Susan Gass. Second Language Research: Methodology and Design. New Jersey: Lawrence Erlbaum Associates Inc, 2005.
- McCormick, Dawn E. and Richard Donato. "Teacher Questions as Scaffolded Assistance in an ESL Classroom." Second and Foreign Language Learning through Classroom Interaction. Ed. J. K. Hall and L. Verplaetse. New Jersey: Lawrence Erlbaum and Associates, 2000.
- McDonough, Kim. "Identifying the Impact of Negative Feedback and Learners' Responses on ESL Question Development." Studies in Second Language Acquisition 27 (2005): 79-108.
- Mehan, Hugh. "Accomplishing Classroom Lessons." Language Use and School Performance. Ed. A. Cicourel et al. New York: Academic Press, 1974.
- Moskowitz, Gertrude. "The Classroom Interaction of Outstanding Foreign Language Teachers." Foreign Language Annals 9 (1976): 135-157.
- Newport, Elissa L., Henry Gleitman, and Lila R. Gleitman. "Mother, I'd rather

- Do It Myself: Some Effects and Non-Effects of Maternal Speech Style.” Talking to Children: Language Input and Acquisition. Ed. C. Snow and C. A. Ferguson. Cambridge: C.U.P., 1977.
- Nicholas, Howard, Patsy Lightbown, and Nina Spada. “Recasts as Feedback to Language Learners.” Language Learning 51 (2001): 719-785.
- Nobuyoshi, Junko, and Rod Ellis. “Focused Communication Tasks and Second Language Acquisition.” ELT Journal 47 (1993): 203-210.
- Nystrom, Nancy J. “Teacher-Student Interaction in Bilingual Classrooms: Four Approaches to Error Feedback.” Classroom Oriented Research. Ed. H. Seliger and M. Long. Rowley: Newbury House Publishers, Inc, 1983.
- Ohta, Amy S. Second Language Acquisition Processes in the Classroom: Learning Japanese. New Jersey: Lawrence Erlbaum Associates Inc, 2001.
- Oliver, Rhonda. “Negative Feedback in Child NS-NNS Conversation.” Studies in Second Language Acquisition 18 (1995): 459-481.
- . “Negotiation of Meaning in Child Interactions.” Modern Language Journal 82 (1998): 376-386.
- . “Age Differences in Negotiation and Feedback in Classroom and Pairwork.” Language Learning 50 (2000): 119-151.
- Oliver, Rhonda, and Alison Mackey. “Interactional Context and Feedback in Child ESL Classrooms.” Modern Language Journal 87 (2003): 519-533.
- Panova, Iliana, and Roy Lyster. “Patterns of Corrective Feedback and Uptake in an Adult ESL Classroom.” TESOL Quarterly 36 (2002): 573-595.
- Richards, Jack. “Error Analysis and Second Language Strategies.” Language Sciences 17 (1971): 12-22.
- . Error Analysis Perspectives on Second Language Acquisition (Applied Linguistics and Language Study). New York: Longman, 1974.
- Sacks, Harvey, Emanuel A. Schegloff, and Gail Jefferson. “A Simplest Systematics for the Organization of Turn-taking for Conversation.” Language 50 (1974): 696-735.

- Seedhouse, Paul. The Interactional Architecture of the Language Classroom: A Conversation Analysis Perspective. Oxford: Blackwell, 2004.
- Sinclair, John, and Malcolm Coulthard. Towards an Analysis of Discourse. London: Oxford University Press, 1975.
- Slimani, Assia. "Evaluation of Classroom Interaction." Evaluating Second Language Education. Ed. C. Alderson and A. Beretta. Cambridge: Cambridge University Press, 1992.
- Snow, Catherine, and Marian Hoefnagel-Hohle. "School Aged Second Language Learners' Access to Simplified Linguistic Input." Language Learning 32 (1982): 411-430.
- Spada, Nina, and Maria. Flohlich. The Communicative Orientation of Language Teaching Observation Scheme: Coding Conventions and Applications. Sydney: National Centre for English Teaching and Research, Macquarie University, 1995.
- Stevick, Earl W. Memory, Meaning, and Method. Rowley: Newbury House, 1978.
- Swain, Merrill. "Communicative Competence: Some Roles of Comprehensible Input and Comprehensible Output in Its Development." Input in Second Language Acquisition. Ed. S. Gass and C. Madden. Massachusetts: Newbury House, 1985.
- . "The Output Hypothesis: Just Speaking and Writing Aren't Enough." Canadian Modern Language Review 50 (1993): 158-164.
- . "Three Functions of Output in Second Language Learning." Principle and Practice in Applied Linguistics: Studies in Honour of H. G. Widdowson. Ed. G. Cook and B. Seidolhofer. Oxford: Oxford University Press, 1995: 125-144.
- . "Focus on Form through Conscious Reflection." Focus on Form in Classroom Second Language Acquisition. Ed. C. Doughty and J. Williams. New York: Cambridge University Press, 1998.
- Swain, Merrill, and Elaine Tarone. "A Sociolinguistic Perspective on Second

Language Use in Immersion Classrooms.” Modern Language Journal 79 (1995): 167-178.

Swain, Merrill, and Sharon Lapkin. “Interaction and Second Language Learning: Two Adolescent French Immersion Students Working Together.” Modern Language Journal 87 (1998): 320-337.

Valdman, Albert. “Learner Systems and Error Analysis.” In Perspective: New Freedom. Ed. G. Jarvis. Illinois: National Textbook Company, 1975.

Vanpatten, Bill. “Attending to Form and Content in the Input.” Studies in Second Language Acquisition 12 (1990): 287-301.

Vanpatten, Bill, and Teresa Cadierno. “Explicit Instruction and Input Processing.” Studies in Second Language Acquisition 15 (1993): 225-243.

Wilson, Stephen. “The Use of Ethnographic Techniques in Educational Research.” Review of Educational Research 47 (1977): 245-265.

石川慎一郎『英語コーパスと言語教育』、東京：大修館書店、2008年。

石黒敏明「教室における Teacher Talk を考える」『英語展望』、東京：英語教育協議会、1988年。

大場衣織「インプット貧困環境におけるリキャストとプロンプトの有効性」『神奈川大学大学院 言語と文化論集 15』、2009年。

---。「学習者の間違いと教師による訂正のフィードバックにおける関係性」『神奈川大学大学院 言語と文化論集 16』、2010年。

小池生夫『第二言語習得研究に基づく最新の英語教育』、東京：大修館書店、1994年。

小池生夫『第二言語習得研究の現在』、東京：大修館書店、2004年。

竹村日出夫『True Colors: Oral Communication I』、東京：数研出版株式会社、2002年。

Appendix A

Examples of Model Dialogues

(1) Topic: "Talking on the Cell"

A: Hello?

B: Hi, (). Is this a good time?

A: Yeah, no problem. What's up ()?

B: I'm here at the station. Where are you right now?

A: I'm waiting for a bus. It should be here by now.

B: Oh, I see. Are you gonna be late?

A: I'm afraid so. I'd say about ten minutes.

(2) Topic: "What's a Good Time for You?"

A: Can I come to your house on Monday?

B: Sure. How about two thirty?

A: OK. I'll see you then.

(3) Topic: "Can I Take a Message?"

A: Hello. This is (). May I speak to ()?

B: I'm afraid he's not in right now.

Appendix B

Examples of Teacher-Student Conversations Including Teacher Corrections

(1)

S: I'm calling again later.

T: I'll call again later. [Recast]

S: I?

T: I'll. [Recast]

S: I'll call later.

(2)

S: Docchimo using? (Do I have to use both of them?)

T: What? [Prompt]

S: Both?

T: Use both. [Recast]

S: Both.

(3)

S: What time do we meet?

T: What time do we meet? [Prompt]

S: Shall we?

T: Should we. [Recast]

S: Should, should.

(4)

S: Shall we going to lunch together?

T: Shall we have lunch together. No going to. [Explicit Correction]

S: (nodding)

T: Okay, shall we have lunch together?

S: That's a good idea.

T: That's a good idea, but good idea is okay. [Explicit Correction]

S: Ah.

T: Okay?

S: Let's go to the restaurant.

T: That restaurant. [Recast]

S: (nodding)

T: Not the restaurant. That restaurant. [Explicit Correction]

S: That? Okay.

(5)

S1: What's your problem?

T: What's your problem change into what's up. [Explicit Correction]

S1: What's up your problem?

T: No problem. Just what's up. [Explicit Correction]

S1: What's up?

T: Yeah.

S1: No problem?

T: Don't write problem. [Explicit Correction]

S1: Don't write ...

S2: Problem dake. (Erase the word "problem" on your note)

T: Yeah, good.

(6)

S: Bake, bake. (You got sunburned)

T: Break? [Prompt]

S: Bake, baked face.

T: Bake, yeah. This is called sunburn. [Explicit Correction]

S: Sunny burn.

T: Sunburn. [Recast]

S: Sunny burn.

T: Sun. Not sunny, sunburn. [Explicit Correction]

S: Sunburn.

神奈川大学大学院
言語と文化論集 特別号

2014年8月 印刷

2014年8月 発行

編集発行 神奈川大学大学院
外国語学研究所
(横浜市神奈川区六角橋 3-27-1)

製作 共立速記印刷株式会社

LANGUAGE AND CULTURE
BULLETIN
OF
THE GRADUATE SCHOOL OF FOREIGN LANGUAGES
KANAGAWA UNIVERSITY
Special number, August, 2014

Doctoral Course, English Language and Literature
(Date of Degree Awarded : March 31, 2014)

A Thesis Submitted for Ph. D. in English Language and Literature

Factors Affecting the Effect of Error Corrections

Iori OBA
