MODIFIED OUTPUT

IN RESPONSE TO CLARIFICATION REQUESTS

AND

SECOND LANGUAGE LEARNING

A thesis
submitted in partial fulfilment
of the requirements for the degree
of
Doctor of Philosophy
at
The University of Waikato
by
MASAYOSHI OGINO

The University of Waikato
2008

ABSTRACT

Modified output, second language (L2) learners' reformulation of their own utterances, has been attracting researchers' interest as an important component of learner interactions, and as a manifestation of interlanguage development and psycholinguistic processing.

The *output hypothesis* (Swain, 1985, 1993, 1995, 2005) claims that the act of production constitutes part of the process of L2 learning in terms of *noticing*, *hypothesis testing* and *metalinguistic functions*. This hypothesis has been used as a theoretical framework to investigate the relationship between modified output and L2 learning (e.g., McDonough, 2001, 2005; Nobuyoshi & Ellis, 1993; O'Relly, Flatiz, & Kromrey, 2001; Takashima & Ellis, 1999). However, the empirical evidence from these studies does not yet appear to confirm unequivocally that the production of modified output facilitates L2 learning.

The present study further explored the impact of modified output on L2 learning, by means of an experimental pre-test, post-test and delayed-post design. The production of modified output was triggered by one type of implicit feedback, clarification requests. The data were collected from 28 undergraduate students who were learning Japanese as a foreign language. The target linguistic feature was the negation of adjectives in Japanese, and a total of 1,011 negations were elicited and analysed. The impact of modified output on L2 learning was measured in two different aspects of potential outcomes of modified output (i.e., grammatical accuracy

and interlanguage development). In addition, the study investigated whether the non-targetlike forms which participants previously modified were then produced in the subsequent situations of use.

The output hypothesis was originally framed in terms of the relationship between output and grammatical accuracy, but the findings of the current study suggest that production of modified output in response to clarification requests may facilitate the progress of interlanguage development towards targetlike use even when its immediate impact on grammatical accuracy may not be observed. Therefore, the present study lends at least partial support to the claim of the output hypothesis.

The results did not clearly demonstrate whether or not production of modified output might sensitise learners to avoid the use of the same non-targetlike form that they have previously modified. This indicates a possibility of the limited role of production of modified output in L2 learning, and suggests that the follow-up feedback to learners' modified output may be necessary to maximise the impact of modified output in facilitating L2 learning.

ACKNOWLEDGEMENTS

TABLE OF CONTENTS

CHAPTER SIX: DISCUSSION AND CONCLUSION

APPENDICES FOR CHAPTER FOUR

APPENDICES FOR CHAPTER FIVE

LIST OF FIGURES

LIST OF TABLES

LIST OF ACRONYMS

A	Regular adjective
Adv	Adverbial Inflector
CR	Clarification Requests
EFL	English as a Foreign Language
ESL	English as a Second Language
IL	Interlanguage
JFL	Japanese as a Foreign Language
JSL	Japanese as a Second Language
L1	First Language
L2	Second Language
MO	Modified Output in response to Clarification Requests
N	Noun
NA	Nominal Adjective
NEG	Negator
NNS	Non-Native Speaker
NONP	Non-Past Tense
NS	Native Speaker
NTL	Non-Targetlike
POL	Politeness
SLA	Second Language Acquisition
TOP	Topic Marker
TL	Targetlike
V	Verb

MODIFIED OUTPUT

IN RESPONSE TO CLARIFICATION REQUESTS

AND

SECOND LANGUAGE LEARNING

A thesis
submitted in partial fulfilment
of the requirements for the degree
of
Doctor of Philosophy
at
The University of Waikato
by
MASAYOSHI OGINO

The University of Waikato
2008

ABSTRACT

Modified output, second language (L2) learners' reformulation of their own utterances, has been attracting researchers' interest as an important component of learner interactions, and as a manifestation of interlanguage development and psycholinguistic processing.

The *output hypothesis* (Swain, 1985, 1993, 1995, 2005) claims that the act of production constitutes part of the process of L2 learning in terms of *noticing*, *hypothesis testing* and *metalinguistic functions*. This hypothesis has been used as a theoretical framework to investigate the relationship between modified output and L2 learning (e.g., McDonough, 2001, 2005; Nobuyoshi & Ellis, 1993; O'Relly, Flatiz, & Kromrey, 2001; Takashima & Ellis, 1999). However, the empirical evidence from these studies does not yet appear to confirm unequivocally that the production of modified output facilitates L2 learning.

The present study further explored the impact of modified output on L2 learning, by means of an experimental pre-test, post-test and delayed-post design. The production of modified output was triggered by one type of implicit feedback, clarification requests. The data were collected from 28 undergraduate students who were learning Japanese as a foreign language. The target linguistic feature was the negation of adjectives in Japanese, and a total of 1,011 negations were elicited and analysed. The impact of modified output on L2 learning was measured in two different aspects of potential outcomes of modified output (i.e., grammatical accuracy

and interlanguage development). In addition, the study investigated whether the non-targetlike forms which participants previously modified were then produced in the subsequent situations of use.

The output hypothesis was originally framed in terms of the relationship between output and grammatical accuracy, but the findings of the current study suggest that production of modified output in response to clarification requests may facilitate the progress of interlanguage development towards targetlike use even when its immediate impact on grammatical accuracy may not be observed. Therefore, the present study lends at least partial support to the claim of the output hypothesis.

The results did not clearly demonstrate whether or not production of modified output might sensitise learners to avoid the use of the same non-targetlike form that they have previously modified. This indicates a possibility of the limited role of production of modified output in L2 learning, and suggests that the follow-up feedback to learners' modified output may be necessary to maximise the impact of modified output in facilitating L2 learning.

ACKNOWLEDGEMENTS

First and foremost, I would like to express my profound gratitude to my chief supervisor, Roger Barnard for his wise guidance, encouragement and patience throughout the ups and downs of the project. He has been a great mentor and educator, who always motivated me with his scholarship.

My thanks also go to Jean Newman, whose analytical approach and critical input has contributed substantially to my study. To Akiko Nakayama, who kindly agreed to join in my supervisory panel in the most difficult final year, I am grateful for her insightful advice and warm support. My thanks are also due to Mike Roberts, who contributed a great deal from the inception of the project.

Special gratitude is also extended to the entire staff at the Department of General and Applied Linguistics, especially to Ray Harlow as a process supervisor and as a chief supervisor while Roger was on leave, and to the graduate advisor, Ian Bruce for his invaluable feedback and constructive comments at the various stages of the project.

I would like to thank the Japanese Section of the Department of Humanies, especially Ken McNeil and Fumiko Nishimura, who were helpful and understanding in my asking their students to take part in my project.

TABLE OF CONTENTS

CHAPTER SIX: DISCUSSION AND CONCLUSION

LIST OF FIGURES

LIST OF TABLES

LIST OF ACRONYMS

A	Regular adjective
Adv	Adverbial Inflector
CR	Clarification Requests
EFL	English as a Foreign Language
ESL	English as a Second Language
IL	Interlanguage
JFL	Japanese as a Foreign Language
JSL	Japanese as a Second Language
L1	First Language
L2	Second Language
MO	Modified Output in response to Clarification Requests
N	Noun
NA	Nominal Adjective
NEG	Negator
NNS	Non-Native Speaker
NONP	Non-Past Tense
NS	Native Speaker
NTL	Non-Targetlike
POL	Politeness
SLA	Second Language Acquisition
TOP	Topic Marker
TL	Targetlike
V	Verb

CHAPTER ONE

INTRODUCTION

1.1   Background

Learners of a second or foreign language occasionally modify their own utterances, −*modified output*− and this is an important component of interaction with input, feedback and negotiation for meaning.

The *output hypothesis* (Swain, 1985, 1993, 1995, 2005) claims that "the act of producing language (speaking or writing) constitutes, under certain circumstances, part of the process of second language learning" (Swain, 2005, p. 471). This hypothesis has provided researchers with a theoretical framework to investigate the relationship between modified output and second language (L2) learning. There are also a number of studies involving modified output in the context of the *input hypothesis* (Krashen, 1985) and the *interaction hypothesis* (Long, 1983, 1996).

Over the two decades since the first proposal of the output hypothesis, a number of studies have emphasised the importance of modified output in L2 learning; the potential impact of this relationship has been discussed in terms of *noticing* (Schmidt, 1990, 1995; Schmidt & Frota, 1986; Swain, 1995), *hypothesis testing* (Swain, 1995), *automaticity* (Anderson, 1982, 1992; de Bot, 1996; DeKeyser, 1997; McLaughlin, 1987), *grammatical encoding* and *monitoring* (Izumi, 2003; Kormos, 2006), and *stimulating syntactic processing* (de Bot, 1996; Izumi, 2000; Pica, Lincoln-Porter,

Paninos, & Linnell, 1996; Swain, 1995). The importance of modified output has also been supported by meta-analysis of the findings of a number of modified output studies up to 2003 (Keck, Iberri-Shea, Tracy-Ventura, & Wa-Mbaleka, 2006).

However, individual studies investigating this relationship have produced mixed results. While some studies demonstrate the benefit of modified output in L2 learning (de la Fuente, 2002; He & Ellis, 1999; Izumi, 2002; Loewen, 2002; McDonough, 2001, 2005; Nobuyoshi & Ellis, 1993; Pica et al., 1996; Van den Branden, 1997), some studies have not shown positive evidence of modified output (Izumi & Bigelow, 2000; Izumi, Bigelow, Fujiwara, & Fearnow, 1999; O'Relly et al., 2001; Takashima & Ellis, 1999). A meta-analysis of interactional research up to 2006 by Mackey and Goo (2007) reported that, unlike that by Keck et al. (2006) noted above, the provision of opportunities for modified output did not show any difference in language development, and they suggest the necessity of more research specifically designed to examine modified output.

These different results, however, are not surprising. Firstly, the production of modified output is a complex phenomenon involving a number of variables including the non-targetlike utterance, a trigger to lead to modified output (e.g., feedback), and provision of an opportunity for modification. Control of these variables differs significantly among studies. Secondly, the variety of results may be related to the variety of target linguistic features which have been investigated in empirical studies. The choice of target linguistic feature might influence the effectiveness of corrective feedback (Ellis, 2007, p. 340), and this may apply to the effectiveness of modified

output as well. Thirdly, the measurements for the impact of modified output and operationalisation of L2 learning employed in each study is another possible factor that may have contributed to different results, which will be discussed in the following section.

1.2 Measuring the impact of modified output on L2 learning

1.2.1 Grammatical accuracy

Grammatical accuracy has been one of the most commonly used constructs to measure the relationship between modified output and L2 learning. Improvements in the grammatical accuracy of use of the targeted linguistic feature after interventions to elicit modified output have been operationalised as evidence of a positive impact of modified output (McDonough, 2001; Nobuyoshi & Ellis, 1993; O'Relly et al., 2001; Takashima & Ellis, 1999; Van den Branden, 1997). Grammatical accuracy seems a reasonable measurement, as the output hypothesis originated from attempts to address the issue of weak accuracy among students of Canadian French immersion programmes (Swain, 1985). However, some limitations of grammatical accuracy as a measurement of L2 learning, especially the impact of an interactional component, have been pointed out (Ellis & Barkhuizen, 2005; Meisel, Clahsen, & Pienemann, 1981; Pienemann, 1998). For example, judgements of accuracy are binary (correct/incorrect) and are based on target language ('native speaker') norms (i.e., the *comparative fallacy*, Bley-Vroman, 1983), and accuracy does not describe "the interlanguage forms that arise as learners approximate to target language norms" (Ellis & Barkhuizen, 2005, p. 92).

1.2.2 Interlanguage development

The analysis of interlanguage development is a measurement which can describe "the underlying developmental system of the L2 learner" (Spada & Lightbown, 1993, p. 208), and therefore it can avoid the comparative fallacy. A growing body of studies has been carried out with a focus on how interaction involving the target feature advances learners through describable developmental stages (Ellis, 2007; Mackey, 1995, 1999, 2006; Mackey, Oliver, & Leeman, 2003; Mackey & Philp, 1998; McDonough, 2001, 2005, 2006; McDonough & Mackey, 2006; Philp, 2003; Silver, 2000; Spada & Lightbown, 1993, 1999).

One of the common characteristics of most of this interactional research is the use of question formation as a target feature in English as a second language (ESL) or a foreign language (EFL). The reasons for the choice of question formation include abundant previous empirical studies (e.g., Cazden, Cancino, Rosansky, & Schumann, 1975; Milon, 1974; Ravem, 1968; Wode, 1981), relative ease of elicitation, and its presence at all stages of learning. However, perhaps the main reason is that the developmental stages of question formation have been identified and established by Pienemann and Johnston (1987) and Pienemann, Johnston, and Brindley (1988).

While the choice of the same target feature among studies makes the findings more generalisable, it also limits the findings to one context in one language. What needs to be explored is whether the findings can be applied to other linguistic features in other languages. One of the pre-requisites to achieve this is the choice of a targeted linguistic feature that is empirically shown to be developed incrementally. A

4

possible target feature in another language is negation in Japanese. Research has revealed that the acquisition of negation in Japanese passes through predictable developmental stages in both first language (Clancy, 1985; Ito, 1990) and L2 learners (Kamura, 2001b; Kanagy, 1991, 2001; Noro, 1995; Takeuchi-Furuya, 1993).

1.2.3 Subsequent use of the non-targetlike utterance that learners have modified

Most of the experimental studies mentioned so far use a pre-test/treatment/post-test design, and measure the changes in accuracy and/or interlanguage development observed in the pre- and post-tests. Another approach is to track the qualitative and quantitative changes of the non-targetlike forms in the subsequent situations of use which individual learners have previously modified. Loewen (2007), referring to research on the incidental focus on form, suggests the comparison between the subsequent production of the targeted linguistic forms and its previous production as "the best measure of the effectiveness" (p. 103). This arises from the difficulty in conducting a pre-test in the study of incidental focus on form without prior planning because the linguistic items involved in the targeted interaction cannot be predicted, but this can be worth applying to the investigation of modified output and L2 learning.

The necessity of tracing the linguistic feature in studies on output has also been suggested (Shehadeh, 2002), but there are only a handful of studies that have tracked down the subsequent performance of those learners who produced modified output and those who did not do so (e.g., Loewen, 2007; Mackey, 1997; McDonough, 2001, 2005). These studies focus on whether or not linguistic forms that have been

5

previously modified/involved in the targeted interactional components are used correctly in subsequent situations of use (i.e. accuracy). What has not yet been explored is the relationship between the types of non-targetlike form that learners modify and how they are used in the subsequent situations of use (e.g., whether they continue to be used, or whether another type of non-targetlike form replaces them). This really involves consideration of interlanguage development as part of the measurement of subsequent use. Also, studies of subsequent use of the target features that learners have previously modified may contribute to clarifying whether the quantity or quality of modified output plays a role in L2 learning. While some researchers (e.g., Krashen, 1994, 1998) see the quantity of modified output as vital to have an impact on L2 learning, some researchers suggest the importance of the quality of modified output in L2 learning (e.g., de Bot, 1996; Ellis, 1999; McDonough, 2001; Shehadeh, 2002). However, what kind of modified output might be related to L2 learning is not clear, and the investigation of the types of non-targetlike form that learners modify is one approach to describing the relationship between the quality of modified output and L2 learning.

1.3   Purpose of the study

The purpose of this study was to explore whether production of modified output facilitates L2 learning. The objective was to look for evidence of the impact of modified output, specifically elicited in response to a type of implicit feedback, clarification requests, on the acquisition of negation of predicate adjectives in the non-past tense in Japanese as a foreign language (JFL). Only a few studies (Doughty & Varela, 1998; Loewen & Nabei, 2007; Spada & Lightbown, 1993) that have

6

considered both accuracy and interlanguage development to measure the impact of interactional components on the same linguistic feature. Therefore, the relationship between modified output and L2 learning will be investigated in terms of grammatical accuracy, and qualitative and quantitative changes of interlanguage in the acquisition of negation of adjectives in JFL. In addition, the types of non-targetlike form that learners modify will be tracked down in their subsequent use. In other words, it aims to shed light on the role of output by investigating the three different aspects of potential outcomes of modified output (i.e., accuracy, interlanguage development and subsequent use).

1.4   Research questions

Based on the background and purpose of the study above, the following three research questions are addressed in the study.

Research Question 1: Is there a positive relationship between the production of modified output in response to clarification requests and grammatical accuracy?

Research Question 2: Is there a positive relationship between the production of modified output in response to clarification requests and interlanguage development?

Research Question 3: Is there a relationship between the non-targetlike forms that the learners modify and their subsequent use in terms of accuracy and interlanguage development?

1.5 Operational definitions

Operational definitions of the present study are presented below.

Modified output is language produced by the learner that is modified from the initial utterance either in response to feedback or without feedback, irrespective of the extent to which the reformulation is targetlike.

A *clarification request* is a type of implicit feedback which indicates a request for clarification of utterances in terms of non-targetlike use. It does not provide learners with further linguistic information that could be incorporated in modified output, and it may be realised in English by such expressions as 'Pardon,' or 'Could you say that again?' and in Japanese, by '*E*' (Pardon?) or '*Moo ichido*' (Once again?)."

Targetlike refers to the kind of linguistic form that is most likely to be used by a competent native speaker of a language in a specific situation of use. In relation to the present study, the term is operationalised as the correct choice of a Japanese negator and an associated morpheme, and the correct inflectional change on any given adjective stem: for example, *samu-ku + na-i* (cold-Adv.+ NEG-NONP, '(it) is not cold'). The use of this term in relation to negation will be developed in further detail in Chapter Three.

Non-targetlike refers to the incorrect choice of a linguistic form from among the range of negation patterns that exist in Japanese. Such an incorrect choice may involve a negator and associated morphemes, an incorrect inflectional change added to any given adjective stem, or some incorrect combination of both: for example,

samu-janai (*janai* consists of an negator and associated morphemes, and it is used for negating nouns and nominal adjectives, not adjectives). The use of a non-targetlike form does not necessarily indicate a breakdown in the communicative competence of the learner, but may rather show interlanguage development.

Accuracy is whether or not an item in the target language is well-formed according to the rule system of the target language; accuracy in the present study is measured by the number of targetlike uses in the attempted uses.

Interlanguage development is qualitative and/or quantitative changes of interlanguage towards targetlike use, which is measured in this study by the progressive reduction in variation and/or frequency of non-targetlike forms of negation of adjectives in Japanese.

Subsequent use refers to the use of the previously modified targeted linguistic feature in the post- and delayed post-tests, and not during the treatment sessions.

The *participants* in this investigation were undergraduates studying JFL at elementary and intermediate classes at the University of Waikato.

1.6　　Significance of the study

The present investigation on modified output is useful for both theoretical and pedagogical reasons. The theoretical significance of the study lies in exploring the role of learners' own modification of non-targetlike utterance in L2 learning through

testing the output hypothesis. The benefit of modified output will be investigated through three different aspects of learner language, while previous research employed one or two measures. The findings on the role of output in interaction approach, which are largely based on European languages, will be tested in a non-European language, Japanese. Also, research employing interlanguage development as a measure will be expanded into negation of adjectives in JFL beyond a commonly used targeted linguistic feature, question formation in ESL. It will also contribute to the knowledge of acquisition of JFL by describing the route of acquisition of the target form, and by confirming or disconfirming the mixed findings of the developmental stages of negation in JSL and JFL.

On the pedagogical side, exploring whether the production of modified output contributes to acquisition is of particular importance in the context of error response strategies in language classrooms. Encouraging learners to modify their own non-targetlike utterance may help them to improve their accuracy or to advance to more targetlike performance. For Japanese language pedagogy, research on the developmental stages of negation will be beneficial in terms of curriculum design and the more appropriate choice of task activities, as there are grammatical structures and morphemes learners might not be able to acquire unless they are developmentally ready (Pienemann, 1998).

1.7   Overview of the study

This chapter has presented the background for the investigation of modified output in the study reported in the following chapters, and suggested an approach to broaden

the empirical base of such interactional research. The three research questions were stated.

In the next chapter, literature reporting research on modified output will be reviewed with reference to the three types of measurement of interactional components. It points out features and issues on research that investigate the relationship between modified output and L2 learning. In Chapter Three, the developmental sequence of the target form, negation in Japanese, is reviewed and examined for its application to interlanguage analysis. Based on Chapter Two and Three, Chapter Four presents the research methodology, which is followed by results and analysis in Chapter Five. Chapter Six discusses the implications of the findings and concludes the study.

CHAPTER TWO

LITERATURE REVIEW

2.0 Introduction

This chapter starts with a review of previous studies on modified output and second language (L2) learning. It is followed by a review focusing on studies investigating the impact of modified output on L2 learning with reference to the three kinds of measurements (2.2) and to data collection methods (2.3).

2.1 Modified output and L2 learning

2.1.1 Roles of output in L2 learning

Over two decades, views about the role of output in second language acquisition (SLA) have shifted from a result of "acquired competence" (Krashen, 1987, p. 16) to "part of the *process* of learning" (Swain, 2005, p. 471, italics in original).

The output hypothesis (Swain, 1985) originated from a study of French immersion programmes in Canadian schools in which students were found to be weak in grammatical accuracy in contrast to their high-level listening skills and communicative fluency. Swain argued that the lack of grammatical accuracy of those immersion students could be attributable to the limited opportunities to produce output or for being pushed to produce output (p. 249). Swain (1995) proposed three functions of output in addition to enhancing fluency: (1) *noticing function* which proposes that the activity of production may prompt second language learners to consciously recognise some of their linguistic problems, (2) *hypothesis-testing*

function which refers to learners' "trial run" (Swain, 2005, p. 476) for comprehensibility or linguistic well-formedness, and (3) *metalinguistic (reflective) function* which claims a mediative role of the use of language to reflect on language in L2 learning.

Several terms have been used to indicate the reformulated output. Among them are *pushed output* (Swain, 1985), *comprehensible output* (Swain, 1985), *enhanced output* (Takashima, 1995), *uptake* (e.g., Lyster & Ranta, 1997), and *modified output* (e.g., Swain & Lapkin, 1995). The term pushed output has been defined as "output that reflects what learners can produce when they are pushed to use the target language accurately and concisely" (Ellis, 2003, p. 349), and Ellis (2003) pointed out the necessity of distinguishing between modified output and pushed output. Pushed output does not necessarily include modification of the initial non-targetlike utterance because it could be a simple repetition, which is illustrated in the example below.

Example 1

Non-native speaker (NNS):   I go cinema

Native speaker (NS):   Uh?

NNS: I go cinema last night

NS: Oh, last night.

(Ellis, 2003, p. 82)

Swain (2005) stated that the term comprehensible output tended to confuse the essence of the output hypothesis, which lies in the process of production of output

13

rather than the output as a product. In other words, being "able to be understood" (Swain, 2005, p. 473) was the focus of research rather than "output that was an improved version of an earlier version in terms of its informational content and/or its grammatical, sociolinguistic, or discourse features" (ibid.)

The term enhanced output is output "grammaticalized as a result of 'pushing'" (Takashima, 1995, p. 174), which focuses on grammatical accuracy in the negotiation of meaning, and is not concerned with the developmental changes that might indicate the progress towards targetlike use.   Example 2 shows output which was produced as a result of learner's modification of the initial non-targetlike utterance in response to teacher's clarification request 'Sorry?'

Example 2

      Student: The prince fall in love at first glance

      Teacher: Sorry?

      Student: The prince falled in love at first glance.

                                               (Takashima & Ellis, 1999, p. 174)

The reformulated output including 'falled' was not grammatically enhanced, but there is a possibility that the attempt to modify the initial output itself, which involves grammatical encoding and monitoring, might contribute to L2 learning (Izumi, 2003; Mackey, 2007a; McDonough, 2001).

In the studies exploring the relationship between output and L2 learning within the framework of the output hypothesis, it seems important to include the output that is still not comprehensible, or not grammatically enhanced in comparison with the initial non-targetlike use. Therefore, a neutral, broad and comprehensive term, 'modified output' will be used to refer to language produced by the learner that is modified from the initial utterance either in response to feedback or without feedback, irrespective of the extent to which the reformulation is targetlike. This approach will widen our understanding of the role of output beyond comprehensible and grammaticalized output.

Uptake is another important construct in interaction research (e.g., Ellis, Basturkmen, & Loewen, 2001; Loewen, 2002, 2004). Uptake is defined by Lyster and Ranta (1997) as "a student's utterance that immediately follows the teacher's feedback and that constitutes a reaction in some way to the teacher's intention to draw attention to some aspect of the student's initial utterance" (p. 49). The discussion that follows establishes the distinction that is made here between uptake and modified output.

Firstly, uptake assumes the provision of feedback while modified output does not do. Therefore, modified output does not exclude learner's output produced without interactional feedback from an external source. The output hypothesis proposes a facilitative role for output "even without implicit or explicit feedback provided from an interlocutor about the learner's output" (Swain & Lapkin, 1995, p. 373), and the term modified output fits well with this hypothesized role of output. It, therefore, leaves scope for future research into the role of self-initiated self-completed repair in

L2 learning in comparison with that of externally prompted repair (i.e., modified output), especially as their mechanisms are assumed to be identical (Kormos, 1996).

Secondly, uptake does not necessarily require the modification of the initial non-targetlike utterance, as it is "*any* student oral response to a feedback move" (Long, 2006, p. 101, italics in original), and the occurrence of uptake can be influenced by whether the learner had an opportunity to respond or not, as Long (2006) pointed out. This feature of uptake leads to the necessity of developing sub-categories, *repair* and *needs-repair* (Lyster & Ranta, 1997), each of which are categorised further into *repetition*, *incorporation of the correct form*, *self-repair*, and *peer-repair*, and *acknowledgement, same error, different error, off target, hesitation,* and *partial repair*. Loewen (2002) developed four subcategories of uptake (repair as a subcategory of uptake in Lyster & Ranta, 1997), *successful uptake, unsuccessful uptake, partially successful uptake* and *inconclusive*. In successful uptake, learners need to repair the linguistic error or incorporate the information received by attempting to use the linguistic item in their own speech (pp. 123-124), while unsuccessful uptake involves "the student simply acknowledging the Response with a *yes*, *mm*, *oh* or a similar acknowledgment token, using the target feature incorrectly, or failing to demonstrate processing of the information received" (p. 124, italics in original). Despite the similarities between uptake and modified output, it should be noted that modified output could include successful uptake, partially successful uptake, and even unsuccessful uptake, and thus they are viewed as distinct constructs. In addition, the aim of the present study, in a broad sense, is to test the output hypothesis, in which the process of modification of non-targetlike utterance

16

constitutes the central construct for the hypothesis, thereby providing another rationale for the use of the term modified output over uptake in this study.

2.1.2   Research on modified output

Research on modified output can be divided into the following categories: (a) variables leading to modified output, (b) types of linguistic modification in response to feedback, (c) the impact of modified output in response to feedback on L2 learning, and (d) the impact of modified output without external feedback.   Figure 2.1 illustrates the overview of research on modified output.

Figure 2.1
Research on modified output
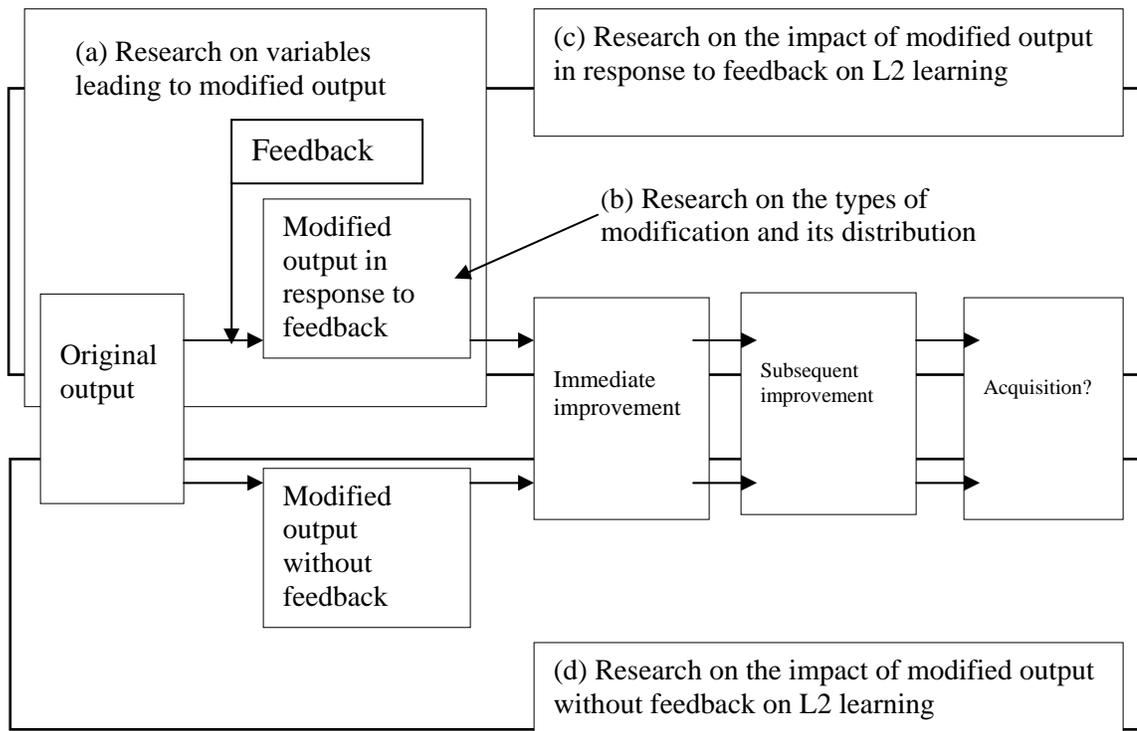
The interaction hypothesis proposed by Long (1983, 1996) generated researchers'
interest in the role of modified output as a component of interaction. A wide range
of research has been conducted to examine whether learners modify their output in
the interaction (see section (a) in Figure 2.1). Among the variables which have been
examined are:

- feedback type leading to modified output (Iwashita, 2001, 2003; Lyster, 2004;
  Lyster & Ranta, 1997; McDonough, 2007; Nassaji, 2007; Pica, 1988; Pica,
  Holliday, Lewis, & Morgenthaler, 1989; Van den Branden, 1997);

- task types (Foster, 1998; Iwashita, 1999; O'Relly et al., 2001; Pica et al.,
  1989; Pica et al., 1996; Shehadeh, 1999);

- type of interlocutor, such as native speaker vs. non-native speaker, and adult
  vs. children (Iwashita, 1999; Pica et al., 1996; Sato & Lyster, 2007; Shehadeh,
  1999);

- proficiency levels (Iwashita, 2001; Pica, 1988; Shehadeh, 1999; Van den
  Branden, 1997; Varonis & Gass, 1985); and

- age and gender of participants (Mackey et al., 2003; Pica, Holliday, Lewis,
  Berducci, & Newman, 1991).

Secondly, there are studies that examined the linguistic categories that learners
modify in response to feedback and its distribution (see section (b) in Figure 2.1). In
other words, researchers look at what kind of linguistic modification (lexical,
morphological, syntactic, phonological, or semantic) learners make in response to
feedback. For example, semantic modified output is observed more frequently than

modified output on syntax, morphology and phonology (Foster, 1998; Pica, 1988). The research in this area is still scarce despite its potential importance for demonstrating the relationship between specific linguistic category and modification (Shehadeh, 2002).

While the two areas discussed above usually do not go further once modified output is produced, the third area is concerned with the impact of modified output on L2 learning (see section (c) in Figure 2.1). Here, the focus of the research is on the changes in learner language before and after the production of modified output, especially on the three functions of output proposed by Swain (1995); noticing function (e.g., Izumi, 2002; Izumi & Bigelow, 2000; Izumi et al., 1999; Swain & Lapkin, 1995), hypothesis-testing function (e.g., Ellis & He, 1999; Mackey, 2002; Nobuyoshi & Ellis, 1993; Shehadeh, 2001, 2003), and metalinguistic function (e.g., LaPierre, 1994; Swain, 1995, 1998; Swain & Lapkin, 2001).

Overall, there appears to be a general consensus on the positive role of output in L2 learning, and the results of meta-analysis of interactional research (Keck et al., 2006) support the role of pushed output in the learning process. Nobuyoshi and Ellis (1993) is one of the earlier studies that demonstrated the impact of modified output on grammatical accuracy, and there are other studies that also support the contributing role of modified output on L2 learning (de la Fuente, 2002; He & Ellis, 1999; Izumi, 2002; Loewen, 2002; McDonough, 2001, 2005; Van den Branden, 1997). However, some studies (Izumi & Bigelow, 2000; Izumi et al., 1999; O'Relly et al., 2001; Takashima & Ellis, 1999) did not show the impact of modified output on

19

L2 learning.   Also, it should be noted that the benefits of modified output claimed on the basis of the results of some studies might be attributable to interaction including modified output (e.g., comprehensible input in Ellis and Takashima, 1999; discourse experience in Ellis and He, 1999) rather than to modified output as a single component of interaction.   Another meta-analysis (Mackey & Goo, 2007) reported interaction without opportunities for modified output contributed more to the acquisition of the targeted linguistic feature than interaction with opportunities for modified output.   The studies in this area are discussed in more detail in section 2.2.

Modified output produced without interactional feedback (i.e., *self-initiated self-completed repair*) could be added as another area of studies on modified output (see section (d) in Figure 2.1).   While output is usually used in a computational metaphor of the input-black box-output paradigm, the term repair is often used in conversation analysis and psycholinguistic studies.   The studies in this area provide valuable insights into the mechanism of production of repair and the organisation and distribution of repair behaviour (see van Hest 1996 for a detailed summary of previous research on self-repair).   Levelt's (1989) speech production model *Perceptual loop theory of monitoring* fits the mechanism of both pushed output (de Bot, 1996) and self-repair (Kormos, 1999, 2006).   Some researchers, including Swain (1995), consider the process of producing modified output identical regardless of the provision of feedback, and argue for the positive role of self-repair on acquisition (Green & Hecht, 1993; Schegloff, Jefferson, & Sacks, 1977; Shonerd, 1994; van Lier, 1988).   On the other hand, there are empirical studies that did not show a significant relationship between self-initiated self-repair and language

development (e.g., McDonough, 2001). This is an interesting area and the differences and similarities between modified output produced with and without feedback need to be explored.

The research reviewed so far is within the interaction approach to second language acquisition, the major constructs of which are input, interaction, feedback and output (Gass and Mackey, 2006). This approach is supported by such theoretical framework as the Input-Interaction-Output (IIO) model (e.g., Gass, 1997; Gass & Selinker, 2001; Long, 1996). The model that Gass (1997) proposes, for example, has five stages to account for how input is converted into output, which are apperception, comprehended input, intake, integration and output. The IIO model "combines the change of input-interaction-output with findings from cognitive psychology, without losing sight of Universal Grammar as a constraint on the SLA processes" (Block, 2003, p. 26). It accommodates such theoretical proposals as information processing model which assumes a shift from controlled to automatic processing (McLaughlin, 1987, 1990), or from declarative knowledge to procedural knowledge (Anderson, 1982, 1992).

In recent years, however, the research on output in L2 learning has been reconsidered (and this is not accounted for in Figure 2.1), and the other main stream has been emerging, which bases its theoretical framework on the sociocultural theory of the mind (Vygotsky, 1978). In this view, learning occurs within the *zone of proximal development* (ZPD), which is "the distance between the actual developmental level as determined by independent problem solving and the level of potential development as

determined through problem solving under adult guidance or in collaboration with more capable peers" (Vygotsky, 1978, p. 86). In applying this perspective to the study of the impact of modified output, modified output is not considered an independent factor that might contribute to L2 learning, but part of social activity that constitutes learning. For example, shifting her perspective on interaction to a sociocultural perspective, Swain has recently avoided using the term output in her work and uses such terms as *verbalization* (Swain, 2000), and *languaging* (Swain, 2006) because this perspective may not fit well with the information processing perspective in which the input the learner receives and the output he/she produces are analysed as separate phenomena (de Bot, 1996).

Nevertheless, the present study places itself within the interaction approach, since the purpose of the study is to test the output hypothesis, which is originally proposed within this approach. The focus of the current research within this approach has also been shifting from whether or not interaction impact learning to such inquiries as "How does interaction create opportunities for learning?" and "What are the relationships among interactional feedback, learner-internal cognitive processes and L2 learning outcomes" (Mackey, 2007b, p. 10). The present study, which explores the specific relationship between modified output and L2 learning, will belong to the latter inquiry and contribute to interpreting the mixed results found in the previous research.

2.2   Research on the relationship between modified output and L2 learning

This section specifically reviews research which explored the impact of modified output on L2 learning, with reference to studies that examine three aspects of learners' language: accuracy, interlanguage development and subsequent use.

2.2.1   Modified output and grammatical accuracy as a measurement

The improvement of grammatical accuracy has been the centre of the investigation of the impact of modified output (e.g., Izumi & Bigelow, 2000; Izumi et al., 1999; McDonough, 2001; Nobuyoshi & Ellis, 1993; O'Relly et al., 2001; Takashima & Ellis, 1999; Van den Branden, 1997) because the output hypothesis originated from the findings of weak grammatical accuracy in the input-rich immersion programme.

Nobuyoshi and Ellis (1993) investigated whether production of modified output in response to clarification requests led to more accurate use of the simple past tense of verbs in English, and whether the improvement was sustained. Three participants in a control group and the other three in an experimental group performed picture jigsaw communication tasks twice with an interval of one week. The results based on accuracy scores (i.e., percentage of correct suppliance divided by the use in obligatory occasions) showed that the participants in the experimental group outperformed those in the control group, which lent support to the output hypothesis outside the context of immersion programmes. It is important to note that several limitations of this study have been pointed out (e.g., Krashen, 1998; O'Relly et al., 2001), which include the small number of participants, the low frequency of

obligatory occasions, and the absence of relevant statistical tests. Nevertheless, this study has stimulated several further studies.

O'Relly et al. (2001) examined the effectiveness of two types of corrective feedback on command forms in Spanish. This study expanded the study of Nobuyoshi and Ellis (1993), supplementing their methodological and procedural limitations by employing a randomised block design, a longer interval before the delayed-post test, and a one-factor analysis of variance. Forty-eight university students of Spanish (intermediate-low, intermediate-mid, and intermediate-high) were assigned into two experimental groups (E1 & E2) and a control group. They had four administrations in which E1 had focused communication tasks by clarification requests (i.e., clarification requests were given only on the error of the target form) and E2 by confirmation checks while the control group had unfocused communication tasks (i.e., no feedback). A post-test consisting of unfocused communication tasks was conducted five weeks after the fourth administration. The results of obligatory occasion analysis[1] showed no statistical significant difference in accuracy scores regarding the production of correct forms among the three groups in the post-test, and, therefore, the impact of modified output on grammatical accuracy was not confirmed. However, the small sample sizes and variability in the mean accuracy scores at the point of pre-test (E1, M = 15.63, SD = 25.07; E2, M = 38.06, SD = 23.29; Control, M = 23.86, SD, 22.64) limited the reliability of statistical tests and generalisability of the findings. This study has a robust research design, but the

---

[1] Obligatory occasion analysis (Brown, 1973) compares the forms used by learners and target language norms, and accuracy scores are calculated by *n* correct suppliance in contexts divided by total obligatory contexts (Ellis & Barkhuizen, 2005).

results suggested the complexity of the relationship among proficiency levels of participants, the importance of the choice of the target form, and the need to obtain sufficient number of tokens of the target form.

The study by Takashima and Ellis (1999) also built upon Nobuyoshi and Ellis (1993) in that learners were given clarification requests to the non-targetlike use of the past-tense forms in English, but it was different in that the interaction between the interlocutor and the learners occurred in a classroom setting involving other learners in contrast to the one-to-one interaction in the original study. Sixty-one Japanese learners of English had a pre-test and three post-tests in which they created oral narratives from two series of four pictures in a language laboratory. Then, they had three treatment sessions in groups of four or five over three weeks, engaging in a picture description activity in each session. Each group had a representative (narrator) who in turn told the story of the group to the class, based on the information he/she had obtained through the group discussion. An experimental group (EG) had a treatment with focused negotiation in which clarification requests were given by a teacher only when a narrator representing a group produced an utterance containing an error in the use of the target form. A control group (CG) was also given clarification requests, but this happened only when a genuine communication breakdown took place, and, therefore, the CG had unfocussed negotiation. Obligatory occasion analysis was used for calculation of accuracy scores for the comparison with the results of Nobuyoshi and Ellis (1993). The results show that unfocused as well as focused negotiation was effective in improving narrators' accuracy of the target features. The narrators who told a picture story to

the class in the CG improved their accuracy while those listeners who only listened without an opportunity to speak showed little improvement. One of the most interesting findings is that listeners in the EG who overheard the modified output produced by narrators also improved their accuracy, while listeners in the CG did not. In other words, the focused negotiation technique in a classroom setting created a situation where modified output was produced, but at the same time it created another variable, modified and enriched input that was heard and could be incorporated into the interlanguage system of other learners. The results of this study showed that modified output in the focused negotiation was not associated with the accuracy in the tests, but demonstrated the impact of modified and enriched input on accuracy.

Issues on grammatical accuracy as a measurement

The previously reviewed studies are concerned with whether the production of modified output facilitated the improvement of grammatical accuracy. There are few studies that demonstrate the relationship between modified output and grammatical accuracy except Nobuyoshi and Ellis (1993), while studies on vocabulary acquisition (e.g., de la Fuente, 2002; He & Ellis, 1999) show some positive impact of modified output. It appears that demonstrating the benefit of modified output on grammar might be "somewhat elusive and hard" (Mitchell & Myles, 2004, p. 176), but it is not clear whether the lack of studies that showed a direct relationship between modified output and grammatical accuracy is due to the difficulty in demonstration, or to the limited or small impact of modified output on accuracy.

Also, grammatical accuracy is not likely to be the only aspect of learner language on which the production of modified output could have some impact. Lightbown (2003) pointed out *"progress* will not necessarily show up as greater accuracy" (p. 5, italics in original), and she illustrated progress not represented by accuracy by referring to two sentences involving question formation in English, 'What's your name?' and '*Why the children want to play?[2]' Although the former is accurate and the latter is inaccurate, the latter is considered to reflect the progress in terms of language development beyond the memorised formulaic expression.

Difficulty in defining the obligatory occasion is another issue (Braidi, 1999). In addition, accuracy in the obligatory occasion analysis does not deal with the use of the target feature in non-obligatory occasions (Braidi, 1999; Ellis, 1994; Mackey & Gass, 2005), which has been argued as a major limitation as a measurement tool. *Targetlike use analysis[3]* (Pica, 1993), a modification of obligatory occasion analysis, deals with overuse of the targeted morphemes by including the suppliance in non-obligatory occasions, but the orientation in the targetlike use remains to be a limitation. Also, another limitation of obligatory occasion analysis relates to the *comparative fallacy* (Bley-Vroman, 1983), and one of the approaches to avoid this issue is interlanguage analysis (also referred to as frequency analysis by Ellis & Barkhuizen, 2005), which is discussed in the next section.

---

[2] Asterisks indicate non-targetlike use.
[3] This takes the over-suppliance of the same form into consideration.

2.2.2   Modified output and interlanguage development as a measurement

The aim of interlanguage analysis is to describe "the stage of development that learners pass through on route to mastery of a linguistic structure" (Ellis & Barkhuizen, 2005, p. 95).   Interlanguage analysis is not concerned with the comparison between learner language and the target language, as is the case with the analysis of grammatical accuracy.   In this analysis, the various linguistic devices that learners use to express a particular grammatical structure are identified at different points in learners' development, which is followed by determination of the stage of acquisition (Ellis, 1994; Ellis & Barkhuizen, 2005).

Spada and Lightbown (1993) was one of the earlier studies that employed interlanguage analysis to measure the impact of interactional components.   Two experimental classes of francophone ESL learners in an elementary school received form-focused instruction and corrective feedback on ESL question formation over a two-week period while a comparison class did not receive the same intervention. The oral production was collected from picture description tasks, and it was analysed in terms of accuracy (i.e., percentage of well-formed question forms) and developmental progress based on the developmental stages[4] proposed by Pienemann and Johnston (1986).   Credits were given for attaining a stage when a student produced two different utterance types in one stage, and the number of students who produced question formation in each of the six developmental stages was compared

---

[4] The six-stage developmental sequence starts with Stage 1 in which single words or fragments such as 'A spot on the dog?' can be acquired.   Stage 2 is SVO with rising intonation (e.g., 'A boy throw the ball?'), and Stage 3 is Fronting (Do-fronting, *wh*-fronting and other fronting), such as 'Do you have three astronaut?'   Acquisition of Stage 2 is the prerequisite of Stage 3 in this developmental sequence. The three example sentences are cited from Spada and Lightbown (1999).   The theoretical framework of this developmental sequence will be discussed in section 3.4 in more detail.

between groups. The comparison class often performed better than the experimental classes in both measurements because of the unexpectedly large amount of focus on form and corrective feedback that the comparative class teacher provided. However, the overall results supported the idea that form-focused instruction and corrective feedback contributed to language development, and demonstrated the changes in the developmental stages as a measurement for investigating the relationship between interactional components and language learning.

Following their study, a growing body of fine-tuned studies have been carried out with a focus on the developmental progress of the target feature involved in interaction. The components of interaction include recast[5] (Mackey & Philp, 1998; McDonough & Mackey, 2006); input, output and negotiation (Silver, 2000); modified output and feedback (McDonough, 2001, 2005); noticing (McDonough & Mackey, 2006; Philp, 2003); corrective feedback (Ellis, 2007; Loewen & Nabei, 2007; McDonough, 2007); and instructions (Spada & Lightbown, 1999).

Mackey and Philp (1998) investigated the impact of recast (not modified output) in negotiated interaction on question formation, using the same six-stage sequence as Spada and Lightbown (1993). They analysed the number of learners who advanced to a higher stage form and the frequency of higher-level structures. The results showed recasts for learners at a higher developmental level facilitated the production of higher stage forms, but not for learners who were not developmentally ready.

---

[5] Recast refers to "more targetlike versions of learners' incorrect utterances, which do not change their meanings" (Mackey, 2007b, p. 15).

Mackey (1999) examined whether conversational interaction facilitated language development of ESL question formation. The focused stages were Stage 4, which include pseudo-inversion, and Yes/No questions, and Stage 5 involving Do/Aux-second. The groups who had engaged in interaction demonstrated sustained stage increase and produced higher stage structures than the group who did not have interaction. The positive impact of the active participation in interaction was shown by the progress of the developmental stages in these studies.

McDonough (2005) also employed the advancement of the developmental stage of ESL question formation to measure the relationship between modified output and language development. She investigated whether or not negative feedback (i.e., clarification requests), or subsequent modified output or their combination, contributed to language development in a pre-test/post-test experimental design. The data used for analysis in McDonough's study were collected from learners (n = 60) who were classified as being at Stage 4 of the developmental sequence proposed by Pienemann and Johnston (1987) and Pienemann, Johnston and Brindley (1988) in the pre-test performance. The number of learners who moved to a higher stage was recorded in each of the four groups (i.e., enhanced and elicited group, elicited group, enhanced no opportunity group, and control group). The analysis indicated that neither modified output nor clarification requests were effective to advance to a higher stage in the developmental sequences, but modified output involving developmentally advanced forms produced in response to clarification requests was effective. In other words, just responding to clarification requests by modification

may not be enough for language development, and the modification needs to be considerably specific.

Features of interlanguage analysis in research on modified output

The studies discussed above investigated the impact of interaction involving modified output on interlanguage through the analysis of change in the developmental stages. The number of learners who moved to a higher stage in an experimental group and a control group were analysed in the experimental studies (e.g., Mackey, 1999; McDonough, 2001; Philp, 2003). When the developmental stages of learners are identified in advance and the targeted learners are within specific stages, this would be useful. Another way is to analyse the number of learners who produced the target form that belongs to each stage category (e.g., Loewen & Nabei, 2007; Spada & Lightbown, 1993). This can be used even when some learners have achieved the highest stage in the sequence.

One of the common features among the studies reviewed above is the use of ESL question formation as the target linguistic form. As mentioned in Chapter One, the rationale for the choice includes abundant empirical studies, easy elicitation, and availability of the forms at all stages of learning, and the developmental stages empirically shown by Pienemann and Johnston (1987) and Pienemann, Johnston, and Brindley (1988).

However, such heavy reliance on this linguistic feature may be a limitation in the research that employs interlanguage analysis to measure the impact of modified

output. Firstly, the generalisation of the results is limited unless investigation is carried out on other target features in other languages. Secondly, target linguistic features that are suitable for interlanguage analysis can be limited because relative difficulty may not be differentiated in some linguistic features when the same framework as Pienemann and Johnston (1987) and Pienemann, Johnston, and Brindley (1988), and its reoriented theory, the Processability Theory (Pienemann, 1998) are used. For example, Spada and Lightbown (1999) pointed out the acquisition of specific possessive determiners in ESL is one such linguistic feature. They involve local and non-local morphology, which fall within the same stage in terms of processing constraints. The production of a phrase such as 'my house is…' requires no exchange of information. This is a local morphology, which is "morphological form determined within the same phrase or constituent" (Spada & Lightbown, 1999, p. 4). In contrast, a sentence 'David loves his mother' is a non-local morphology, and it requires "holding information from one sentence constituent in mind while selecting the correct form of a morpheme" (ibid.). No prediction is made between these in the Processability Theory although there are some differences in terms of processing constraints. Also, Loewen and Nabei (2007) suggested that Stage 4 in the question form in the developmental sequence (Pienemann et al., 1988) was too broad to reflect the changes by a small number of instances of corrective feedback in a half hour treatment. Charters (2005), in her PhD dissertation on nominal structures in Mandarin SLA, found most of these structures are not differentiated in the Processability Theory and fall within the same processing category. The choice of the target linguistic feature is, therefore, important.

Another common feature of research on modified output with interlanguage analysis is the use of *emergence* as a criterion for evidence of reaching a higher stage, rather than other possible criteria, such as the use of most frequent form (e.g., Cancino, Rosansky, & Schumann, 1978) and the first non-imitative use (e.g., Wode, 1978) (see Braidi, 1999 for explanation of each criterion). The emergence criterion was developed by Meisel, Clahsen and Pienemann (1981) as part of their multi-dimensional model of SLA, based on the assumption that "not every change in a learner's production necessarily represents a move to a new developmental stage" (Braidi, 1999, p. 22). In principle, one usage is considered sufficient as evidence of relaxing part of processing constraint, but formulaic production needs to be excluded because it does not involve genuine processing. The distinction between formulaic and non-formulaic is hard to make, and, therefore, lexical variation needs to be shown, that leads to 'two usages.' Therefore, the studies along this line were based and developed upon a two productive usages at one time measure (Pienemann & Johnston, 1987). Some researchers (e.g., Ellis & Barkhuizen, 2005), however, consider a quantification based approach safer than the emergence approach because of the difficulty in differentiating the two. Also, Tarone (1998) considers the emergence criteria for acquisition glossing over the basic fact of interlanguage variation, which could be shown by the occurrence of emergence at different points in time on different tasks.

An area to be investigated further may be the changes in the lower stage forms in the developmental sequence. Some studies (Loewen & Nabei, 2007; McDonough, 2005; Spada & Lightbown, 1993) have reported them, but more detailed analysis will

33

provide some insight into the language development, which might not be captured through the focus on the emergence and movement towards a higher stage. When the developmental stages of the target linguistic feature are not clearly identified, it may be possible to use the changes in the use of non-targetlike forms in order to measure the impact of interaction including modified output. However, it would be necessary to have a control group, since preceding and following developments are required to be known to explain certain structural properties of learner performance (Meisel et al., 1981).

As discussed, interlanguage analysis has demonstrated the impact of interactional components, including modified output on L2 learning, which might not be revealed with an analysis of accurate use (e.g., obligatory occasion analysis and targetlike use analysis). Ellis and Barkhuizen (2005) pointed out that "[i]nterlanguage analysis may not be an alternative to accuracy analysis, but may be best used as a complementary analysis alongside it" (p. 109). However, few studies have analysed both accuracy of use and developmental progress on the same target feature. Spada and Lightbown (1993) carried out accuracy analysis and interlanguage analysis, and showed the parallel improvement both in accuracy and interlanguage in each of the comparison class and the experimental classes which received form-focused instruction and corrective feedback. As noted earlier in this section, classroom interaction data in their study revealed that the teacher of the comparative class gave a considerable amount of correctly formed target form and consistent corrective feedback, which resulted in the comparative class often outperforming the experimental classes. The difference in the impact of form-focused instruction and

corrective feedback on accuracy and interlanguage development was not the aim of their study, and the relationship between accuracy and interlanguage development remains unclear. This aspect merits further research, and the approach that makes the most of the strengths of each analysis would contribute to clarification of the impact of modified output on L2 learning.

2.2.3 Modified output and analysis of subsequent use as a measurement

Research discussed in the previous sections employed a pre-test/treatment/post-test design and examined the changes in learners' performance before and after the treatment in terms of accuracy or developmental progress. A more direct approach to measuring the impact of modified output on L2 learning is to explore the relationship between modified output and the subsequent use of the linguistic feature that has been modified.

McDonough and Mackey (2006) showed two interesting examples regarding the relationship between learners' response to recast and subsequent use within a couple of turns. One is that a learner who immediately repeated the phrases targeted by recast still produced the same type of non-targetlike use. The other is that another learner, who simply acknowledged a recast by saying 'yeah,' produced a targetlike form in the immediate subsequent turn. These illustrate the complexity of the relationship between feedback, production, and subsequent use, but it would be of importance to explore the use of the target feature beyond immediate turns to describe the relationship between interactional components and L2 learning.

A small number of studies on interaction have tracked subsequent use of a certain linguistic feature beyond immediate turns (e.g., recast and interaction by Mackey and Philp, 1998; interaction by Mackey, 1999; corrective feedback by Nassaji and Swain, 2000; language-related episode by Williams, 2001; recast by Loewen, 2002, 2007), and the need for more studies has been suggested (Loewen, 2007; McDonough & Mackey, 2006; Mitchell & Myles, 2004; Shehadeh, 2002).   The section below first reviews research investigating the subsequent use of the linguistic features that are involved in interaction, then discusses research specifically dealing with the subsequent use of features that are modified.

Subsequent use of linguistic features involved in interaction

Nassaji and Swain (2000) were concerned with whether or not the target feature involving random help was correctly used in the subsequent use.   This is a case study investigating the acquisition of English articles (i.e., *a*, *an, the,* and the zero article) by two Korean learners of English from a Vygotskian perspective.   One of the learners received corrective feedback provided within the zone of proximal development (ZPD) in four forty-minute tutorials.   The other learner (non-ZPD student) received feedback randomly provided within the regulatory scale proposed by Aljaafreh and Lantolf (1994).   The results of process-product analysis in the student-specific, task-related cloze tests demonstrated explicit feedback (Level 6-12 in the regulatory scale) led to higher accuracy of the article usage (62.5%) in the subsequent use than implicit feedback (level 2-5) did (37.5%).   The analysis of subsequent use demonstrated changes in the accurate use of target feature after receiving feedback.

Williams (2001) investigated the effectiveness of spontaneous attention to form, and examined whether and how the linguistic item that was focused on in the language-related episodes (LREs) was used in subsequent production. Data analysis was based on a total of 303 successful LREs with grammatical or lexical issues of eight ESL learners, and tailor-made tests given two weeks after. The tailor-made tests took various forms (e.g., multiple-choice discourse completion and fill-in-the-blanks tasks), depending on the LREs of each participant. Learning was operationalised as an increase in the accurate use of the targeted items in subsequently elicited situations. Although the occurrences of subsequent use of the items that were targeted in LRE were very low (17% out of 178 LREs targeting lexical items), it was found that LREs were associated with accurate performance in the test. This study suggested some potential of subsequent use analysis but also showed "difficulty of capturing evidence of delayed integration and use" (p. 336).

Adams (2003) investigated the effects of learners noticing differences, and traced the written reformulations produced by L2 Spanish learners. A control group (i.e., task repetition group) only had a story writing task based on pictures as a pre-test and a post-test. The same story tasks were used for both of the tests but the task conditions were different in that the pre-test was collaborative writing (pair work) while the post-test was individual work. A noticing group had a noticing session in Spanish (10 minutes) in which each dyad had an opportunity to compare their original story and a version reformulated by a native or near-native speaker of Spanish. They were instructed to verbalize their thoughts about the differences between the two versions. The other noticing group had the same noticing session, but this was immediately

followed by a stimulated recall session with the researcher (i.e., noticing with stimulated recall group). Approximately half of the total reformulations (n = 746) were grammatical structures (e.g., verbal morphology, preposition, gender) and the other half were lexically based. Learners' use of the original non-targetlike form was traced to the post-test and there were 445 cases in which learners attempted to use the form in the post-test. Because the focus of the study was 'noticing' through the comparison between original and reformulated writing, the target items of reformulation were not pre-selected, and the types of non-targetlike form and their changes in the post-tests were not reported. Sixty-one percent of forms which learners in the noticing group and the noticing with stimulated recall group had reported as 'noticed' were used in a more targetlike manner in their subsequent use. The results showed that noticing facilitated more than task repetition alone and that noticing with stimulated recall facilitated more than noticing alone. The descriptions of the forms which learners reported as noticed but not used in a more targetlike manner were not provided. The investigation into those forms might provide more insight in L2 learning.

Loewen (2007) examined the prior and subsequent use of linguistic features targeted in form-focused episodes (FFEs), defined as "all the discourse related to the targeted linguistic form" (p. 104). One of his motivations arose from the limitation of employing a quasi-experimental design in research examining an incidental focus on form, in which a targeted linguistic feature is not pre-determined and the use of pre-test is not possible (unlike when examining planned focus on form). Also, it was a way not to rely on uptake, a common measurement for assessing recast, which may

not be evidence of L2 learning (Loewen & Philp, 2006; Long, 2006; McDonough & Mackey, 2006; Russell & Spada, 2006; Williams, 2001). The data were collected from classroom observations of three ESL classes (n = 33), and the data analysis was based on the 121 FFEs of prior and subsequent use by learners who were originally involved in the FFE. The mean accuracy score improved from 47% (before FFEs) to 74% (after FFEs), which indicated the possibility of immediate effects of incidental focus on form on the subsequent use. FFEs collected did not have pre-determined target features and the aspects of language targeted in FFEs were varied (i.e., grammar, vocabulary and pronunciation), and, therefore, the effectiveness of recasts on different aspects of linguistic features was not within the scope of the study. With regard to the relationship between uptake and subsequent use, 78.9% of successful uptake had correct use in subsequent use, but 66.7% of unsuccessful uptake was correct in subsequent use as well. Loewen pointed out the low frequency of language episodes followed by a subsequent use as a limitation of subsequent use analysis. There were only 22 instances of uptake (19 successful uptake and 3 unsuccessful uptake) in the data gathered during four and half hours of classroom observation, and he recommended the combined use of subsequent use analysis with other measurements such as individualised tests. It was also reported that out of 121 FFEs, only three instances (i.e., preposition, plural, and pronunciation) were accompanied by both prior and subsequent use. Studies in laboratory settings with more controlled elicitation of the prior and subsequent use may contribute to generalisability of their findings. What needs to be explored is how this kind of tracing of the developmental changes in interaction can be applied to the studies that investigate the impact of modified output.

Subsequent use of modified linguistic feature

There are fewer studies that tracked down the subsequent use of linguistic features that learners modified. An unpublished study by Mackey (1997), cited in Mackey (1999) and McDonough (2001), investigated the subsequent use of question form by ESL learners (n = 4) who produced five or more instances of modified output of the target form and those (n = 4) who demonstrated a pattern of not modifying responses. The results show that the production of modified output might be associated with the subsequent production of higher stage forms although the sample size was small and the criteria for the classification of the two groups was arbitrary (McDonough, 2001).

McDonough (2001) identified some modifiers (participants who produced modified output more than median of the group) did not develop their interlanguage as measured by a movement towards a higher stage in ESL question formation. She classified the linguistic features participants modified into seven categories (aux insertion, aux changes, inversion, main verb changes, *wh*-question words, subject changes, question tags), and examined the difference between modifiers who developed their interlanguage and those who did not in the post-tests. It was found that the amount of modified output was comparable between the groups, but the modifiers who developed produced more modified output involving insertion of auxiliary verbs than those modifiers who did not develop. This study suggested the relationship between how learners modified a non-targetlike form and interlanguage development in the subsequent situations of use.

Swain and Lapkin (2007) reported the learning process of a Grade 7 French immersion student called Neil from a sociocultural perspective. They tracked the changes of his language through multitask activities that consisted of five stages over two weeks. Stage 1 was a story writing task based on a set of drawings as a pre-test. Neil compared his own writing with a reformulated one in Stage 2 (noticing), which was followed by Stage 3 of a stimulated recall of videotaped Stage 2. Stage 4 was another writing task, in which he wrote an original story, and Stage 5 was a subsequent interview for Neil to reflect on his learning. This study aimed at demonstrating "ways in which Neil's L2 learning is distributed across his immediate physical and social setting" (p. 83), and aimed at measuring the impact of interactional components. Also, the study tracked the use of target linguistic features that received feedback and were modified, and the change of Neil's language through mediation was successfully demonstrated.

Research discussed in this section is concerned with whether the targeted linguistic feature involved in a certain interaction was used in subsequent use, and whether or not it was correctly used. Take an irregular past tense verb, 'ate' in English, for example. The previous studies would have investigated whether the non-targetlike use of the irregular past-tense form involved in the targeted component of interaction (e.g., recast) is used correctly or not in the subsequent use (Figure 2.2).

Figure 2.2
Accuracy as a focus of analysis in subsequent use

Language Related Episode	Subsequent use
Error→feedback→'ate'	Did the learner use 'ate' correctly or not?
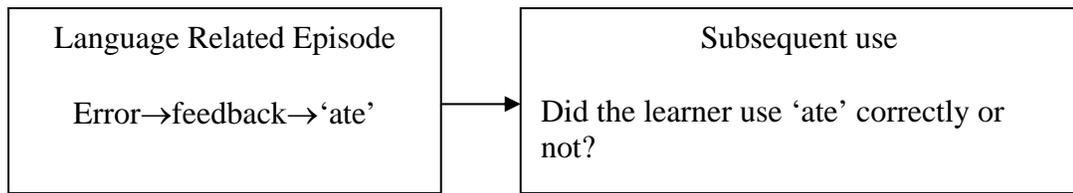
In other words, accuracy in the subsequent use has been the focus, and little attention has been paid to the relationship between the non-targetlike forms learners modify and their developmental aspect in the subsequent use (Figure 2.3).

Figure 2.3
Interlanguage development as a focus of analysis in subsequent use

Language Related Episode	Subsequent use
'*eated'→feedback→'ate' '*ated'→feedback→'ate' What was the type of non-targetlike form ('eated' or 'ated') involved in the language related episode?	Did the learner stop using 'eated,' which he/she had previously modified by 'ate,' or did he/she continue to use it? How about the instance of 'ated'? Is there a relationship between the type of non-targetlike form and their subsequent use?

Maybe, learners who produced modified output on 'eated' (over-generalised –*ed*) are likely to use 'eated' again in the subsequent use while learners who produced modified output on 'ated' (double marking) are not likely to use the same 'ated' in the subsequent use.

Nassaji (2007) pointed out not differentiating the types of linguistic feature targeted by the feedback might contribute to producing the conflicting results on the

effectiveness of interactional feedback. For example, learners perceive morphosyntactic features, less accurately, in comparison with lexis and phonology (Carpenter, Jeon, MacGregor, & Mackey, 2006; Mackey, Gass, & McDonough, 2000; Sheen, 2006). In the category of morphosyntactic features, ESL learners noticed recasts to question formation more than to plurals or past tense because of its perceptual saliency and communicative value (Mackey, 2006). Similarly, there seems to be a need for more research which focuses on the types of non-targetlike form learners modify in order to explore the impact of modified output on L2 learning.

Further investigation with this approach may reveal a more direct relationship between modified output and L2 learning. Questions that need to be asked would include 'Are there any differences among the types of non-targetlike use that are modified?' 'Are there any relationships between the types of non-targetlike use and modification?' and 'Are there any relationships between the types of non-targetlike use learners modify and their subsequent use?'

2.3   Data collection in research on modified output and L2 learning

This section reviews the research literature relating to modified output in relation to features of and issues related to data collection. Specifically, the section will focus on elicitation instruments (2.3.1), frequency of modified output (2.3.2) and the isolated components of interaction (2.3.3).

2.3.1   Elicitation instrument

One of the methodological issues of production-based research is how to elicit the pre-selected linguistic feature in ways that address avoidance strategies (ways through which learners try to avoid using a form in an obligatory occasion). It is possible that creating obligatory occasions for the use of target feature causes interruptions and unnaturalness in communication (Hopkins & Nettle, 1994; Whitlow, 2001), and, therefore, it needs to be developed with consideration of theoretical, practical and pedagogical aspects (see Izumi & Bigelow, 2001 for discussion of the balance). Eliciting modified output is more challenging than that because it cannot be directly manipulated as an independent variable. Also, it requires at least four steps to elicit one token of modified output: provision of obligatory occasions, the initial output with non-targetlike form, noticing the necessity for modification of the initial output, and production of modified output. The choice of an appropriate language activity could promote the occurrences of obligatory occasions, and appropriate feedback, such as focused feedback, can be used to lead to modified output, although it is not possible to control whether learners use non-targetlike forms, or whether and how they respond to feedback.

Implicit methodological techniques

One of the ways to induce modified output is the use of *implicit methodological techniques* (Ellis, 2003), in which feedback is given on learner's use of the target feature while the focus of the task is communication. This technique has been successfully demonstrated in the form of focused feedback using clarification requests (McDonough, 2001; Nobuyoshi & Ellis, 1993; O'Relly et al., 2001;

Takashima & Ellis, 1999).   For example, Nobuyoshi and Ellis (1993) used a

variation called *methodologically focused communication task* for the experimental

group in which a target linguistic feature became a focus of the task from the

viewpoint of teacher/researcher while the focus of participants was still placed on

communication rather than form.   Participants in the experimental group were given

clarification requests by their teacher when they did not use a past tense form

correctly and when they did not use a past tense form on an obligatory occasion.

The control group did not receive any clarification requests after making errors when

using the target feature.   This successfully created the difference in terms of

receiving clarification requests and producing modified output.

Clarification requests as a trigger for modified output

Clarification requests [6] are one type of the implicit feedback, which request a

clarification in some aspects of utterance.   Functions of clarification requests from

listener to the speaker are twofold: signals about the misapprehension or

inappropriateness, and plea for reformulation (Saxton, Houston-Price, & Dawson,

2005).   This feedback does not provide learners with further linguistic information,

unlike other types of feedback such as recasts or metalinguistic feedback.   It has

been shown that clarification requests are effective in leading to modified output,

compared to other types of feedback such as recasts or confirmation checks (Linnell,

---

[6] Several other terms have been used in the literature, such as *indicators* (Varonis & Gass, 1985), *signals* (Pica et al., 1989) and *elicitation* (Nassaji, 2007).   Lyster (2004) classifies clarification requests as a type of *prompts* that includes repetitions, metalinguistic clues and elicitations, all of which are common in withholding correct forms and offer an opportunity for self-repairing.

1995; Lyster, 2004; Lyster & Ranta, 1997; McDonough, 2007; Pica, 1988; Pica et al., 1989).

Clarification requests can be divided into two subcategories, depending on whether the location of problematic utterance is indicated or not. General clarification requests (Example 3) do not specify problematic words/phrases that require clarification, while specific clarification requests (Example 4) highlight the problem.

Example 3: General clarification request

      Speaker: Knights have horse, they do?

      Listener: What?

Example 4: Specific clarification request

      Speaker: Knights have horse, they do?

      Listener: They have what?

                            (both examples cited from Saxton et al., 2005, p. 397)

Specific clarification requests elicit "higher levels of revisions" (Saxton et al., 2005) in comparison with general clarification requests, but the latter have been used in studies on modified output (e.g., McDonough, 2005; Nobuyoshi & Ellis, 1993; O'Relly et al., 2001; Takashima & Ellis, 1999). General clarification requests are ambiguous in terms of the location of problem in the utterance, and they do not provide linguistic information which learners can incorporate into the production of modified output. This is considered to make it possible to investigate the impact of

modified output on L2 learning, excluding the other variables as much as possible.

Lyster (2004) posits three possible advantages learners can gain through receiving prompts including clarification requests: (1) more opportunity for output practice and for proceduralization of newly acquired rule-based representations, (2) drawing the learner's attention to feedback and creating conditions for conscious awareness, which adds efficiency to working memory, and (3) coordination of retrieval of previously disconnected elements (p. 410). Clarification requests could also emphasize the connection between form and meaning (Leeman, 2007), and the benefit of receiving clarification requests has been demonstrated (McDonough, 2005; McDonough, 2007; McDonough & Mackey, 2006; Nobuyoshi & Ellis, 1993).

In the field of L1 acquisition, the role of clarification requests has been studied for more than 25 years (see Saxton et al., 2005, for a comprehensive review of L1 clarification request research). Saxton et al. (2005) looked into the grammar-correcting function of clarification requests. They argue that the role of clarification requests (*clarification questions* in their terms) should not be underestimated, stating that clarification questions "bear the potential to cue recall of previously acquired grammatical forms" (p. 393). The *prompt hypothesis* (Saxton, 2000) predicts a grammar-correcting function for clarification requests, and suggests a possibility that clarification requests can direct children's attention to the ungrammatical aspects of their own utterance if they have linguistic knowledge of the correct grammatical form.

Some L2 interactional studies have found that clarification requests are associated with L2 development (McDonough, 2005; McDonough, 2007; McDonough & Mackey, 2006). However, the effectiveness of clarification requests is still not conclusive. The indirectness and vagueness of clarification requests have been suggested as making it difficult for learners to notice the motivation of such requests (Ammar & Spada, 2006; O'Relly et al., 2001). The need for additional studies for identification of linguistic features that interactional feedback positively affects, and for learners' response to feedback (Mackey, 2007b; McDonough, 2007), and research on the impact of clarification requests on L2 learning is one of the sub-areas for investigation.

2.3.2 Frequency of modified output

Frequency of modified output reported in the previous studies varied. For example, Krashen (1998) argued the frequency of modified output is too low to have some impact on L2 learning, referring to about four instances of modified output per hour in Pica (1988), 12 instances per hour in Van den Branden (1997) and about one per hour in Lyster and Ranta (1997). Shehadeh (2002) argued that frequency of modified output is not necessarily low, referring to research which reported a much higher frequency of modified output (30 instances of modified output in Iwashita (2001) and 60 in Shehadeh (1999)), and pointed out the difference in learners' proficiency levels might be related to the difference in the frequency of modified output. However, Loewen (2002), whose participants were intermediate level,

reported clarification requests led to only 44 uptakes[7] (32 successful uptakes, 12 unsuccessful uptakes) in a total of 1,373 focus on form episodes in thirty-two hours of meaning-focused ESL lessons. These results demonstrate the complex relationship between frequency of modified output and other variables such as proficiency levels and nature of tasks. It is important to note, however, that the results of experimental studies which investigated the impact of modified output on a specific target feature are often based on a low frequency of modified output (e.g., Mdn of correctly reformulated output = 4 in Nobuyoshi & Ellis, 1993; Mdn = 4 in Ellis & Takashima, 1999; Mdn for question formation = 4, Mdn of past-tense = 3 in McDonough, 2001).

There are two positions regarding the relationship between the frequency of modified output and its impact on L2 learning. Krashen (1998) pointed out that the scarcity of output, especially comprehensible output, makes it difficult for output to contribute to linguistic competence. Another position considers that quality is more important than quantity. Ellis (1999) argued that pushed output might be scarce but it "may afford qualitative opportunities to notice specific features that are problematic to learners" (p. 13). Shehadeh (2002) disputed the scarcity argument by Krashen and took the same position as Ellis. McDonough (2001) empirically supports Ellis' claim by demonstrating that 3.0 tokens of modified output (median of the participants who produced modified output above the median of the group) were associated with language development. Also, she found some learners who produced modified

---

[7] The linguistic focus of the episodes covers grammar, vocabulary and pronunciation.

output (modifiers) did not advance to a higher stage in the developmental sequence of question formation while some did advance. There was no apparent difference in linguistic knowledge and quantity of modified output between learners [8] who advanced to a higher stage and those who did not. The difference was found in the quality of modified output, and the modifiers who advanced produced significantly more modified output involving insertion of auxiliary verbs (Example 5) than those who did not advance.

Example 5: Modified output involving insertion of auxiliary verbs

> Learner:　　　　　Where it have a good view?
>
> Native speaker:　Where it is?　Sorry?
>
> Learner:　　　　　Where does it have a good view?
>
> 　　　　　　　　　　　　　　　　　　　　　　　(McDonough, 2001, p. 119)

This suggests a relationship between how learners modified the non-targetlike form and interlanguage development, and supports the importance of the quality of modified output. However, there is still a lack of studies that have investigated the relationship between the quality of modified output and L2 learning. This corresponds with the necessity of research on the subsequent use of modified linguistic feature (section 2.2.3), and, therefore, the investigation of the non-targetlike forms learners modify is an interesting area for further research.

---

[8] This refers to learners who produced modified output above the group median.

2.3.3 Isolation of modified output from other components of interaction

Production of modified output involves other components of interaction. However, most of the studies on modified output do not differentiate it from other components. For example, the results of Takashima and Ellis (1999) were based on the comparison of the impact of modified output combined with other interactional components (i.e., clarification requests, noticing, and comprehensible input).

The study by McDonough (2001, 2005) is one of a few studies that isolated and measured the impact of the two components of interaction, clarification requests and modified output. This was achieved by the interlocutor's continuous talk, often by changing the topic, straight after the feedback, as shown in Example 6.

Example 6

    Learner: What we do with it?

    Native speaker: What we do? Uh let's see well we could talk about the purpose if you want

                                                        (McDonough, 2005, p. 86)

This group (n = 17) was given a total of 72 feedbacks (repetition of the erroneous utterance) but all the opportunities to produce modified output in response to this were prevented. The results did not demonstrate the relationship between clarification requests and L2 language learning. It is not reported whether learners became aware of the intention of the manipulation. The success of this kind of technique will depend on the appropriate choice of target form and the data elicitation

task, without which it would be difficult to keep talking after giving clarification requests so as not to give an opportunity for modification.   Also, it should be noted that isolating the utterances in interaction without looking at language development holistically has attracted criticism (e.g., van Lier, 1996), as "modified output does not occur in vacuum – it occurs as a response to input and to the opportunity to interact" (He & Ellis, 1999, p. 131).

2.4   Summary

This chapter has outlined research on modified output, which indicated that studies on modified output need to seek to overcome the limitations that the seminal study by Nobuyoshi and Ellis (1993) faced.   Also, this chapter reviewed how the relationship between modified output and L2 learning has been measured.   Three types of measurements, grammatical accuracy, interlanguage and subsequent use, were reviewed and the features of each type of analysis were discussed.   What became clear is the potential of the combined use of those analyses to describe the impact of modified output on L2 learning.   There is a lack of research into the various types of developmental non-targetlike patterns that learners may or may not modify and how the use of such patterns may change subsequently.

Relating to data collection in previous studies, the validity of the use of clarification requests to lead to modified output is discussed.   Also, the difficulty in eliciting sufficient tokens of modified output is identified, as this needs to be considered in the development of data elicitation instrument and research design.

In the next chapter, the acquisition of negation in Japanese, the target linguistic feature of the study, is reviewed with reference to the developmental stages based on empirical studies and the Processability Theory (Pienemann, 1998).

CHAPTER THREE

ACQUISITION OF NEGATION OF ADJECTIVES IN JAPANESE

3.0   Introduction

As discussed, the main linguistic feature targeted in *interaction–development studies* (McDonough, 2007, p. 324) has been question formation in English, and it is necessary to investigate another linguistic feature in another language for the generalisability of the findings and for expanding the interaction-development studies.   The target feature needs to be investigated in terms of grammatical accuracy and interlanguage development.   In order to measure the impact of interactional components on interlanguage, the target linguistic feature needs to satisfy two features for describing interlanguage development.   Firstly, that the acquisition of the linguistic feature has been empirically demonstrated to be developmental, and secondly, that its developmental stages have been identified.   Based on these criteria, negation of predicate adjectives in non-past tense in Japanese was chosen as a target feature for the present study.

This chapter starts with a brief grammatical sketch of negation in Japanese, which describes types of adjective (3.1.1), and rules of constructions of negation (3.1.2). This is followed by a review of empirical research on the acquisition of negation (3.2) and of the findings related to developmental stages of the negation of predicate adjectives (3.3 & 3.4). It will also be discussed which of the two developmental stages – those based on empirical findings or those predicted in theoretical framework

(i.e., Processability Theory, Pienemann, 1998) – are appropriate for investigating the relationship between modified output and L2 learning.

3.1   Grammatical sketch of Japanese

Japanese is an agglutinating language, taking an SOV (subject, object and verb) order.   Not only verbs but also adjectives change forms based on tense, negation and other register features including formality.   Unlike English and other European languages, there is no correspondence between the person (i.e., the first-, second-, and third-person), the number (singular and plural) and the predicate forms.

3.1.1   Types of adjective in Japanese

Adjectives in Japanese are categorised into two types on the basis of the morphological classes: *regular adjectives* and *nominal adjectives* [9] .   Regular adjectives end with a non-past tense morpheme *'i,'* and they are often called *i-adjectives*.   Nominal adjectives are often called *na-adjectives* since they require *na* when they are used as a noun modifier.   They are sometimes referred to as *quasi-adjectives* or *adjectival nouns* because the conjugation ending patterns are different from regular adjectives and are identical with those of nouns (Tsujimura, 1996) when used as a predicate followed by a copula *desu* (formal) or *da* (informal).   However, unlike nouns, they cannot become the sentential topic, subject, or object by

---

[9] It varies in the previous studies whether or not the negation of nominal adjectives and nouns were included in the same predicate category for analysis.   For example, Kanagy (1991) put them in the same category because both take the same inflection patterns, while other studies (Hansen-Strain, 1993; Kamura, 2001b) analysed them separately and Kamura found similar developmental patterns between them.   Hayashi (1999) suggests the necessity of further investigation in this point because her research on the regression of negation shows no difference in accuracy in Japanese as a second language between regular adjectives and nominal adjectives.

themselves (Maynard, 1990). There are some nominal adjectives that end with *i*, although the number is small. This often confuses learners, but there are no special rules to distinguish them from regular adjectives. The term *adjectives* will be used hereafter to refer to regular adjectives in contrast to *nominal adjectives*. Table 3.1 shows negation patterns of all the four categories of predicate.

Table 3.1
Non-past negation patterns in terms of predicate category

Predicate category	Formal style	Informal style
Verb (Class 1[1])	(I) do not/will not write	
kakimasu (formal) *kaku* (informal) '(I) write/will write'	V[2] irrealis[3] + masen *kaki masen*	V irrealis + *nai* *Kaka nai (desu)*
Verb (Class 2)	(I) do not/will not see	
mimasu (formal) *miru* (informal) '(I) see/will see'	V irrealis + *masen* *mi masen*	V irrealis + *nai* *mi nai (desu)*
Regular adjective	(It) is not cold	
samui desu (formal) *samui* (informal) '(It) is cold'	A root + *ku arimasen* *samu ku arimasen*	A root + *ku nai* *samu ku nai (desu)*
Nominal adjective	(It) is not quiet	
shizuka desu (formal) *shizuka da* (informal) '(It) is quite'	NA root + *dewa arimasen*[4] *shizuka dewa arimasen* NA root + *ja arimasen* *shizuka ja*[5] *arimasen*	NA root + *dewa nai* *shizuka dewa nai (desu)* NA root + *ja nai* *shizuka ja nai (desu)*
Noun	(It) is not an airplane	
hikooki desu (formal) *hikooki da* (informal) '(It) is an airplane'	N + *dewa arimasen* *hikooki dewa arimasen* N + *ja arimasen* *hikooki ja arimasen*	N + *dewa nai* *hikooki dewa nai (desu)* N + *ja nai* *hikooki ja nai (desu)*

Note 1. There are three classes of verbs, and Class 3 verbs (irregular verbs) are not included in the table.
Note 2. V = verb, N = noun, NA = nominal adjective, and A = adjective
Note 3. Irrealis: the non-infinite form in negation (*mizen-kei* in Japanese grammar terminology). For verbal roots ending in consonants (Class 1), it is the root plus '*a.*' For verbal roots ending in vowels (Class 2), it is the same as the root (Sano, 1998, p. 4).
Note 4. Negative morphemes appear as part of copula in the case of nouns and nominal adjectives.
Note 5 Ja is a contraction of the more formal *dewa* (Kanagy, 1991).

Most adjectives can be used both in an attributive position and as a predicate.

(1) *Samu-i fuyu*   'Cold winter'

     cold-winter

(2) *Fuyu wa samu-i*   'Winter is cold'

     winter-TOP-cold-NONP

Japanese adjectival phrases consist of a root and an inflectional ending which forms the stem of the adjectives, together with optional auxiliaries (see Figure 3.1).

Figure 3.1
Adverbial phrases

                 Root — Inflectional ending (+ Auxiliary)
                     └_____┘
                     Stem

                                  (adapted from Shibatani, 1990, p. 224)

The predicate alone can constitute a complete sentence, with the subject and other elements omitted when they can be understood from context (Kanagy, 1991, p. 28), and, therefore, *samu-i* (cold-NONP) is a complete sentence meaning '(It) is cold.' Like verbs and auxiliary verbs[10], adjectives are inflected according to tense, mood, negation and politeness levels (i.e., formal or informal). The inflectional paradigms of adjectives (tense and negation) in an informal style and formal style are presented in Table 3.2.

---

[10] Auxiliary verbs do not inflect alone (Shibatani, 1990).

Table 3.2

Inflectional paradigms of adjectives in an informal style and formal style

Non-past tense		Past-tense	
samu-i cold-NONP[1]	'(It) is cold'	*samu-kat-ta* cold-PAST	'(It) was cold'
samu-i desu cold-NONP-POL	'(It) is cold'	*samu-kat-ta desu*' cold-PAST-POL	(It) was cold'
samu-ku na-i (desu) cold-Adv NEG-NONP(-POL)	'(It) is not cold'	*samu-ku na-kat-ta (desu)* cold-Adv NEG-PAST (-POL)	'(It) was not cold'
samu-ku arimase-n[2] cold-Adv POL-NEG	'(It) is not cold'	*samu-ku arimasen deshita* cold--Adv NEG-POL-PAST	'(It) was not cold'

Note 1. NONP = non-past, which "expresses present, habitual, and future actions and events" (Tohsaku, 1994, p. 187). NEG = negator, POL = politeness/formality, Adv = adverbial inflector (adapted from Kanagy, 1991; Tsujimura, 1996, p. 146).

Note 2. *Samu-ku na-i desu* and *Samu-ku arimase-n* ('(It) is not cold') are "identical in meaning and politeness level" (Tohsaku, 1994, p. 133).

3.1.2 Rules of construction of negation in Japanese

Negation in Japanese is post-verbal while that in English is pre-verbal. The negative construction in Japanese involves a negative morpheme which is always bound to other morphemes, such as tense (non-past or past) and politeness level (formal or informal). Following the previous studies on negation (Kamura, 2001b; Kanagy, 1991, 2001), the term *negation pattern* is used hereafter to refer to the bound morphemes of negation and others. The negation patterns are different among the predicate categories. For example, *ku-na-i* (hereafter, shown as *kunai* as a negation pattern) is used for negating adjectives in an informal style while *ja-na-i* (hereafter *janai*) is used for nominal adjectives and nouns.

Construction of negation of predicate adjective follows three steps: (1) identification of *samu-i* as an adjective, (2) replacement of the *i* (non-past tense morpheme) with an

adverbial inflector *ku,* and (3) addition of a negation pattern *nai* (a negative morpheme *na* + a non-past tense marker *i*). The correct negation form is *samu-ku-na-i,* '(It) is not cold' shown in Figure 3.2. Several variations have been reported in both L1 and L2 studies, and they include non-targetlike patterns caused by the overgeneralised use of a negation pattern to other predicate categories, or by the failure to inflect the stem.

Figure 3.2
Tree structure of 'samu-ku + na-i'

samu-ku + na-i (cold-Adv.+ NEG-NONP, '(it) is not cold')

```
                        AP
                       /  \
                      /    \
                A stem      Aux
                   |          |
                   |          |

              samu-ku    +    na-i
```
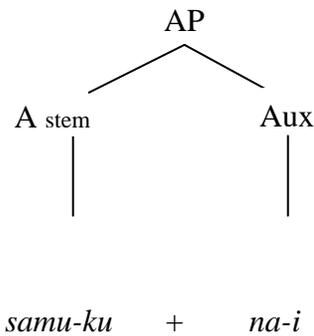
(The plus sign '+' indicates the boundary between the inflectional ending and the auxiliary suffix proper)

3.2   Acquisition of negation in Japanese

3.2.1   Acquisition of negation in L1

The acquisition of negation in Japanese has been found to be developmental (e.g., Clancy, 1985; Ito, 1981, 1990; Okada, 2002; Sano, 1998). L1 Japanese children start with one-word stage in which expressions for negation (*nai*), rejection (*iya*), prohibition (*dame*), non-existence (*nai/inai*) and denial (*chigau*) are randomly used (Clancy, 1985). This is followed by a stage, X + *nai* pattern (Clancy, 1985) where they add the negation pattern *nai* irrespective of predicate categories (Clancy, 1985; Ito, 1990). This is similar to the external negation in L1 English Klima & Bellugi (1966) described as Stage 1 (e.g., 'No mitten,' or 'Wear mitten no'). In the next stage, multiple uses of unanalysed negation patterns affixed to unanalysed predicate categories are observed, followed by a stage when the correct negation patterns are affixed to the appropriate stem form.

3.2.2   Acquisition of negation in L2 Japanese

A number of longitudinal and cross-sectional studies on negation in L2 Japanese have been conducted (Hansen, 1999; Hansen-Strain, 1993; Hayashi, 1999; Kamura, 2001a, 2001b, 2003; Kamura & Sakoda, 2001; Kanagy, 1991, 1994, 2001; Komori & Sakano, 1988; Noro, 1995; Takeuchi-Furuya, 1993). The findings include that the acquisition of the negation of adjectives is more difficult than that of other predicate categories (Hansen-Strain, 1993; Kamura, 2001b; Kanagy, 1991; Takeuchi-Furuya, 1993), and that acquisition of past tense negation, particularly that of adjectives, is reported to be more difficult than that of non-past tense (Hansen-Strain, 1993; Hatanaka, 1996, cited in Kanagy, 2001; Kamura & Sakoda, 2001; Kanagy, 1991).

Hansen (1993) investigated the attrition patterns of negation of 24 American high school students over a three month break, and her data did not support that JSL learners followed a common sequence in the acquisition of negation because of an enormous individual variation in the use of negation patterns.   However, the majority of research in this area supports that L2 Japanese learners go through stages in the process of acquisition of negation (Kamura, 2001a, 2001b, 2003; Kamura & Sakoda, 2001; Kanagy, 1991, 1994, 2001; Komori & Sakano, 1988; Noro, 1995; Takeuchi-Furuya, 1993).   Kanagy's (1991) pioneering research on L2 negation in Japanese demonstrated clear developmental sequences in the acquisition of negation, and that negation patterns are acquired in the following order: (*nai*)[11] *> janai > masen > kunai* (*kujanai*) > (*masen deshita*).   Also, Kanagy (2001) presented three developmental stages of negation common to all the categories (nouns including nominal adjectives, verbs, and adjectives).   Stage 1 is characterised by *externalised negation*, in which predicates are unanalysed and a fixed negation pattern is attached irrespective of categories.   Stage 2 is *variable negation* that is characterised by one or more unanalysed negation patterns with partial or no analysis of the predicate categories. Stage 3 is *internalised negation* when appropriate negation patterns are attached to inflected stems.

Kamura (2001b) described the developmental stages of each predicate category, based on the data collected from five individual interviews (every four to six weeks apart) with eleven L2 Chinese learners of JSL.   The variation of non-targetlike negation patterns showed a gradual decrease for verb negation (i.e., *dictionary form

---

[11] Brackets indicate the use of the negation pattern by only a few learners.

+ *nai* > **janai* > confusedly conjugated patterns > targetlike use), and so did negation

of nouns and nominal adjectives (**nai* & **kunai* > targetlike use).

3.3   Development of negation of adjectives in Japanese

Clancy (1985) proposed a hypothetical developmental sequence of negation of

adjectives in L1 Japanese, based on her data and other research. In the first stage,

children negate adjectives by adding *nai* to the adjective's stem and produce a non-

targetlike form such as **atsui-nai.* Clancy stated that the use of *nai* instead of the

correct negation pattern, *kunai,* is associated with perceptual salience, as *ku* is

"semantically opaque formative sandwiched between the recognizable adjectival root

and *nai*" (p. 403). In the second stage, the correct negation pattern, *kunai*, is

acquired, but it is affixed to the uninflected stem of adjectives (**atsui-kunai*) rather

than inflected stem of adjectives (*atsu-kunai*). This is followed by the third stages, in

which the analysis of negation patterns and inflection of adjectives (i.e., segmentation

of adjectives, root + *i*) are achieved.

The first stage of negation of adjectives in L2 Japanese has produced mixed results in

terms of the first stage. Kanagy (1991) found the use of *nai* as the first stage,

followed by more types of negation patterns in her JSL study[12], which is similar to

the first stage after one-word stage in L1 research. On the other hand, Kamura

(2001b) found only a small number of adjective + *nai* patterns. She also found a

trend of gradual decrease in the variation of non-targetlike patterns of adjectives.

---

[12] Kanagy (1991) was concerned with the acquisition of negation patterns among all the categories, and was not concerned with the development by each category.

The more language development progresses, the fewer the variation of non-targetlike patterns learners commonly use becomes, and there was a disappearing order, *nai > *janai > targetlike use. She considers the use of multiple types of non-targetlike pattern as evidence that L2 learners test hypotheses from an early stage of learning negation. Noro's (1995) data, collected from a Chinese-speaking child, showed only two instances of *nai* in 44 non-targetlike negation patterns out of 273 instances of negation of adjectives. The number of instances of *nai* in other predicate categories was also very few: two out of 15 non-targetlike negation patterns in the negation of nominal adjectives, one out of seven in the negation of nouns, and nine out of 38 in verb negation. Matsumoto (1999, cited in Kamura, 2001b) reported that the X + *nai* pattern was not observed at all in her data of a longitudinal study of a Chinese-speaking child of JSL.

The findings on *janai* are consistent in that it is likely to remain until the end of the second stage. However, the order of disappearance between *janai* and uninflected or incorrectly inflected or stem + *kunai* is not conclusive. *Janai* is a pattern for negating nouns and nominal adjectives, but non-targetlike use occurs when *janai* is used for negating adjectives. *Kunai* is the correct pattern for negating adjectives, but non-targetlike use involving inflectional errors such as *samu-i-kunai* (i.e., failure to drop *i*) for *samu-kunai* is observed. Kanagy (1994, 2001) reported the overuse of *janai* in the early stages and that it disappeared earlier than *kunai*. Noro (1995) found that *janai* emerged at the same time as the correct negation pattern, and it decreased earlier than *kunai*. On the other hand, Kamura (2001b) found *janai* continued to be used until the last stage. Takeuchi-Furuya (1993) reported that

janai appeared earlier than *kunai* and disappeared later than *kunai*. These results indicate that dropping *i* is more difficult than simply adding *janai* to the uninflected stem of adjectives. They also suggest that learners might find it more difficult to distinguish adjectives from noun/nominal adjectives than inflecting *i* to *ku*.

The different results in previous studies might be associated with differences in longitudinal and cross-sectional approaches, learning environment (i.e., JSL & JFL), and the learners' prior knowledge of negation in all or some categories among studies, and Table 3.3 presents the developmental stages of negation of adjectives, based on the findings common in the previous studies (Kamura, 2001b; Kanagy, 1991, 2001; Komori & Sakano, 1988; Noro, 1995; Takeuchi-Furuya, 1993).

Table 3.3
Developmental stages of negation of adjectives in Japanese

Stage	Feature	Explanation	Examples
Stage 3	Targetlike pattern	The choice of the negation pattern is correct and the inflection of the stem is correct.	*samu kunai*
Stage 2	Absence of the non-targetlike pattern *nai*	*Nai* is not used any longer but other types of non-targetlike pattern are still used. (There might be another stage between Stage 2 and 3, which features the disappearance of *janai* or *kunai*, but the findings are mixed regarding the disappearing order.)	Variants except *samu(i) nai* e.g., *samu(i) janai* *samu(i) kunai* *samu(i) masen* *samu(i) kujanai*
Stage 1	Use of the non-targetlike pattern *nai*	Adj root + (*i*) *nai* are used. Targetlike patterns as well as other types of non-targetlike pattern may be used.	*samu(i) nai* other variants

Most of the studies that investigated the developmental stages of negation used frequency of types of non-targetlike use for identifying developmental stages. It is

suggested that the use of frequency to determine the developmental stages can be misleading because of variability of learner language, and the use of statistical techniques has been suggested as one way to deal with it (Ellis & Barkhuizen, 2005). However, the sample size in these studies is usually small because negation is divided into three or four predicate categories, and it is often difficult to carry out statistical tests. Also, it should be noted that no studies on L2 Japanese negation used emergence criteria, probably because learners in previous studies have been observed to use targetlike negation patterns as well as non-targetlike negation patterns from the beginning of the data collection. Therefore, some caution is needed when the stages proposed in these studies are used as a base to measure the impact of interaction on L2 learning.

3.4 Applicability of the developmental stages of negation of adjectives predicted by the Processability Theory

As discussed in Chapter Two, interactional research often used the developmental stages identified by Pienemann and Johnston (1987) and Pienemann, Johnston, and Brindley (1988), in order to measure the impact of interaction in L2 learning. These studies are the foundation and predecessor of Pienemann's (1998) Processability Theory (PT), which aims to determine and explain what causes the development of linguistic competence to follow a describable route. Its basis is the architecture of human language processing which constrains the availability of psycholinguistic resources required for carrying out linguistic processing. PT proposes a hierarchy of language processing procedures, on which basis, the processability of linguistic structures can be predicted**.**

However, Pienemann (1998) stated "the learning of the morphological form of the affix constitutes a task that is different from managing the information distribution in the affixation process where diacritic features have to be exchanged within different grammatical structures" (p. 154). Kawaguchi (2005), in reference to a study by Di Biase and Kawaguchi (2002), also stated "we assumed that agglutinating morphology is controlled by principles different from syntactic principles, hence we excluded the development of agglutinating morphology from the scope" (p. 257). Thus the developmental stages of negation of adjectives in Japanese, which involves morphological change with the appropriate combination of stem and affix, may not be within the scope of original PT.

Nevertheless, it seems meaningful to illustrate the predicted stages of negation of adjectives, to point out some issues in employing it for achieving the aim of the current study, and to discuss the possibility of employing the stages for measuring the impact of modified output, because the results would be more persuasive if the same framework as previous interactional research used is employed.

Adjective morphology in Japanese is similar to verbal morphology in Japanese, and, therefore, the hypothesized developmental stages of negation of predicate adjectives (non-past tense in informal style) in Table 3.4 are based and developed on the stages of verbal morphology that Kawaguchi (2000) identified and empirically tested. The procedural skills the learner needs to build up at each stage are shown in the first column, and the types of grammatical information exchanges of each stage are displayed in the second column. The original hierarchy that Pienemann (1998)

66

proposed has six processing procedures but Stages 5 (S-procedure/word order rules –

saliency procedure) and 6 (Subordinate-clause) were not applicable to morphology as

Stage 4 (inter-phrasal morpheme) is the highest morphological stage (Kawaguchi,

2000, p. 242) and the negation of adjectives does not involve those processing

procedures.

Table 3.4

Hypothesized hierarchy for negation of adjectives in L2 Japanese predicted by the Processability Theory

Stage & processing procedure	Exchange of information	Negation of adjectives
4. Simplified sentence procedure (word order rules + saliency)	Inter-phrasal information (Inter-phrasal morphemes)	*samu-ku-na-ke-re-ba* (*If it is not cold*)
3. Phrasal procedure	Phrasal information (Phrasal morphemes)	*samu-kunai* *samu-ku-arimasen*
2. Category procedure	No information exchange (Lexical morphemes)	*samu-i-desu* **samu-i-kunai* (analysed or unanalysed) **samu-i-janai* **samu-i-nai* **samu-i-masen*
1. Word/lemma	No information exchange	*samu-i, desu, ka, nai, janai, arimasen, kunai,* etc.

Note. Stage 2 for negation of adjectives includes ungrammatical phrases, which are marked with asterisks.

Negation of adjectives in Japanese involves two processing procedures, category

procedure and phrasal procedure in PT. Ellis (2007), who investigated effects of

feedback on the regular past tense '-ed' and the comparative '-er' in English, would

be helpful to illustrate the difference between them. While attaching the

morphological marker '-ed' to the verb does not involve information exchanges (i.e.,

category procedure), the use of the comparative '-er' does (i.e., phrasal procedure). The procedure of each lower level is considered to be the pre-requisite for the functioning of the higher level, and category procedure is lower than phrasal procedure. Therefore, the past tense '-ed' is predicted to be acquired before the comparative '-er.'

The morphological principle of negation of adjectives in Japanese goes through the category procedure illustrated by the example of regular past tense '-ed,' followed by the phrasal procedure illustrated by the comparative '-er.' In other words, the construction of negation of adjectives requires attaching a negation pattern to the adjective (no information exchange), followed by inflection of the stem (information exchange).

If these stages predicted within PT are used to measure the impact of interactional components, the research will share the same framework in terms of developmental sequence based on processing procedures and the findings can contribute to an understanding of the relationship between interaction and interlanguage development. However, there are at least three limitations with the predicted stages for interlanguage analysis within the context of negation of Japanese adjectives.

Firstly, it does not predict the relative difficulty among the combination of adjectives and non-targetlike negation patterns (i.e. which pattern, *janai* or *kunai* emerges or disappears first). This is because their processing procedures fall within the same stage. Secondly, the difficulty order between inflected stem and uninflected stem

has not been observed in empirical studies. In PT, the inflection of stem from A-NONP to A-Adv (e.g., *samu-i* to *samu-ku*) is considered to involve information exchange and it is a phrasal procedure. On the other hand, the use of an uninflected stem is a category procedure where no exchange of information is required. However, the empirical studies have not suggested that an inflected stem with a negation pattern emerged after an uninflected stem with a negation pattern. For example, Kamura (2001b) did not find any patterns, such as **samui-janai > *samu-janai > samu-kunai*. The prediction of PT might be supported with more substantial data, or other modules need to be found to explain agglutinating phenomena, such as stem-suffix combinations with morphophonemic variation as Kawaguchi (2005) suggested. The third issue is the criterion used to determine whether a learner reaches a new stage. The interactional research mentioned above chose 'emergence' as a criterion as evidence of reaching a higher stage. However, it has been found that negation of adjectives in L2 Japanese starts with targetlike use as well as non-targetlike use (section 3.3), which indicates that learners are not constrained by processing procedure, and emergence criteria are not appropriate.

In conclusion, the developmental stages predicted by PT have provided conceptual framework for many studies (e.g., Mackey, 1999; McDonough, 2005), and there were *prima facie* grounds to suppose it could be applicable to the investigation of negation of adjectives. However, within the limited scope of the present study and for the reasons stated above, it seems to be inappropriate for achieving the purpose of the current study.

3.5   Summary

This chapter briefly illustrated constructions of adjectives and negation in Japanese, following this by a review of the literature relating to the acquisition of negation in Japanese.   The developmental stages of negation of adjectives found in empirical studies were described, and the developmental sequence was presented, based on the common findings in the previous studies.   They are characterised by the reduction of variation of non-targetlike negation pattern in the order of *nai* > *janai* > targetlike use.   Developmental stages of negation of adjectives predicted by the Processability Theory were also discussed.   Although this theory has been supported by empirical evidence in a number of languages, it was found that some kind of modification to the theory might be necessary to deal with a morphological linguistic feature in an agglutinating language – the negation of adjectives in Japanese.   Therefore, the developmental stages based on empirical findings were chosen rather than theory-driven developmental stages for measuring the impact of production of modified output.   Chapter Four following restates the research questions and explains the data collection and analysis procedures used in the present study.

CHAPTER FOUR

METHODOLOGY

4.0   Introduction

This chapter restates the research questions and presents the associated hypotheses (4.1 & 4.2). The target feature and participants in this study are described (4.3 & 4.4). The experimental design, data collection procedures, testing instrument and treatment sessions are explained in detail in 4.5 to 4.8 respectively. Data coding and scoring are presented (4.9).

4.1   Research questions

The present study investigates whether or not the production of modified output facilitates L2 learning. Modified output produced specifically in response to clarification requests is referred to as 'MO' in order to differentiate from modified output in general. The impact of MO on L2 learning is measured in the three aspects of learner language discussed in Chapter Two: grammatical accuracy, language development and subsequent use of the types of non-targetlike negation patterns learners previously modified. The research questions addressed are:

Research Question 1: Is there a positive relationship between the production of MO and grammatical accuracy?

Research Question 2: Is there a positive relationship between the production of MO and interlanguage development?

Research Question 3: Is there a relationship between the types of non-targetlike forms that the learners modify and their subsequent use in terms of accuracy and interlanguage development?

4.2   Hypotheses

The following predictions are generated in relation to the impact of MO within the framework of the output hypothesis.

Hypothesis 1.1: The experimental group will show greater gains in grammatical accuracy than the control group.

Hypothesis 1.2: There will be a relationship between the production of MO in the treatment sessions and gains of grammatical accuracy scores in the post-test and delayed post-test.

Hypothesis 2.1a: The variations of non-targetlike negation patterns will diminish in both the experimental and the control groups.

Hypothesis 2.1b: The decrease in the variation of non-targetlike negation patterns will be greater in the output of the experimental group than in that of the control group

Hypothesis 2.2: There will be a difference in the use of types of non-targetlike pattern between the experimental group and the control group as a result of differences associated with the production of MO.

Hypothesis 2.3: The experimental group will outperform the control group in terms of the number of participants who use higher stage forms.

Hypothesis 3.1: Participants who produce MO in the treatment sessions will be sensitised to avoid the same non-targetlike pattern in subsequent situations of use.

Hypothesis 3.2: There will be a relationship between the types of non-targetlike negation pattern participants modify and their subsequent use.

4.3   Target feature

Negation of predicate adjectives in the non-past tense in Japanese was chosen as a target feature of the study for the three reasons. Firstly, the acquisition of negation has been found to be developmental, and the impact of production of MO might be described as changes of interlanguage within the developmental sequence. Secondly, acquisition of negation of adjectives takes longer than that of verbs, nouns and nominal adjectives (section 3.2.2). Therefore, the period of being at the same stage in the negation of adjectives was considered to be longer than that in other categories, and the impact of MO, if any, may be more observable than other predicate categories. Thirdly, the target features of interlanguage analysis in this context have been limited to English (mostly question formation with a few exceptions such as and regular past tense '-ed' and the comparative '-er' by Ellis, (2007)). Also, there are very few studies that have investigated the relationship between modified output and L2 learning in Japanese languages, while Japanese has been targeted in interactional research (e.g., Ishida, 2004; Iwashita, 2003; Long, Inagaki, & Ortega, 1998; Loschky, 1994). The use of negation in Japanese as a target feature in the same research context will contribute to the generalisability of the findings in previous studies.

4.4 Participants

4.4.1 Recruitment of participants

Participants were recruited from among students who enrolled in Japanese courses at the University of Waikato in New Zealand. Those students who agreed to take part in the project were asked to sign a consent form indicating their agreement to having their production data audio-recorded and to its use for research purposes. A questionnaire survey was conducted to get some background information about participants. Copies of the letter of request for participation, consent form and questionnaire are included as Appendices 4.A, 4.B and 4.C.

The present study employed a randomised block design. A total of 52 participants took the pre-test, and they were matched on the basis of the accuracy scores on the target form in the pre-test, and each member of the pair was randomly assigned to an experimental group or to a control group. Interactional research employing interlanguage analysis often assigns participants to groups based on the targeted developmental stage (e.g., McDonough, 2001), but the present study chose accuracy for the purpose of comparison with the studies that have used accuracy as a measurement for the impact of MO (Nobuyoshi & Ellis, 1993; O'Relly et al., 2001; Takashima & Ellis, 1999). The participants were recruited from beginner and intermediate courses, and twelve participants who scored 90%[13] or higher in the pre-test were excluded from the data analysis. It was also considered whether to exclude participants whose accuracy score was less than 20%, or who did not use the target

---

[13] The accuracy percentage has been used widely as a criterion of acquisition in L1 and L2 since Brown (1973), but it varies among research, and these cut-off accuracy percentages (e.g., 90% in Dulay & Burt, 1974; 75% in Ellis, 1988) are arbitrary.

form more than a certain number of times (e.g. three times) to minimise the possible variability of performance. However, it was decided not to, as the total number of participants was small and participants were matched with the results of accuracy scores in the pre-test. More importantly, greatest improvement in the interlanguage was expected among those participants. Another twelve participants were excluded from the analysis because they did not complete all of the required sessions or withdrew half way through. The spoken data of 28 participants (14 each in the experimental and control group) were analysed for the study.

4.4.2   Background of participants

Some of the key characteristics of the participant sample obtained from the questionnaire are presented in Table 4.1. The participants consisted of 14 male and 14 female participants (by coincidence), and their ages (except for three participants who did not give their ages) ranged from 17 to 46 years, with an average of 23.2 years (SD = 8.1). Fifteen participants were from JAPA132 (second level of the beginners' stream, five months into their study at university), three from JAPA102 (second level of the main stream, pre-requisite of five year or equivalent study prior to university, five months into their study at university), and two from JAPA 232 (fourth level of the beginners' stream, one year and five months into their study at university). The other eight participants were from JAPA101 (first level of the main stream, two weeks into their study at university), the data collection for which was conducted four months later than the others.

The average length of Japanese language learning at secondary school was one and a half years for both groups (SD = 2.14 for the experimental group and 2.07 for the control group), and that at university was 8.9 months (SD = 6.47) for the experimental group and 5.6 months (SD = 4.60) for the control group. The first languages of the participants were English, Taiwanese, Chinese, Croatian and French. Twelve participants reported having been to Japan previously, and two of the participants had stayed in Japan longer than two months.

Table 4.1
Summary of participant information

Experimental group

Participant[1]	Course	Country of origin	First Language	Sex	Age	Been to Japan?	School[2]	University[3]
1	101	NZ	English	F	18	No	4 years	1 month
2	132	NZ	English	M	23	No	5 years	7 months
3	232	NZ	English	M	19	No	0	18 months
4	101	China	Chinese	F	17	No	4 years	1 month
5	101	NZ	English	F	19	Yes (5months)	3 years	1 month
6	132	Taiwan	Chinese	M	21	Yes (1 month)	0	7 months
7	132	France	French	M	21	Yes (2 weeks)	0	7 months
8	102	China	Chinese	F	-	No	0	18 months
9	102	NZ	English	F	18	Yes (1 month)	5 years	18 months
10	132	China	Chinese	M	22	No	0	7 months
11	232	Taiwan	Taiwanese	F	24	Yes (2 weeks)	0	18 months
12	132	NZ	English	M	20	No	0	7 months
13	132	China	Chinese	F	22	No	0	7 months
14	132	NZ	English	M	22	No	0	7 months

Control Group

1	101	NZ	English	F	18	Yes	2 years	1 month
2	102	NZ	English	F	-	No	0	18 months
3	101	NZ	English	F	18	Yes (2 weeks)	5 years	1 month
4	101	NZ	English	M	18	Yes (2 weeks)	5 years	1 month
5	132	Taiwan	Chinese	M	20	No	1 years	7 months
6	132	Korea	Korean	F	27	Yes (2 weeks)	0	7 months
7	132	NZ	English	F	46	No	4 years	7 months
8	132	China	Chinese	M	21	No	0	7 months
9	101	NZ	English	M	18	Yes (2 weeks)	4 years	1 month
10	132	NZ	English	M	41	Yes (1 month)	0	7 months
11	132	NZ	English	M	18	Yes (2 weeks)	0	7 months
12	101	NZ	English	M	24	Yes (2months)	0	1 month
13	132	Korea	Korean	M	44	No	0	7 months
14	132	Croatia	Croatian/English	F	-	No	0	7 months

Note 1. The participants in each group are numbered and listed based on the accuracy scores in the pre-test.

Note 2. The column 'School' shows the length of learning Japanese at secondary school.

Note 3. The column 'University' shows the length of learning Japanese at university.

4.5   Experimental design

The design used here was a quasi-experimental design, including a pre-test, post-test and delayed post-test with randomisation of groups (see Figure 4.1 following for an overview).   The framework of a cross-sectional approach was built upon the previous studies examining modified output with an implicit methodological technique (section 2.3.1).   The aspect of the longitudinal approach was developed on the studies on developmental stages of negation in Japanese (e.g., Kamura, 2001b; Kanagy, 1991).   The experimental sequence of the present study was over a period of two months, totalling approximately 44 hours.   The independent variables are the treatment (i.e. clarification requests: two levels between subjects), and the timing of the tests (pre-, post-, and delayed post-tests: three levels within subjects).   The dependent variables were accuracy scores, use of non-targetlike negation patterns, and developmental changes in the use of negation patterns.

4.6   Data collection procedures

4.6.1   Pre-test

Computer-administered picture description speaking activities were used as the pre-test, which is described in section 4.7.

4.6.2   Treatment sessions

Two treatment sessions were conducted.   The first treatment session was held two weeks after the pre-test, and the second treatment session was carried out one week after the first treatment session.   Each participant in both the experimental group and the control group had an interview of approximately 20 minutes as the first treatment,

78

and a 15-minute session of mechanical drills as the second treatment. The difference in the treatment session between the groups was that the experimental group had a methodologically focused interview and drills while the control group had methodologically unfocused ones in which no clarification requests were given to non-targetlike forms. The details of the procedures of treatment sessions are reported in section 4.8.

4.6.3   Post-test and delayed post-test

A post-test was conducted one week after the second treatment, and a second post-test (i.e., delayed post-test) was held three or four weeks after the first post-test. The main aim of the delayed test was to examine whether any change in the post-tests was temporary or sustained. Also, the change of interlanguage might be observed not in the post-test but in the delayed post-test as shown in previous studies (e.g., Ellis, Loewen, & Erlam, 2006; Mackey, 1999).

Figure 4.1
Overview of the research design

Pre-test (Week 1)
Computer-administered picture description activities

⇩　　⇩

Pair matching based on the results of the pre-test and random assignment into two groups

⇩　　⇩

Experimental Group	Control Group

⇩　　⇩

Treatment session 1 (Week 3) A methodologically focused interview Clarification requests were given when a non-targetlike use on the target feature was observed.	Treatment session 1 (Week 3) A methodologically unfocused interview Clarification requests were not given except after genuine communication breakdowns.

⇩　　⇩

Treatment session 2 (Week 4) Methodologically focused mechanical drills Clarification requests were given when a non-targetlike use on the target feature was observed.	Treatment session 2 (Week 4) Methodologically unfocused mechanical drills Clarification requests were not given except after genuine communication breakdowns.

⇩　　⇩

Post-test (Week 5) Computer-administered picture description activities

⇩

Delayed Post-test (Week 8 or 9) Computer-administered picture description activities

4.7   Testing instrument

A set of two computer-administered speaking activities was used for the pre-test.   It was carried out in a language laboratory equipped with an individual microphone and a computer screen for each participant.   The participants individually performed two picture description activities.   They were instructed to reply orally to 40 questions (20 questions to elicit negation of adjectives as a target form, 10 distractor questions to elicit affirmative adjectives, and another 10 questions to elicit non-adjectives).   All the instructions were provided in English, but they heard the questions in Japanese as pictures were presented on a computer screen.   Participants could go to the next question by clicking the mouse, but the computer automatically presented the next question and a picture after one minute.   The activities consisted of vocabulary and structures that had been taught in class prior to the pre-test.   Complete transcriptions of the questions used for the tests are given in Appendix 4.D1, and part of instructions and pictures are found in Appendix 4.D2.   It took approximately 20 minutes to complete and the performance was audio-recorded.   The response of the participants was transcribed and the researcher and a research assistant individually coded them.

An identical test was used for pre-test, post-test and delayed post-tests.   The use of the same test may possibly affect validity because of task repetition effects; if participants repeat the same task twice or more, they are likely to take advantage of the familiarity with the task format, thus paying more attention to form and improving performance.   Bygate (2001), for example, reports a participant's lexical selection, grammatical items, and her ability to self-correct were better on repetition

of the task. Also, task repetition can be considered from the perspective of automatisation. Repeated use of a structure can lead to smoother and faster access to the integration of knowledge, and help the transfer from controlled to automatic processing. After consideration, it was felt that such potential threats to validity were outweighed by the greater reliability ensured by the use of tests consisting of the identical lexis and syntactic structures, and the necessity to construct tasks within the participants' limited proficiency in Japanese. The production of the target forms on the same lexical item was compared directly between the tests, which solved one of the critical issues of the comparability of the three tests, and subtle developmental change of negation patterns could be investigated.

Changing the order of questions or the lexical items in the pre-, post- and delayed post-tests was one of the ways to minimise the practice effects, but it was not adopted because there was a possibility that the changes (e.g., question order) in the tests might influence the types of negation patterns to be used. For example, learners might use the negation pattern for nouns when they needed to use negation patterns for adjectives merely because they used noun negation one question before. It would have been desirable if the order among verbs, adjectives and nouns/nominal adjectives had been consistent across tests with different lexical items, but this was not possible because of the limited vocabulary learners were familiar with.

All the participants went through the same procedures in the present study, and were not informed of the repetition of the tests to minimise the possibility of their revising the target forms during the interval. Frequency distribution measures (quantitative

changes in interlanguage), which the present study employed as one way of analysis (section 4.9.3), have been pointed out to be "highly sensitive to the tasks or to the segment of discourse" (Ishida, 2004), and, the use of the identical tests was a way to deal with the sensitivity of the tasks.

One of the possible concerns with beginner level participants was the low frequency of the production of MO because their internalised linguistic knowledge might be just sufficient to perform given tasks but not sufficient to notice the use of non-targetlike forms and modify them. Kanagy (1991) reported participants in the lower proficiency levels in her study occasionally failed to comprehend the interview questions, and she suspected it might have affected the production of the target form. A series of pilot studies for the present study, which used past-tense form of adjectives in Japanese as a target form, also showed that low proficiency learners did not produce many instances of modified output per error, probably because they did not have sufficient familiarity with the target form even though it had been taught in class. However, beginner level participants were still considered appropriate because this is the level in which the development of negation over all the four predicate categories makes greater progress than learners at intermediate or advanced levels, as shown in Kanagy (2001). Therefore, the potential lack of vocabulary and linguistic knowledge that might affect their performance was dealt with by restricting the vocabulary and structures in the data elicitation instrument to what they were familiar with, although this resulted in limiting the variation of activities in the treatment sessions.

In order to ensure consistency in the procedures and the validity of the instrument for the main study, the testing activities were piloted with a small number of learners of Japanese (n = 2) and native speakers of Japanese (n = 2) and the results showed that the activities successfully elicited negation of adjectives.

4.8 Treatment sessions

The aim of the two treatment sessions for the experimental group was to elicit modified output on the target form by manipulating clarification requests while the control group was not given feedback for the production of modified output. Two different elicitation instruments, an interview as a meaning-focused activity and mechanical drills as a form-focused activity were employed for the purpose of maximising the frequency of production of modified output.

4.8.1 First treatment session: A structured interview

The first treatment session was a methodologically focused interview (see section 4.8.1.3) for the experimental group and unfocused interview for the control group between participants and the researcher who acted as an interlocutor. This session was held in a faculty office at the university where participants were enrolled as students, approximately two weeks after the pre-test. The researcher interviewed each participant individually at a previously arranged time, seated across from the participant at a desk, with a small tape recorder for the purpose of audio recording. This arrangement was the same with the second treatment session.

Communicative task-based activities (e.g., picture description tasks or information gap tasks) have been extensively used in the studies on interaction, including modified output, because these can create interaction focusing on meaning. However, the difficulty has been pointed out in constructing a communicative task that could elicit the use of negation in Japanese (Kanagy, 1991). Spontaneous conversation with a native speaker was another option but Kanagy (1991) reported one-hour conversation with a native speaker elicited only one or two negations in her preliminary assessment of the data elicitation instrument. This is supported by the analysis of interview data of 45 beginner and intermediate learners of Japanese in KY corpus[14] by Ogino (2006), which found only 18 instances of the use of negation of adjectives.

Structured interviews (often with pictures) have been used in the studies of negation in Japanese (Hansen, 1999; Hayashi, 1999; Kamura, 2001a, 2001b; Kanagy, 1991), and it was considered to be the most appropriate technique to elicit negation of predicate adjectives for the present study as well. The advantage of the oral interview is its structured format by which the interlocutor can control elicitation of the target form to some extent despite the common difficulties in controlling learner's output in task-based activities. One of the limitations was that the results gained could be different from those obtained through naturalistic settings and classroom

---

[14] KY corpus (version 1.2) was developed by Kamata, O., and Yamauchi, H., and it consists of transcripts of spoken corpora obtained through with 90 learners of Japanese (15 novice, 30 intermediate, 30 advanced and 15 superior levels) in the Oral Proficiency Interview (OPI). OPI was developed by American Council on the Teaching of Foreign Languages (ACTFL), and a qualified OPI tester interviews learners. Each interview consists of warm-up, level checks, probes using task-based role playing and wind-down.

settings (Kanagy, 1991, p. 87). Also, a methodological concern was the possibility of participants noticing the aim of the study by priming negative responses, and it was dealt with by the mixture of questions that could induce affirmative responses (Kanagy, 1991; Kamura, 2001). No pictures were used in the interview in the current study so that the participants had more freedom to respond without being constrained by the information of cue pictures, and this freedom was considered to help participants produce natural and spontaneous language.

The interview consisted of a total of 40 questions that were categorised under such topics as 'University Life' and 'New Zealand.' Twenty-five questions aimed at inducing negation of adjectives while the other 15 questions, including *wh*-questions, were distractors that did not aim at eliciting negation. In addition to distractors, further questions that were naturally and spontaneously derived from the response to initial questions were interspersed, following a technique by Kanagy (1991). The aim was to shift the focus from form to meaning, for the purpose of minimising the methodological concern regarding the participants' awareness of the aim of the tasks. Questions were carefully elaborated with some flexibility so that the negation of adjectives could be induced as often as possible. The interview schedule was the same for the two groups, but the questions actually asked varied among the participants because one reply could contain answers to several questions. For example, there were cases in which participants included information in their reply that was related to the questions to be asked. In that case, those questions were not asked to keep the natural flow of the interview. Few questions by the interlocutor

included negation to avoid the possibility of participants using them as a model or further comprehensible input for modification of their initial output.

A series of pilot studies were conducted with learners of Japanese at the University of Waikato to develop appropriate techniques that could elicit the negation of adjectives and lead to modified output for achieving the aim of the present study. Based on the results of the pilot studies, three features were incorporated into the structured interview: explicit instructions, follow-up priming questions, and methodologically focused and unfocused interviews.

4.8.1.1 Explicit instructions

One of the features of this interview was that participants were explicitly instructed to respond to each question in a full sentence. That is, if the question was a *wh-* question (e.g., 'What do you like?'), the participants needed to reply by saying, for instance, 'I like sushi' rather than just saying 'Sushi.' In the case of Yes-No question (e.g., 'Is it expensive?') and if the answer was affirmative, they needed to say 'Yes, it is,' and in the case of a negative response, to say 'No' with a negative sentence (e.g., 'No, it is not'). This format seems unnatural and forced, but explicit and specific instructions were required to avoid one-word answers of 'Yes' or 'No.'

4.8.1.2 Follow-up priming questions

Priming questions related to the initial question were asked to elicit negation forms from the participants. Whether the participants used an affirmative sentence or a negative sentence as a response to each question was up to each participant.

Therefore, questions were carefully phrased to elicit a negative reply in a series of one set of questions, as in the example, below:

Example 6 (taken from the current study)

    Interlocutor: *Mokuyoobi wa isogashii desu ka.*

                        Are you busy on Wednesday?

    Daine[15]:   *Maa, maa ne.   Yojikan gurai.*

                        So, so.   I have about four classes.

    Interlocutor: *Aa, soo.   Soshite, shuumatsu wa isogashii desu ka.*

                        Is that right?   Are you busy on the weekend?

    Diane: *Amari isogashi kunai[16] n desu.*

                        I am not very busy.

When the question 'Are you busy on Wednesday?' was asked, there were at least two possible responses, that is, 'No, I am not' or 'Yes, I am.'   If the former reply using negation was induced successfully, the interlocutor went on to the next question.   If the latter was used and negation was not used in the reply, the participant was given a follow-up priming question the topic (subject) of which was replaced with other words (e.g., 'Are you busy on the weekend?').   If a negative reply was still not given to this question, other follow-up priming questions (but no more than three) were asked (e.g., 'Are you busy tonight?').   There was still a possibility that participants

---

[15] All names in the examples have been changed.
[16] The underlined word indicates the use of negation.

replied with an antonym, such as 'No, I am free' instead of using negation, and this was not controllable.

4.8.1.3 Methodologically focused and unfocused interviews

A methodologically focused interview was developed for the present study based upon implicit methodological techniques (section 2.3.1). Clarification requests in this interview were general clarification requests as opposed to the specific clarification requests. They were given only to the experimental group by means of a standard formula – for example, *Sumimasen* (Excuse me?), or *E* (Pardon?) –, in a way which did not provide further linguistic information about the target form to the participant. They were exclusively given as the interlocutor responses to any instance of production of a non-targetlike pattern irrespective of whether the interlocutor actually understood the participant's utterance or not. No further clarification requests were given when a participant failed to produce a targetlike pattern in response to the initial general clarification request. The interlocutor did not give any feedback and indication about the grammaticality of their MO, and the interview was carried out as naturally as possible. The control group had a methodologically unfocused interview, in which the participants did not receive any clarification requests on the target form at all except when a genuine communication breakdown occurred.

4.8.2 Second treatment session: Mechanical drills

The importance of conducting treatment activities more than twice has been suggested in order to ascertain the effectiveness of treatment (Takashima, 1995), but

the results of the pilot study suggested the difficulty in using a structured interview twice or more because of the learners' limited vocabulary and grammatical knowledge. Therefore, as a second treatment, mechanical drills were employed to elicit as many negations of adjectives as possible in a different context from the first treatment session.

Mechanical drills are a feature of the Audiolingual Method with its theoretical foundations in behaviourism and structural linguistics. Thus their use does not fit into the interactionist view of communicative interaction particularly because they do not provide opportunities for participants to produce language in context which is considered to be necessary for L2 acquisition. However, the opportunity to negotiate meaning while performing a drill is not necessarily precluded. It at least provides an opportunity for interaction between the participant and the interlocutor to occur, in which participants are required to produce output, to receive feedback, consider the intent of the clarification requests, to reflect on their own utterance, and to respond to the feedback in some way to complete the task.

It has been pointed out that this kind of exercises limits learner output in terms of length and complexity (e.g., Ellis, 2005), but the purpose of the present study did not necessarily require participants to produce lengthy and complex language. One of the advantages of the use of a mechanical drill in the context of the current study was that it could keep participants from using antonyms of an adjective instead of using negation, which was often observed in the pilot study.

The experimental group had a methodologically focused drill which is an application of the methodologically focused communication tasks, while the control group had unfocused drills.   The procedures for the drill activity are outlined below.

Step 1: Vocabulary check

A set of picture flash cards (verbs, nouns, nominal adjectives and adjectives) were shown (see the list of words in Appendix 4.E) and participants were instructed to say the words described by the picture in Japanese.   The flash cards the picture of which they could not describe were excluded.   Three cards from three predicate categories of verbs, nouns and nominal adjectives were used for Round 1 and 2 to minimise the issue of vocabulary knowledge.   Ten cards of adjectives were presented and as many adjectives as each participant was familiar with were used for the drill activity.   The words used in the pre-test and post-tests were not included in this session to avoid practice effects of specific words.

Step 2:   Mechanical drill: Round 1 (audio-recorded)

Nine flashcards (3 verbs, 3 nouns and 3 nominal adjectives) and several cards of adjectives (ranging between 3 and 9) were shown one by one and participants were instructed to say the affirmative forms and negative form – for example, *tabemasu, tabemasen* (I eat; I don't eat), *samui, samu kunai* (It's cold; It's not cold).   The experimental group had a methodologically focused mechanical drill, and clarification requests were given when participants used a non-targetlike form on the negation of adjectives (not on negations of other predicate categories).   The order of presentation of the flash cards was randomised among predicate categories so that the

91

participants were required to produce a variety of negation patterns randomly. The control group also had a mechanical drill but clarification requests were not given on the non-targetlike pattern participants used.

Step 3: Revision of how to make negations among the four predicate categories

The rules for creating the negation forms of the four predicate categories were explicitly revised by using a brief PowerPoint presentation (2-3 minutes). The aim was to minimise the possible variables related to revision between the pre-test and post-test (e.g., some participants might revise the target forms or ask questions regarding the target form in class).

Step 4: Mechanical drill: Round 2 (audio-recorded)

Step 2 above was repeated as a reinforcement activity to have the participants produce more negations, (and to receive more clarification requests and to have them produce more modified output for the experimental group). The improvement observed in Round 2 in comparison with Round 1 could be attributed to the mixture of effects of variables, such as revision in Step 3 or task repetition for both the experimental group and the control group, and clarification requests and production of modified output for the experimental group. The results of a Mann-Whitney test showed that the difference in accuracy scores between Round 1 and Round 2 in each group was not significant ($Z = -.096$, $p = .946$ for the experimental group, $Z = -.231$, $p = .839$ for the control group).

4.9   Data coding and scoring

The procedures for data coding for the treatment activities are described first (4.9.1), followed by those used to code the testing data (4.9.2 - 4.9.4).

4.9.1   Response moves to clarification requests in the treatment sessions

There are a number of categories of response move triggered by feedback such as modified output, repetition, acknowledgement, topic continuation, inability to respond and ignorance of feedback (Sato & Lyster, 2007), but response moves to clarification requests in the treatment sessions were classified into the four categories based and developed upon McDonough (2001), as shown in Figure 4.2. The categories include modified output and repetition, which Saxton et al. (2005) treated as two broad categories[17]. Modified output produced before clarification requests were given was excluded from coding[18], as it was not within the scope of the study.

Figure 4.2
Four categories of response move to clarification requests

---
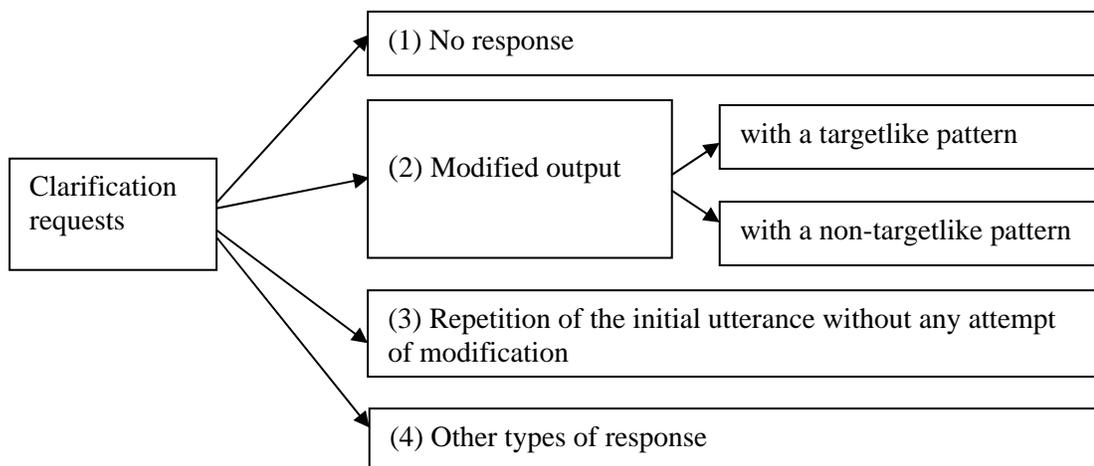
[17] The term *revision* was used for modified output in Saxton et al. (2005).
[18] There were only two such cases.

(1) No response: This is considered to be an indication of not noticing the clarification requests although it is also possible participants tactically ignored the clarification requests. No data of this category were found in the present data.

(2) Modified output: MO was further classified into either MO with a targetlike pattern or MO with a non-targetlike pattern. The term *targetlike pattern* refers to the correct choice of negation pattern with the correct inflections of the adjective stem (Example 7). *Non-targetlike pattern* refers to an incorrect choice of the negation pattern (e.g., *samu janai for *samu kunai,* 'it is not cold') or the incorrect inflectional change (e.g., *samui kunai), or combination of both (*samui janai) as shown in Example 8. It is important to note that MO with a non-targetlike pattern was included in the analysis because it is possible that process of reformulation, regardless of targetlike or non-targetlike modification, is beneficial for L2 learning.

Example 7: *MO with a targetlike pattern*

Interlocutor:	*Nihongo no shukudai wa ooi desu ka.*
	Do you have a lot of Japanese homework?
Catlin:	*Ooi?*
	A lot?
Interlocutor:	*Ooi wa many, a lot.*
	Ooi means many, a lot.
Catlin:	*Iie, *oo … iku nai desu …*
	Not much.

Interlocutor: *N*

                  Pardon?

Catlin: *Ooi ku ... e ... *ooi janai ... oo kunai...*

                  Not much, not much, not much.

Example 8: MO with a non-targetlike pattern

Interlocutor: *Ookii desu ka.*

                  Is it big?

Emily: *Amari *ookii janai.*

                  Not so big.

Interlocutor: *Soo, mooichido onegai shimasu.*

                  Well, once again please?

Emily: *Amari *ookii dewa arimasen desu ne.*

                  Not so big, isn't it?

(3) Repetition: This category refers to a repetition of the initial non-targetlike utterance without any attempt at modification to which a clarification request was given. This is not regarded as modified output as it does not involve reformulation (McDonough, 2001, 2005), but it was recorded as an independent category in order to observe the changes in the subsequent use for the comparison with MO.

Example 9

Interlocutor: *Mandarin wa muzukashii desu ka.*

                  Is Mandarin difficult?

Helen:	*Iie, muzukashi ... n ... *muzukashii janai.*
	Difficult … n … not difficult.
Interlocutor:	*N, mooichido.*
	Once again?
Helen:	**Muzukashii janai.*
	Not difficult.

(4) Other types of response: This category includes change of vocabulary, giving up answering half way, and use of non-Japanese language.

Example 10

Interlocutor:	*Furansu-jin wa ooi desu ka.*
	Are there many French people?
Nick:	*Furansu-jin *ooi kunai desu*
	There are not many French people.
Interlocutor:	*N.*
	Pardon?
Nick:	There isn't so many French here. (Answered in English.)

4.9.2   Accuracy score in the tests

Accuracy scores were used to measure the impact of MO on L2 learning. The utterances that included the use of negation of predicate adjectives were transcribed according to standard orthography, and then the attempted use of negation of adjectives was identified and recorded as either targetlike use or non-targetlike use. The accuracy score was calculated based on the number of targetlike items supplied by the participants divided by the total number of attempted uses of the target form. The participants who produced only a small number of target forms could have been excluded from analysis to minimise distortion of percentage scores as carried out in Izumi and Bigelow (2000), but it was decided not to exclude these participants so as to observe their developmental changes that were likely to be large. A research associate acting as a second rater crosschecked the transcriptions, and identified the attempts of the target form. The inter-rater reliability was determined by simple percentage agreement, which was 99%.

Obligatory occasion analysis (Brown, 1973) and targetlike use analysis (Pica, 1983) were not used because of difficulty in determining obligatory occasions for the use of negation of adjectives. For example, participants can avoid the use of negation of adjectives to a question 'Is the car expensive?' by using an antonym (i.e., 'It is cheap') even though the answer is 'No.' Identical lexical items were not excluded, but counted as one token. Although there was an issue of artificial inflation of accuracy scores due to including multiple occurrences (Ishida, 2004; Mackey & Gass, 2005), the number of the instances of those cases was only one out of the total of

1,011 negative responses collected in the tests, and its influence on the results was considered minimal.

4.9.3   Interlanguage development in the tests

This study focused on the interlanguage development represented by quantitative and qualitative changes in the use of non-targetlike patterns in order to measure the impact of MO on L2 learning.

4.9.3.1 Categories of non-targetlike negation patterns

Each non-targetlike negation pattern identified was first classified into six categories as shown in Table 4.2, developed upon from the categories in the previous studies (e.g., four categories in Kamura, 2001b; six categories in Kanagy, 1991; five categories in Noro, 1995).   Formal and informal forms (e.g., *samu ku arimasen* versus *samu kunai* for 'it is not cold') were coded into the same category.   The inflected and uninflected stems were not differentiated but classified into the same category.   It was decided not to differentiate the inflected and uninflected stems because of the small sample size in the current study.   The difference between inflected stems (e.g., *samui*) and uninflected stems (e.g., *samu*) followed by negation patterns is that the former (e.g., **samui janai*) is the external negation without analysis of the predicate category and the choice of negation pattern, while the latter (e.g., **samu janai)* is the external negation with some kind of analysis which is shown by dropping the non-past tense morpheme *i.*   It appears that the production of **samu janai* is more difficult than **samui janai* in terms of processing.   Although more substantial data might reveal some patterns of development, no studies have

identified if there is any difference between the two within the developmental

sequence.

Table 4.2
Categories of non-targetlike negation patterns of adjectives in Japanese

Categories	Non-targetlike Negation patterns	Explanation
nai	*nai*	An incorrect negation pattern that has the suffix *nai* (non-past negator). e.g. *samu* (*i*) *nai*
	arimasen	An incorrect negation pattern that has the suffix *arimasen*. e.g. *samu* (*i*) *arimasen*
masen	*masen*	An incorrect negation pattern that has the suffix *masen*, (a suffix for verb negation). It includes *samu* (*i*) *masen, samu* (*i*) *ja masen, samu* (*i*) *ku masen,* and *samu* (*i*) *ku ja masen.*
janai	*dewa nai*	An incorrect negation pattern that has the suffix (*i*) *dewa nai* (a suffix for noun/nominal adjective negation). It includes a formal style, (*i*) *dewa arimasen.* e.g., *samu (i) dewa nai/dewa arimasen*
	ja nai	An incorrect negation pattern that has the suffix (*i*) *ja nai. Ja* is a contracted form of *dewa* in informal speech. It includes a formal style, (*i) ja arimasen.* e.g., *samu (i) janai/ja arimasen*
kujanai	*ku dewa nai*	An incorrect negation pattern that has the suffix (*i*) *ku dewa nai.* This is the combination of (*i*) *ku* + noun/nominal adjective negation pattern (informal). It includes a formal style, (*i*) *ku dewa arimasen.* Also, it includes ku janai/ ku ja arimasen, which has *ja,* a contracted form of *dewa* e.g., *samu (i) ku dewa nai/arimasen, samu (i) ku ja nai/ arimasen*
kunai	*ku nai (incorrect)*	A correct negation pattern that has the suffix *ku nai.* This category includes the patterns in which the choice of the negation pattern is correct, but the non-past morpheme *i* is not dropped before adding *ku nai.* It includes a formal style, (*i) ku arimasen.* e.g. *samu i ku nai/arimasen*
Others	others	Non-targetlike use that does not belong to the categories above, including pronunciations and the stem of the adjective itself. Also, it includes stopping the phrase half way through and switching to non-Japanese language.

4.9.3.2 Developmental stages

Three developmental stages presented in section 3.3 were used for the analysis of developmental stages. Table 4.3 below is a duplication of Table 3.3 for the reader's convenience.

Table 4.3
Developmental stages of negation of adjectives in Japanese

Stage	Feature	Explanation	Examples
Stage 3	Targetlike pattern	The choice of the negation pattern is correct and the inflection of the stem is correct.	*Samu kunai*
Stage 2	Absence of the non-targetlike pattern *nai*	*Nai* is not used any longer but other types of non-targetlike pattern are still used. (There might be another stage between Stage 2 and 3, which features the disappearance of *janai* or *kunai*, but the findings are mixed regarding the disappearing order.)	*Variants except samu(i) nai e.g., *Samu(i) janai *Samu(i) kunai *Samu(i) masen *Samu(i) kujanai*
Stage 1	Use of the non-targetlike pattern *nai*	Adj root + (*i*) *nai* are used. Targetlike patterns as well as other types of non-targetlike pattern may be used.	*Samu(i) nai* other variants

Reaching a stage was operationalised as the production of the negation pattern from a stage category on at least two different lexical items in order to differentiate analysed from unanalysed use. The number of participants who produced negations from each stage were analysed between the experimental group and the control group in order to measure the impact of MO on interlanguage development. The number of participants who moved to a higher stage was not analysed because it was observed that the majority of participants were at Stage 3, which indicates that they were already free from processing constraint, and movement to a higher stage in the post-tests was not expected.

In their consideration of interlanguage analysis, Ellis and Barkhuizen (2005) emphasised the importance of checking the assumptions of previous studies concerning developmental stages, and the data of the control group served for this purpose in the present study.

4.9.4 Subsequent use analysis

As defined in section 1.5, *subsequent use,* here, refers to the use of the previously modified targeted linguistic feature in the post- and delayed post-tests. The types of non-targetlike pattern each participant in the experimental group modified in the treatment sessions were first identified, and these were then tracked to find whether or not the non-targetlike pattern was used in the post-test and delayed post-test.

Development of interlanguage in this analysis was operationalised as the non-use of the same type of previously modified non-targetlike negation pattern in the post-tests. It was hypothesised that if there was some impact of MO on the development of interlanguage, the specific non-targetlike pattern would be used less in the post-tests, based on the assumption of the gradual decrease of the patterns in frequency and variation. Also, it was assumed that if there were no or little impact of MO, participants would continue to use the same non-targetlike pattern in both of the post-tests. Therefore, the degree of impact of MO was hypothesised to be the strongest when the non-targetlike pattern was not used in either of the post-tests, and the weakest when the non-targetlike pattern was used in both of the post-tests. When a previously modified non-targetlike pattern was used in the post-test, but not in the delayed post-test, there is a possibility that it might be related to the delayed effects of

MO.   The use of a previously modified non-targetlike pattern only in the delayed post-test might suggest temporary impact.

In sum, there are four possible combinations of the use and non-use of the same type of previously modified non-targetlike pattern in the post-tests, and each combination is illustrated in Table 4.4.   The plus sign ('+') shows that the non-targetlike pattern would occur in the indicated test, and the minus sign ('-') shows the non-use of the non-targetlike pattern.   The table also displays the hypothesised degree of impact of MO.

Table 4.4
Possible combinations of the subsequent use/non-use of the same type of previously modified non-targetlike pattern and hypothesised degree of impact of MO

Description	Post-test	Delayed post-test	Hypothesised degree of impact of MO
Non-use in both of the post-tests	-	-	Strong impact
Use in the delayed post-test only	-	+	Temporary impact
Use in the post-test only	+	-	Delayed impact
Use in the both of the post-tests	+	+	No impact

It is known that some linguistic features may disappear at one stage but reappear later on (Meisel et al., 1981), and learners who used only the targetlike form at one stage may start to use non-targetlike form at the next stage, followed by another stage of the use of the targetlike form *(U-shaped behaviour*, Kellerman, 1985).   Therefore, the interpretation of the results needs much caution.

4.10   Summary

This chapter restated the research questions and hypotheses of the study, and explained the research methodology employed in the present study.   Chapter Five presents the results of analysis of spoken data collected from 28 learners of Japanese

CHAPTER FIVE

RESULTS

5.0   Introduction

This chapter presents the results of analysis of spoken data obtained from the total of 28 university students who participated in this study.   Section 5.1 analyses the oral production data collected from interviews and mechanical drills in the treatment sessions.   Because the control group did not receive clarification requests on non-targetlike forms, the sections 5.1.1 to 5.1.4 specifically report the results related to clarification requests and modified output only in the experimental group.   Data from the pre-test, post-test and delayed post-test are presented and discussed in section 5.2.   Sections 5.3, 5.4 and 5.5 respectively present the results of analyses for each of the three research questions and their associated hypotheses which were presented in Chapter Four.

5.1   Oral production data in the treatment sessions

The oral production of participants in the treatment activities was coded in terms of the amount of the attempted use of negation, targetlike patterns and non-targetlike patterns.   It may be worth repeating at this stage the operational definitions given in Section 1.5 in Chapter One: *targetlike* is operationalised as the correct choice of a Japanese negator and an associated morpheme, and the correct inflectional change on any given adjective stem; *non-targetlike* refers to the incorrect choice of a linguistic form from among the range of negation patterns that exist in Japanese.   The use of a

non-targetlike form does not necessarily indicate a breakdown in the communicative competence of the learner, but may rather show interlanguage development.

Descriptive statistics are provided in Table 5.1. It also displays the percentage accuracy scores, and the number of clarification requests given to non-targetlike patterns in the experimental group (see Appendices 5.A and 5.B for the results of each participant in each activity).

Table 5.1
Oral production data from the treatment sessions

	Experimental group (n = 14)				Control group (n = 14)			
	Total	M	SD	Mdn	Total	M	SD	Mdn
Attempted use	220	15.7	4.86	15.0	246	17.6	4.85	19.0
Targetlike use	139	9.9	5.90	10.5	131	9.4	5.26	10.5
Non-targetlike use	81	5.8	5.78	4.5	115	8.2	4.81	8.0
Accuracy score (%)	n/a	61.4	30.15	67.9	n/a	48.2	24.96	49.8
Clarification requests	70	5.0	4.79	3.5	n/a	n/a	n/a	n/a

The treatment activities (one interview and two mechanical drills) elicited a total of 220 negations of adjectives in the experimental group and 246 in the control group. A *t*-test was carried out to compare the mean of attempted use of negation for each group. Alpha was set at .05 for all statistical tests in the analysis of treatment data and testing data. All the statistical tests were performed using the *Statistical Package for the Social Sciences* (SPSS) version 12.1. The results indicated that there were no significant differences between the groups: t (26) = 1.013, p = .321. The mean accuracy scores between the groups were also compared and they were not significantly different, t (26) = -1.263, p = .218.

Figure 5.1 summarises the data of the experimental group, which will be discussed in the following sections.

Figure 5.1
Elicitation and production of MO (Experimental Group)

Attempted use of negation of adjectives
220

Non-targetlike pattern	Targetlike pattern
81	70

Clarification requests given	Clarification requests not given
70	11

MO	Repetition	Others
29	34	7

Targetlike pattern
18

nai	masen	janai	kujanai	kunai	others
5	1	6	0	5	1

Non-targetlike pattern
11

nai	masen	janai	kujanai	kunai	others
2	0	4	0	3	2
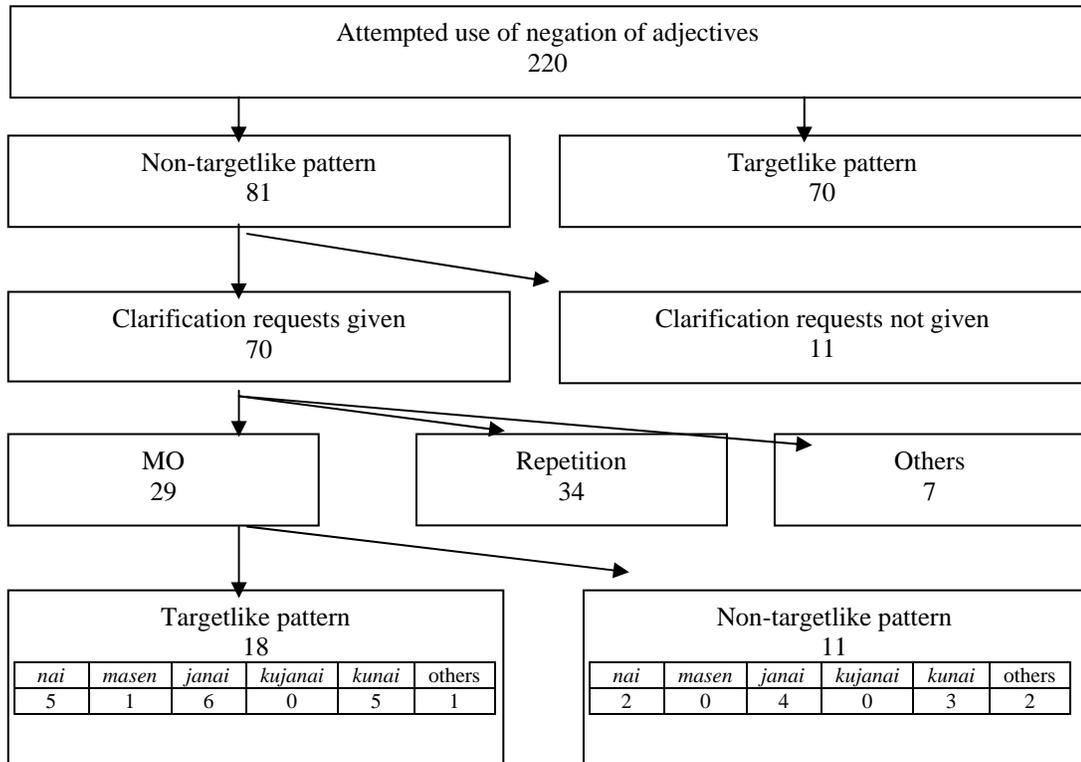
5.1.1 Types of non-targetlike pattern that received clarification requests

The experimental group produced 81 non-targetlike patterns and clarification requests were given to 70 of them (Mdn = 3.5, range 16).   The control group produced a total of 115 non-targetlike patterns, and no feedback, including clarification requests, was provided; therefore, the participants in the control group did not have opportunities to produce modified output following feedback.   It should be noted that the provision of clarification requests depended upon the production of non-targetlike patterns by each participant.   There were 11 non-targetlike patterns to which clarification requests were not given either because the interlocutor failed to do so, or because it could have significantly hindered the natural flow of the conversation.

Table 5.2 shows the types and frequency of non-targetlike patterns that received clarification requests.   As illustrated in section 4.9.3.1, non-targetlike patterns were coded into six categories (i.e., *nai, masen, janai, kujanai, kunai* and others).   *Nai* is a negation pattern used for verb negation in an informal register, and it is included in all the other negation patterns in an informal register.   *Masen* is also a pattern for verb negation but for a formal register.   *Janai* is a pattern for negating nouns and nominal adjectives (informal), and *kujanai* is a mixture of *janai* and *kunai,* and incorrect for any negation.   *Kunai* is a correct pattern for negating adjectives but referred to when the inflection of the stem was incorrect (i.e., *samui-kunai* for *samu-kunai*).   The category 'Others' includes the patterns that do not belong to the categories above and those phrases that were interrupted half way through and switched to non-Japanese language.

Table 5.2

Types and frequency of non-targetlike patterns which received clarification requests (Experimental Group)

| | NTL produced | NTL which received clarification requests | | | | |
		Total	%	M	SD	Mdn
nai	10	10	14.3	0.7	1.20	0
masen	5	5	7.1	0.4	0.93	0
janai	45	40	57.0	2.9	3.42	2
kujanai	1	0	0.0	0.0	0.00	0
kunai	14	12	17.1	0.9	0.86	1
others	6	3	4.3	0.2	0.43	0

Note. NTL = non-targetlike pattern

The results show that *janai* was the most frequently used non-targetlike pattern. Participants seemed to be often confused about the predicate category (i.e., verb, noun, nominal adjective, adjective) of the lexical item they were going to negate, and they tended to use *janai,* which is a negation pattern for negating a noun/nominal adjective. *Janai* received the largest number of clarification requests, followed by *kunai* and *nai* (see Appendices 5.C1-C3 for the individual data in each activity, and 5.C4 for the summary of all the activities by individual participants).

5.1.2 Types of response move to clarification requests

All the 70 clarification requests given to the non-targetlike patterns prompted a response from the participants in the experimental group (i.e., 100% response rate) although it was entirely up to them whether or not they responded to clarification requests. Response moves were classified into four categories: no response, modified output in response to clarification requests (MO), repetition of the initial non-targetlike pattern (Repetition), and other types of response move (Others). The category 'no-response' was discarded from the analysis, as there were no instances

observed in the data. The results are presented in Table 5.3 (see Appendix 5.D for the data in each activity).

Table 5.3
Types of response move to clarification requests (Experimental Group)

	Total	%	M	SD	Mdn
MO	29	41.4	2.1	2.23	1.5
Repetition	34	48.6	2.4	2.41	1.5
Others	7	10.0	0.5	0.76	0.0

The results showed that 29 clarification requests led to the production of MO and the group median was 2.1 (range 8), and the most frequently used response was Repetition (34 instances, Mdn = 2.4, range 9). Both MO and Repetition are considered to be produced as a response to clarification requests.

It is reasonable to suggest that production of MO indicates that the participants noticed the intention of clarification requests (i.e., a request for grammatical reformulation). Internal evidence in the recordings of interaction, such as pause length following the provision of clarification requests, lends support to this assumption. Repetition might be an indication of non-recognition of ungrammaticality of their initial utterance, but it is also possible that they just did not know any other way of expressing it despite their recognition of necessity of grammatical reformulation. It should be noted, however, that these are not beyond speculation, as no data was collected on the relationship between response moves and participants' interpretation of clarification requests to provide insights into this aspect. This issue will be referred to in detail in Section 6.2.4.

5.1.3 Types of negation patterns in each response move

Table 5.4 shows the frequency of each type of negation pattern according to the types of response move to clarification requests.

Table 5.4
Types of non-targetlike negation pattern by response move (Experimental Group)

		MO			Repetition			Others		
		Total	%	Number of participants	Total	%	Number of participants	Total	%	Number of participants
nai	(n = 10)	7	70.0	4	3	30.0	2	0	0.0	0
masen	(n = 5)	1	20.0	1	3	60.0	2	1	20.0	1
janai	(n = 40)	10	25.0	7	25	62.5	11	5	12.5	4
kujanai	(n = 0)	0	0.0	0	0	0.0	0	0	0.0	0
kunai	(n = 12)	8	66.7	8	3	25.0	3	1	8.3	1
others	(n = 3)	3	100.0	1	0	0.0	0	0	0.0	0

Note. MO = modified output (irrespective of targetlike or non-targetlike pattern)

One of the noticeable trends is the types of response move when clarification requests followed *janai*. Nearly two-thirds (62.5%) of clarification requests to *janai* were responded to with Repetition, in contrast to 30.0% for *nai* and 25.0% for *kunai*. Also, 11 out of 14 participants responded with Repetition, when clarification requests followed *janai*. Only 25.0% of clarification requests after *janai* led to the production of MO while *nai* led to the most MO (70%), followed by *kunai* (66.7%). How participants responded to clarification requests could closely be related to their prior knowledge of the target feature. However, participants in the present study were assumed to have prior explicit knowledge about the use of the target form, and, therefore, the results are considered to indicate a possibility that how learners respond

to clarification requests might be associated with the types of non-targetlike pattern to which clarification requests were given (see Appendices 5.E1-E3 for the detailed data).   This is discussed further in section 6.2.3.

5.1.4   MO with targetlike patterns and non-targetlike patterns

MO was further classified into either targetlike or non-targetlike patterns.   Table 5.5 shows that 18 tokens of MO (62.1%) were produced with a targetlike pattern while 11 tokens (37.9%) ended with MO with a non-targetlike pattern, out of a total of 29 occurrences of MO.   (See Appendix 5.F1-F3 for the data of each of the activities and 5.G for the summary of MO by individual participants).

Table 5.5
MO with targetlike and non-targetlike patterns (Experimental Group)

	Total	%	M	SD	Mdn
MO with targetlike pattern	18	62.1	1.3	1.38	1.0
MO with non-targetlike pattern	11	37.9	0.8	1.05	0.5

When MO was produced, it was reformulated with a targetlike pattern more often than with a non-targetlike pattern, but the production rate of MO with a targetlike pattern in response to the total of 70 clarification requests was 25.7%.   This is consistent with findings in the previous studies: 24.7% in Nobuyoshi and Ellis (1993), 28% in Lyster and Ranta (1997), 21.3% in O'Relly et al. (2001) and 31% in

Nassaji[19] (2007), and indicates clarification requests may not be very facilitative in leading learners to targetlike modification.

Table 5.6 shows the distribution with percentages of MO with targetlike and non-targetlike patterns according to the types of negation pattern (see Appendix 5.H for the data of individual participants).

Table 5.6
Distribution of MO with targetlike and non-targetlike patterns by types of negation pattern (Experimental Group)

| | | MO with a targetlike pattern | | | MO with a non-targetlike pattern | | |
		Total	%	Number of participants	Total	%	Number of participants
nai	(n = 7)	5	71.4	2	2	28.6	1
masen	(n = 1)	1	100.0	1	0	0.0	0
janai	(n = 10)	6	60.0	4	4	40.0	3
kujanai	(n = 0)	0	0.0	0	0	0.0	0
kunai	(n = 8)	5	62.5	5	3	37.5	3
others	(n = 3)	1	33.3	1	2	66.7	2
Sum	(n = 29)	18	62.1	n/a	11	37.9	n/a

Janai received the largest number of clarification requests, 62.5% of which were followed by Repetition as shown in section 5.1.3. However, of the modifications of *janai* which were prompted by a clarification request, 60% were targetlike; this was close to the percentage of targetlike MO for *kunai*. This suggests the production of MO with a targetlike pattern might be less associated with the types of pattern than the noticing of clarification requests. In other words, noticing the interlocutor's motivation for clarification requests to *janai* might be harder than for other types of

---

[19] The feedback used in Nassaji (2007) was termed *elicitations,* which consists of five subtypes of elicitation; unmarked elicitation, marked elicitation, marked elicitation + prompt, marked elicitation + enhanced prompt, and elliptical elicitation. Clarification requests are included in unmarked elicitation.

pattern, but when learners understand the illocutionary intent for grammatical modification, they are able to produce MO with a targetlike pattern, as they can do for other types of pattern.

5.1.5   Summary of oral production data

There was no significant difference between groups in the number of negation responses elicited by the activities.   Clarification requests were given to 70 out of a total of 81 non-targetlike patterns produced by participants in the experimental group. Clarification requests prompted response moves, either production of MO, or Repetition, from the participants in the experimental group, and the response rate was 100%.   Repetition of the initial non-targetlike utterance occupied the largest response move (48.6%), but clarification requests gave participants an opportunity for reconsidering their own utterance, and 41.6% of clarification requests triggered the production of MO.   Even though clarification requests did not provide participants with any linguistic information, 62.1% of MO was targetlike.

In terms of types of pattern, *janai* received the most clarification requests (i.e., *janai* was the incorrect negation pattern that was most frequently used).   However, nearly half of the clarification requests after *janai* met the response of Repetition, and the production rate of MO after *janai* was much lower than that of *nai* and *kunai.*   This finding suggests how participants responded to clarification requests (i.e., whether or not they noticed the intention of clarification requests) might be associated with the types of pattern to which clarification requests were given.   On the other hand, it was found that whether MO was targetlike or non-targetlike was not largely influenced by

the types of pattern. Therefore, the results indicate a relationship between noticing of clarification requests and the types of pattern which clarification requests were given.

5.2. Testing data: Use of the target forms

A total of 1,011 negations were elicited in the pre-test, post-test and delayed post-test, and the results reported in the rest of this chapter are based on the analysis of these data. Descriptive statistics are provided in Table 5.7. The distribution of scores was not normal and the median was used as the measure of central tendency. The median will be used unless otherwise stated in the rest of chapter.

Table 5.7
Numbers of negations produced in the tests

		Pre-test	Post-test	Delayed post-test
Experimental	Total	99	187	186
Group	M	7.1	13.4	13.3
(n =14)	SD	4.12	3.89	3.71
	Mdn	7.5	14.0	13.5
	Range	1-13	5-18	5-19
Control	Total	148	190	201
Group	M	10.6	13.6	14.4
(n =14)	SD	5.73	5.35	5.69
	Mdn	11.0	15.5	17.0
	Range	2-20	4-20	4-20

As illustrated in Chapter Four, participants were matched on the basis of their accuracy scores in the pre-test, and then randomly assigned into the two groups. A non-parametric Mann-Whitney test was carried out to confirm the pre-treatment equivalence of the two groups, and the results indicated the difference in the median accuracy scores was not significant, $Z = -.407$, $p = .701$ at the point of pre-test.

The number of negations produced in the pre-test by the experimental group was smaller by 49 than that by the control group. One of the possible contributing factors to this was that the experimental group had three participants who produced only one negation while the control group had none. Also, the experimental group had only four participants who produced more than ten negations compared to the control group (eight participants). The results of Mann-Whitney tests revealed that the differences in the number of attempted uses in each session between the groups were not significant, $Z = -1.684$, $p = .093$; post-test, $Z = -.531$, $p = .595$; delayed post-test, $Z = -1.346$, $p = .178$.

The rest of this chapter reports the findings as they relate to the research questions and associated hypotheses.

5.3. Results for Research Question 1: Evidence of the relationship between the production of MO and grammatical accuracy

The first research question and hypotheses were:

Research Question 1: Is there a positive relationship between the production of MO and grammatical accuracy?

Hypothesis 1.1: The experimental group will show greater gains in grammatical accuracy than the control group.

Hypothesis 1.2: There will be a relationship between the production of MO in the treatment sessions and gains of grammatical accuracy scores in the post-test and delayed post-test in the experimental group

5.3.1 Hypothesis 1.1

The evidence of impact of MO on grammatical accuracy was examined through the comparison of accuracy scores between the experimental group and the control group in the three tests, which are shown in Table 5.8. The changes in the mean accuracy scores are graphically presented in Figure 5.2. The individual data sets of the three tests is provided in Appendices 5.I1 and 5.I2.

Table 5.8
Group accuracy scores (%) in the tests

		Pre-test	Post-test	Delayed post-test
Experimental Group (n =14)	M	26.2	67.8	75.3
	SD	30.77	22.56	28.77
	Mdn	12.5	76.8	84.0
Control Group (n =14)	M	27.9	61.1	67.8
	SD	30.73	38.75	34.44
	Mdn	19.2	72.4	80.4

Figure 5.2
Mean accuracy scores
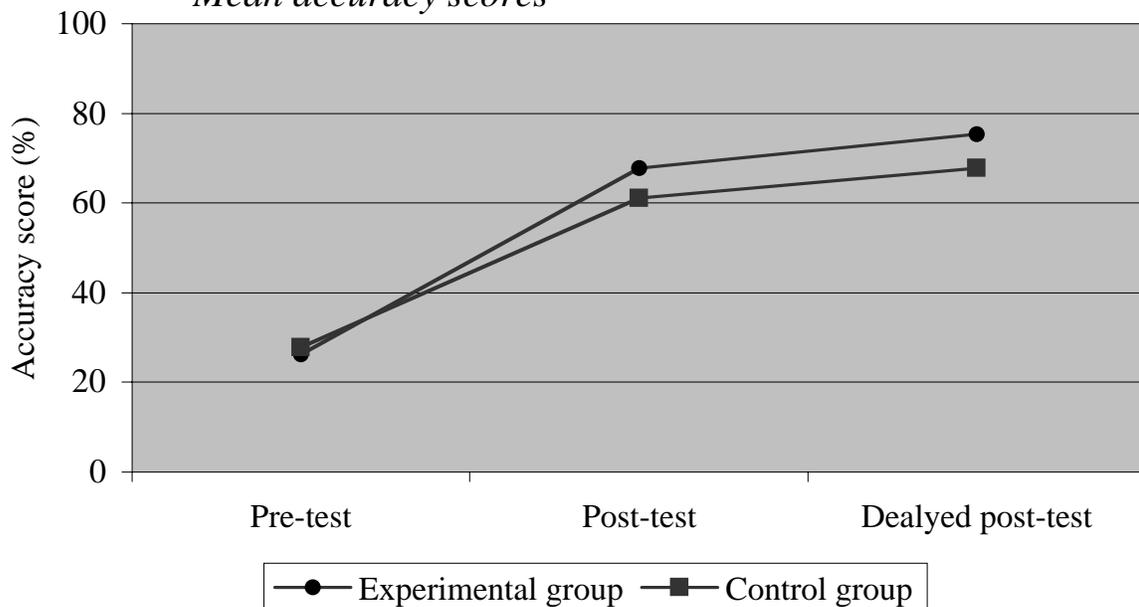
Both the experimental group and the control group showed a similar steep rate of improvement in the post-test. The high scores in the post-test were not only maintained but were improved further in the delayed post-test in both groups. The gains in the experimental group in accuracy score were 41.6% (post-test score minus pre-test score) and 49.1% (delayed post-test score minus pre-test score) while those of the control group were 33.2% and 39.9% respectively (see Appendix 5.J for gains of individual participants). One of the possible reasons for the improvement in the post-test is task repetition effects as discussed in Chapter Four. Another possible source is the benefit of treatment activities themselves in which participants individually interacted with a native speaker interlocutor.

A 2 (group) x 3 (tests) mixed design ANOVA was conducted to evaluate the impact of MO on grammatical accuracy in the three tests. The results indicated that there was no significant main effect of group, F (1, 26) = .165, p > .05. There was a significant main effect of test, F (1.33, 34.56) = 42.37, p < .001. Bonferroni corrected post hoc tests showed that the main effects reflect a significant difference (p < .001) between the pre-test and the post-test, and between the pre-test and the delayed post-test, but not between the post-test and the delayed post-test (p > .05). There was no significant interaction between test and group, F (1.33 34.56) = .482, p > .05.

Hypothesis 1.1, which predicted the experimental group would outperform the control group in the gains of grammatical accuracy score, was not supported.

5.3.2 Hypothesis 1.2

The second hypothesis predicted a relationship between the production of MO and gains of grammatical accuracy scores in the experimental group. A correlation was carried out to determine the strength of relationship between the production of MO in the treatment sessions and gain scores in accuracy in the post-test and delayed post-test. The two subcategories of MO (i.e., MO with a targetlike pattern, and MO with a non-targetlike pattern) were also separately calculated. This is because the process of reformulation could contribute to interlanguage development regardless of whether the MO is targetlike or non-targetlike. None of the Spearman's rank correlation coefficients reached statistical significance (see Appendices 5.K1 and K2 for the summary of the data, and Appendix 5.L for the results), and, therefore, no significant relation was indicated between gains and each of the three types of MO.

Gain scores and modifiers/non-modifiers

An additional analysis was carried out to further investigate the relationship between MO and accuracy gains. Following McDonough (2001), participants in the experimental group were classified as modifiers (with scores above the median of the group) and non-modifiers (with scores at or below median), according to the group median scores of the three types of MO (i.e., MO irrespective of targetlike and non-targetlike pattern, MO with targetlike pattern, and MO with non-targetlike pattern). For example, the group median score of MO with targetlike pattern was 1, and those who produced more than 1 token of MO with targetlike pattern were categorised as modifiers while those who produced 1 or less were assigned as non-modifiers.

Modifiers with targetlike pattern (n = 4) produced a median of 2.5 instances and non-modifiers (n = 10) produced a median of 1.0.

The aim of this classification was to maximise differences between those who produced MO more and those who produced less. The results of the Mann-Whitney tests in Table 5.9 showed the difference was significant between modifiers and non-modifiers for each of three MO types and confirmed that the groups represented different populations.

Table 5.9
Instances of MO produced by modifiers and non-modifiers (Experimental Group)

| | Modifiers (above median) | | | | Non-modifiers (at or below median) | | | | | |
	N	Total	Mdn	Range	N	Total	Mdn	Range	Z	p
MO	7	24	2.0	2-8	7	5	1.0	0-1	-3.282	.001
MO with TL	4	12	2.5	2-5	10	6	1.0	0-1	-2.983	.003
MO with NTL	7	11	1.0	1-3	7	0	0.0	0	-3.435	.001

Note. MO = MO with TL and MO with NTL, TL = targetlike pattern, NTL = non-targetlike pattern

As shown in Table 5.10, each median for the gain scores of accuracy in the post-test and delayed post-test for modifiers was considerably higher than for non-modifiers, (see Appendices 5.M1-M3 for the individual data of each group). The median gain scores for both post-test and delayed post-test for the modifiers and non-modifiers were compared by Mann-Whitney tests. The results (Table 5.10) indicated the differences between the two groups were not significant in gain scores for each of the three types of MO. Therefore, the hypothesis concerning the relationship between

MO and accuracy gains was not supported even though the two groups were intentionally created for differentiation of those who produced more and less MO.

Table 5.10
The median gain scores by modifiers and non-modifiers (Experimental Group)

	Modifiers		Non-modifiers			
	Mdn	range	Mdn	range	Z	*p*
MO	Modifier (n = 7)		Non-modifier (n = 7)			
Post-test gain	55.6	4.4 - 78.6	39.5	-2.3 - 78.6	-1.023	.306
Delayed post-test gain	46.4	-12.7 - 100.0	35.9	15.4 - 100.0	-.578	.563
MO with TL	Modifier (n = 4)		Non-modifier (n = 10)			
Post-test gain	42.9	4.4 - 75.0	40.9	-2.3 - 78.6	-.142	.887
Delayed post-test gain	56.0	-12.7 -100.0	41.2	15.4 - 100.0	-.071	.943
MO with NTL	Modifier (n = 7)		Non-modifier (n = 7)			
Post-test gain	41.7	4.4 – 78.6	26.7	- 2.3 - 78.6	-.767	.443
Delayed post-test gain	50.0	-12.7 -100.0	35.9	25.0 - 100.0	-.322	.748

Note. Post-test gain = gain from pre-test to post-test, delayed post-gain = gain from pre-test to delayed-post test

5.3.3 Summary of the findings for Research Question 1

The results for the comparison of accuracy scores between the groups did not support Hypothesis 1.1, which predicted that the experimental group would show greater gains in accuracy scores than the control group. This is entirely consistent with the findings by Takashima and Ellis (1999) and O'Relly et al. (2001), but different from the earlier, smaller scale study carried out by Nobuyoshi and Ellis (1993). The results of correlation and the further analysis did not support the second hypothesis of the relationship between production of MO and gain scores in grammatical accuracy. The impact of MO was not demonstrated in accuracy.

5.4. Results for Research Question 2: Evidence of a relationship between the production of MO and interlanguage development

The second research question asked is whether or not there is a positive relationship between the production of MO and interlanguage development. In relation to this question, the following hypotheses on three aspects of interlanguage were tested.

Hypothesis 2.1a: The variations of non-targetlike negation patterns will diminish in both the experimental group and the control groups.

Hypothesis 2.1b: The decrease in the variation of non-targetlike negation patterns will be greater in the output of the experimental group than in that of the control group.

Hypothesis 2.2: There will be a difference in the changes of the use of types of non-targetlike pattern across the tests between the experimental group and the control group, as a result of differences associated with the production of MO.

Hypothesis 2.3: The experimental group will outperform the control group in terms of the number of participants who use higher stage forms.

5.4.1. Hypothesis 2.1a and 2.1b: Changes in the number of types of non-targetlike negation pattern

As illustrated in Chapter Three, it was predicted that participants would produce fewer types of non-targetlike pattern as interlanguage developed, and that the decrease in the number of types of pattern in the experimental group would be greater than that in the control group because of the impact of the production of MO.

Table 5.11 shows the numbers of one-type, two-type and three-type users in each group across the three tests. Some participants used only one type of non-targetlike pattern in each test and they are referred to as one-type users in the table below. It should be noted that a learner might use only one type of negation pattern but it might be different in each of the three tests. Participants who used two types of non-targetlike negation pattern are called two-type users, and those who used three are called three-type users.

Table 5.11
Numbers of one-type users, two-type users and three-type users in each test

	Experimental Group			Control Group		
	Pre-test	Post-test	Delayed post-test	Pre-test	Post-test	Delayed post-test
One-type users	5	2	2	3	2	2
Two-type users	4	2	2	4	1	1
Three-type users	0	2	1	2	0	1

The data shows a trend that that the number of one-type, two-type and three-types users decreased across the tests in both the experimental and the control groups. There were no substantial differences in the change of the number of participants who used one, two or three non-targetlike patterns.

Table 5.12 shows the types of non-targetlike negation pattern produced by each of one-type, two-type and three-type users, since Table 5.11 does not give information about whether participants used the same type of non-targetlike pattern used across the tests.

Table 5.12

Types of non-targetlike pattern used by one-type users, two-type users and three-type users, and the number of participants in each test

	Experimental Group			Control Group		
	Pre-test (n = 14)	Post-test (n = 14)	Delayed post-test (n = 8)	Pre-test (n = 14)	Post-test (n = 9)	Delayed post-test (n = 10)
Types of pattern used by one-type users						
nai	2			1		
masen	1					
janai	4	2	3	4	4	5
kujanai	1		1		1	
kunai	1	5		1		1
Variation of types of pattern used	5	2	2	3	2	2
Types of pattern used by two- type users						
nai-masen	1					
nai-janai	1			2		
nai-kujanai				1		
nai-kunai			2			
janai-masen	1					
janai-kujanai		2		2		
janai-kunai	2	3		1	4	3
kunai-masen			1			
Variation of types of pattern used	4	2	2	4	1	1
Types of pattern used by three-type users						
nai-masen-janai				1		
nai -masen-kunai		1				
nai- janai-kunai				1		1
janai-kunai-masen		1				
janai-kunai-kujanai			1			
Variation of types of pattern used	0	2	1	2	0	1

As predicted, fewer types of non-targetlike pattern were used in the post-tests than in the pre-test, which is a common trend in both of the groups. Thus, the first part of the hypothesis on the decrease in the variation of the types of non-targetlike negation pattern, based on the findings of JSL (Kamura, 2001b), was supported in JFL as well. One common trend between the experimental group and the control group is that *janai* continued to be used from the pre-test to post-tests by one-type users. However, there are no clear differences between the two groups, and the decrease was not greater in the output of the experimental group than that of the control group. Therefore, the second part of the hypothesis regarding the decrease in the experimental group being greater than that of the control group was not supported.

5.4.2   Hypothesis 2.2: Changes in the types of non-targetlike negation pattern

This section reports the analysis of the changes in the types of non-targetlike negation

pattern over the three tests.   Changes across the three tests within the group and

differences between the groups are described, and contributing factors for the changes

and differences are analysed with reference to the production of MO.   Tables 5.13

and 5.14 display the frequency of each type of non-targetlike pattern, group mean,

and percentage in each of the test.   The data for the control group are followed by

those for the experimental group for discussion purpose in this section.   The data for

individual participants and use of each type of pattern are presented in Appendices

5.N and 5.O.

Table 5.13
Non-targetlike negation patterns in the control group (n = 14)

	Pre-test			Post-test			Delayed post-test		
Pattern	Total	M	%	Total	M	%	Total	M	%
nai	12	0.9	12.6	0	0.0	0.0	1	0.1	1.4
masen	1	0.1	1.1	0	0.0	0.0	0	0.0	0.0
janai	65	4.6	68.4	65	4.6	84.4	59	4.2	79.7
kujanai	3	0.2	3.2	1	0.1	1.3	0	0.0	0.0
kunai	5	0.4	5.3	7	0.5	9.1	7	0.5	9.5
others	9	0.6	9.5	4	0.3	5.2	7	0.5	9.5
Total	95	6.8	100.0	77	5.5	100.0	74	5.3	100.0

Note. Except for *kunai*, the patterns listed are all non-targetlike patterns for constructing the
negation of adjectives.

Table 5.14
Non-targetlike negation patterns in the experimental group (n = 14)

	Pre-test			Post-test			Delayed post-test		
Pattern	Total	M	%	Total	M	%	Total	M	%
nai	8	0.6	12.3	2	0.1	3.8	2	0.1	6.7
masen	13	0.9	20.0	2	0.1	3.8	1	0.1	3.3
janai	32	2.3	49.2	15	1.1	28.8	12	0.9	40.0
kujanai	1	0.1	1.5	7	0.5	13.5	3	0.2	10.0
kunai	7	0.5	10.8	25	1.8	48.1	11	0.8	36.7
others	4	0.3	6.2	1	0.1	1.9	1	0.1	3.3
Total	65	4.6	100.0	52	3.7	100.0	30	2.1	100.0

5.4.2.1   Changes across the tests within group

Frequency of types of non-targetlike pattern

Figures 5.3 and 5.4 graphically present the distribution of each type of non-targetlike

pattern across the tests by group.

Figure 5.3:
*Percentages of each type of negation pattern
in the control group*
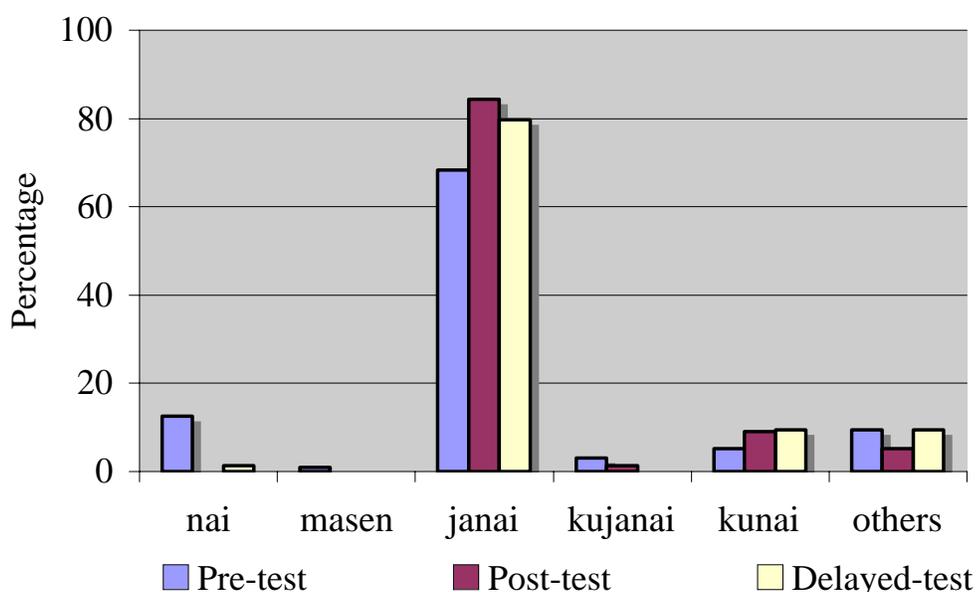
In the control group, there was little change in the percentage of each type of pattern

across the three tests (Figure 5.3).   The use of *janai* was the most frequent

throughout the tests (68.4% in the pre-test, 84.4% in the post-test and 79.7% in the

delayed post-test), and it is worth remarking that these results are consistent with the

findings that showed the predominant use of *janai* (Kamura, 2001b; Kanagy, 1991;

Noro, 1995).   *Nai* was the second largest (12.6%), following *janai* in the pre-test, but

it almost disappeared in the post-tests.   The use of *kunai* only increased slightly, but

the distribution itself was small, from 5.3% in the pre-test to 9.1% in the post-test and

9.5% in the delayed post-test.

In contrast, in the experimental group, some interesting changes in the distribution of

each type of pattern across the tests were observed (Figure 5.4).

Figure 5.4:
Percentages of each type of negation patter
in the experimental group
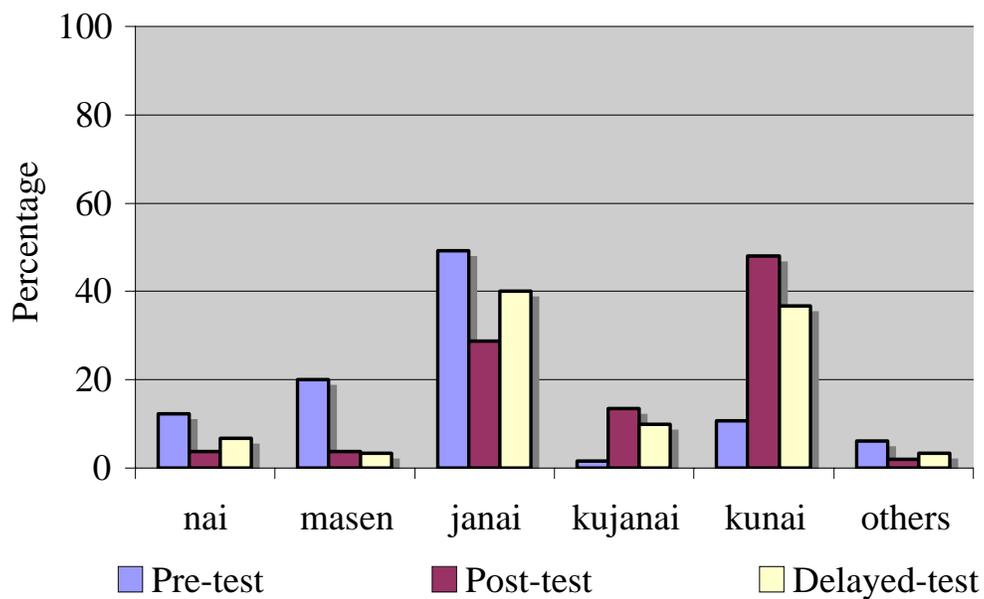
Janai showed the largest distribution (49.2%) in the pre-test, but it dropped to 28.8%

in the post-test. In the delayed-post-test, the distribution of *janai* became largest

again. Another major change was *kunai*, which increased from 10.8% in the pre-test

to 48.1% in the post-test and outnumbered *janai*. The distribution of *kunai* remained

high (36.7%) in the delayed post-test, following *janai* (40.0%). There were some

changes with *masen, nai,* and *kujanai* in the post-test. Both *nai* and *masen* decreased in the post-test, but the distributions of each type were small. *Kujanai* increased in the post-test, and it may signal a developmental change in interlanguage (Kanagy, 1991), as it is a mixture of analysis of predicate category (use of the adverbial inflector, *ku*) and non-analysis of negation patterns (choice of incorrect negation pattern, *janai*). However, its distribution was small in comparison with other types of negation pattern.

Number of participants by types of negation patterns

Figures 5.5 shows the changes in the number of participants who used each type of pattern in the control group. The number of participants who used each type of non-targetlike negation pattern did not change much except *nai*, which decreased from six in the pre-test to zero in the post-test. *Janai* was used by the majority of participants throughout the tests, and *kunai* was the only non-targetlike pattern which showed a steady increase in the number of participants who used it.

Figure 5.5:
Number of participants by types of negation pattern
in the control group (n = 14)

127
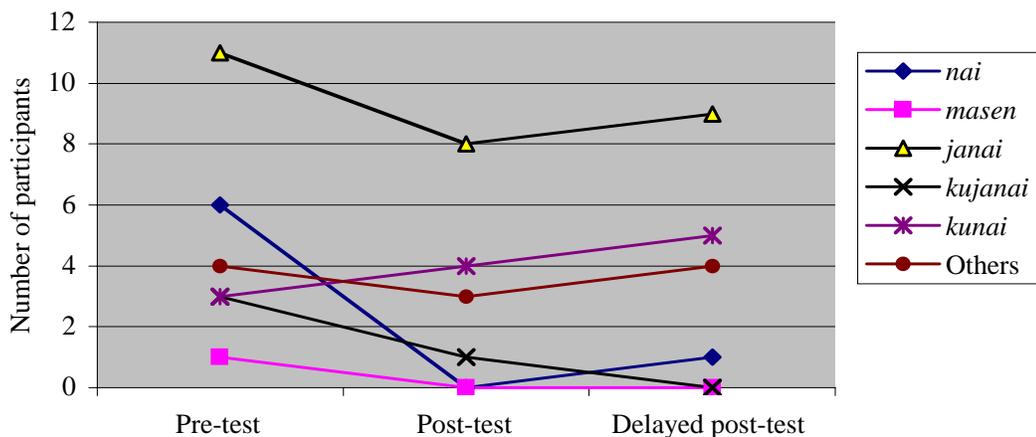
Figure 5.6 shows that the number of *janai* users in the experimental group was the same (n = 8) between the pre-test and post-test despite the distributional decrease of frequency from 49.2% in the pre-test to 28.8% in the post-test, as shown in Table 5.14. The number of *janai* users halved in the delayed post-test although the frequency of *janai* increased from 28.8% in the post-test to 40.0% in the delayed post-test. Another dynamic change is the number of *kunai* users, which showed a sharp increase from three in the pre-test to ten in the post-test, and a sharp drop to four in the delayed post-test. The number of *nai* users dropped, which was a similar trend to that in the control group.

Figure 5.6:
Number of participants by types of negation pattern in the experimental group (n = 14)
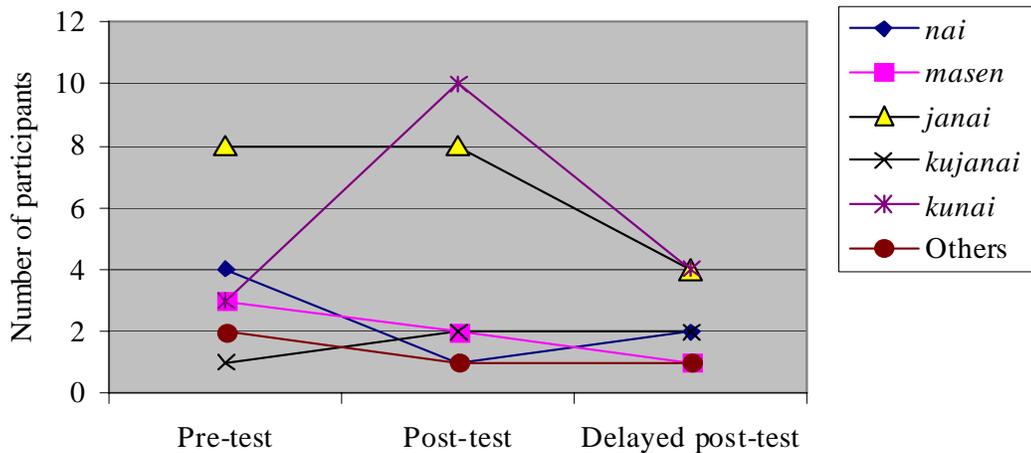
5.4.2.2　　Differences between groups across tests

While the changes in the distribution of types of non-targetlike negation pattern and in the number of users of each pattern across the tests were limited in the control group, there were more dynamic and substantial changes in the experimental group.

The dominant use of *janai* across the tests was common to the both groups, but one of the major differences between them was a decrease in the use of *janai* and an increase in *kunai* in the post-test in the experimental group.　These changes may be associated with the impact of the production of MO.　In other words, the positive impact of the production of MO might suppress the use of *janai* (a non-targetlike pattern for negating adjectives) and facilitate participants to more use of *kunai* (a correct negation pattern for adjectives but the inflectional change is still not targetlike) in the post-test.　Also, the relatively high percentage of *kunai* in the delayed post-test in the experimental group may indicate the sustained impact of MO.

It is important to note that both the experimental and the control groups significantly improved accuracy scores across the tests, and more importantly, they showed a similar improvement (section 5.3.1).　On the other hand, there were some differences in learners' interlanguage reflected by the frequency of non-targetlike patterns and the number of participants who used each type of pattern between the groups.　These contrasting results between the experimental group and the control group may imply an impact of MO on interlanguage development but little impact on grammatical accuracy.

The next sections discuss the contributing factors for the decrease in *janai* and increase in *kunai*, and their relationship with MO.

5.4.2.3   Contributing factors for the decrease of the non-targetlike pattern *janai*, and production of MO

Table 5.15 shows data about the use of *janai* in the three tests.   Despite the relatively large percentage decrease in the frequency of *janai* from the pre-test (49.2%) to the post-test (28.8%), the number of participants who used *janai* (n = 8) did not change from the pre-test to the post-test.   On the other hand, the mean occurrences of *janai* by the participants who used it changed from 4.0 in the pre-test to 1.9 in the post-test. In short, the distribution of *janai* decreased in the post-test not because of the decrease in the number of *janai* users, but because of the decrease in frequency of non-targetlike instances of *janai*.

Table 5.15
Janai users in the three tests

Participant	Experimental Group (n = 14)			Control Group (n =14)		
	Pre-test	Post-test	Delayed-test	Pre-test	Post-test	Delayed-test
Number of users	8	8	4	11	8	9
Frequency of *janai*/all the non-targetlike patterns	32/65 (49.2%)	15/52 (28.8%)	12/30 (40.0%)	65/95 (68.4%)	65/77 (84.4%)	59/74 (79.7%)
M by users	4.0	1.9	3.0	5.1	8.1	6.6
SD	2.45	1.46	3.37	4.21	5.17	6.56
Mdn by users	4.0	1.0	1.5	3.0	7.5	5.0

 Note. Mean and medians are calculated, based on the frequency by participants who used *janai*.

An analysis of the membership of *janai* users in the experimental group, however, revealed the membership was different, as shown in Table 5.16.

Table 5.16
Group membership and frequency of janai users in each test

Participant	Experimental Group (n = 14)			Control Group (n =14)		
	Pre-test (n = 8)	Post-test (n = 8)	Delayed-test (n = 4)	Pre-test (n = 11)	Post-test (n = 8)	Delayed-test (n = 9)
P1				1		
P2	1	1	1			
P3		1				
P4	1	1		7	11	11
P5	3	5	8	3		1
P6	8	1	2	13	3	1
P7	3			6		
P8	5			6	2	1
P9	6			1	7	5
P10				12	14	17
P11	5	3	1	2	16	16
P12		2		2	8	6
P13				3	4	1
P14		1				
Total	32	15	12	65	65	59

There were eight participants who used *janai* in the pre-test and five of them (P2, P4, P5, P6, P11) used *janai* in both the pre-test and post-test. Three participants (P7, P8, P9) used it in the pre-test but did not use it in the post-test while another three (P3, P12, P14) who did not use *janai* in the pre-test used it in the post-test.

The main factor for the decrease of the frequency of *janai* was that the three participants (P7, P8, P9) who used *janai* in the pre-test did not use it in the post-test. They produced 14 (43.8%) in a total of 32 tokens of *janai* in the pre-test. The decreased use of *janai* is attributable to the difference in personnel using this non-targetlike pattern.

An additional analysis on the three participants revealed that none of them produced MO on *janai* at all, although two of them produced MO with a targetlike pattern and MO with a non-targetlike pattern on *kunai* (see Appendix 5.P).   No trend was found on MO among the three participants, and the relationship between the non-use of *janai* in the post-test and production of MO was not confirmed.

5.4.2.4   Contributing factors for the increase of the non-targetlike pattern *kunai,* and production of MO

Use of *kunai* was assumed to be an indication of development in terms of choosing the right pattern although the inflection of stem was still non-targetlike.   Table 5.17 shows that the main factor of the increase of *kunai* was the increase in the number of users of *kunai*.   It increased from three in the pre-test to ten in the post-test, while mean frequency of *kunai* users changed from 2.3 in the pre-test to 2.5 in the post-test.

Table 5.17
Kunai users in the three tests

Participant	Experimental Group (n = 14)			Control Group (n = 14)		
	Pre-test	Post-test	Delayed-test	Pre-test	Post-test	Delayed-test
Number of users	3	10	4	3	4	5
Frequency of *kunai/*all the non-targetlike patterns	7/65 (10.8%)	25/52 (48.1%)	11/30 (36.7%)	5/95 (5.3%)	7/77 (9.1%)	7/74 (9.5%)
M by user	2.3	2.5	2.8	1.7	1.8	1.4
SD	1.53	1.18	1.71	1.15	0.96	0.89
Mdn by user	2.0	3.0	2.5	1.0	1.5	1.0

Production of MO might be associated with the change in participants' interlanguage from non-use of *kunai* in the pre-test to use of *kunai* in the post-test, and this was further examined.   Table 5.18 shows the membership of *kunai* users in the three tests.

Table 5.18
Group membership and frequency of kunai users in each test

Participant	Experimental group (n = 14)			Control group (n = 14)		
	Pre-test (n =3)	Post-test (n = 10)	Delayed-test (n = 4)	Pre-test (n =3)	Post-test (n = 4)	Delayed-test (n = 5)
P1	2	3				
P2	1					
P3		1		3		1
P4	4	1			3	1
P5				1		
P6		2	1			
P7		4	3			
P8		3			1	
P9		3			1	1
P10		3	2			
P11				1		
P12		4	5			3
P13		1			2	1
P14						
Total	7	25	11	5	7	7

The experimental group had eight participants who did not use *kunai* in the pre-test but used it in the post-test.   Half of those did not use *kunai* in the delayed post-test. On the other hand, there were only four of those in the control group, and only one did not use it in the delayed post-test.

Table 5.19 lists the eight participants in the experimental group who did not use *kunai* in the pre-test but used it in the post-test, and the three participants who did not use *kunai* either in the pre-test or in the post-test.   The frequency of MO on *kunai* and its breakdown are presented.

133

Table 5.19
MO, MOTL and MONTL by the use of kunai in the pre-test and post-test (Experimental Group)

Participants who did not use *kunai* in the pre-test but used it in the post-test				Participants who did not use *kunai* either in the pre-test or in the post-test			
Participant	MO	MOTL	MONTL	Participant	MO	MOTL	MONTL
P3	2	2	0	P5	6	3	3
P6	1	1	0	P11	2	1	1
P7	2	1	1	P14	1	0	1
P8	2	1	1				
P9	0	0	0				
P10	8	5	3				
P12	1	1	0				
P13	2	2	0				
M	2.3	1.6	0.6	M	3.0	1.3	1.7
SD	2.43	1.51	1.06	SD	2.65	1.53	1.15
Mdn	2.0	1.0	0.0	Mdn	2.0	1.0	1.0

Note. MO = modified output in response to clarification requests, MOTL = MO with a targetlike pattern, MONTL = MO with a non-targetlike pattern

The median for each of MO, MOTL and MONTL was compared using Mann-Whitney tests, and the difference was not significant, Z = -.433, p = .776 for MO, Z = -.323, p = .776 for MOTL, Z = -1.652, p = .133 for MONTL. Therefore, the relationship between the production of MO and the use of *kunai* in the post-test was not confirmed.

In sum, the control group showed little difference in the distribution of the types of non-targetlike negation pattern over the three tests. In other words, participants in the control group became able to produce more targetlike pattern than non-targetlike patterns (section 5.3.1), but the types of non-targetlike pattern they used changed little over the tests. The experimental group showed a sharp increase in the use of *kunai*, which seems to be attributable to the increase in the number of participants who used *kunai.*

A possible explanation is that the impact of MO facilitated those who did not use *kunai* in the pre-test to use it the post-test. Although additional analyses did not confirm the relationship between these changes and the production of MO, the changes observed in the experimental group in the post-test might be associated with factors involving MO.

5.4.3 Hypothesis 2.3: Changes in the developmental stages

The aim of this section is to examine the hypothesis which predicted that the experimental group would outperform the control group in terms of the number of participants who used higher stage forms. The negative responses of each participant were analysed individually and each response was assigned to one of the three stages. As described in section 4.9.3.2, two different lexical items were used to determine the attainment at each stage. The number of participants who produced negation patterns at each stage in each test is presented in Table 5.20.

Table 5.20
Number of participants by stage categories in each test

	Stage 1 Use of the pattern *nai*	Stage 2 Absence of the pattern *nai*	Stage 3 Targetlike pattern
Experimental Group (n = 14)			
Pre-test	2	9	7
Post-test	1	10	13
Delayed-test	0	6	14
Control Group (n = 14)			
Pre-test	3	10	7
Post-test	0	8	11
Delayed-test	0	5	12

Note. There were cases in which some participants did not have 'two different usages' to attain any of the stages, and they were not assigned to any stage. Also, some of the total number of participants of the three stages for each test exceeds the total number of participants in each group (n = 14), as participants could be assigned to all the stages if they used the forms for each stage.

135

Although the present study used a randomised block design based on accuracy scores in the pre-test, the number of participants from each stage at the point of the pre-test happened to be similar between the two groups. Seven participants in each group produced targetlike patterns (Stage 3) at least twice in the pre-test. The number of participants who did not use *nai* but used other types of non-targetlike pattern (Stage 2) was nine in the experimental group and ten in the control group. There were only two in the experimental group and three in the control group who used *nai* (Stage 1) in the pre-test. The differences between the groups are not substantial and no impact of MO was observed in the language development in the experimental group. Between the pre-test and post-test, six more students in the experimental group and four in the control group advanced to Stage 3 (use of targetlike pattern). All the participants in the experimental group reached Stage 3, as one more participant in the experimental group advanced to Stage 3 in the delayed post-test. The control group also had another participant who advanced to Stage 3 in the delayed post-test, but there were still two participants who did not produce Stage 3 forms.

The results did not show evidence of the experimental group outperforming the control group in terms of the number of participants produced negation patterns from a higher stage, and, therefore, did not support the Hypotheses 2.3.

5.4.4 Summary of the findings for Research Question 2

The relationship between MO and interlanguage development was examined in terms of the number of types of non-targetlike negation pattern, types of pattern, and developmental stages. There appeared to be some differences in the types of pattern

in interlanguage development reflected by the frequency of non-targetlike patterns between the experimental group and the control group. Therefore, a positive relationship between the production of MO and interlanguage development can be tentatively suggested.

5.5   Results for Research Question 3: Evidence of relationship between the MO and subsequent use

The third research question addressed whether or not there is a relationship between the non-targetlike forms that the learners modify and their subsequent use in terms of accuracy and interlanguage development. The focus of the analysis is on the data of the experimental group, as the control group were not given clarification requests to trigger modified output. The following hypotheses were tested.

Hypothesis 3.1: Participants who produce MO (i.e., modified output in response to clarification requests) in the treatment sessions will be sensitised to avoid the same non-targetlike pattern in subsequent situations of use.

Hypothesis 3.2: There will be a relationship between the types of non-targetlike negation pattern that participants modified and their subsequent use.

Section 5.5.1 summarises the data for the analysis for Research Question 3. Section 5.5.2 examines the subsequent use of the non-targetlike negation patterns that participants modified in the treatment activities. The relationship between the types of pattern participants modified and their subsequent use is discussed in sections 5.5.3 and 5.5.4. Section 5.5.5 summarises the findings.

5.5.1 Data for the analysis of subsequent use of MO

Out of 70 responses to clarification requests in treatment sessions, 34 were Repetition (i.e., repetition of the initial non-targetlike pattern) and 29 were MO, as reported in section 5.1. In some cases, participants produced MO relating to the same type of non-targetlike negation patterns more than once (e.g., a participant who, on three different occasions, was given three clarification requests on the use of *janai,* produced two instances of MO with a targetlike pattern), and each case was counted as one instance of MO. The total number of instances of MO by types of non-targetlike pattern and by different participants was 15, and that of Repetition was 18.

Individual participants' responses to clarification requests on the same non-targetlike pattern were often inconsistent. While there were participants who continued to respond to clarification requests to the same type of non-targetlike pattern with MO, another participant who was given three clarification requests on the use of *janai* responded with three types of response move, MO, Repetition and other types of response (e.g., change of vocabulary, giving up answering half way, and use of non-Japanese language, shown as *Others* hereafter).

It was considered necessary to separate instances of a single type of response move from such multiple types of response move, and the 15 instances of MO were divided into two categories by whether or not they involved response moves other than MO (i.e., Repetition and Others). There were seven instances in which participants modified the non-targetlike negation pattern as shown by the examples of Participant

A and B in Table 5.21.   There were eight other instances that involved multiple types

of response move including MO (Participant C in Table 5.21).

Table 5.21
Single type of response move of MO (Participants A & B) and multiple types of response move including MO (Participant C)

Examples of single type of response move of MO

	Initial non-targetlike use		Response to clarification requests	Category of response moves
Participant A	*shiroi janai* (It is not white)	→	*shiro kunai*	MO with a targetlike pattern
Participant B	*shiroi janai* (It is not white)	→	**shiro masen*	MO with a non-targetlike pattern

An example of multiple types of response move including MO

	Initial non-targetlike use		Response to clarification requests	Category of response moves
Participant C	*shiroi janai* (It is not white)	→	*shiro kunai*	MO with a targetlike pattern
	*shiroi janai (It is not white)	→	**shiroi masen*	MO with a non-targetlike pattern
	*shiroi janai (It is not white)	→	**shiroi janai*	Repetition

The data analysed for Research Question 3 are summarised in Figure 5.7.   Also, see

Appendix 5.Q for the summary of individual data.

Figure 5.7
Summary of data for Research Question 3
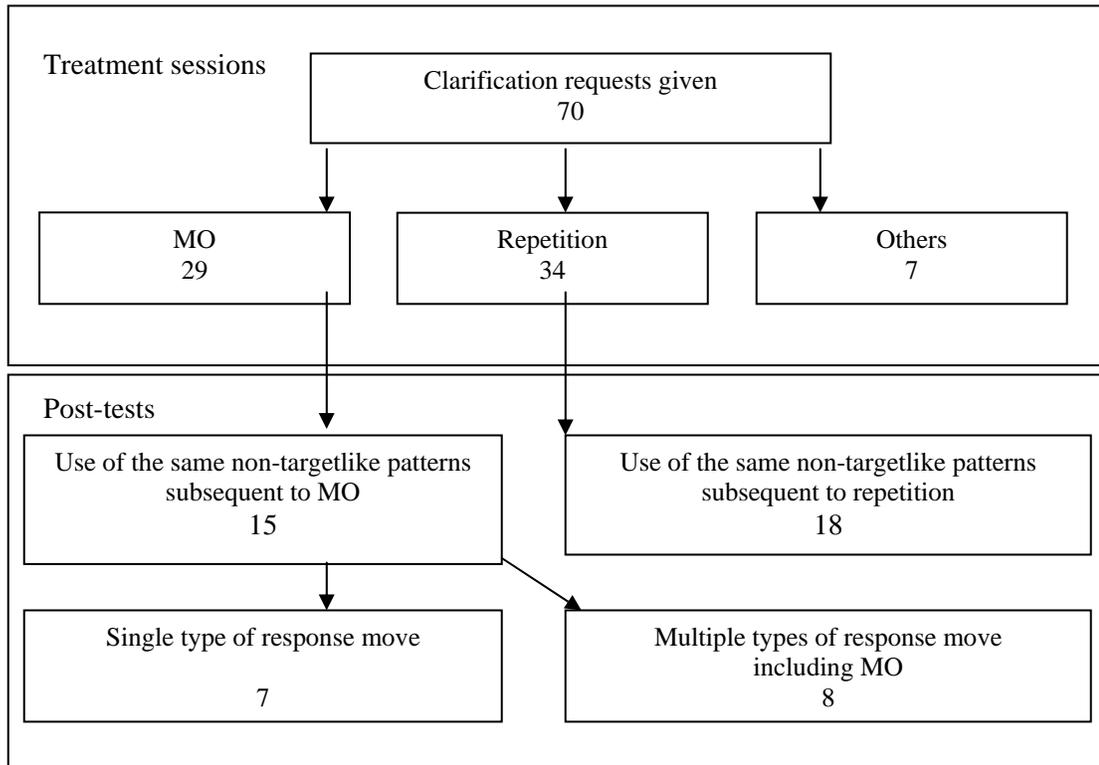
5.5.2   Hypothesis 3.1: Sensitisation to avoid the same non-targetlike pattern

In order to test Hypothesis 3.1, subsequent use of the following three aspects were investigated: (1) MO and Repetition, (2) MO with a targetlike pattern and MO with a non-targetlike pattern, and (3) MO as a single type of response move and multiple types of response move including MO.

5.5.2.1   MO and Repetition

The following results are based on the seven instances of subsequent use by a single type of response move (i.e., the data do not include multiple types of response move relating to a non-targetlike pattern).   Table 5.22 displays the number of instances of

subsequent use of the non-targetlike negation patterns that participants had modified in the treatment activities, by combining the two post-tests. Also, the data for use subsequent to Repetition as a single type of response move are shown for comparison. (See the explanation about the hypothesised degree of impact of MO in section 4.9.4.)

Table 5.22
Use of the same type of non-targetlike pattern subsequent to MO and to Repetition by combination of the post-tests

Subsequent use of the same type of pattern which participants modified	Hypothesised degree of impact of MO	Types of response in the treatment sessions	
		MO (n = 7)	Repetition (n = 9)
Non-use in both of the post-tests	Strong impact	4	6
Use of the delayed post-test only	Temporary impact	1	0
Use in the post-test only	Delayed impact	2	2
Use in the both of the post-tests	No impact	0	1

Note. MO includes both MO with a targetlike pattern and MO with a non-targetlike pattern.
    NTL = non-targetlike

Statistical tests were not carried out due to the small sample size, but trends can be seen from Table 5.22. Out of the seven instances of MO as a single type of response move, there were four instances of non-use of the same non-targetlike negation patterns in both of the post-tests, which participants previously modified. Although there are instances in which participants used the same types of pattern in the post-test (n = 2) and in the delayed post-test (n = 1), no instances were found in which participants continued to use the same type of non-targetlike pattern they previously modified, in each of the post-tests.

This seems to support the hypothesis that some positive impact of MO would sensitise participants to avoid the use of the non-targetlike negation pattern that they had previously modified. However, similar results were shown in the subsequent use of the non-targetlike patterns that participants repeated in response to clarification requests. Out of the nine instances of Repetition as a single type of response move, there were six instances in which the types of pattern responded to with Repetition were not used in either of the post-tests. Thus, the results suggested that participants who had previously produced MO were no more likely to use the same non-targetlike pattern in subsequent use than those who repeated the initial non-targetlike pattern.

5.5.2.2 MO with a targetlike pattern and with a non-targetlike pattern

It was investigated whether there was a difference in the use subsequent to MO with a targetlike pattern and with a non-targetlike pattern. Table 5.23 displays the breakdown of the seven instances of subsequent use to MO.

Table 5.23
Use subsequent to MO with a targetlike pattern and MO with a non-targetlike pattern

	MO with TL (n = 5)	MO with NTL (n = 2)
Non-use of in both of the post-tests	2	2
Use in the delayed post-test only	1	0
Use of in the post-test only	2	0
Use of in both of the post-tests	0	0

Note. MO with TL is the instance in which participants produced only MO with a targetlike pattern, and it excludes all the other types of response move, including MO with a non-targetlike pattern.

There were two instances in which the type of pattern that participants modified with a targetlike pattern was not used in either of the post-tests. However, there were also two instances in which participants produced MO with a non-targetlike pattern but did not use the pattern in either of the post-tests (i.e., participants could not get it right although they tried to modify it).

A possible explanation is that the process of production of MO itself may be more important than whether MO is targetlike or not (Mackey, 2007b, p. 22). Even though MO was produced with a non-targetlike pattern, it might have facilitated language development by contributing to the sensitising of participants to avoidance of the use of the same non-targetlike pattern in subsequent use. However, the results can only be considered as indicative due to the small number of instances. In addition, the possibility cannot be discounted that the avoidance of a certain type of pattern in the post-tests following MO may not necessarily be evidence of language development, but it may rather be attributable to other learning factors.

5.5.2.3   A single type of response move and multiple types of response move

This section discusses the relationship between the types of response move (a single type of response move and multiple types of response move) and subsequent use. Table 5.24 presents the number of subsequent uses and non-uses of the types of pattern participants modified by response move (see Appendix 5.R for more detailed data). The data for Repetition as a single type of response move is also included for comparison.

Table 5.24

Non-use and use of the negation pattern in the post-tests by types of response move in the treatment sessions

| Treatment sessions | Frequency and percentage | | | | | |
| | Post-test | | | Delayed post-test | | |
	Non-use	Use (non - targetlike)	Total	Non-use	Use (non - targetlike)	Total
MO as a single type of response move	5 (71.4%)	2 (28.6%)	7 (100%)	6 (85.7%)	1 (14.3%)	7 (100%)
Multiple types of response move including MO	1 (12.5%)	7 (87.5%)	8 (100%)	4 (50.0%)	4 (50.0%)	8 (100%)
Repetition as a single type of response move	5 (55.6%)	4 (44.5%)	9 (100%)	7 (77.8%)	2 (22.2%)	9 (100%)

There is a trend that participants who responded to clarification requests with multiple types of response were more likely to use the same non-targetlike pattern they had previously modified in subsequent use than participants who responded with MO with a single response move. The same types of non-targetlike negation pattern that were responded to with multiple types of response move, including MO with a targetlike pattern, were used in the post-tests in seven out of eight instances of multiple types of response move.

For example, Excerpts 1, 2 and 3 show that Mary (Participant 13) produced three non-targetlike patterns involving *masen* in the treatment sessions, and responded to each of the three clarification requests with three different types of response move (i.e., MO, Repetition and Others). This variation may be an indication of the

instability of learner's interlanguage, and she still used the same *masen* in the post-test.

Excerpt 1 from the interview: MO with a targetlike pattern

Interlocutor: *Nyuujiirando de kuruma wa takai desu ka.*

Are cars expensive in New Zealand?

Mary: *Taka … taka … ku … *maesn.*

(They are) not expensive.

Interlocutor: *N.*   ←**Clarification request**

Pardon?

Mary: *Iie … taka … kunai.*

It is not high.   ←**MO with a targetlike pattern**

Excerpt 2 from the interview: Repetition of the initial non-targetlike pattern

Interlocutor: *Chuugoku, takai desu ka.*

Are (houses) expensive in China?

Mary: *Uchi de *takai masen.   Iie, *takai … masen.*

Houses, not expensive.   No, not expensive.

Interlocutor: *N.*   ←**Clarification request**

Pardon?

Mary: *Iie, uchi ga, uchi wa *takai masen.*   ←**Repetition**

No, houses, houses are not expensive.

Excerpt 3 from the interview: Others (i.e., other types of response)

Interlocutor:	*Chuugoku wa kuruma ga takai desu ka. Kuruma …*	
	Are cars expensive in China? Cars …	
Mary:	*Kuruma …*	
	Cars …	
Interlocutor:	*Kuruma, takai desu ka.*	
	Are cars expensive?	
Mary:	*Iie, *taku … masen.*	
	No, they are not expensive.	
Interlocutor:	*N, mooi chido.*	←**Clarification request**
	Once again?	
Mary:	*Kuruma wa … kuruma … ga arimasen*	
	There are no cars. ← **Response other than MO or repetition**	

On the other hand, there was only one instance in which the type of non-targetlike pattern responded to with multiple types of response move was not used in the post-test. It may be that the participants who produced MO as a single type of response move had a clear idea about the illocutionary intent of clarification requests, and, therefore, 71.4% of the non-targetlike patterns that were responded to only with MO were not used in the post-test. On the other hand, multiple types of response move might reflect participants' uncertainty about the use of the pattern (i.e., trial and error approach) as well as uncertain interpretation about the motivation for clarification requests. However, this is not beyond speculation, as the investigation of learners'

internal cognitive processes is not within the scope of the present study and no empirical data was collected to substantiate the possibility above.

To summarise the findings, there was not sufficient evidence to support Hypothesis 3.1. Those participants who modified non-targetlike patterns in the treatment sessions were less likely to use the same non-targetlike pattern in subsequent use. However, whether participants responded to clarification requests with MO or Repetition did not make a substantial difference to subsequent use, which makes the findings much less strong. The difference in the impact on subsequent use between MO with a targetlike pattern and MO with a non-targetlike pattern was not confirmed. However, there was some clear difference in subsequent use when participants responded to with MO only or with MO and other types of response move.

5.5.3 Hypothesis 3.2: Types of non-targetlike pattern that were modified and their subsequent use

Hypothesis 3.2 concerns the relationship between the types of non-targetlike pattern that participants modified and their subsequent use. Table 5.25 summarises data relating to the use of the same type of non-targetlike pattern in the post-tests that participants had previously modified. This includes the frequency of MO with a targetlike pattern (third row) and the subsequent use of the non-targetlike negation pattern in the post-test and delayed post-test (second, fourth and fifth rows).

Table 5.25

Subsequent use of the types of pattern responded to with targetlike MO in the treatment sessions

Participant (n = 5)	P10	P12	P2	P3	P13
Patterns modified	*nai*	*nai*	*kunai*	*kunai*	*kunai*
MO with TL in treatment	4	1	1	1	1
Post-test	-	-	-	1	1
Delayed post-test	1	-	-	-	-

Note. TL = targetlike pattern

Participant 10 (shown as P10 in the table) had four tokens of MO with a targetlike pattern (i.e., P10 was able to modify the non-targetlike pattern *nai* into a targetlike pattern four times in response to clarification requests), and did not use it in the post-test but used it again in the delayed post-test. There were two instances in which participants (P2 & P12) modified a non-targetlike pattern into a targetlike pattern in the treatment activities and did not use the same non-targetlike pattern in either of the post-tests. The other two instances were related to *kunai,* and two participants (P3 and P13) were able to produce MO with a targetlike pattern on *kunai*, but used it again only in the post-test. The results on subsequent use to MO with a targetlike pattern did not show a trend in the subsequent situations of use.

Table 5.26 displays the types of pattern of MO that resulted in non-targetlike patterns in subsequent use. The non-targetlike patterns involved were also *nai* and *kunai.* They were used in neither of the post-tests, which indicates that participants (P5 & P14) did not use the same non-targetlike pattern although they were not able to modify it into a targetlike pattern in the previous situation of use. It is not clear what

caused the non-use of the pattern, but production of MO with a non-targetlike pattern is one of the possible factors.

Table 5.26
Subsequent use of the types of pattern responded to with non-targetlike MO in the treatment sessions

Participant (n = 2)	P5	P14
Patterns modified	*nai*	*kunai*
MO with NTL in treatment	2	1
Post-test	-	-
Delayed post-test	-	-

Note. NTL = non-targetlike pattern

Data on use subsequent to Repetition are also given in Table 5.27 to provide a wider picture of use subsequent to clarification requests.   This shows that the types of pattern responding to clarification requests with Repetition were *nai, masen, janai* and *kunai*, and their subsequent use in the post-tests appears rather random.   *Janai* and *masen* were not listed at all in the tables of MO with a targetlike pattern (Table 5.25) and MO with a non-targetlike pattern (Table 5.26), but four instances out of eight in Table 5.27 had *janai*.

Table 5.27
Subsequent use of the types of pattern responded to with Repetition in the treatment sessions

Participant (n = 9)	P11	P13	P10	P2	P3	P8	P14	P5	P6
Patterns modified	*nai*	*nai*	*masen*	*janai*	*janai*	*janai*	*janai*	*kunai*	*kunai*
Repetition in treatment	1	2	2	1	3	1	1	1	1
Post-test	-	2	-	1	-	-	1	-	2
Delayed post-test	-	-	-	1	-	-	-	-	1

No overall trend was found in subsequent use of the types of pattern responded to with MO with a targetlike pattern, MO with a non-targetlike pattern or Repetition. However, the non-targetlike patterns that were responded to either by MO with a

149

targetlike pattern or a non-targetlike pattern were *nai* and *kunai,* whilst *masen*, *janai* and *kujanai* were not found. On the other hand, Table 5.27 (Repetition) includes all the types of non-targetlike pattern except *kujanai*. The relationship between types of pattern that are modified and their subsequent use may be associated with the relative ease in correctly interpreting the motivation for clarification requests when they were given in the case of *nai* and *kunai.*

5.5.4   Use subsequent to MO involving at least one targetlike pattern

The analysis in section 5.5.3 has focused on the use subsequent to MO as a single type of response move, which excludes the subsequent use of the types of non-targetlike pattern responded to with multiple types of response move. In order to broaden the view, this section discusses all the instances that involved at least one token of MO with a targetlike pattern, irrespective of other types of response move, such as MO with a non-targetlike pattern or repetition.

The results are summarised in Table 5.28, listed by types of pattern. Types of response move are indicated by TL (MO with a targetlike pattern), NTL (MO with a non-targetlike pattern), R (Repetition), and O (Others: other types of response), followed by the frequency in the third row. The figures in the fourth and fifth row of the table show the frequency of the use of pattern in each of the post-tests. There were no instances of MO on *kujanai.*

Table 5.28

Use subsequent to MO with TL in the treatment, listed by types of pattern

Participant (n = 2)	P10	P12
Patterns modified	*nai*	*nai*
Response moves	TL 4	TL 1
Post-test	-	-
Delayed post-test	1	-

Participant (n = 1)	P13
Patterns modified	*masen*
Response moves	TL 1
	R 1
	O 1
Post-test	1
Delayed post-test	-

Participant (n = 4)	P5	P6	P10	P11
Patterns modified	*janai*	*janai*	*janai*	*janai*
Response moves	TL 3	TL 1	TL 1	TL 1
	NTL 1		NTL 2	
	R 8	R 2	R 3	R 1
	O 1			O 2
Post-test	5	1	-	3
Delayed post-test	8	2	-	1

Participant (n = 5)	P2	P3	P13	P7	P8
Patterns modified	*kunai*	*kunai*	*kunai*	*kunai*	*kunai*
Response moves	TL 1	TL 1	TL 1	TL 1	TL 1
				NTL 1	NTL 1
				R 1	
Post-test	-	1	1	4	3
Delayed post-test	-	-	-	3	-

Note. TL = targetlike pattern, NTL = non-targetlike pattern, R = repetition,
 O = other types of response

The features of subsequent use by each type of pattern are summarised below.

Nai: Two participants produced MO in response to clarification requests on *nai* in the

treatment sessions, and both of them modified the non-targetlike pattern by using a

targetlike pattern. One of them (P12) did not use *nai* either in the post-test or in the

151

delayed post-test, which might be related to the impact of MO. The other participant (P10) used *nai* again in the delayed post-test. P10 had four instances of MO with TL, which means P10 used the non-targetlike negation pattern, *nai,* four times, and a clarification request was given each time, and the participant modified the initial non-targetlike pattern involving *janai* by a targetlike pattern four times. This participant did not use the same type of pattern (i.e., *janai*) in the post-test, but used it in the delayed post-test. One possible explanation is that the production of MO facilitated language development, reflected by non-use of the non-targetlike pattern in the delayed post-test, but the impact did not last long enough up to the delayed post-test. However, more samples are required to confirm this.

Masen: Only one participant (P13) produced MO with a targetlike pattern when clarification requests on *masen* were given, and the participant used it in the post-test, but not in the delayed post-test. This participant used Repetition and other moves in response to clarification requests on the use of *masen* as well as producing MO with a targetlike pattern. The delayed impact of MO might be associated with this, but it is not possible to speculate beyond this, as there was only one instance.

Janai: While one participant (P10) refrained from continuing to use *janai* in the post-tests, three of the four participants still used this in both of the post-tests despite their production of MO with a targetlike pattern in the treatment. This is considered to indicate no impact of MO on L2 learning, specifically in relation to the use of *janai*. The instance of P11 illustrates it well. P11 responded to clarification requests by one token of MO with a targetlike pattern, one Repetition and two Others. However,

P11 continued to use the same non-targetlike pattern three times in the post-test and once in the delayed post-test. This suggests that the impact of MO on subsequent use might be associated with the type of non-targetlike pattern in which MO was produced. This, on the other hand, could indicate some positive impact of MO on subsequent use, because there is only one instance (P7 on *kunai*) besides *janai* in which the same non-targetlike pattern was repeated in both of the post-tests. Also, it should be noted that all four participants responded to clarification requests to *janai* with multiple response moves, including MO with a targetlike pattern. In other words, the response move to clarification requests to *janai* was not consistent within the same participant.

Kunai: There were five participants who produced MO in response to clarification requests to *kunai* in the treatment sessions. While two participants (P2 and P11) used it in the both of post-tests, three continued to use the pattern in the post-test and one of the three used it in the delayed post-test.

To summarise the results, it cannot be shown that any specific type of non-targetlike negation pattern was less likely to be used subsequent to modification. However, it was found that *janai* was more likely to be used than other types of pattern even after the production of MO with a targetlike pattern in response to clarification requests to *janai*. *Janai* seems to be more resistant to change than other non-targetlike patterns, which raises a question about why only *janai* might be used even by learners who can modify the non-targetlike pattern involving *janai* into the targetlike pattern.

Hypothesis 3.2, that predicted a relationship between the types of pattern participants modified and their subsequent use, was supported, but it can only be considered as indicative because of the small sample size.

5.5.5 Summary of the findings for Research Question 3

The first hypothesis was not strongly supported. Although participants were less likely to use the same type of non-targetlike negation pattern in the subsequent use that they had previously modified, the results were inconclusive because the same trend was shown in the subsequent use in the types of pattern responded to with Repetition. Also, when participants responded to clarification requests with only MO, it was more likely that the non-targetlike pattern was not used in the subsequent use than multiple types of response move including MO.

The second hypothesis was supported, as a relationship was found between the production of MO and the use of *janai* subsequent to it. This suggests the possibility that the impact of MO on subsequent use may be influenced by the types of pattern that learners modify. Participants might continue to use a certain type of non-targetlike negation pattern in the post-tests, even though they have previously produced MO with a targetlike pattern on the very same pattern, but they may not subsequently repeat another type of pattern which they have modified with a targetlike pattern.

The answer to Research Question 3, which addressed the relationship between the non-targetlike negation patterns that the learners modify and their subsequent use, is affirmative, but further study is required to confirm these findings.

5.6   Summary

This chapter described the data collected in the treatment sessions and three tests. The next chapter reviews the main findings and discusses them in light of the relationship between the production of MO and L2 learning.

CHAPTER SIX

DISCUSSION AND CONCLUSION

6.0   Introduction

This chapter reviews the major findings of the study (6.1) and discusses the relationship between modified output in response to clarification requests (MO) and L2 learning (6.2).   It concludes with the limitations (6.3), implications of the findings for theoretical perspectives (6.4) and directions for future research (6.5).

6.1.   Review of major findings

The purpose of the study was to investigate whether or not the production of MO facilitates L2 learning, specifically the learning of Japanese as a foreign language, within the framework of the output hypothesis.

The first research question addressed the relationship between the production of MO and grammatical accuracy.   The accuracy scores of the experimental group improved slightly more than the control group, but the difference between the two groups was not statistically significant.   Also, no relationship was confirmed between the production of MO and gains in accuracy scores within the experimental group.

The second research question concerned the impact of MO on interlanguage development, which was measured by the changes in the number of types of non-targetlike negation pattern, occurrences of types of pattern and developmental stages. No significant difference was observed in the number of types and developmental

stages between the groups, but some noticeable differences were observed between them in the occurrence of each type of non-targetlike pattern (i.e., decrease in *janai* and increase in *kunai* in the experimental group in the post-test). Additional analyses of the relationship between the production of MO and individual profiles of interlanguage did not confirm that the difference was associated with production of MO, but a relationship between MO and interlanguage was suggested.

As regards the third research question, this study went beyond the comparison of group performance between pre-test and post-tests, and explored the subsequent use of the types of non-targetlike pattern that individual participants previously modified. A trend was suggested that production of MO was likely to sensitise participants to avoid the use of the same type of non-targetlike pattern that they previously modified in the treatment sessions, but it was not conclusive due to the small size of the sample and a similar trend observed with Repetition (i.e., repetition of the initial non-targetlike pattern). However, it was found that a specific type of non-targetlike negation pattern (i.e., *janai*), which was previously modified by a targetlike pattern, continued to be used in the subsequent situations of use, and this is considered to demonstrate a relationship between the production of MO and L2 learning. The summary of findings is presented in Table 6.1.

Table 6.1
Summary of the findings

Research Question 1	
Experimental group (EG) outperformed Control group (CG) in accuracy score	not supported
Relationship between MO gains and accuracy within EG	not supported

Research Question 2	
Fewer variation of types of non-targetlike pattern in the two groups	supported
Greater decrease in the variation of non-targetlike pattern in EG	not supported
Changes in the occurrence of each type of non-targetlike patterns	partially supported
EG outperformed CG in terms of the number of participants who produced higher stage forms in the developmental sequence	not supported

Research Question 3	
Sensitisation to avoid the same type of pattern previously modified	not supported
Relationship between the types of pattern modified and their subsequent use	supported

In sum, the present study did not demonstrate conclusively that participants who were given opportunities for the production of MO improved in terms of grammatical accuracy. However, participants in the experimental group progressed their interlanguage towards targetlike use more than those who were not given opportunities to produce MO, which may indicate the impact of MO. The relationship between the production of MO and L2 learning was not confirmed in terms of sensitising avoidance of the same non-targetlike pattern which has been

modified, but a relationship between the production of MO and subsequent use was shown by one specific type of non-targetlike pattern which is less likely to be amenable to change. These findings provide some insights into the relationship between modified output and L2 learning, which are discussed in the following sections.

6.2 Discussion of the relationship between MO and L2 learning

This section discusses the relationship between MO and L2 learning with reference to grammatical accuracy and interlanguage (6.2.1), quantity and quality of MO (6.2.2), factors contributing to production of modified output (6.2.3) and subsequent use (6.2.4).

6.2.1 Impact of MO on grammatical accuracy and interlanguage development

Grammatical accuracy and interlanguage development are two different measurements (section 2.2.2), and parallel improvement in both measurements across the tests was not necessarily expected. However, as noted in section 2.2.2, only a few interactional studies (e.g., Doughty & Varela, 1998; Loewen & Nabei, 2007; Spada & Lightbown, 1993) have used both grammatical accuracy and interlanguage development as measurements relating to one target feature and the relationship between them has not been clearly established.

The analyses of grammatical accuracy (Research Question 1) and interlanguage development (Research Question 2) between the experimental and control groups showed some contrasting results. The present study demonstrated evidence of non-

parallel improvement between the two measurements, as participants in the control group used more targetlike patterns than non-targetlike patterns in the post-tests compared to the pre-test, but the non-targetlike patterns they used were not substantially different between the pre-test and post-tests. In other words, the control group showed an improvement in grammatical accuracy with few substantial changes in interlanguage across the tests. On the other hand, the experimental group showed some changes in interlanguage measured by the use of non-targetlike negation patterns across the tests, which may suggest the impact of production of MO on interlanguage development. It should be remembered that both groups similarly improved accuracy scores in both of the post-tests, and the difference between the groups was not significant.

The question is why the impact of MO was not observed in grammatical accuracy, and why it was observed in interlanguage development. One explanation is that the production of MO may be more facilitative of interlanguage than of grammatical accuracy over a relatively short-term developmental time frame. Another explanation is that MO is facilitative of both grammatical accuracy and interlanguage, but that the impact of MO is hard to demonstrate merely in terms of accuracy. For example, Bygate (2001) found accuracy less open to the influence of such external interventions as task repetition, and a conservative measure (errors per *t*-unit) was suggested as one of the possible reasons. The contrasting results between grammatical accuracy and interlanguage may be due to the possibility that the picture description speaking task was insufficiently sensitive as a data elicitation instrument, or its procedure of scoring needed more fine tuning.

It is difficult to determine with certainty which explanation is more plausible because of the limitations of the present study, but the findings in the current study at least suggest that the impact of production of modified output on L2 learning, if any, may be observed more in interlanguage development than grammatical accuracy. This may be supported by the results of previous studies, as only one study (Nobuyoshi & Ellis, 1993) has clearly demonstrated a positive relationship between modified output and grammatical accuracy while other studies have not done so (e.g., O'Relly et al., 2001; Takashima & Ellis, 1999). Further research is necessary to explore the relationship between grammatical accuracy and interlanguage, and their relationship with the production of MO.

6.2.2   Quantity and quality of MO

The experimental group received only a small number of clarification requests (Mdn = 3.5, range = 0-16), and produced a smaller number of instances of modified output (Mdn = 1.5, range = 0-8). Despite the limited nature of the treatment, which could be insufficient to accentuate the difference between the experimental group and the control group, some differences were observed in interlanguage development between the two groups. These findings support the notion that modified output could be associated with L2 learning even though its frequency is low (Ellis, 1999; McDonough, 2001; Shehadeh, 2002), as opposed to such claims that modified output is "too scarce to make a real contribution to linguistic competence" (Krashen, 1998, p. 180).

It is possible, however, that the changes in the interlanguage in the experimental group were more associated with the combined impact of clarification requests and production of modified output rather than modified output itself, as they were measured not separately but in combination. Also, a greater frequency of modified output than that in the present study may demonstrate a stronger impact on learners' performance including grammatical accuracy: however, this remains inconclusive and requires further research which can manipulate the frequency of production of modified output in finer detail.

6.2.3   Factors contributing to production of modified output

Swain (2005) stated that "the output hypothesis is about what learners did when 'pushed,' what processes they engaged in" (p. 473), emphasising the importance of the process of the production of modified output in L2 learning. In relation to this, the findings discussed in section 6.2.2 raise a question: What factors observed in the process of production of MO can be associated with L2 learning, especially the changes of interlangauge observed in the experimental group, if a low frequency of MO is facilitative of interlanguage development? The factors contributing to production of modified output include (1) obligatory occasions for the use of negation of adjectives, (2) production of initial non-targetlike negation pattern, (3) receiving clarification requests, (4) noticing of clarification requests, (5) interpretation of clarification requests (6) noticing the gap between own utterance and targetlike use and (7) modified output. Each of the seven factors will be discussed in the order in which they would occur in the process of production of modified output. It should be noted that factors (4), (5) and (6) which involve participants' internal cognitive

processes are not within the focus of the present study and no data was collected. Nevertheless, they are included in the following discussion to gain a wider picture of the relationship between modified output and L2 learning.

(1) Obligatory occasions and (2) production of initial non-targetlike negation pattern

The number of obligatory occasions and production of initial non-targetlike negation pattern between the experimental and control groups was not significantly different, and they were unlikely to contribute to the difference in the changes in interlanguage development. (The findings related to sensitisation to the avoidance of the non-targetlike form suggested that the types of the initial non-targetlike pattern which participants initially produced and then modified in response to clarification requests may interact with the impact of modified output in L2 learning, and this will be discussed in section 6.2.4.)

(3) Receiving clarification requests

The impact of clarification requests on L2 learning has been inconclusive in the previous studies (positive results in McDonough, 2005; McDonough & Mackey, 2006, and less positive results in Nobuyoshi & Ellis, 1993; O'Relly et al., 2001; Takashima & Ellis, 1999). The present study used only a few kinds of formulaic expressions for requesting clarification. Their common feature was that they did not provide linguistic information that participants could incorporate for modification of the initial non-targetlike form, and the difference in terms of the quality of clarification requests can be considered to be minimal. There was no significant difference in the gain in accuracy scores between those who received group median

or more clarification requests and those who received fewer clarification requests than the group median. Thus, the quantity of clarification requests that participants received does not seem to impact upon the participants' performance in the current study. This leaves the possibility that the factors involved in the production of MO which might be associated with the changes in interlanguage are not the quality or quantity of clarification requests, but participants' internal cognitive processes, in other words, whether or not participants noticed them, and how they interpreted and responded to them.

(4) Learners' noticing of clarification requests

Noticing has been claimed to be important in L2 learning (Ellis, 1994; Long, 1996; Schmidt, 1995; Schmidt & Frota, 1986; Skehan, 1998), but there are different levels of noticing and its definition and operationalisation varies among researchers since noticing (as well as other terms in SLA) has been used "in a general way in claims about the utility of interaction" (Mackey, 2007b, p. 25). For example, Schmidt (1994) defined noticing (of the input) as "registration of the occurrence of a stimulus event in conscious awareness and subsequent storage in long term memory" (p. 179). However, according to Robinson (1995), Schmidt's noticing is the closest to *detection*[20] in Tomlin and Villa (1994), which is "the cognitive registration of sensory stimuli" (p. 192). Robinson (1995) proposed a definition to reconcile the different views on whether or not conscious noticing is necessary for L2 learning: "detection plus rehearsal in short-term memory prior to encoding in long-term

---

[20] *Detection* is one of the three components of attention in Tomlin and Villa (1994), along with *alertness* (a overall, general readiness to deal with incoming stimuli or data) and *orientation* (committing attentional resources to sensory stimuli).

memory" (p. 296). Philp (2003), who investigated the variables that constrain noticing of recasts, shares the same view of noticing as "step beyond detection" (p. 102) with Robinson (1995).

Noticing of clarification requests in the context of the present study, however, is much more limited than noticing as noted above in that it does not involve focal awareness on the specific aspect of input which could be incorporated into the participants' further output or comparing their own utterance with targetlike use, but involves mere recognition of clarification requests as a prompt that indicates the existence of communication breakdown.

Investigation of noticing in general faces methodological problems as it cannot be directly accessed, and researchers have employed on-line measures (e.g., think-aloud protocols, Leow, 1997) and off-line measures such as diary entries (Schmidt & Frota, 1986), questionnaire response (Mackey, 2006; Robinson, 1997), verbal reports such as cued immediate recall (Egi, 2007; Philp, 2003) and stimulated recalls (Mackey, 2006; Mackey et al., 2000). The difficulty in accessing learners' internal processing is also applied to noticing of corrective feedback, but the production of modified output (as well as the repetition of the initial non-targetlike utterance) is assumed to indicate a manifestation of learners' noticing of clarification in the process of production of modified output. Noticing of a clarification request does not help participants to obtain more linguistic input, but initiates an opportunity to reflect on their own utterances and the process of production of modified output, which might not occur otherwise. Therefore, this is considered to be an important process that

might be associated with the changes in interlanguage of participants in the experimental group.

(5) Interpretation of clarification requests

The overlap and mismatch between the intent of the provider of the feedback and the learners' interpretation of feedback has been investigated and described (Carpenter et al., 2006; Egi, 2007; Gass & Mackey, 2006; Han & Kim, 2007; Mackey et al., 2000; Philp, 2003; Roberts, 1995; Takashima, 1995), and the present study suggested a possible discrepancy between the interpretation of clarification requests and the types of non-targetlike pattern. Although this cannot be confirmed without a measure of participants' perception (e.g., stimulated recall), it was noticeable that clarification requests to *janai* seem to show a larger mismatch between the intent of the interlocutor and participants' interpretation signified by response move than that to other types of non-targetlike pattern. These requests were responded to with Repetition (62.5%) more than with MO (25.0%), while clarification requests to other types of non-targetlike pattern were responded to with more MO than with Repetition. The questions needing to be asked are: What differentiated participants' interpretation of clarification requests, especially in terms of leading response to clarification requests to *janai* to Repetition rather than MO? and How is the outcome of the interpretation associated with L2 learning?

A number of factors could constrain participants' interpretation of clarification requests, and the following three kinds of variables[21] need to be considered in the context of the present study: feedback-related variables, language-related variables and learner-related variables. The first of these can be dealt with briefly, but the other two require more detailed explanations. Firstly, feedback-related variables (e.g., types or length of feedback, number of changes in feedback) do not appear to have played an important role in terms of influencing participants' interpretation of clarification requests in the current study, because only one type of feedback (i.e., clarification requests) was used, and as such, the clarification requests did not provide further linguistic information to participants.

The second set of variables are language-related, and they include the nature of the linguistic feature to which feedback is given (e.g., morphosyntax, lexis, semantics or phonology), and the degree of difference between the non-targetlike form and targetlike form. The current study focused on one specific linguistic feature, negation of adjectives in Japanese, and, therefore, there was no difference in terms of the nature of the targeted linguistic feature. Within the same targeted linguistic feature, however, the interpretation of clarification requests may be influenced by the types of non-targetlike pattern that were targeted by clarification requests. The non-targetlike negation pattern, *janai,* may have some linguistic features which prevent participants from noticing the motivation of the requests for grammatical

---

[21] Another variable is related to the provider (e.g., teacher) of clarification requests. For example Han (2001) showed the relationship between how the provider of feedback understands the source of communicative breakdown and learner's interpretation of feedback. However, it is not considered here, as the interlocutor in the present study focused on one specific target feature.

167

reformulation.   Previous research (Kamura, 2001b; Takeuchi-Furuya, 1993) has

found that *janai* is a non-targetlike pattern that does not disappear until a later stage

of acquisition of negation of adjectives (section 3.3), and the present study showed

that *janai* was the most frequently used non-targetlike pattern and thus received most

clarification requests in the treatment sessions (section 5.5.1).   Also, *janai* seems to

be resistant to change by the impact of MO (section 5.5.4).

It seems reasonable to assume that the more perceptually salient the non-targetlike

form a learner uses is, the easier it becomes for the learner to notice the intention of

the clarification request, but *janai* is not necessarily more salient than other types of

negation patterns.   Beginning L2 learners as well as L1 Japanese children are likely

to pay more attention to the ends of intonation units and *ja* in *janai* is not recognized

at an early stage (Clancy, 1985), and *janai* seems perceptually less salient than *nai*.

Also, perceptual salience may not account for *janai* inviting more tokens of repetition

of the initial non-targetlike utterance than modification, because 66.7% of

clarification requests to *kunai*, another less salient pattern[22], led to MO, in contrast to

25.0% of *janai*.   It should be noted that the frequency of *kunai* in the input, which is

used only for negating adjectives, is lower than that of *janai* which is used for

negating nouns and nominal adjectives (section 4.8.1).   Thus, this study does not

provide evidence to confirm the association between language-related variables

including the nature of the targeted linguistic feature and interpretation of

clarification requests.

---

[22] *Ku* in *kunai* could be filtered out between a recognizable adjective root and a negator *nai* (Clancy, 1985).

Learners' interpretation of other types of implicit feedback, such as recast, can be constrained by feedback length, and the shorter the feedback is, the more accurately learners perceive it (Philp, 2003). Although the length of clarification requests is unlikely to influence learners' interpretation, the shorter the learners' utterance which holds the non-targetlike negation pattern is, the easier it would be for learners to spot the non-targetlike use of the utterance when a clarification request is given. Conversely, the longer the learner's utterance is, the more difficult it becomes for the learner to match his/her perception of clarification requests with the intent of the provider of clarification requests. This is demonstrated by Excerpt 4, which shows that a participant modified two parts of the initial utterance (a non-targetlike negation pattern *dewa arimasen* and a targetlike form *to omou* meaning 'I think') in response to clarification requests, probably because he was unsure of whether a clarification request was given to *dewa arimasen* or *to omou,* although he noticed the intention of clarification requests which resulted from the ill-formedness of his own utterance.

Excerpt 4

Interlocutor:	*Nihongo no guramaa wa muzukashii desu ka.*	
	Is Japanese grammar difficult?	
Xian:	*Iie, muzukashii … *dewa arimasen to omou.*	
	No, I think it is not difficult.	
Interlocutor:	*N.*	← **Clarification request**
	Pardon?	
Xian:	*Muzukashii *janai desu.* ← **MO with non-targetlike pattern**	
	It is not difficult.	

However, the amount and complexity of the language that participants needed to retrieve from their initial utterance in order to respond to a clarification request was very limited in the current study, restricted in fact to short phrases or simple sentences. Therefore, there are not many such instances in the present study, and the influence of language-related variables in general seems to be relatively limited.

The third set of variables that may be associated with the interpretation of clarification requests are learner-related (e.g., proficiency levels, developmental readiness, motivation, aptitude, L1 influence, and cognitive factors such as working memory capacity). Among them, research has suggested the relationship between interpretation of feedback and developmental readiness (Gass & Mackey, 2006; Han, 2002; Ishida, 2004; Mackey, 1999, 2006; Mackey et al., 2000; Mackey & Philp, 1998; Nassaji & Swain, 2000; Philp, 2003; Trofimovich, Ammar, & Gatbonton, 2007). For example, Mackey and Philp (1998) found feedback (i.e., recasts in their study) did not lead learners to acquire forms that they were not developmentally ready to acquire. Philp (2003) suggests that feedback (i.e., recasts) might be effectively used when it corresponds with learners' developmental level, and that "the learner is biased by her own IL [interlanguage] grammar and potential immediate developments beyond it" (p. 115).

Developmental readiness indicates a psycholinguistic state when a learner has access to the linguistic and procedural skills required to produce a specific language structure (Mansouri & Duffy, 2005, p. 84). This is associated with the Processability Theory (Pienemann, 1998), which claims a hierarchy of processing skills in which

acquisition of each lower level procedure is a prerequisite for the functioning of higher level. Thus, the relationship between developmental readiness and feedback can be explained by the subset of the Processability Theory, the *teachability hypothesis* (Pienemann, 1984, 1998; Pienemann & Johnston, 1987; Pienemann et al., 1988). It claims that "stages of acquisition cannot be skipped through formal instruction, and that instruction will be beneficial if it focuses on structures from 'the next stage'" (Pienemann, 1998, p. 13).

The relationship between developmental readiness and learners' interpretation of clarification requests, however, has not been investigated in the context above, probably because clarification requests are a prompt that does not contain linguistic information which learners are directed to notice and incorporate into their interlanguage. It may be that, like the relationship between recast and developmental readiness, whether or not learners interpret clarification requests as a cue for grammatical reformulation is associated with their developmental readiness.

In order to demonstrate this relationship, it is necessary to compare the developmental readiness of participants who responded to clarification requests to *janai* with MO with that of participants who responded to it with Repetition. Participants often responded to clarification requests to *janai* with multiple types of response move. There were only four participants who responded to clarification requests to *janai* with only Repetition, and one of the four produced Stage 1 forms only, and the other three did not produce Stage 1 forms but Stage 2 forms. No participants responded to clarification requests to *janai* only with MO, and comparison was not possible with

this approach. However, *janai* is the negation pattern for nouns and nominal adjectives, and the use of *janai* for negating adjectives is considered to indicate that the learner fails to distinguish adjectives from nouns and nominal adjectives. This seems to suggest that the developmental readiness of those learners is lower than those who chose the correct negation pattern, *kunai,* but did not inflect the adjective stem correctly. However, this assumption needs further empirical validation based on the investigation on learners' explicit and implicit knowledge on the different types of predicate.

It should be noted that developmental readiness is not the sole variable that could influence learners' interpretation of clarification requests. For example, proficiency levels of learners have been reported to be associated with the noticing of recasts (Philp, 2003), and the relationship between noticing of recast and such cognitive factors as phonological memory, attention control and analytical ability has been suggested (Trofimovich et al., 2007). Therefore, it is possible that noticing of clarification requests may be associated with those, but it is beyond the scope of the current study and future research needs to investigate it.

(6) Noticing the gap between own utterance and targetlike use

Production of modified output in the present study is considered to be an indication that participants noticed[23] that there was a gap between appropriate targetlike use and

---

[23] There are at least two different levels of noticing: *noticing a gap* (Schmidt & Frota, 1986) and *noticing a hole* (Doughty, 2001; Swain, 2000). A gap indicates the difference between the target language and learners' interlanguage (Muranoi, 2007), and this can be called *noticing mismatches* (Doughty, 2001, p. 225). A hole refers to the difference between what learners want to say and what

what they said. In order to notice the gap, learners are required to have "sufficient and coordinated working and long-term memory resources to enable the cognitive comparison" (Doughty, 2001, p. 225). Although noticing the gap does not necessarily predict that the gap is filled with more targetlike use (de Bot, 1996), this cognitive comparison is considered to contribute to L2 learning because it may involve "not simply awareness of and noticing of an L2 stimulus but recall of a prior, related stimulus to which it is compared" (DeKeyser, 2007, p. 309), restructuring of interlanguage knowledge and the formation of new form-meaning connections (Williams, 2005). This noticing the gap, which Swain (1995) claims may involve the cognitive processes that consolidate learners' existing knowledge, seems to be associated with the changes in interlanguage development in the present study.

(7) Modified output

Modified output is the product of a number of processes after the provision of obligatory occasions. There are two types of modified output, targetlike or non-targetlike, and it has been suggested that the process of production of modified output may be more important than whether the modified output is targetlike or non-targetlike (Gass, 1997; Mackey, 2007b; Swain, 2005). The results did not clearly show whether or not the impact of MO on L2 learning could be different between targetlike modified output and non-targetlike modified output because of the small sample size.

---

they can say (Swain, 1995) and noticing a hole indicates that learners noticed that "they do not know how to express precisely the meaning they wish to convey *at the very moment of attempting to produce it* (Swain, 2000, p. 100, italics in original). Swain (2005) suggested noticing a hole may be a step forward to noticing the gap.

There are other contributing factors, and one of them is whether or not learners have opportunities for responding to clarification requests. Treatment in the present study was designed to provide participants in the experimental group with those opportunities, for the purpose of minimising it as a contribution factor to production of modified output. Another important factor is participants' prior knowledge. As noted in section 5.1.3, participants had been taught the target feature in class and were assumed to have explicit knowledge. However, a picture-description task employed in the present study is considered to be a better measure of implicit L2 knowledge rather than of explicit knowledge, which leaves the possibility of prior explicit knowledge as a contributing factor.

In sum, among the factors contributing to production of MO in the present study, the findings suggest the internal cognitive processes after receiving clarification requests and before the production of MO might be associated with the changes of interlanguage in the experimental group more than the quality and quantity of clarification requests given to participants. The claim of the output hypothesis is that "the act of producing language (speaking or writing) constitutes, under certain circumstances, part of the process of second language learning" (Swain, 2005, p. 471). In the light of this outcome, the present study suggests the importance and necessity of further investigation into the relationship between modified output, internal cognitive processes and L2 learning, and provides some empirical support for the theoretical significance the research in this area.

6.2.4 MO, subsequent use, and L2 learning

6.2.4.1 Sensitisation and non-sensitisation

A tendency that emerged from the study was that the production of MO appeared to sensitise learners to avoid the same non-targetlike pattern that had been previously modified. However, in relation to this trend, the findings were inconclusive. The main reason for this, in addition to the small sample size, was that the analysis of subsequent use of the types of non-targetlike pattern which participants responded to with repetition of the initial non-targetlike form (Repetition) also showed a similar trend as that observed with MO.

Repetition is considered to require less mental activity than production of MO, and one of the possible reasons is related to a learner's incorrect interpretation of clarification requests and/or an incorrect judgement about the well-formedness of his/her own utterance. In other words, it could represent a mismatch between the intent of the interlocutor and the perception of the learner. Repetition may lead to stabilisation rather than to non-use of the same pattern previously repeated, and there are some individual instances which show that the impact of Repetition did not sensitise participants to the non-use of the non-targetlike pattern they previously modified. For example, one participant, who produced non-targetlike patterns involving *janai*, responded to clarification requests with eight Repetitions. She still used another five *janai* in the post-test and another eight in the delayed post-test. Such evidence suggests that Repetition did not sensitise her to avoidance of the same non-targetlike pattern in subsequent situations of use. Nevertheless, the overall results did not strongly indicate that participants who responded with Repetition of

the non-targetlike pattern then went on to use the same non-targetlike pattern in subsequent use. There was a trend in the post-tests that fewer non-targetlike patterns that were responded to with Repetition were used.

Repetition could involve retrieval and conscious reflection of the initial non-targetlike utterance, and the failure to perceive the intent of feedback does not necessarily deny the benefit of feedback (Egi, 2007; Mackey, 2007a). The reprocessed output, represented in the form of Repetition, may still "represent the leading edge of a learner's interlanguage" (Swain, 1995, p. 131). Thus, it is possible that Repetition as a response to clarification requests may sensitise learners to avoidance of the repeated use of the non-targetlike utterance in subsequent situations of use. Also, Repetition may turn learners' attention to the targeted form in the input in the further interaction, which may contribute to L2 learning as well.

The present study suggests that learners are no more likely to be sensitised to avoid the use of the non-targetlike pattern if they produce MO or Repetition. In this respect, it seems that the output hypothesis may not be supported, but this conclusion should be interpreted as merely indicative. Firstly, Swain (1995) clearly stated that the three functions of output which may trigger L2 learning do not necessarily operate in any production of output, and it is possible that some conditions in the present study limited the operation of those functions. Secondly, MO and Repetition could not be compared with subsequent use of the types of non-targetlike pattern where clarification requests were given but received no response, because all the clarification requests in the present study were responded to with one of the three

types of response moves (i.e., MO, Repetition, and Others), which is one of the limitations of the current study. Also, statistical tests would be necessary to confirm these findings in any future research, which would require a larger sample data.

6.2.4.2 Types of modified non-targetlike pattern and subsequent use

Some researchers have claimed that immediate uptake in response to feedback does not necessarily indicate L2 learning (Gass, 2003; Loewen, 2007; Lyster, 1998; Mackey & Philp, 1998; Nassaji, 2007; Nicholas, Lightbown, & Spada, 2001). The findings of the present study suggest that even with immediate MO with targetlike use learners may still continue to use the same non-targetlike pattern in the subsequent use. When clarification requests were given to a non-targetlike pattern, some participants seem to have noticed that they had chosen an incorrect negation pattern, and replaced it with *kunai* after dropping *i* (non-past tense morpheme) in the stem of adjective. However, in subsequent use in the post-tests, some of them reverted to the use of the same non-targetlike pattern. The benefit of production of modified output in L2 learning has been argued to lie in promoting automaticity, hypothesis testing, noticing and grammatical encoding and monitoring, and the results suggests that hypothesis testing and noticing triggered by production of MO might be constrained by some factors, possibly including learner's internal cognitive processes.

This pattern of linguistic behaviour was revealed specifically with *janai* while it occurred on fewer occasions in other types of non-targetlike negation pattern. It appears to indicate that the processes involved in the production of modified output

177

on *janai* had little noticeable impact on learners' interlanguage development in relation to this group of negative morphemes. However, it is important to note that the results of the current study suggest this may apply only to MO on certain variants of the target form (i.e., *janai*), and that MO on other non-targetlike pattern may facilitate L2 learning. It may be that the impact of production of MO on the subsequent situations of use is influenced by the very types of non-targetlike form which learners initially modify.

The question is why *janai* appears to be resistant to change. Also, why is it the only non-targetlike negation pattern that showed a clear trend of consistent use even after the production of MO with a targetlike pattern? What in *janai* is different from other types of non-targetlike pattern that may be more amendable to change?

One possible approach to answer these questions might be suggested by the types of response move by the three participants who continued to use *janai* in the post-tests even after the production of MO with a targetlike pattern in the treatment sessions (section 5.5.4). These participants commonly used multiple types of response move for clarification requests to *janai*, which seems to reflect their uncertainty about their use of *janai* and interpretation of the motivation for clarification requests. Tarone (1983) suggested *variation*, "shift *within* the performance of any given individual" (p. 73, italics in original), is a manifestation of learners' underlying capability, and it seems to be applied to the variation of individual learner's response to clarification requests to the same non-targetlike pattern. Learners may produce MO with targetlike pattern as a result of hypothesis testing, but they cannot be confident in the

use of the modified form in the subsequent situations of use without further confirmation that the modification was more targetlike than the initial use. Swain and Lapkin (1995) have suggested the importance of the provision of feedback which could improve learners' incorrect hypotheses and inappropriate generalizations (that appear in the process of production of output). This may apply not only to the initial non-targetlike use but to its modified version, produced as a result of noticing the gap.

Evidence of such forms as *janai* indicates a limitation of the impact of modified output as an isolated interactional component on L2 learning, but suggests the conditions under which the three functions of output may not work may include the type of non-targetlike form which learners modify. As discussed earlier, whether learners modify a non-targetlike form is influenced by a variety of factors including noticing and interpretation of feedback and noticing the gap, which is associated with the developmental readiness of learners. Also, persistent use of *janai* after the production of modified output addresses the importance of interaction which does not end with the production of modified output, but involves follow-up feedback to the modified output. In other words, production of modified output might be more beneficial in L2 learning if learners are provided with further feedback which could confirm/disconfirm their hypothesis that the original utterance needs to be modified. The updated version of interaction hypothesis (Long, 1996) emphasises the role of negotiation for meaning that connects "input, internal learner capacities, particularly selective attention, and output in productive ways" (p. 452), and the findings in the current study may indicate some support to the interaction hypothesis.

Replication of this type of research with other grammatical forms and other linguistic features (e.g., syntax, lexis or semantics) is necessary to investigate whether or not such MO-resistant variants as *janai* in negation of adjectives are found in other linguistic features. If this is the case, what needs to be explored is commonalities among those learners in terms of developmental readiness, common features among those MO-resistant forms, and how MO-resistance can be accounted for. Also, it would be meaningful to investigate learners' perceptions when they produce modified output with targetlike use and non-targetlike use, and how the difference in perceptions impacts L2 learning.

6.3   Limitations of the study

In addition to the small sample size (n = 28), there are a number of limitations in the present study, which are related to research method (6.3.1) and participants (6.3.2).

6.3.1   Limitations related to research method

One of the limitations of the study relating to research method is the use of a single testing instrument (i.e., computer-administrated picture description task) rather than multiple testing instruments. This was largely due to the limited grammatical knowledge and vocabulary of the participants, which constrained the variation of tasks they could perform. The findings would be more persuasive if multiple testing instruments were used.

Implicit methodological techniques were applied to the interview and mechanical drills to elicit modified output on the pre-selected target feature, and participants in

the experimental group were given a number of clarification requests when they used non-targetlike patterns. This was carried out as naturally as possible, but one of the participants received 18 clarification requests on the targeted feature in the treatment sessions. This is not out of line with the data of clarification requests reported in previous studies (7-21 in Nobuyoshi & Ellis, 1993; 0-13 in Takashima & Ellis, 1999), but receiving so many clarification requests appeared to become rather artificial and unnatural, as pointed out by Ellis (1999). Thus, some further development of the MO elicitation instrument is desirable, and it would be necessary to measure the participants' perceptions about clarification requests.

Another limitation is the length of data collection and timing of the post-tests. The present study was conducted over the time-span of two months, a time period that is somewhat shorter than descriptive studies that investigated the developmental stages of negation in Japanese (e.g., six months in Kamura, 2001b; seven months in Kanagy, 2001). A longer period of data collection is likely to describe changes in learner interlanguage better. However, since this study was experimental in design, the first post-test needed to be conducted within a relatively short period of time in order to measure the early impact of the intervention and to eliminate confounding variables. The first post-test, therefore, was held one week after the second treatment session, following the design employed in previous experimental studies (Nobuyoshi & Ellis, 1993; O'Relly et al., 2001; Takashima & Ellis, 1999). However, the results may be attributable not to the intervention but to uncontrolled factors such as further instruction in class. As a methodological variation for future research, it would be of interest to conduct a post-test straight after the treatment sessions, and to conduct

more delayed-post tests within the same length of time or in an extended period of time.

Both meaning-focused and form-focused activities were used in the treatment sessions in order to elicit as many tokens of modified output as possible. For the purposes of the present study, the modified output in each case was analysed as if it were of the same kind, based on the assumption that the process leading to the production of MO is unitary. However, there might be some differences in the impact on L2 learning between those two kinds of MO, especially in terms of the availability of attentional resources. Future research may need to investigate the impact of each type of MO separately.

The present study narrowly focused on the changes in the negation of predicate adjectives in the non-past tense. The development of negation of adjectives in Japanese seems to be closely associated with development of negation of verbs, nominal adjectives and nouns. This would apply to the other tense (i.e., the past tense). Investigations including other categories and the other tense would be worthy of pursuing as a future agenda. Also, the impact of MO might be more observable in other linguistic features such as syntax or phonology than the inflectional morphology that was the target of the present study. One of the differences between morphological inflection and other linguistic features is that transfer from L1 to L2 is not as strong at such a level as syntax because of difficulty in constructing the relation between L1 and L2 (MacWhinney, 2005, p. 55). This may be associated with the findings in research on modified output, and targeting

multiple linguistic features involving these would be interesting. Muranoi (2007) pointed out that the majority of target features used to test the output hypothesis have been more or less simple. This trend is probably associated with the possible difficulty in eliciting complex linguistic features, but the relationship between modified output on complex structures and L2 learning needs to be investigated.

6.3.2   Limitations related to participants

A large range in accuracy scores in the pre-test and general speaking skills among participants might be another contributing factor to the results. Although participants who scored 90% or higher in the pre-test were not included in the data set, both the experimental group and the control group had a wide range of accuracy scores between 0% and near 90%. This makes it difficult for statistical tests to detect the differences between the groups.

The relationship between the participants' first language and production of MO was not investigated. A possible impact of the first language morphosyntactic knowledge on L2 morphosyntactic acquisition has been suggested (e.g., Bley-Vroman, 1989; MacWhinney, 1995), and future studies might attend to such linguistic issues. Also, factors related to differences in learning styles and types of learners were not investigated in the present study. Production of modified output as a result of being pushed involves risk-taking and could cause "embarrassment in the eyes of others" (Batstone, 2002, p. 4), and the way learners respond to feedback might be different not only in a different task but in a different culture (e.g., some learners from a certain culture might consciously avoid the opportunity for modifying the non-targetlike

utterance so as not to lose face). Nobuyoshi and Ellis (1993) suggested that a participant in their study who produced only two tokens of modified output despite being given 21 clarification requests might be a functionally-oriented learner as opposed to a structurally-oriented learner. Although it would not be easy to distinguish one type from the other, this is an important area that needs to be investigated further.

Participants in the present study were recruited from a single institution, and, therefore, the trend in the use of non-targetlike negation patterns observed in the study might reflect the teaching method, teaching material, and the order of introduction of the negation patterns (see Kawaguchi, 2000). This obviously limits the generalisability of the findings from this study, as with previous studies (e.g., McDonough, 2001). In order to address this limitation, it would be necessary to collect data in multiple institutions.

6.4   Implications of the findings of the study

Despite the limitations discussed above, there are a number of implications that arise from this study in relation to interlanguage development. These include: accuracy and interlanguage development, learners' interpretation of clarification requests, types of modified forms, the elicitation of modified output, follow-up feedback to modified output, and the acquisition of negation of adjectives in Japanese.

6.4.1 Grammatical accuracy and interlanguage development

One of the implications is that studies which seek to measure only grammatical accuracy to demonstrate evidence of impact of modified output may be worth revisiting because the impact of production of modified output may not be observed in accuracy but in progressive changes in learners' interlanguage. In this study, this was shown by the experimental group, which demonstrated substantial changes in the use of non-targetlike form while the difference in the improvement in grammatical accuracy scores was not significant between the experimental and the control group.

6.4.2 Factors relating to learners' interpretation of clarification requests

Another implication that arises from the study is the necessity of research that explores what differentiates the learners' interpretation of clarification requests either as requests to modify or as requests to repeat. The present study showed that, with certain types of non-targetlike pattern, learners are likely to interpret the intent of clarification requests as a request to repeat. Tentative findings here suggest that the relationship between clarification requests and learners' interpretations of them may relate to the particular developmental readiness of learners. However, the relationship between the interpretation of clarification requests and learner developmental readiness is an area that has not been examined to any extent, making it a possible area of exploration for future research. Observations relating to the types of learner response move (MO or Repetition) are limited solely to the negation of adjectives in Japanese. Therefore, any further investigation of this aspect of learner noticing would need to examine whether this applies to other linguistic features in Japanese and in other languages.

A recent study by Gass and Lewis (2007) investigated the difference in perception of interactional feedback between heritage learners, "individuals who are living in a second language environment but who were raised in a home where a language other than the second language is spoken" (p. 80), and non-heritage learners. It was found that both groups perceived morphosyntactic feedback less accurately than lexical and phonological feedback, and there was a distinctive difference in the perception of semantics between the two groups of learners. The present study was not concerned with this aspect, but it is possible that language background of learners may be associated with the perception of clarification requests. Further research is needed in this area, which could include investigation of how the difference in perception in the two groups impacts the production of modified output and L2 learning.

6.4.3. Types of non-targetlike form that learners modify and their subsequent use

This study proposed an approach to exploring the types of non-targetlike pattern that learners modify and the use and non-use of the same type of form in subsequent situations. This is considered to be a more direct way to explore the relationship between modified output and L2 learning than comparison in performance in the pre-test and post-test(s), but the results of the present study are not conclusive and only suggest a possible relationship between them. Thus, another implication is to explore whether such a relationship exists in areas other than negation of adjectives in Japanese, and in other languages, and then how developmental readiness of learners is associated with it.

186

Question formation in English might be a good target feature for this purpose, because it has been empirically investigated in interaction-development research and the developmental sequence has been established. For example, there might be a trend that learners who modify a non-targetlike use 'Why the woman waiting?' (Stage 3 in Pienemann and Johnston's (1986) sequence) by 'Why is the woman waiting?' might not use the same type of non-targetlike form in subsequent situations of use. On the other hand, the same learners who modify 'Where is he come from?' (Stage 5, cited from McDonough, 2005) by 'Where does he come from?' may produce the same type of non-targetlike form in subsequent situations of use, because the developmental stage is higher than their developmental readiness. Further investigation of the non-targetlike forms that learners notice and modify, and their impact on the learners' subsequent use potentially provides another insight into the role of modified output in L2 learning.

6.4.4   The elicitation of modified output

The trigger to elicit modified output was limited to general clarification requests, but future research needs to investigate the difference in the impact of modified output produced in response to general clarification requests and specific clarification requests (section 2.3). The latter is more explicit in terms of the location of the problematic source, and this might facilitate learners to pinpoint the non-targetlike use and to produce modified output, and its impact on L2 learning might differ from modified output in response to general clarification requests.

Also, the comparison between modified output in response to feedback and without feedback from external sources (i.e., self-initiated self-completed repair) is an important area for further research. For example, exploring the relationship between the types of pattern and self-initiated self-completed repair, and its comparison with that of modified output in response to feedback will lead to a better understanding of the mechanism in noticing, production of modified output and their impact on L2 learning. While insights may be obtained from data gathered in tightly controlled experimental settings, ethnographical research in L2 classrooms may also provide another perspective on this phenomenon.

6.4.5   MO, follow-up feedback and subsequent use

Lack of follow-up feedback to participants' modified output might be associated with the consistent use of a non-targetlike pattern in subsequent situations of use despite the previous production of MO. Production of MO is considered to be a reflection of learner's hypothesis testing (Swain, 1995), which involves formulating, testing, confirming, modifying, and rejecting hypothesis (Muranoi, 2007). The validity of hypothesis testing utterance cannot be established without feedback specifically directed to the modified output (Ellis, 2003), and learners will be left in the dark unless that feedback contains some clues that learners can pick up. The present study illustrated this point by showing that the participants frequently repeated their initial output without modification, or produced random and variable modification, because there was no follow-up move by the interlocutor which might have clarified the purpose of the initial feedback. It is beyond the scope of the present study to investigate whether or not further feedback to MO would impact upon learners'

188

interlanguage, because further feedback to modified output was not given to the participants in the current study. Therefore, this is another question that awaits an answer in future research. In order to confirm this, research needs to demonstrate that learners who produce modified output on a certain form and receive the follow-up feedback outperform learners who produce modified output but do not receive the follow-up feedback.

6.4.6   Acquisition of negation of adjectives in Japanese

Kamura (2001b) found the developmental sequence in her JSL study differs from that in L1 research, and suggested the necessity of research in learning JFL to confirm whether her findings in JSL were consistent with JFL. The present study added to the empirical base of research on acquisition of negation in the JFL context in which most of the learners do not have opportunities to use the target language outside the L2 classroom. The developmental stages found in the present study support part of the findings in Kamura which include the disappearance of *nai* at an early stage, persistent use of *janai* for a longer period, and a trend of gradual decrease in the variation of non-targetlike negation patterns. This indicates learners might go through a similar path in the acquisition of negation of adjectives whether they are in a JFL or a JSL environment, and the developmental sequence might be different from that in L1. This gives rises to a theoretical implication for more research on the acquisition of negation in Japanese in terms of the developmental sequence between L1 and L2.

6.5   Directions for future research

One of the directions for future research is to continue further exploration of "the specific developmental contribution" (Mackey & Oliver, 2002, p. 474) of modified output in L2 learning. The focus of interactional research has been turning from whether or not interaction impacts on L2 learning to 'how interaction facilitates L2 learning' (Mackey, 2007b; Mackey et al., 2000; Mackey & Oliver, 2002).   In contrast to the findings of the impact of input or negotiation for meaning in L2 learning, however, it is still unclear whether the production of modified output does facilitate L2 learning, let alone how it could facilitate L2 learning.

Difficulty in obtaining sufficient data of modified output relating to specific target linguistic features, and the methodological difficulty in isolating modified output from other interactional factors remain challenges both to overcome.   More importantly, production of modified output cannot be separated from other factors, such as interpretation of feedback and noticing the gap (although the opposite seems possible), and this would be obviously one of the limitations in research on modified output as an isolated interactional factors.   Referring to the limitation of studies of noticing in interaction in terms of researchers' direct inaccessibility to internal processing, Mackey (2006) stated "noticing may be more productively viewed along a continuum rather than as a fixed occurrence" (p. 424).   This may apply to research on modified output, especially because no modified output is produced without noticing.   In addition, focusing on an isolated interactional component has the potential to overlook what is happening in the learner's mind, and its real impact on L2 learning.   The following interaction between and the interlocutor and Kate

(Participant 5 in the experimental group) illustrates the importance of holistic analysis of discourse in interaction.

Excerpt 5

1	Interlocutor:	*Tesuto wa ooi desu ka.*	
		Do you have many tests?	
2	Kate:	*Amari *oi dewa arimasen desu.*	← **Non-targetlike use**
		Not many.	
3	Interlocutor:	*Moo ichido*	← **Clarification request**
		Once again?	
4	Kate:	*Amari *ooi dewa arimasen desu ne.*	← **Repetition of Line 2**
		Not many.	
5	Interlocutor:	*Soo desu ka. Nihongo no jisho o motte imasu ka.*	
		Is that right? Do you have a Japanese dictionary?	
6	Kate*:*	*Hai, arimasu.*	
		Yes, I do.	
7	Interlocutor:	*Sono jisho wa atarashii desu ka.*	
		Is the dictionary new?	
8	Kate:	*Atarashii dewa arimasen.*	← **Same type of non-targetlike negation pattern as in Lines 2 & 4**
		It is not new.	
9	Interlocutor:	*N.*	← **Clarification requests**
		Pardon?	

10	Kate:	*Furui no jisho desu ne.*	← **Non-use of negation**
		It's an old dictionary.	
11	Interlocutor:	*Aa, soo desu ka.*	
		Is that right?	

Kate used a non-targetlike negation pattern, *janai* (Line 2), and the interlocutor gave a clarification request (Line 3), to which she responded with a repetition of the initial utterance (Line 4). In Line 8, Kate used another *janai* on a different adjective, and she was given another clarification request (Line 9). This time, instead of repeating the initial non-targetlike pattern, Kate used an antonym, 'old,' and did not use negative form of 'new' (Line 10). The reason she used the antonym is not clear since no measures such as stimulated recalls were used, but the repeated clarification requests to *janai* appeared to sensitise her to avoid the repeated use of *janai*. This may have triggered Kate to pay more attention to further input involving negation (although it was made unavailable in the treatment sessions at least), and might contribute to L2 learning. Also, Kate might need appropriate feedback to her use of *janai* in order "to narrow the range of possible hypotheses that can account for the data" (Carroll & Swain, 1993, p. 358). Since no follow-up feedback was given in the present study, she might start to question her use of *janai* and try to find some clues in the following interaction (i.e., rasing awareness of language and attention to form, Long, 1996), getting out of the darkness without further direct assistance. Her learning might be facilitated as a result of receiving clarification requests, repeating the same type of non-targetlike use, and not receiving follow-up feedback.

Mackey (2007b) argued the importance of being open to the potential impact of interactional feedback on interlanguage change "even if learners do not perceive feedback as providing information about grammaticality, and even if they do not have a sufficient command of the L2 to modify their utterances, or choose not to modify them for some other reason" (p. 24). In order to be able to account for this potential influence of interactional feedback, it would seem that it is necessary to investigate the interaction surrounding the modified output as well as the modified output within the interaction.

Another direction, therefore, is to extend the investigation beyond modified output as an interactional component. It is obvious that production of modified output always involves other components, and it only exists in an interaction. The subsequent use and non-use of a linguistic feature that learners have previously modified might be constrained by the very linguistic feature involved. Also, the production of modified output is likely to be constrained by learners' interpretation of given feedback, which might be further constrained by the non-targetlike form they initially use. In classroom settings, learners are engaged in interaction, which consists of input, output, feedback, modified output, and modified input. Thus, each component is interwoven, and cutting out one component might not reflect what constitutes L2 learning (Batstone, 2007; Ellis & He, 1999; Ohta, 2001; van Lier, 1996).

Research within the framework of a socio-cultural perspective looks into L2 learning from holistic perspectives (e.g., Foster & Ohta, 2005; Nassaji & Swain, 2000; Ohta, 2000; Swain & Lapkin, 1998; Swain & Lapkin, 2007), in which modified output as

an interactional component is absorbed into collaborative dialogue, and output is seen as part of "a cognitive tool that mediates our thinking" (Swain, 2005) beyond production. Swain (2006) uses a term *languaging,* which refers to "the process of making meaning and shaping knowledge and experience through language" (p. 98), moving away from the input-output metaphor.

Since "there is more than one route to L2 development through interaction" (Mackey, 2007b, p. 24), research in the future needs to approach the routes from a variety of angles. The combination of more detailed analysis between modified output and L2 learning, and a more holistic approach will broaden our understanding of the role of the production of modified output.

6.6   Conclusion

Notwithstanding its limitations, the present study lends at least partial support to the claim of Swain's (1985, 1993, 1995, 2005) output hypothesis. This hypothesis was originally framed in terms of the impact of modified output on grammatical accuracy, but the findings of the current study suggest that production of modified output in response to clarification requests may facilitate the progress of interlanguage development towards targetlike use even when its immediate impact on grammatical accuracy is not observed. Kanagy (2001) presented a question: "What exactly 'triggers' a change in learner interlanguage, causing them to shift from one way of expressing an L2 form, to another way?" (p. 85), and the present study suggested that the production of modified output in response to clarification requests can be one of the triggers.

It was indicated that the changes in interlanguage might be associated with learners'
internal cognitive factors (e.g., noticing and interpretation of clarification requests)
which are involved in "the activity of producing the target language" (Swain, 2005).
Also, the types of non-targetlike form that learners modify might be associated with
these factors, thus suggesting their importance when learners engage in the
production of modified output. This gives support to the claim "what goes on
between the first output and the second … is part of the process of second language
learning" (Swain & Lapkin, 1995, p. 386), and validates the significance of research
which investigates the relationship between learners' internal cognitive factors and
L2 learning. The factors involved in the production of modified output may account
for the findings that changes in interlanguage may occur even with a small number of
occurrences of modified output, but the present study does not provide us with the
data about which factors are related to it and further research is awaited.

Swain (1995) proposed, as a task for future research, to investigate the conditions
under which production of modified output is demonstrated as facilitative and non-
facilitative. The present study has contributed to this by demonstrating evidence that
the production of MO on a particular non-targetlike form did not facilitate targetlike
use in subsequent situations of use. This indicates a possibility of the limited role of
production of modified output in L2 learning, and that the follow-up feedback to
learners' modified output may play a supplementary but essential role in maximising
the impact of modified output in facilitating L2 learning.

LIST OF REFERENCES

Aljaafreh, A., & Lantolf, J. P. (1994). Negative feedback as regulation and second language learning in the zone of proximal development. *The Modern Language Journal, [Special issue on sociocultural theory and L2 learning], 78*, 465-483.

Ammar, A., & Spada, N. (2006). One size fits all? Recasts, prompts and the acquisition of English possessive determiners. *Studies in Second Language Acquisition, 28*, 543-574.

Anderson, J. R. (1982). Acquisition of cognitive skill. *Psychological Review, 89*, 369-406.

Anderson, J. R. (1992). Automaticity and the Act* theory. *American Journal of Sociology, 105*, 165-180.

Batstone, R. (2002). Contexts of engagement: A discourse perspective on 'intake' and 'pushed output'. *System, 30*, 1-14.

Batstone, R. (2007). A role for discourse frames and learner interpretation in focus on form. *New Zealand Studies in Applied Linguistics, 13*, 22-35.

Bley-Vroman, R. (1983). The comparative fallacy in interlanguage studies: The case of systematicity. *Language Learning 33*, 1-17.

Bley-Vroman, R. (1989). What is the logical problem of foreign language learning? In S. M. Gass & J. Schachter (Eds.), *Linguistic perspectives on second language acquisition* (pp. 41-68). Cambridge: Cambridge University Press.

Block, D. (2003). *The social turn in second language acquisition*. Washington, D.C. : Georgetown University Press.

Braidi, S. (1999). *The acquisition of second language syntax.* New York: Arnold.

Brown, R. (1973). *A first language: The early stages*. Cambridge, Mass: Harvard University Press.

Bygate, M. (2001). Effects of task repetition on the structure and control of oral language. In M. Bygate, P. Skehan & M. Swain (Eds.), *Researching pedagogic tasks: Second language learning, teaching and testing* (pp. 23-48). Harlow, England: Longman.

Cancino, H., Rosansky, E. J., & Schumann, J. H. (1978). The acquisition of English negatives and interrogatives by native Spanish speakers. In E. M. Hatch (Ed.), *Second language acquisition: A book of reading* (pp. 207-230). Rowley, MA: Newbury House.

Carpenter, H., Jeon, K. S., MacGregor, D., & Mackey, A. (2006). Learners' interpretations of recasts. *Studies in Second Language Acquisition, 28*, 209-236.

Carroll, S., & Swain, M. (1993). Explicit and implicit negative feedback: An empirical study of the learning of linguistic generalizations. *Studies in Second Language Acquisition, 15*, 357-386.

Cazden, C., Cancino, E., Rosansky, E., & Schumann, J. (1975). *Second language acquisition sequences in children, adolescents and adults: Final report*. Washington, D.C.: National Institute of Education.

Charters, H. (2005). *The second language acquisition of Mandarin Nominal Syntax*. Unpublished doctoral dissertation, University of Auckland, Auckland.

Clancy, P. (1985). The acquisition of Japanese. In D. Slobin (Ed.), *The crosslinguistic study of language acquisition, Vol 1, The data* (pp. 373-524). Hillsdale, NJ: Erlbaum.

de Bot, K. (1996). The psycholinguistics of the output hypothesis. *Language Learning, 46*, 529-555.

de la Fuente, M. J. (2002). Negotiation and oral acquisition of L2 vocabulary: The roles of input and output in the receptive and productive acquisition of words. *Studies in Second Language Acquisition, 24*, 81-112.

DeKeyser, R. (1997). Beyond explicit rule learning: Automatizing second language morphosyntax. *Studies in Second Language Acquisition, 19*, 195-221.

DeKeyser, R. (Ed.). (2007). *Practice in a second language: Perspectives from applied linguistics and coming psychology*. New York: Cambridge University Press.

Di Biase, B., & Kawaguchi, S. (2002). Exploring the typological plausibility of processability theory. *Second Language Research, 18*, 274-302.

Doughty, C. (2001). Cognitive underpinnings of focus on form. In P. Robinson (Ed.), *Cognition and second language instruction*. Cambridge: Cambridge University Press.

Doughty, C., & Varela, E. (1998). Communicative focus on form. In C. Doughty & J. Williams (Eds.), *Focus on form in classroom second language acquisition* (pp. 114-138). New York: Cambridge University Press.

Dulay, H., & Burt, M. K. (1974). Natural sequence in child second language acquisition. *Language Learning, 24*, 37-53.

Egi, T. (2007). Recasts, learners' interpretations, and L2 development. In A. Mackey (Ed.), *Conversational interaction in second language acquisition: A collection of empirical studies* (pp. 249-267). Oxford: Oxford University Press.

Ellis, R. (1988). The effects of linguistic environment on the second language acquisition of grammatical rules. *Applied Linguistics, 9*, 257-274.

Ellis, R. (1994). *The study of second language acquisition*. Oxford: Oxford University Press.

Ellis, R. (1999). *Learning a second language through interaction*. Amsterdam: John Benjamins.

Ellis, R. (2003). *Task-based language learning and teaching*. Oxford: Oxford University Press.

Ellis, R. (2005). Principles of instructed language learning. *System, 33*, 209-224.

Ellis, R. (2007). The differential effects of corrective feedback on two grammatical structures. In A. Mackey (Ed.), *Conversational interaction in second*

language acquisition: A collection of empirical studies (pp. 340-360). Oxford: Oxford University Press.

Ellis, R., & Barkhuizen, G. (2005). *Analysing learner language*. Oxford: Oxford University Press.

Ellis, R., Basturkmen, H., & Loewen, S. (2001). Learner uptake in communicative ESL lessons. *Language Learning, 51*, 281-318.

Ellis, R., & He, X. (1999). The roles of modified input and output in the incidental acquisition of word meanings. *Studies in Second Language Acquisition, 21*, 285-301.

Ellis, R., Loewen, S., & Erlam, R. (2006). Implicit and explicit corrective feedback and the acquisition of L2 grammar. *Studies in Second Language Acquisition, 26*, 399-432.

Foster, P. (1998). A classroom perspective on the negotiation of meaning. *Applied Linguistics, 19*, 1-23.

Foster, P., & Ohta, A. S. (2005). Negotiation for meaning and peer assistance in second language classrooms. *Applied Linguistics, 26*, 402-430.

Gass, S. M. (1997). *Input, interaction, and the second language learner*. Mahwah, NJ: Lawrence Erlbaum.

Gass, S. M. (2003). Input and interaction. In C. Doughty & M. H. Long (Eds.), *The handbook of second language acquisition* (pp. 224-255). Oxford: Blackwell.

Gass, S. M., & Lewis, K. (2007). Perceptions about interactional feedback: Differences between heritage language learners and non-heritage language learners. In A. Mackey (Ed.), *Conversational interaction in second language acquisition: A collection of empirical studies* (pp. 79-99). Oxford: Oxford University Press.

Gass, S. M., & Mackey, A. (2006). Input, interaction and output: An overview. *AILA review, 19*, 3-17.

Gass, S. M., & Selinker, L. (2001). *Second Language Acquisition: An introductory course*. Mahwah, N.J. : Lawrence Erlbaum Associates.

Green, P., & Hecht, K. (1993). Pupil self-correction in oral communication in English as a foreign language. *System, 21*, 151-163.

Han, Z.-H. (2001). Fine-tuning corrective feedback. *Foreign Language Annals, 34*, 582-599.

Han, Z.-H. (2002). A study of the impact of recasts on tense consistency in L2 output. *TESOL Quarterly, 36*, 543-572.

Han, Z.-H., & Kim, J. H. (2007). Recasts in communicative EFL classes: do teacher intent and learner interpretation overlap? In A. Mackey (Ed.), *Conversational interaction in second language acquisition: A collection of empirical studies* (pp. 269-297). Oxford: Oxford University Press.

Hansen, L. (1999). Not a total loss: The attrition of Japanese negation over three decades. In L. Hansen (Ed.), *Second language attrition in Japanese contexts*. New York: Oxford University Press.

Hansen-Strain, L. (1993). *Language loss over a break in instruction: Negation in the L2 Japanese of American High School Students.* Paper presented at the Proceedings of the fourth conference on second language research in Japan, Niigata: International University of Japan.

Hatanaka, A. (1996). *Shokyu nihongo gakushuusha no hiteikei no shuutoku ni kansuru kenkyuu [An examination of the acquisition of negation by beginning learners of Japanese].* Unpublished master's thesis, Ochanomizu Women's University, Tokyo.

Hayashi, B. (1999). Testing the regression hypothesis: The remains of the Japanese negation system in Micronesia. In L. Hansen (Ed.), *Second language attrition in Japanese contexts* (pp. 154-168). New York: Oxford University Press.

He, X., & Ellis, R. (1999). Modified output and the acquisition of word meanings. In R. Ellis (Ed.), *Learning a second language through interaction* (pp. 115-132). Amsterdam: John Benjamins.

Hopkins, D., & Nettle, M. (1994). Second language acquisition research: A response to Rod Ellis. *ELT Journal, 48*, 157-161.

Ishida, M. (2004). Effects of recasts on the acquisition of the aspectual form *-te i*(*ru*) by learners of Japanese as a foreign language. *Language Learning, 54*, 311-394.

Ito, K. (1981). Two aspects of negation in child language. In P. S. Dale & D. O. Ingram (Eds.), *Child language: An international perspective* (pp. 105-114). Baltimore, MD.: University Park Press.

Ito, K. (1990). *Kodomo no kotoba: Suutoku to souzou [Children's language: Acquisition and creation]*. Tokyo: Keisoo Shoboo.

Iwashita, N. (1999). Tasks and learners' output in nonnative-nonnative interaction. In K. Kanno (Ed.), *Studies on the acquisition of Japanese as a second language* (pp. 31-52). Amsterdam: John Benjamins.

Iwashita, N. (2001). The effect of learner proficiency on interactional moves and modified output in nonnative-nonnative interaction in Japanese as a foreign language. *System, 29*, 267-287.

Iwashita, N. (2003). Negative feedback and positive evidence in task-based interaction: Different effects on L2 development. *Studies in Second Language Acquisition, 25*, 1-36.

Izumi, S. (2000). *Promoting noticing and SLA: An empirical study of the effects of output and input enhancement on ESL relativization.* Unpublished doctoral dissertation, Georgetown University, Washington, DC.

Izumi, S. (2002). Output, input enhancement, and the noticing hypothesis: An experimental study on ESL relativization. *Studies in Second Language Acquisition, 24*, 541-577.

Izumi, S. (2003). Comprehension and production processes in second language learning: In search of the psycholinguistic rationale of the output hypothesis. *Applied Linguistics, 24*, 168-196.

Izumi, S., & Bigelow, M. (2000). Does output promote noticing and second language acquisition? *TESOL Quarterly, 34*, 239-278.

Izumi, S., & Bigelow, M. (2001). Methodological and theoretical issues in testing the effects of focus on form. *TESOL Quarterly, 35*, 181-189.

Izumi, S., Bigelow, M., Fujiwara, M., & Fearnow, S. (1999). Testing the output hypothesis: Effects of output on noticing and second language acquisition. *Studies in Second Language Acquisition, 21*, 421-452.

Kamura, N. (2001a). Acquisition study of Japanese negation by Chinese-speaking learners: Focusing on past tense. *Nihongo Kyoiku [Journal of Japanese Language Teaching], 110*, 72-81.

Kamura, N. (2001b). Nihongo no hiteikei no shuutoku: Chuugoku bogowasha ni taisuru juudantekina hatsuwa chousa ni motozuite [Developmental sequences of negation in Japanese by adult Chinese-speaking learners]. *Acquisition of Japanese as a Second Language, 4*, 63-80.

Kamura, N. (2003). The acquisition of expressions of negation in Japanese by adult Chinese speaking learners -based on oral interviews-. *Acquisition of Japanese as a second language, 6*, 52-69.

Kamura, N., & Sakoda, K. (2001). Gakushuusha no goyoo o umidasu gengo shori no sutoratejii (2) - Hiteikei 'jyanai' no baai [A language processing strategy that produces learners' errors (2) - The case of the negative form 'JYANAI'-]. *Hiroshima Daigaku Nihongo Kyooiku Kenkyuu, 11*, 43-48.

Kanagy, R. (1991). *Developmental sequences in the acquisition of Japanese as a foreign language: The case of negation.* Unpublished doctoral dissertation, University of Pennsylvania.

Kanagy, R. (1994). Developmental sequences in learning Japanese: A look at negation. *Issues in Applied Linguistics: Special Issue on Applied Linguistics from an East Asian Perspective, 5*, 255-277.

Kanagy, R. (2001). Developmental sequences, second language acquisition and Japanese language pedagogy. In H. Nara (Ed.), *Advances in Japanese language pedagogy* (pp. 47-92). Columbus: National East Asian Languages Resource Centre.

Kawaguchi, S. (2000). Acquisition of Japanese verbal morphology: Applying processability theory to Japanese. *Studia Linguistica, 54*, 238-248.

Kawaguchi, S. (2005). Argument structure and syntactic development in Japanese as a second language. In M. Pienemann (Ed.), *Cross-linguistic aspects of processability theory* (pp. 253-298). Amsterdam: John Benjamins Publishing Company.

Keck, C. M., Iberri-Shea, G., Tracy-Ventura, N., & Wa-Mbaleka, S. (2006). Investigating the empirical link between task-based interaction and acquisition: A meta-analysis. In J. M. Norris & L. Ortega (Eds.), *Synthesizing research on language learning and teaching* (pp. 91-130). Amsterdam: John Benjamins.

Kellerman, E. (1985). If at first you do succeed... In S. M. Gass & C. Madden (Eds.), *Input in second language acquisition*. Rowley, Mass.: Newbury House.

Klima, E., & Bellugi, V. (1966). Syntactic regularities in the speech of children. In J. Lyons & R. J. Wales (Eds.), *Psycholinguistic Papers* (pp. 183-208). Edinburgh: Edinburgh University Press.

Komori, S., & Sakano, N. (1988). Shuudan tesuto ni yoru shokyuu bunpou no kakutoku ni tsuite. *Journal of Japanese Language Teaching, 65*, 126-128.

Kormos, J. (1999). Monitoring and self-repair in L2. *Language Learning, 49*, 303-342.

Kormos, J. (2006). *Speech Production and Second Language Acquisition* Mahwah, New Jersey: Lawrence Erlbaum

Krashen, S. (1985). *The input hypothesis: Issues and implications*. New York: Longman.

Krashen, S. (1987). *Principles and practice in second language acquisition*. New York: Prentice Hall International.

Krashen, S. (1998). Comprehensible output? *System, 26*, 175-182.

LaPierre, D. (1994). *Language output in a cooperative learning setting: Determining its effects on second language learning.* Unpublished master's thesis, University of Toronto, Toronto, Ontario, Canada.

Leeman, J. (2007). Feedback in L2 learning: Responding to errors during practice. In *Practice in a second language: Perspectives from applied linguistics and cognitive psychology*. New York: Cambridge University Press.

Leow, R. (1997). Attention, awareness, and foreign language behavior. *Language Learning, 47*, 467-505.

Levelt, W. (1989). *Speaking: From intention to articulation*. Cambridge: Cambridge University Press.

Lightbown, P. (2003). SLA research in the classroom/SLA for the classroom. *Language Learning Journal, 28*, 4-13.

Linnell, J. (1995). Can negotiation provide a context for learning syntax in a second language? *Working Papers in Educational Linguistics, 11*, 83-102.

Loewen, S. (2002). *The occurrence and effectiveness of incidental focus on form in meaning-focused ESL lessons.* Unpublished doctoral dissertation, University of Auckland, Auckland.

Loewen, S. (2004). Uptake in incidental focus on form in meaning-focused ESL lessons. *Language Learning, 54*, 153-188.

Loewen, S. (2007). The prior and subsequent use of forms targeted in incidental focus on form. In S. Fotos & H. Nassaji (Eds.), *Form-focused instruction and teacher education: Studies in honour of Rod Ellis* (pp. 101-106). Oxford: Oxford University Press.

Loewen, S., & Nabei, T. (2007). Measuring the effects of oral corrective feedback on L2 knowledge. In A. Mackey (Ed.), *Conversational interaction in second language acquisition: A collection of empirical studies* (pp. 361-377). Oxford: Oxford University Press.

Loewen, S., & Philp, J. (2006). An in-depth analysis of recasts in the adult L2 classroom. *Modern Language Journal, 90*, 536-556.

Long, M. (1983). Native speaker/non-native speaker conversation in the second language classroom. In M. Clarke & J. Handscombe (Eds.), *On TESOL '82* (pp. 207-225). Washington, DC: TESOL.

Long, M. (1996). The role of the linguistic environment in second language acquisition. In W. Ritchie & T. Bhatia (Eds.), *Handbook of second language acquisition* (pp. 413-468). San Diego: Academic Press.

Long, M. (2006). *Problems in SLA*. Mahwah, New Jersey: Lawrence Erlbaum

Long, M., Inagaki, S., & Ortega, L. (1998). The role of implicit negative feedback in SLA: Models and recasts in Japanese and Spanish. *The Modern Language Journal, [Special issue on sociocultural theory and L2 learning], 82*, 357-371.

Loschky, L. (1994). Comprehensible input and second language acquisition: What is the relationship? *Studies in Second Language Acquisition, 16*, 303-323.

Lyster, R. (1998). Recasts, repetition, and ambiguity in L2 classroom discourse. *Studies in Second Language Acquisition, 20*, 51-81.

Lyster, R. (2004). Differential effects of prompts and recasts in form-focused instruction. *Studies in Second Language Acquisition, 26*, 399-432.

Lyster, R., & Ranta, L. (1997). Corrective feedback and learner uptake: negotiation of form in communicative classrooms. *Studies in Second Language Acquisition, 19*, 37-66.

Mackey, A. (1995). *Stepping up the pace: Input, interaction and interlanguage development: An empirical study of questions in ESL.* Unpublished doctoral dissertation, University of Sydney.

Mackey, A. (1997). *Interactional modifications and the development of questions in English as a second language* Unpublished manuscript, Michigan State University, East Lansing, Michigan.

Mackey, A. (1999). Input, interaction, and second language development: An empirical study of question formation in ESL. *Studies in Second Language Acquisition, 21*, 557-587.

Mackey, A. (2002). Beyond production: Learners' perceptions about Interactional processes. *International Journal of Educational Research (Special issue on the role of interaction in instructed language learning), 37*, 379-394.

Mackey, A. (2006). Feedback, noticing and instructed second language learning. *Applied Linguistics, 27*, 405-430.

Mackey, A. (2007a). Interaction as practice. In R. DeKeyser (Ed.), *Practice in a second language: Perspectives from applied linguistics and cognitive psychology* (pp. 85-109). New York: Cambridge University Press.

Mackey, A. (2007b). Introduction: The role of conversational interaction in second language acquisition. In A. Mackey (Ed.), *Conversational interaction in second language acquisition: A collection of empirical studies* (pp. 1-26). Oxford: Oxford University Press.

Mackey, A., & Gass, S. M. (2005). *Second Language Research: Methodology and design*. Mahwah, New Jersey: Lawrence Erlbaum

Mackey, A., Gass, S. M., & McDonough, K. (2000). How do learners perceive interactional feedback? *Studies in Second Language Acquisition, 22*, 471-497.

Mackey, A., & Goo, J. (2007). Interaction research in SLA: A meta-analysis and research synthesis. In A. Mackey (Ed.), *Conversational interaction in second language acquisition: A collection of empirical studies* (pp. 408-452). Oxford: Oxford University Press.

Mackey, A., & Oliver, R. (2002). Interactional feedback and children's L2 development. *System, 30*, 459-477.

Mackey, A., Oliver, R., & Leeman, J. (2003). Interactional input and the incorporation of feedback: An exploration of NS-NNS and NNS-NNS adult and child dyads. *Language Learning, 53*, 35-66.

Mackey, A., & Philp, J. (1998). Conversational interaction and second language development: Recasts, responses and red herrings? *The Modern Language Journal, 82*, 338-356.

MacWhinney, B. (2005). A unified model of language acquisition. In J. Kroll & A. De Groot (Eds.), *Handbook of Bilingualism: Psycholinguistic Approaches*. Oxford: Oxford University Press.

MacWhinney, B. (Ed.). (1995). *The CHILDES Project: Tools for analysing talks* (2nd ed.). Hillsdale, NJ: Lawrence Erlbaum.

Mansouri, F., & Duffy, L. (2005). The pedagogic effectiveness of developmental readiness in ESL grammar instruction. *Applied Linguistics Association of Australia, 28*, 81-99.

Matsumoto, K. (1999). Jidou nihongo gakushuusha no "hitei hyougen" no shuutoku - ichi chuugokujin no ni nenkan no oudanchousa o toushite [The acquisition of Japanese negation by a Chinese school boy: The result of a two-year longitudinal study]. *Handbook of the First Conference of the Japanese Society for Language Sciences, JCHAT* 9-12.

Maynard, S. (1990). *An introduction to Japanese grammar and communication strategies*. Tokyo, Japan: The Japan Times.

McDonough, K. (2001). *Exploring the relationship between modified output and L2 learning*. Unpublished doctoral dissertation, Georgetown University, Washington, DC.

McDonough, K. (2005). Identifying the impact of negative feedback and learner's responses on ESL question development. *Studies in Second Language Acquisition, 27*, 79-103.

McDonough, K. (2006). Interaction and syntactic priming: English L2 speakers' production of dative constructions. *Studies in Second Language Acquisition, 28*, 179-207.

McDonough, K. (2007). Interactional feedback and the emergence of simple past activity verbs in L2 English. In A. Mackey (Ed.), *Conversational interaction in second language acquisition: A collection of empirical studies* (pp. 323-338). Oxford: Oxford University Press.

McDonough, K., & Mackey, A. (2006). Responses to recasts: Repetitions, primed production and linguistic development. *Language Learning, 56*, 693-720.

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

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

Meisel, J., Clahsen, H., & Pienemann, M. (1981). On determining developmental stages in natural second language acquisition. *Studies in Second Language Acquisition, 3*, 109-135.

Milon, J. (1974). The development of negation in English by a second language learner. *TESOL Quarterly, 8*, 137-143.

Mitchell, R., & Myles, F. (2004). *Second language learning theories* (second ed.). London: A Hodder Arnold Publication.

Muranoi, H. (2007). Output practice in the L2 classroom. In R. DeKeyser (Ed.), *Practice in a second language: Perspectives from applied linguistics and cognitive psychology*. New York: Cambridge University Press.

Nassaji, H. (2007). Elicitation and reformulation and their relationship with learner repair in dyadic interaction. *Language Learning, 57*, 511-548.

Nassaji, H., & Swain, M. (2000). A Vygotskian perspective on corrective feedback in L2: The effect of random versus negotiated help on the learning of English articles. *Language Awareness, 9*, 34-51.

Nicholas, H., Lightbown, P., & Spada, N. (2001). Recasts as feedback to language learners. *Language Learning, 51*, 719-758.

Nobuyoshi, J., & Ellis, R. (1993). Focused communication tasks and second language acquisition. *ELT Journal, 47*, 203-210.

Noro, I. (1995). Daini gengo ni okeru hiteikei no shutoku katei - chugokujin no kodomo no jirei kenkyuu [The acquisition of negation in Japanese as a second language: A case study of a Chinese speaking child]. *Shizuoka daigaku kyouiku gakubu kenkyu houkoku (jinbun-shakai kagaku hen), 45*, 1-12.

O'Relly, L. V., Flatiz, J., & Kromrey, J. (2001). Two modes of correcting communicative tasks: Recent findings. *Foreign Language Annals, 34*, 246-257.

Ogino, M. (2006). *The developmental sequence of negation in Japanese in KY corpus*.Unpublished manuscript.

Ohta, A. (2001). *Second language acquisition processes in the classroom: Learning Japanese*. Mahwah, New Jersey: Lawrence Erlbaum

Ohta, A. S. (2000). Rethinking interaction in SLA: Developmentally appropriate assistance in the zone of proximal development and the acquisition of L2 grammar. In J. Lantolf (Ed.), *Sociocultural theory and second language learning* (pp. 51-78). Oxford: Oxford University Press.

Okada, K. (2002). Acquisition of adjective negation by a young Japanese child. *Hokkaido Tokai Daigaku Kiyou, Jinbun Shakai Kagaku kei [Hokkaido Tokai University Bulletin: Humanities and Social Sciences], 15*, 53-70.

Philp, J. (2003). Constraints on "noticing the gap": Nonnative speakers' noticing of recasts in NS-NNS interaction. *Studies in Second Language Acquisition, 25*, 99-126.

Pica, T. (1983). Methods of morpheme quantification: Their effect on the interpretation of second language data. *Studies in Second Language Acquisition, 6*, 69-78.

Pica, T. (1988). Interlanguage adjustments as an outcome of NS-NNS negotiated interaction. *Language Learning, 38*, 45-73.

Pica, T. (1993). Adult acquisition of English as a second language under different conditions of exposure. *Language Learning, 33*, 465-497.

Pica, T., Holliday, L., Lewis, N., Berducci, D., & Newman, J. (1991). Language learning through interaction: What role does gender play? *Studies in Second Language Acquisition, 13*, 343-376.

Pica, T., Holliday, L., Lewis, N., & Morgenthaler, L. (1989). Comprehensible output as an outcome of linguistic demands on the learner. *Studies in Second Language Acquisition, 11*, 63-90.

Pica, T., Lincoln-Porter, F., Paninos, D., & Linnell, J. (1996). Language learners' interaction: How does it address the input, output and feedback needs of L2 learners? *TESOL Quarterly, 30*, 59-84.

Pienemann, M. (1984). Psychological constraints on the teachability of languages. *Studies in Second Language Acquisition, 6*, 186-214.

Pienemann, M. (1998). *Language processing and second language development: Processability theory*. Amsterdam: John Benjamins.

Pienemann, M., & Johnston, M. (1986). An acquisition-based procedure for second language assessment. *Australian Review of Applied Linguistics, 9*, 92-122.

Pienemann, M., & Johnston, M. (1987). Factors influencing the development of language proficiency. In D. Nunan (Ed.), *Applying second language acquisition research* (pp. 45-141). Adelaide, Australia: National Curriculum Resource Centre, Adult Migrant Education Program.

Pienemann, M., Johnston, M., & Brindley, G. (1988). Constructing an acquisition-based procedure for second language assessment. *Studies in Second Language Acquisition, 10*, 217-243.

Ravem, R. (1968). Language acquisition in a second language environment. *International Review of Applied Linguistics, 6*, 175-185.

Roberts, M. (1995). Awareness and the efficacy of error correction. In R. Schmidt (Ed.), *Attention and awareness in foreign language learning (Technical report, No. 9)* (pp. 162-182). Honolulu: University of Hawai'i, Second Language Teaching and Curriculum Center.

Robinson, P. (1995). Attention, memory, and the "noticing" hypothesis. *Language Learning, 45*, 283-331.

Robinson, P. (1997). Individual differences and the fundamental similarity of implicit and expect adult second language learning. *Language Learning, 47*(45-99).

Russell, J., & Spada, N. (2006). The effectiveness of corrective feedback for the acquisition of L2 grammar: A meta-analysis of the research. In J. Norris & L. Ortega (Eds.), *Synthesizing research on language learning and teaching* (pp. 133-164). Amsterdam: John Benjamins.

Sano, T. (1998). On morphological maturation. Retrieved January 21, 2006, from the World Wide Web: http://coe-sun.kuis.ac.jp/public/paper/outside/sano1.pdf

Sato, M., & Lyster, R. (2007). Modified output of Japanese EFL learners: variable effects of interlocutor versus feedback types. In A. Mackey (Ed.), *Conversational interaction in second language acquisition: A collection of empirical studies* (pp. 123-142). Oxford: Oxford University Press.

Saxton, M. (2000). Negative evidence and negative feedback: Immediate effects of the grammaticality of child speech. *First Language, 20*, 221-252.

Saxton, M., Houston-Price, C., & Dawson, N. (2005). The prompt hypothesis: Clarification requests as corrective input for grammatical errors. *Applied Psycholinguistics, 26*, 393-414.

Schegloff, E., Jefferson, G., & Sacks, H. (1977). The preference for self-correction in the organization of repair in conversation. *Language, 53*, 361-382.

Schmidt, R. (1990). The role of consciousness in second language learning. *Applied Linguistics, 11*, 129-158.

Schmidt, R. (1994). Implicit learning and the cognitive unconscious. In N. Ellis (Ed.), *Implicit and explicit learning of languages* (pp. 165-209). London: Academic Press.

Schmidt, R. (1995). Consciousness and foreign language learning: A tutorial on the role of attention and awareness in learning. In R. Schmidt (Ed.), *Attention and awareness in foreign language learning*. Honolulu: University of Hawai'i Press.

Schmidt, R., & Frota, S. (1986). Developing basic conversational ability in a second language: A case study of an adult learner of Portuguese. In R. Day (Ed.), *Talking to learn: Conversation in second language acquisition* (pp. 237-326). Rowley, MA: Newbury House.

Sheen, Y. (2006). Exploring the relationship between characteristics of recasts and learner uptake. *Language Teaching Research, 8*, 263-300.

Shehadeh, A. (1999). Non-native speakers' production of modified comprehensible output and second language learning. *Language Learning, 49*, 627-675.

Shehadeh, A. (2001). Self- and other-initiated modified output during task-based interaction. *TESOL Quarterly, 35*, 433-457.

Shehadeh, A. (2002). Comprehensible output, from occurrence to acquisition: An agenda for acquisitional research. *Language Learning, 52*, 597-647.

Shehadeh, A. (2003). Learner output, hypothesis testing, and internalizing linguistic knowledge. *System, 31*, 155-171.

Shibatani, M. (1990). *The languages of Japan*. New York: Cambridge University Press.

Shonerd, H. (1994). Repair in spontaneous speech: A window on second language development. In V. John-Steiner, C. Panofsky & L. Smith (Eds.), *Sociocultural approaches to language and literacy: an interactionist perspective* (pp. 82-108). Cambridge: Cambridge University Press.

Silver, R. (2000). Input, output, and negotiation: Conditions for second language development. In B. Swierzbin, F. Morris, M. E. Anderson, C. A. Klee & E. Tarone (Eds.), *Selected Proceedings of the Second Language Research Forum* (pp. 345-371).

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

Spada, N., & Lightbown, P. (1993). Instruction and the development of questions in L2 classrooms. *Studies in Second Language Acquisition, 15*, 205-224.

Spada, N., & Lightbown, P. (1999). Instruction, first language influence, and developmental readiness in second language acquisition. *The Modern Language Journal, 83*, 1-22.

Swain, M. (1985). Communicative competence: Some roles of comprehensible input and comprehensible output in its development. In S. M. Gass & C. Madden (Eds.), *Input in second language acquisition* (pp. 235-253). Rowley, Mass.: Newbury House.

Swain, M. (1993). The output hypothesis: Just speaking and writing aren't enough. *The Canadian Modern Language Review, 50*, 158-164.

Swain, M. (1995). Three functions of output in second language learning. In G. Cook & B. Seidlhofer (Eds.), *Principle and practice in applied linguistics: Studies in honour of Henry G. Widdowson* (pp. 125-144). Oxford: Oxford University Press.

Swain, M. (1998). Focus on form through conscious reflection. In C. Doughty & J. Williams (Eds.), *Focus on form in classroom second language acquisition* (pp. 64-81). New York: Cambridge University Press.

Swain, M. (2000). The output hypothesis and beyond: mediating acquisition through collaborative dialogue. In J. Lantolf (Ed.), *Sociocultural theory and second language learning* (pp. 97-114). Oxford: Oxford University Press.

Swain, M. (2005). The output hypothesis: Theory and Research. In E. Hinkel (Ed.), *Handbook of Research in Second Language Teaching and Learning* (pp. 471-483). Mahwah, New Jersey: Lawrence Erlbaum

Swain, M. (2006). Languaging, agency and collaboration in advanced language proficiency. In H. Brynes (Ed.), *Advanced language learning: The contributions of Halliday and Vygotsky* (pp. 95-108). London: Continuum.

Swain, M., & Lapkin, S. (1995). Problems in output and the cognitive processes they generate: A step towards second language learning. *Applied linguistics, 16*, 371-391.

Swain, M., & Lapkin, S. (1998). Interaction and second language learning: Two adolescent French immersion students working together. *The Modern Language Journal, 82*, 320-337.

Swain, M., & Lapkin, S. (2001). Focus on form through collaborative dialogue: Exploring task effects. In M. Bygate, P. Skehan & M. Swain (Eds.),

Researching pedagogic tasks: Second language learning, teaching and testing (pp. 99-118). London: Longman.

Swain, M., & Lapkin, S. (2007). The distributed nature of second language learning: Neil's perspective. In S. Fotos & H. Nassaji (Eds.), *Form-focused instruction and teacher education: Studies in honour of Rod Ellis*. Oxford: Oxford University Press.

Takashima, H. (1995). *A study of focused feedback, or output enhancement, in promoting accuracy in communication activities.* Unpublished doctoral dissertation, Temple University Japan, Tokyo.

Takashima, H., & Ellis, R. (1999). Output enhancement and the acquisition of the past tense. In R. Ellis (Ed.), *Learning a second language through interaction* (pp. 173-188). Amsterdam: John Benjamins.

Takeuchi-Furuya. (1993). *A study of naturalistic acquisition of Japanese by two native speakers of English in Japan.* Unpublished doctoral dissertation, Temple University, Tokyo.

Tarone, E. (1983). On the trial of interlanguage system. *Applied Linguistics 4*, 142-163.

Tarone, E. (1998). Research on **i**nterlanguage variation: Implications for language testing. In L. F. Bachman & A. D. Cohen (Eds.), *Interfaces between second language acquisition and language testing research* (pp. 71-89). New York: Cambridge University Press.

Tohsaku, Y. (1994). *Yookoso! An invitation to contemporary Japanese, Volume 1.* New York: McGraw-Hill.

Tomlin, R., & Villa, V. (1994). Attention in cognitive science and second language acquisition. *Studies in Second Language Acquisition, 16*, 185-204.

Trofimovich, P., Ammar, A., & Gatbonton, E. (2007). How effective are recasts? The role of attention, memory, and analytical ability. In A. Mackey (Ed.), *Conversational interaction in second language acquisition: A collection of empirical studies* (pp. 171-195). Oxford: Oxford University Press.

Tsujimura, N. (1996). *An introduction to Japanese linguistics*. Cambridge, Massachusetts: Blackwell Publishers Ltd.

Van den Branden, K. (1997). Effects of negotiation on language learners' output. *Language Learning, 47*, 589-636.

van Hest, E. (1996). *Self-repair in L1 and L2 production*. Tilburg: Tilburg University Press.

van Lier, L. (1988). *The classroom and the language learner*. London: Longman.

van Lier, L. (1996). *Interaction in the language curriculum: Awareness, autonomy and authenticity*. London & New York: Longman.

Varonis, E., & Gass, S. M. (1985). Non-native/non-native conversations: A model for negotiation of meaning. *Applied Linguistics, 6*, 71-90.

Vygotsky, L. (1978). *Mind in society: The development of higher mental processes*. Cambridge, MA: Harvard University Press.

Whitlow, J. (2001). Comments on Shinichi Izumi and Martha Bigelow's "Does output promote noticing in second language acquisition?": Some methodological and theoretical considerations. *TESOL quarterly, 35*, 177-181.

Williams, J. (2001). The effectiveness of spontaneous attention to form. *System, 29*, 325-340.

Williams, J. (2005). Form-focused instruction. In E. Hinkel (Ed.), *Handbook of Research in Second Language Teaching and Learning* (pp. 671-691). Mahwah, New Jersey: Lawrence Erlbaum.

Wode, H. (1978). The L1 vs. L2 acquisition of English interrogation. *Working Papers on Bilingualism, 15*, 37-57.

Wode, H. (1981). *Learning a second language*. Tübingen: Narr.

APPENDICES FOR CHAPTER FOUR

Appendix 4.A: *Letter of request for participation in a research project*

Department of General and Applied Linguistics
Faculty of Arts & Social Sciences
Te Kura Kete Aronui
The University of Waikato
Private Bag 3105
Hamilton
New Zealand

THE UNIVERSITY OF
WAIKATO
Te Whare Wānanga o Waikato

26 February 2007

Request for Participation in a Research Project

Project Name: Learning of Japanese as a foreign language
Researcher: Masa Ogino (PhD candidate) **Email**: masayoshi@xtra.co.nz
Chief Supervisor: Prof. Ray Harlow   **Email:** rharlow@waikato.ac.nz

Introduction
I am currently conducting research into the learning of Japanese as a foreign language
for my PhD research. You are invited to consider participating in this study. This
letter will describe the purpose and nature of the study and your right as a participant
in the study. If you decide to participate, you will be asked to sign and date in the
consent form.

Explanation of the study
We will be looking at how we learn Japanese and I anticipate that this research will
provide valuable insight into the mechanism involved in learning Japanese and
foreign languages. I cannot be specific about the focus of my research at this stage
because it might affect the result. I will explain the project more fully after I have
collected all the data.

If you agree to participate in the project, you will be asked to attend five 20 minute-
sessions over two months. You will carry out one speaking activity during your class
time and another four speaking activities outside of class time over eight weeks. As
part of study you will also complete a questionnaire about your history of learning
Japanese. The first session will be held during a lecture time and you will be asked to
put your preferred time for the 2nd, 3rd, 4th and 5th sessions. A tape-recorder will be
used to record what you are saying during all the activities.

Schedule of the project

Session	*Date*	*Venue*
1st session	2/3 (Friday) 1:00-2:00	KB.01
2nd session	Your preferred time on 14/3 (Wed), 15/3 (Thurs) or 16/3 (Fri)	JK3.01
3rd session	Your preferred time on 21/3 (Wed), 22/3 (Thurs) or 23/3 (Fri)	JK3.01
4th session	Your preferred time on 29/3 (Thurs) or 30/3 (Fri)	JK3.01
5th session	Your preferred time on 26/4 (Thurs) or 27/4 (Fri)	JK3.01

Confidentiality

All information collected will be treated with the strictest confidentiality. The data will be used anonymously, and the reporting of the information from the research will not identify you or your institutional affiliation. The data will be stored indefinitely under secure conditions as the results of the research may be used in future publications.

Your participation

Participating in this study is strictly voluntary. That means you do not have to be a part of the study. Your decision to participate will in no way affect your grade in the Japanese course you are enrolled in. If at any point you change your mind and no longer want to participate, you can tell the researcher.

You will not be paid for participation in this study, but those who have completed all the five sessions will go into a draw for a $25 book voucher.

If you are uncomfortable with any aspect of the research, I would welcome the opportunity to discuss those concerns. Please contact me at masayoshi@xtra.co.nz. If your concerns are not all to be resolved to your satisfaction, you can withdraw from the study at any time.

Attached is a letter to ask for your participation by one of my supervisors, Dr Roger Barnard.

<u>What to do now</u>

Please come to the lecture (JAPA101) at A.G.11 at 1:00 on 2/3 (Friday), and we will move to KB01. More details will be given, questions about this project will be answered, and the first session will be held there.

Thanks for your time and help.

Yours sincerely,

Masayoshi Ogino (e-mail address: masayoshi@xtra.co.nz)
PhD candidate
Department of General and Applied Linguistics
The University of Waikato

Chief Supervisor: Prof. Ray Harlow (e-mail address: rharlow@waikato.ac.nz)
Department of General and Applied Linguistics
The University of Waikato

Appendix 4.B: *Consent form*

Research Project

Participant Consent Form

I agree to participate in the research being conducted by Masayoshi Ogino, PhD student at the University of Waikato. I have received an information sheet outlining the study. I understand that this will involve five 20-minute-sessions. I can withdraw at any time and refuse to answer any questions.

Name: ....

Signature: ....

Email ....

Telephone (if e-mail is not available)

....

Date: ....

Questionnaire for Participants in the Project

This information will not be used in a way which will identify any individual. We are asking for names just to help us keep track of the data, but none of you will be identified by name.

1. Name   _____   _____

2. Age: _____ years

3. Native country: _____

4. First language (Learned as a child and most fluent now)

   _____

5. Other languages: _____

   Fluency level: beginner/intermediate/advanced

   _____

   Fluency level: beginner/intermediate/advanced

   _____

   Fluency level: beginner/intermediate/advanced

6. Have you been to Japan before?
   - ☐ Yes
   - ☐ No

   If yes, how long?   _____Years _____Months

7. Japanese study:
   - ☐ intermediate   _____years
   - ☐ secondary school   _____years
   - ☐ university   _____years

Thank you for your time and co-operation

Appendix 4.D1: *Transcriptions of questions used in testing activities*

Questions for Task 1	Translation
1. Kore wa hon desu ka.	Is this a book?
2. Kyoo wa tenki ga ii desu ka.	Is the weather fine today?
3. Terebi wa omoshiroi desu ka.	It the T.V. programme funny?
4. Tonari no kyooshitsu wa urusai desu ka.	Is the classroom next door noisy?
5. Kyoo wa atsui desu ka.	Is it hot today?
6. Kono mondai wa yasashii desu ka.	Is this question easy?
7. Nipponman wa nihonjin desu ka.	Is Nipponman Japanese?
8. Nipponman wa eigo o hanashimasu ka.	Does Nipponman speak English?
9. Nipponman wa se ga takai desu ka.	Is Nipponman tall?
10. Kore wa omoi desu ka.	Is this heavy?
11. Kore mo omoi desu ka.	Is this also heavy?
12. Kono hito wa wakai desu ka.	Is this person young?
13. Honda san no uchi wa chikai desu ka.	Is Mr Honda's house nearby?
14. Yamada san no uchi wa semai desu ka.	Is Mr Yamada's house small?
15. Honda san no uchi mo semai desu ka.	Is Mr Honda's house also small?
16. Nipponman wa sushi ga suki desu ka.	Does Nipponman like sushi?
17. Nipponman wa wasabi mo suki desu ka.	Does Nipponman also like wasabi?
18. Kore wa mazui desu ka.	Is this yuk (taste)?
19. Kono suupu mo mazui desu ka.	Is this soup yuk?
20. Kono hon wa usui desu ka.	Is this book thin?

Note:
15 adjectives
    10 adjective negations
     5 adjective affirmatives
5 distractors
    1 verb affirmative
    1 noun affirmative
    1 noun negative
    1 nominal adjective affirmative
    1 nominal adjective negation

Questions for Task 2	Translation
1. Sono appato wa ookii desu ka.	Is the apartment house large?
2. Sono apaato wa atarashii desu ka.	Is the apartment new?
3. Tonari no biru mo atarashii desu ka.	Is the next building also new?
4. Sono apaato wa hikui desu ka.	Is the apartment house low?
5. Tonari no biru mo hikui desuka.	Is the neighbouring building also low?
6. Sono apaato wa yasui desu ka.	Is the apartment house cheap?
7. Sono apaato wa akai desu ka.	Is the apartment house red?
8. Tonari no biru mo akai desu ka.	Is the neighbouring building also red?
9. Apaato no so ba ni ki ga arimasu ka.	Is there a tree near the apartment house?
10. Sono ki wa takai desu ka.	Is the tree tall?
11. Hashi ga arimasu ka.	Are there any bridges?
12. Hashi wa nagai desu ka.	Is the bridge long?
13. Kawa wa kitanai desu ka.	Is the river dirty?
14. Michi wa hiroi desu ka.	Is the street wide?
15. Sushi baa ga arimasu ka.	Are there any sushi bars?
16. Sushi baa wa tooi desu ka.	Is the sushi bar expensive?
17. Ki wa sukunai desu ga, kuruma mo sukunai desu ka.	There are only a few trees, but are there only a few cars, as well?
18. Kyoo wa hare desu ka.	Is the weather fine today?
19. Kyoo wa samui desu ka.	Is it cold today?
20. Ima san-ji desu ka.	Is it three o'clock, now?

Note:
15 adjectives
   10 adjective negations
    5 adjective affirmatives
5 distractors
   2 verb affirmatives
   1 verb negations
   1 noun affirmative
   1 noun negative

TASK 1

Describing a town

IMPORTANT
Your answer to each question should be very clear.

If your answer to a question is "Yes", you need to make your answer clear by saying "Yes, I am/do".

If "no", you need to say "No, I am not/do not".
For example, if you are asked "Did you have breakfast?", and if your answer is "no", you need to say "No, I didn't" in Japanese.

When you have answered a question, click once and you will hear the next question.

If you do not answer the question within one minute, you will hear the next question automatically.

To hear a question again, click on this icon.

Click to go to the next slide.

Are you ready to do the task "Describing a town" now?

If you have a question, put your hand up and wait.

If you are ready, press the button on the tape-deck so that your reply is audio-recorded. Please press it now.

Please click to see the town and to start the task.
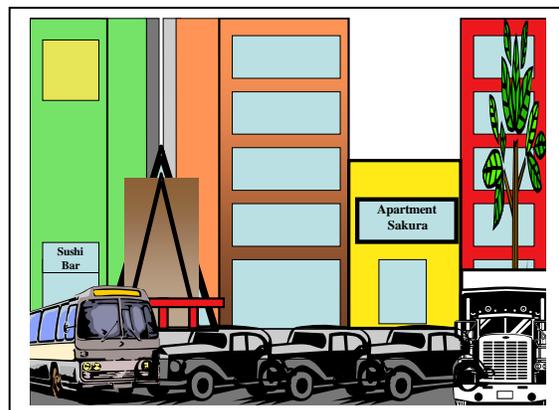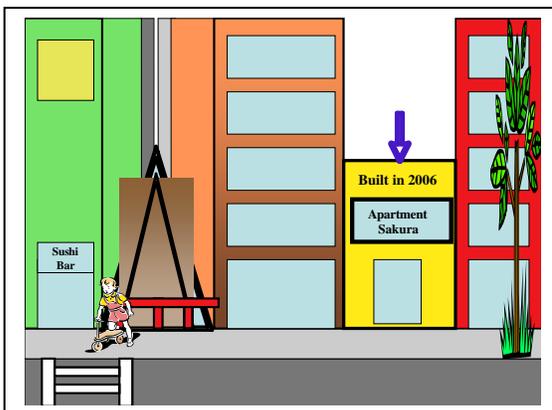
Appendix 4.E: *List of words used in the second treatment sessions*

Verbs	1. *mimasu*	to see
	2. *kikimasu*	to listen
	3. *kakimasu*	to write
	4. *nomimasu*	to drink
	5. *tabemasu*	to eat
	6. *benkyooshimasu*	to study
Nouns	1. *pen*	pen
	2. *hon*	book
	3. *keshigomu*	eraser
	4. *isu*	chair
	5. *tsukue*	desk
	6. *enpitsu*	pencil
Nominal-adjectives	1. *yuumei*	famous
	2. *suki*	like
	3. *kirai*	dislike
	4. *shizuka*	quiet
	5. *nigiyaka*	lively
	6. *benri*	convenient
Adjectives	1. *tanoshii*	enjoyable
	2. *tsumaranai*	boring
	3. *atatakai*	warm
	4. *suzushii*	cool
	5. *oishii*	delicious
	6. *amai*	sweet
	7. *isogashii*	busy
	8. *shiroi*	white
	9. *aoi*	blue
	10. *kuroi*	black

APPENDICES FOR CHAPTER FIVE

Appendix 5.A1
Individual oral production data from the interview of treatment sessions (Experimental group)

	Attempt	TL	NTL	Accuracy	CR given	No response	MOTL	MONTL	Repetition	Others
P1	3	3	0	100.0	0	0	0	0	0	0
P2	3	3	0	100.0	0	0	0	0	0	0
P3	7	5	2	71.4	2	0	2	0	0	0
P4	2	0	2	0.0	2	0	0	1	1	0
P5	11	1	10	9.1	6	0	0	1	4	1
P6	2	1	1	50.0	1	0	0	0	1	0
P7	6	4	2	66.7	2	0	1	0	1	0
P8	4	1	3	25.0	2	0	1	1	0	0
P9	1	1	0	100.0	0	0	0	0	0	0
P10	4	0	4	0.0	3	0	0	1	2	0
P11	5	2	3	40.0	3	0	0	1	1	1
P12	2	0	2	0.0	2	0	0	0	1	1
P13	4	1	3	25.0	3	0	1	0	1	1
P14	2	0	2	0.0	0	0	0	0	0	0
Sum	56	22	34	587.2	26	0	5	5	12	4
M	4.0	1.6	2.4	41.9	1.9	0.0	0.4	0.4	0.9	0.3
SD	2.63	1.60	2.50	39.57	1.66	0.00	0.63	0.50	1.10	0.47
Mdn	3.5	1.0	2.0	32.5	2.0	0.0	0.0	0.0	1.0	2.0

Note. TL = targetlike pattern, NTL = non-targetlike pattern, MOTL = modified output with targetlike pattern, MONTL = modified output with non-targetlike pattern

Appendix 5.A2
Individual oral production data from the interview of treatment sessions (Control group)

	Attempt	TL pattern	NTL pattern	Accuracy
P1	7	3	4	42.9
P2	3	0	3	0.0
P3	1	0	1	0.0
P4	3	1	2	33.3
P5	8	4	4	50.0
P6	4	2	2	50.0
P7	5	3	2	60.0
P8	5	3	2	60.0
P9	2	1	1	50.0
P10	10	0	10	0.0
P11	3	1	2	33.3
P12	2	0	2	0.0
P13	3	0	3	0.0
P14	4	2	2	50.0
Sum	60	20	40	429.5
M	4.3	1.4	2.9	30.7
SD	2.52	1.40	2.25	24.93
Mdn	3.5	1.0	2.0	38.1

Appendix 5.A3

Individual oral production data from Round 1 in the mechanical drills of the treatment sessions (Experimental group)

	Attempt	TL pattern	NTL pattern	Accuracy	CR given	No response	MOTL	MONTL	Repetition	Others
P1	6	6	0	100.0	0	0	0	0	0	0
P2	8	7	1	87.5	1	0	0	0	1	0
P3	3	3	0	100.0	0	0	0	0	0	0
P4	6	6	0	100.0	0	0	0	0	0	0
P5	7	1	6	14.3	5	0	2	0	3	0
P6	5	5	0	100.0	0	0	0	0	0	0
P7	4	3	1	75.0	1	0	0	1	0	0
P8	6	5	1	83.3	1	0	0	0	1	0
P9	8	8	0	100.0	0	0	0	0	0	0
P10	6	0	6	0.0	6	0	3	0	2	1
P11	10	6	4	60.0	3	0	1	0	1	1
P12	3	1	2	33.3	2	0	0	0	2	0
P13	6	3	3	50.0	3	0	0	0	3	0
P14	2	2	0	100.0	0	0	0	0	0	0
Sum	80	56	24	1003.4	22	0	6	1	13	2
M	5.7	4.0	1.7	71.7	1.6	0.0	0.4	0.1	0.9	0.1
SD	2.20	2.48	2.20	34.77	1.99	0.00	0.94	0.27	1.14	0.36
Mdn	6.0	4.0	1.0	85.4	1.0	0.0	0.0	0.0	0.5	1.0

Appendix 5.A4

Individual oral production data from Round 1 in the mechanical drills of the treatment sessions (Control group)

	Attempt	TL pattern	NTL pattern	Accuracy
P1	7	7	0	100.0
P2	9	8	1	88.9
P3	5	5	0	100.0
P4	7	4	3	57.1
P5	8	3	5	37.5
P6	8	7	1	87.5
P7	8	6	2	75.0
P8	8	1	7	12.5
P9	6	6	0	100.0
P10	4	1	3	25.0
P11	9	3	6	33.3
P12	5	2	3	40.0
P13	2	0	2	0.0
P14	8	5	3	62.5
Sum	94	58	36	819.4
M	6.7	4.1	2.6	55.5
SD	2.05	2.54	2.21	34.17
Mdn	7.5	4.5	2.5	59.8

Appendix 5.A5
Individual oral production data from Round 2 in the mechanical drills of the treatment sessions (Experimental group)

	Attempt	TL pattern	NTL pattern	Accuracy	CR given	No response	MOTL	MONTL	Repetition	Others
P1	6	6	0	100.0	0	0	0	0	0	0
P2	10	9	1	90.0	1	0	1	0	0	0
P3	7	4	3	57.1	3	0	0	0	3	0
P4	5	5	0	100.0	0	0	0	0	0	0
P5	7	2	5	28.6	5	0	1	2	2	0
P6	7	4	3	57.1	3	0	1	0	2	0
P7	2	2	0	100.0	0	0	0	0	0	0
P8	6	6	0	100.0	0	0	0	0	0	0
P9	9	9	0	100.0	0	0	0	0	0	0
P10	5	0	5	0.0	5	0	2	2	1	0
P11	8	8	0	100.0	0	0	0	0	0	0
P12	3	2	1	66.7	1	0	1	0	0	0
P13	4	2	2	50.0	2	0	1	0	0	1
P14	5	2	3	40.0	2	0	0	1	1	0
Sum	84	61	23	989.5	22	0	7	5	9	1
M	6.0	4.4	1.6	70.7	1.6	0.0	0.5	0.4	0.6	0.1
SD	2.22	2.90	1.86	32.88	1.83	0.00	0.65	0.74	1.01	0.27
Mdn	6.0	4.0	1.0	78.3	1.0	0.0	0.0	0.0	0.0	0.0

Appendix 5.A6
Individual oral production data from Round 2 in the mechanical drills of the treatment sessions (Control group)

	Attempt	TL pattern	NTL pattern	Accuracy
P1	7	6	1	85.7
P2	10	6	4	60.0
P3	5	5	0	100.0
P4	7	2	5	28.6
P5	8	5	3	62.5
P6	7	7	0	100.0
P7	8	5	3	62.5
P8	8	0	8	0.0
P9	6	6	0	100.0
P10	4	2	2	50.0
P11	8	4	4	50.0
P12	5	1	4	20.0
P13	2	0	2	0.0
P14	7	4	3	57.1
Sum	92	53	39	776.4
M	6.6	3.8	2.8	55.5
SD	2.03	2.36	2.22	34.17
Mdn	7.0	4.5	3.0	58.6

Appendix 5.B1
Summary of the individual oral production data from the treatment sessions (Experimental group)

Participant	Attempt	TL	NTL	Accuracy	CR	MO	MOTL	MONTL	Repetition	Others	MO (%)	MOTL (%)
P1	15	15	0	100.0	0	0	0	0	0	0	0.0	0.0
P2	21	19	2	92.5	2	1	1	0	1	0	50.0	50.0
P3	17	12	5	76.2	5	2	2	0	3	0	40.0	40.0
P4	13	11	2	66.7	2	1	0	1	1	0	50.0	0.0
P5	25	4	21	17.3	16	6	3	3	9	1	37.5	18.8
P6	14	10	4	69.0	4	1	1	0	3	0	25.0	25.0
P7	12	9	3	80.6	3	2	1	1	1	0	66.7	33.3
P8	16	12	4	69.4	3	2	1	1	1	0	66.7	33.3
P9	18	18	0	100.0	0	0	0	0	0	0	0.0	0.0
P10	15	0	15	0.0	14	8	5	3	5	1	57.1	35.7
P11	23	16	7	66.7	6	2	1	1	2	2	33.3	16.7
P12	8	3	5	33.3	5	1	1	0	3	1	20.0	20.0
P13	14	6	8	41.7	8	2	2	0	4	2	25.0	25.0
P14	9	4	5	46.7	2	1	0	1	1	0	50.0	0.0
Sum	220	139	81	860.1	70.0	29.0	18.0	11.0	34.0	7.0	521.3	297.8
M	15.7	9.9	5.8	61.4	5.0	2.1	1.3	0.8	2.4	0.5	37.2	21.3
SD	4.86	5.90	5.78	30.15	4.8	2.2	1.4	1.1	2.4	0.8	21.5	16.5
Mdn	15.0	10.5	4.5	67.9	3.5	1.5	1.0	0.5	1.5	0.0	38.8	22.5

Note. MO (%) = production rate of modified output, MOTL (%) = production rate of modified output with targetlike pattern

Appendix 5.B2
Summary of the individual oral production data from the treatment sessions (Control group)

Participant	Attempt	TL	NTL	Accuracy
P1	21	16	5	76.2
P2	22	14	8	49.6
P3	11	10	1	66.7
P4	17	7	10	39.7
P5	24	12	12	50.0
P6	19	16	3	79.2
P7	21	14	7	65.8
P8	21	4	17	24.2
P9	14	13	1	83.3
P10	18	3	15	25.0
P11	20	8	12	38.9
P12	12	3	9	20.0
P13	7	0	7	0.0
P14	19	11	8	56.5
Sum	246	131	115	675.1
M	17.6	9.4	8.2	48.2
SD	4.85	5.26	4.81	24.96
Mdn	19.0	10.5	8.0	49.8

Appendix 5.C1

Types of pattern which received clarification requests in the interview (Experimental group)

Participant	TL	NTL	accuracy	CR	*nai*	*masen*	*janai*	*kujanai*	*kunai*	(others)
P1	3	0	100.0	0						
P2	3	0	100.0	0						
P3	5	2	71.4	2					1	1
P4	0	2	0.0	2			2			
P5	1	10	9.1	6			6			
P6	1	1	50.0	1			1			
P7	4	2	66.7	2					2	
P8	1	3	25.0	2					2	
P9	1	0	100.0	0						
P10	0	4	0.0	3		2				1
P11	2	3	40.0	3			2			1
P12	0	2	0.0	2			2			
P13	1	3	25.0	3		3				
P14	0	2	0.0	0						

Note. TL = targetlike pattern, NTL = non-targetlike pattern, CR = clarification requests

Appendix 5.C2

Types of pattern which received clarification requests in Round 1 (Experimental group)

Participant	TL	NTL	accuracy	CR	*nai*	*masen*	*janai*	*kujanai*	*kunai*	(others)
P1	6	0	100.0	0						
P2	7	1	87.5	1			1			
P3	3	0	100.0	0						
P4	6	0	100.0	0						
P5	1	6	14.3	5			5			
P6	5	0	100.0	0						
P7	3	1	75.0	1					1	
P8	5	1	83.3	1			1			
P9	8	0	100.0	0						
P10	0	6	0.0	6	3		2		1	
P11	6	4	60.0	3	1		2			
P12	1	2	33.3	2			2			
P13	3	3	50.0	3	2		1			
P14	2	0	100.0	0						

Appendix 5.C3
Types of pattern which received clarification requests in Round 2
(*Experimental group)*

Participant	TL	NTL	accuracy	CR	*nai*	*masen*	*janai*	*kujanai*	*kunai*	(others)
P1	6	0	100.0	0						
P2	9	1	90.0	1					1	
P3	4	3	57.1	3			3			
P4	5	0	100.0	0						
P5	2	5	28.6	5	2		2		1	
P6	4	3	57.1	3			2		1	
P7	2	0	100.0	0						
P8	6	0	100.0	0						
P9	9	0	100.0	0						
P10	0	5	0.0	5	1		4			
P11	8	0	100.0	0						
P12	2	1	66.7	1	1					
P13	2	2	50.0	2			1		1	
P14	2	3	40.0	2			1		1	

Appendix 5.C4
Summary of types of pattern which received clarification requests in treatment
sessions (Experimental group)

Participant	NTL	CR	*nai*	*masen*	*janai*	*kujanai*	*kunai*	(others)
P1	0	0	0	0	0	0	0	0
P2	2	2	0	0	1	0	1	0
P3	5	5	0	0	3	0	1	1
P4	2	2	0	0	2	0	0	0
P5	21	16	2	0	13	0	1	0
P6	4	4	0	0	3	0	1	0
P7	3	3	0	0	0	0	3	0
P8	4	3	0	0	1	0	2	0
P9	0	0	0	0	0	0	0	0
P10	15	14	4	2	6	0	1	1
P11	7	6	1	0	4	0	0	1
P12	5	5	1	0	4	0	0	0
P13	8	8	2	3	2	0	1	0
P14	5	2	0	0	1	0	1	0
Sum	81	70	10	5	40	0	12	3
M	5.8	5.0	0.7	0.4	2.9	0	0.9	0.2
SD	5.78	4.79	1.20	0.93	3.42	0	0.86	0.43
Mdn	4.5	3.5	0	0	2	0	1	0

Appendix 5.D:
Response moves to clarification requests in each of treatment activities
(Experimental group)

| | Interview | | Mechanical Drills | | | |
| | | | Round 1 | | Round 2 | |
	Sum	%	Sum	%	Sum	%
MO	10	38.5	7	31.8	12	54.5
Repetition	12	46.2	13	59.0	9	40.9
Others	4	15.4	2	9.0	1	4.5
Total	26	n/a	22	n/a	22	n/a

Appendix 5.E1
Types of pattern modified (MOTL + MONTL) in response to clarification requests
(Experimental group)

Participant	MO	*nai*	*masen*	*janai*	*kujanai*	*kunai*	others
P1	0						
P2	1					1	
P3	2					1	1
P4	1			1			
P5	6	2		4			
P6	1			1			
P7	2					2	
P8	2					2	
P9	0						
P10	8	4		3			1
P11	2			1			1
P12	1	1					
P13	2		1			1	
P14	1					1	
Total	29	7	1	10	0	8	3

Note. MO = modified output in response to clarification requests, TL = targetlike pattern, NTL = non-targetlike pattern

Appendix 5.E2
Types of pattern repeated in response to clarification requests
(Experimental group)

Participant	Repetition	*nai*	*masen*	*janai*	*kujanai*	*kunai*	others
P1	0						
P2	1			1			
P3	3			3			
P4	1			1			
P5	9			8		1	
P6	3			2		1	
P7	1					1	
P8	1			1			
P9	0						
P10	5		2	3			
P11	2	1		1			
P12	3			3			
P13	4	2	1	1			
P14	1			1			
Total	34	3	3	25	0	3	0

Appendix 5.E3
Types of pattern: Other types of response in response to clarification requests (Experimental group)

Participant	Other types of response	*nai*	*masen*	*janai*	*kujanai*	*kunai*	others
P1	0						
P2	0						
P3	0						
P4	0						
P5	1			1			
P6	0						
P7	0						
P8	0						
P9	0						
P10	1					1	
P11	2			2			
P12	1			1			
P13	2		1	1			
P14	0						
Total	7	0	1	5	0	1	0

Appendix 5.F1
MO with targetlike and non-targetlike pattern by each activity
(Experimental group)

Interview

	Sum	M	SD	%	Mdn
MO with TL	5	0.4	0.63	50.0	0.0
MO with NTL	5	0.4	0.50	50.0	0.0

Appendix 5.F2:
Round 1 in mechanical drills (Experimental group)

	Sum	M	SD	%	Mdn
MO with TL	6	0.4	0.94	85.6	0.0
MO with NTL	1	0.1	0.27	14.3	0.0

Appendix 5.F3
Round 2 in mechanical drills (Experimental group)

	Sum	M	SD	%	Mdn
MO with TL	7	0.5	0.65	58.3	0.0
MO with NTL	5	0.4	0.74	41.7	0.0

Note. TL = targetlike pattern, NTL = non-targetlike pattern

Appendix 5.G
MO, MO with TL, and MO with NTL in the treatment sessions (Experimental group)

	MO (MOTL + MONTL)				MO with TL				MO with NTL			
	Int	R1	R2	Sum	Int	R1	R2	Sum	Int	R1	R2	Sum
P1	0	0	0	0	0	0	0	0	0	0	0	0
P2	0	0	1	1	0	0	1	1	0	0	0	0
P3	2	0	0	2	2	0	0	2	0	0	0	0
P4	1	0	0	1	0	0	0	0	1	0	0	1
P5	1	2	3	6	0	2	1	3	1	0	2	3
P6	0	0	1	1	0	0	1	1	0	0	0	0
P7	1	1	0	2	1	0	0	1	0	1	0	1
P8	2	0	0	2	1	0	0	1	1	0	0	1
P9	0	0	0	0	0	0	0	0	0	0	0	0
P10	1	3	4	8	0	3	2	5	1	0	2	3
P11	1	1	0	2	0	1	0	1	1	0	0	1
P12	0	0	1	1	0	0	1	1	0	0	0	0
P13	1	0	1	2	1	0	1	2	0	0	0	0
P14	0	0	1	1	0	0	0	0	0	0	1	1
Sum	10	7	12	29	5	6	7	18	5	1	5	11
M	0.7	0.5	0.9	2.1	0.4	0.4	0.5	1.3	0.4	0.1	0.4	0.8
SD	0.73	0.94	1.23	2.23	0.63	0.94	0.65	1.38	0.50	0.27	0.74	1.05
Mdn	1.0	0	0.5	1.5	0	0	0	1	0	0	0	0.5

Note. Int = interview session, R1 = Round 1 in the mechanical drills,
R2 = Round 2 in the mechanical drills, TL = targetlike pattern, NTL = non-targetlike pattern

Appendix 5.H
Types of pattern modified sorted by accuracy gain score (between pre-test and post-test) (Experimental group)

Participant	Gain	*nai*		*masen*		*janai*		*kujanai*		*kunai*		(others)	
		TL	NTL	TL	NTL	TL	NTL	TL	NTL	TL	NTL	TL	NTL
P1	-2.3												
P2	26.7									1			
P3	19.0									1		1	
P4	39.5						1						
P5	4.4		2			3	1						
P6	46.7					1							
P7	41.7									1	1		
P8	78.6									1	1		
P9	78.6												
P10	75.0	4				1	2						1
P11	55.6					1							1
P12	12.5	1											
P13	66.7			1						1			
P14	40.0										1		
Sum	n/a	5	2	1	0	6	4	0	0	5	3	1	2

237

Appendix 5.I1
Testing data (Experimental group)

	Pre-test				Post-test				Delayed post-test			
	a	b	c	d	a	b	c	d	a	b	c	d
P1	13	11	2	84.6	17	14	3	82.4	16	16	0	100.0
P2	6	4	2	66.7	15	14	1	93.3	19	17	2	89.5
P3	9	6	3	66.7	14	12	2	85.7	17	17	0	100.0
P4	10	5	5	50.0	19	17	2	89.5	17	17	0	100.0
P5	5	2	3	40.0	9	4	5	44.4	11	3	8	27.3
P6	12	4	8	33.3	15	12	3	80.0	13	9	4	69.2
P7	8	2	6	25.0	15	10	5	66.7	14	10	4	71.4
P8	12	0	12	0.0	14	11	3	78.6	13	13	0	100.0
P9	9	0	9	0.0	14	11	3	78.6	13	13	0	100.0
P10	7	0	7	0.0	12	9	3	75.0	14	11	3	78.6
P11	5	0	5	0.0	18	10	8	55.6	15	5	10	33.3
P12	1	0	1	0.0	8	1	7	12.5	8	2	6	25.0
P13	1	0	1	0.0	12	8	4	66.7	11	11	0	100.0
P14	1	0	1	0.0	5	2	3	40.0	5	3	2	60.0
Sum	99	34	65	n/a	187	135	52	n/a	186	147	39	n/a
M	7.1	2.4	4.6	26.2	13.4	9.6	3.7	67.8	13.3	10.5	2.8	75.3
S.D	4.12	3.27	3.41	30.76	3.89	4.60	1.94	22.56	3.71	5.46	3.29	28.76
Mdn	7.5	1.0	4.0	12.5	14.0	10.5	3.0	76.8	13.5	11.0	2.0	84.0

Appendix 5.I2
Testing data (Control group)

	Pre-test				Post-test				Delayed post-test			
	a	b	c	d	a	b	c	d	a	b	c	d
P1	18	16	2	88.9	20	20	0	100.0	20	20	0	100.0
P2	14	11	3	78.6	18	18	0	100.0	8	8	0	100.0
P3	6	3	3	50.0	5	5	0	100.0	4	3	1	75.0
P4	16	8	8	50.0	18	4	14	22.2	19	7	12	36.8
P5	7	3	4	42.9	7	5	2	71.4	7	6	1	85.7
P6	20	7	13	35.0	18	15	3	83.3	17	16	1	94.1
P7	10	3	7	30.0	4	4	0	100.0	5	5	0	100.0
P8	12	1	11	8.3	15	11	4	73.3	19	17	2	89.5
P9	15	1	14	6.7	16	8	8	50.0	19	12	7	63.2
P10	13	0	13	0.0	14	0	14	0.0	17	0	17	0.0
P11	8	0	8	0.0	17	1	16	5.9	17	1	16	5.9
P12	4	0	4	0.0	8	0	8	0.0	17	7	10	41.2
P13	3	0	3	0.0	17	11	6	64.7	17	11	6	64.7
P14	2	0	2	0.0	13	11	2	84.6	15	14	1	93.3
Sum	148	53	95	n/a	190	113	77	n/a	201	127	74	n/a
M	10.6	3.8	7	27.9	13.6	8.1	6	61.1	14.4	9.1	5	67.8
S.D	5.73	4.95	4.42	30.72	5.35	6.49	5.69	38.75	5.69	6.12	6.14	34.45
Mdn	11.0	2.0	5.5	19.2	15.5	6.5	3.5	72.4	17.0	7.5	1.5	80.4

Note. a: attempt of using target form   b: targetlike use
c: non-targetlike use   d: accuracy rate (%)

Appendix 5.J
Gains in the post-test and delayed post-test

	Experimental Group (n = 14)			Control Group (n = 14)	
	Gains in the post-test	Gains in the delayed post-test		Gains in the post-test	Gains in the delayed post-test
P1	-2.3	15.4	P1	11.1	11.1
P2	26.7	22.8	P2	21.4	21.4
P3	19.0	33.3	P3	50.0	25.0
P4	39.5	50.0	P4	-27.8	-13.2
P5	4.4	-12.7	P5	28.6	42.9
P6	46.7	35.9	P6	48.3	59.1
P7	41.7	46.4	P7	70.0	70.0
P8	78.6	100.0	P8	65.0	81.1
P9	78.6	100.0	P9	43.3	56.5
P10	75.0	78.6	P10	0.0	0.0
P11	55.6	33.3	P11	5.9	5.9
P12	12.5	25.0	P12	0.0	41.2
P13	66.7	100.0	P13	64.7	64.7
P14	40.0	60.0	P14	84.6	93.3
M	41.6	49.1	M	33.2	39.9
SD	26.17	33.46	SD	31.55	31.27
Mdn	42.0	41.2	Mdn	36.0	42.0

Note. Gain in the post test = gain between pre-test and post-test
    Gain in the delayed-post test = gain between pre-test and delayed post-test

Appendix 5.K1
Frequency of MO, MOTL, MONTL by higher and lower gain scorers (pre-post)

Participants	Group	Gain	MO	MOTL	MONTL
		(Post-test - pre-test)			
P8	Higher gain	78.6	2	1	1
P9	Higher gain	78.6	0	0	0
P10	Higher gain	75.0	8	5	3
P13	Higher gain	66.7	2	2	0
P11	Higher gain	55.6	2	1	1
P6	Higher gain	46.7	1	1	0
P7	Higher gain	41.7	2	1	1
P14	Lower gain	40.0	1	0	1
P4	Lower gain	39.5	1	0	1
P2	Lower gain	26.7	1	1	0
P3	Lower gain	19.0	2	2	0
P12	Lower gain	12.5	1	1	0
P5	Lower gain	4.4	6	3	3
P1	Lower gain	-2.3	0	0	0
M		41.6	2.1	1.3	0.8
SD		27.16	2.23	1.38	1.05
Mdn		40.9	1.5	1.0	0.5

Note. MO = modified output, MOTL = modified output with targetlike pattern,
MONTL = modified output with non-targetlike pattern

Appendix 5.K2
Frequency of MO, MOTL, MONTL by higher and lower gain scorers (pre-delayed)

Participants	Group	Gain	MO	MOTL	MONTL
		(Delayed post-test - Pre-test)			
P8	Higher gain	100.0	2	1	1
P9	Higher gain	100.0	0	0	0
P13	Higher gain	100.0	2	2	0
P10	Higher gain	78.6	8	5	3
P14	Higher gain	60.0	1	0	1
P4	Higher gain	50.0	1	0	1
P7	Higher gain	46.4	2	1	1
P6	Lower gain	35.9	1	1	0
P3	Lower gain	33.3	2	2	0
P11	Lower gain	33.3	2	1	1
P12	Lower gain	25.0	1	1	0
P2	Lower gain	22.8	1	1	0
P1	Lower gain	15.4	0	0	0
P5	Lower gain	-12.7	6	3	3
M		49.1	2.1	1.3	0.8
SD		34.72	2.23	1.38	1.05
Mdn		41.2	1.5	1	0.5

Appendix 5.L
Spearman's rank correlation coefficients
Relationship between gains in the post-tests and MO in the treatment sessions

Variables	MO		MO with TL		MO with NTL	
	r_s	p	r_s	p	r_s	p
Gains in post-test	.200	.494	.053	.856	.157	.592
Gains in delayed post-test	.079	.789	-.089	.763	.095	.748

Note. MO = MOTL + MONT, TL = targetlike pattern, NTL = non-targetlike pattern

Appendix 5.M1
Gain scores by modifies and non-modifiers of MO

Modifiers	MO	Gains in the post-test	Gains in the delayed test	Non-modifiers	MO	Gains in the post-test	Gains in the delayed test
P3	2	19.0	33.3	P1	0	-2.3	15.4
P5	6	4.4	-12.7	P2	1	26.7	22.8
P7	2	41.7	46.4	P4	1	39.5	50.0
P8	2	78.6	100.0	P6	1	46.7	35.9
P10	8	75.0	78.6	P9	0	78.6	100.0
P11	2	55.6	33.3	P12	1	12.5	25.0
P13	2	66.7	100.0	P14	1	40.0	60.0
M	3.4	48.7	54.1	M	0.7	34.5	44.2
SD	2.51	28.43	41.23	SD	0.49	25.97	29.22
Mdn	2.0	55.6	46.4	Mdn	1.0	39.5	35.9

Note. Modifiers: above the group median of 1.5, non-modifiers: at or below group median of 1.5

Appendix 5.M 2
Gain scores by modifies and non-modifiers of MOTL

Modifiers	MOTL	Gains in the post-test	Gains in the delayed test	Non-modifiers	MOTL	Gains in the post-test	Gains in the delayed test
P3	2	19.0	33.3	P1	0	-2.3	15.4
P5	3	4.4	-12.7	P2	1	26.7	22.8
P10	5	75.0	78.6	P4	0	39.5	50.0
P13	2	66.7	100.0	P6	1	46.7	35.9
				P7	1	41.7	46.4
				P8	1	78.6	100.0
				P9	0	78.6	100.0
				P11	1	55.6	33.3
				P12	1	12.5	25.0
				P14	0	40.0	60.0
M	3.0	41.3	49.8	M	0.6	41.8	48.9
SD	1.41	34.83	50.09	SD	0.52	25.74	30.07
Mdn	2.5	42.9	56.0	Mdn	1.0	40.9	41.2

Note. Modifiers: above the group median of 1.0, non-modifiers: at or below group median of 1.0

Appendix 5.M3
Gain scores by modifies and non-modifiers of MONTL

Modifier	MONTL	Gains in the post-test	Gains in the delayed test	Non-modifier	MO with NTL	Gains in the post-test	Gains in the delayed test
P4	1	39.5	50.0	P1	0	-2.3	22.8
P5	3	4.4	-12.7	P2	0	26.7	33.3
P7	1	41.7	46.4	P3	0	19.0	50.0
P8	1	78.6	100.0	P6	0	46.7	35.9
P10	3	75.0	78.6	P9	0	78.6	100.0
P11	1	55.6	33.3	P12	0	12.5	25.0
P14	1	40.0	60.0	P13	0	66.7	100.0
M	1.6	47.8	50.8	M	0.0	35.4	52.4
SD	0.98	25.17	35.68	SD	0.00	29.62	33.67
Mdn	1.0	41.7	50.0	Mdn	0.0	26.7	35.9

Note. Modifiers: above the group median of 0.5, non-modifiers: at or below group median of 0.5

Appendices 5.N: Individual data on the patterns in the tests
Appendices 5.Ns present the individual data on non-targetlike negation patterns produced in the three tests. The order of participants in the table is sorted based on the accuracy scores in each test, placing the highest accuracy scorers at the top and the lowest at the bottom. The amount of correct use, errors, accuracy scores and the number of types are also included. The category "others" is included for indication purpose only and it is not included in the count of the number of patterns used. The value '0' for the number of each negation pattern is not given in order to visually show the distribution of patterns with reference to accuracy scores in the pre-test.

Appendix 5.N1
Individual data in pre-test, sorted according to accuracy score
(Experimental group)

Participant	correct use	errors	accuracy	number of patterns	*nai*	*masen*	*janai*	*kujanai*	*kunai*	(others)
P1	11	2	84.6	1					2	
P2	4	2	66.7	2			1		1	
P3	6	3	66.7	1	3					
P4	5	5	50.0	2			1		4	
P5	2	3	40.0	1			3			
P6	4	8	33.3	1			8			
P7	2	6	25.0	1			3			(3)
P8	0	12	0.0	2		7	5			
P9	0	9	0.0	2	3		6			
P10	0	7	0.0	2	1	5				(1)
P11	0	5	0.0	1			5			
P12	0	1	0.0	1	1					
P13	0	1	0.0	1		1				
P14	0	1	0.0	1					1	

Appendix 5.N2
Individual data in post-test, sorted according to accuracy score
(Experimental group)

Participant	correct use	errors	accuracy	number of patterns	*nai*	*masen*	*janai*	*kujanai*	*kunai*	(others)
P2	14	1	93.3	1			1			
P4	17	2	89.5	2			1		1	
P3	12	2	85.7	2			1		1	
P1	14	3	82.4	1					3	
P6	12	3	80.0	2			1		2	
P8	11	3	78.6	1					3	
P9	11	3	78.6	1					3	
P10	9	3	75.0	1					3	
P7	10	5	66.7	1					4	(1)
P13	8	4	66.7	3	2	1			1	
P11	10	8	55.6	2			3	5		
P5	4	5	44.4	1			5			
P14	2	3	40.0	2			1	2		
P12	1	7	12.5	3		1	2		4	

Appendix 5.N3

Individual data in delayed post-test, sorted according to accuracy score (Experimental group)

Participant	correct use	errors	accuracy	number of patterns	*nai*	*masen*	*janai*	*kujanai*	*kunai*	(others)
P1	16	0	100.0	0						
P3	17	0	100.0	0						
P4	17	0	100.0	0						
P8	13	0	100.0	0						
P9	13	0	100.0	0						
P13	11	0	100.0	0						
P2	17	2	89.5	1			1			(1)
P10	11	3	78.6	2	1				2	
P7	10	4	71.4	2	1				3	
P6	9	4	69.2	3			2	1	1	
P14	3	2	60.0	1				2		
P11	5	10	33.3	1			1			
P5	3	8	27.3	1			8			
P12	2	6	25.0	2		1			5	

Appendix 5.N4
Individual data in pre-test, sorted according to accuracy score (Control group)

Participant	correct use	errors	accuracy	number of patterns	*nai*	*masen*	*janai*	*kujanai*	*kunai*	(others)
P1	16	2	88.9	2			1	1		
P2	11	3	78.6	1	3					
P3	3	3	50.0	1					3	
P4	8	8	50.0	2			7	1		
P5	3	4	42.9	2			3		1	
P6	7	13	35.0	1			13			
P7	3	7	30.0	1			6			(1)
P8	1	11	8.3	3	4	1	6			
P9	1	14	6.7	2	1		1			(3)
P10	0	13	0.0	2	1		12			
P11	0	8	0.0	3	2		2		1	(3)
P12	0	4	0.0	1			2			(2)
P13	0	3	0.0	1			3			
P14	0	2	0.0	2	1			1		

Appendix 5.N5
Individual data in post-test, sorted according to accuracy score (Control group)

Participant	correct use	errors	accuracy	number of patterns	*nai*	*masen*	*janai*	*kujanai*	*kunai*	(others)
P1	20	0	100.0	0						
P2	18	0	100.0	0						
P3	5	0	100.0	0						
P7	4	0	100.0	0						
P14	11	2	84.6	0						(2)
P6	15	3	83.3	1			3			
P8	11	4	73.3	2			2		1	(1)
P5	5	2	71.4	1				1		(1)
P13	11	6	64.7	2			4		2	
P9	8	8	50.0	2			7		1	
P4	4	14	22.2	2			11		3	
P11	1	16	5.9	1			16			
P10	0	14	0.0	1			14			
P12	0	8	0.0	1			8			

Appendix 5.N6:
Individual data in delayed post-test, sorted according to accuracy score (Control group)

Participant	correct use	errors	accuracy	number of patterns	*nai*	*masen*	*janai*	*kujanai*	*kunai*	(others)
P1	20	0	100.0	0						
P2	8	0	100.0	0						
P7	5	0	100.0	0						
P6	16	1	94.1	1			1			
P14	14	1	93.3	0						(1)
P8	17	2	89.5	1			1			(1)
P5	6	1	85.7	1			1			
P3	3	1	75.0	1					1	
P13	11	6	64.7	2			1		1	(4)
P9	12	7	63.2	2			5		1	(1)
P12	7	10	41.2	3	1		6		3	
P4	7	12	36.8	2			11		1	
P11	1	16	5.9	1			16			
P10	0	17	0.0	1			17			

Appendix 5.O1
Membership and frequency nai in the three tests

Participant	Experimental Group (n = 14)			Control Group (n =14)		
	Pre-test (n = 4)	Post-test (n = 1)	Delayed-test (n = 2)	Pre-test (n = 6)	Post-test (n = 0)	Delayed-test (n = 1)
P1						
P2				3		
P3	3					
P4						
P5						
P6						
P7			1			
P8				4		
P9	3			1		
P10	1		1	1		
P11				2		
P12	1					1
P13		2				
P14				1		
Total	8	2	2	12	0	1

Appendix 5.O2
Membership and frequency of masen in the three tests

Participant	Experimental Group (n = 14)			Control Group (n =14)		
	Pre-test (n = 3)	Post-test (n = 2)	Delayed-test (n = 1)	Pre-test (n = 1)	Post-test (n = 0)	Delayed-test (n = 0)
P1						
P2						
P3						
P4						
P5						
P6						
P7						
P8	7			1		
P9						
P10	5					
P11						
P12		1	1			
P13	1	1				
P14						
Total	13	2	1	1	0	0

Appendix 5.O3
Membership and frequency of kujanai in the three tests

	Experimental Group (n = 14)			Control Group (n =14)		
Participant	Pre-test (n = 1)	Post-test (n = 2)	Delayed-test (n = 2)	Pre-test (n = 3)	Post-test (n = 1)	Delayed-test (n = 0)
P1				1		
P2						
P3						
P4				1		
P5					1	
P6			1			
P7						
P8						
P9						
P10						
P11		5				
P12						
P13						
P14	1	2	2	1		
Total	1	7	3	3	1	0

Appendix 5.P
Participants who used janai in the pre-test but did not use it in the post-test
(Experimental group)

Participant	MO	MOTL		MONTL	
P7	2	1	*kunai*	1	*kunai*
P8	2	1	*kunai*	1	*kunai*
P9	0	0	n/a	0	n/a

Note. MO = modified output in response to clarification requests, TL = targetlike pattern, NTL = non-targetlike pattern

Appendix 5.Q

Individual data of pre-test, post-test, delayed post-test and treatment sessions (Experimental group)

Participant	Test	nai	masen	janai	kajanai	kunai	others	Targetlike use	Non-targetlike use	Accuracy %	Number of types of pattern	MOTL	MONTL	Repeat	Others
1	Pre-test					2		2	11	84.6	1				
	Treatment														
	Post-test					3		3	14	82.4	1				
	Delayed-test							0	16	100.0	0				
2	Pre-test			1		1		2	4	66.7	2				
	Treatment			1R		1TL						1		1	
	Post-test			1				1	14	93.3	1				
	Delayed-test			1			1	2	17	89.5	1				
3	Pre-test	3						3	6	66.7	1				
	Treatment			3R		1TL	1TL					2		3	
	Post-test					1		2	12	85.7	2				
	Delayed-test							0	17	100.0	0				
4	Pre-test			1		4		5	5	50.0	2				
	Treatment			1NTL 1R									1	1	
	Post-test			1		1		2	17	89.5	2				
	Delayed-test							0	17	100.0	0				
5	Pre-test			3				3	2	40.0	1				
	Treatment	2NTL		3TL 1NTL 8R 1O		1R						3	3	9	1
	Post-test			5				5	4	44.4	1				
	Delayed-test			8				8	3	27.3	1				
6	Pre-test			8				8	4	33.3	1				
	Treatment			1TL 2R		1R						1		3	
	Post-test			1		2		3	12	80.0	2				
	Delayed-test			2	1	1		4	9	69.2	3				
7	Pre-test			3			3	6	2	25.0	1				
	Treatment					1TL 1NTL 1R						1	1	1	
	Post-test					4	1	5	10	66.7	1				
	Delayed-test	1				3		4	10	71.4	2				

Participant	Test	*nai*	*masen*	*jamai*	*kujanai*	*kunai*	others	Targetlike use	Non-targetlike use	Accuracy %	Number of types of pattern	MOTL	MONTL	Repeat	Others
8	Pre-test		7	5				12	0	0.0	2				
	Treatment			1R		1TL 1NTL						1	1	1	
	Post-test					3		3	11	78.6	1				
	Delayed-test							0	13	100.0	0				
9	Pre-test	3		6				9	0	0.0	2				
	Treatment														
	Post-test					3		3	11	78.6	1				
	Delayed-test							0	13	100.0	0				
10	Pre-test	1	5				1	7	0	0.0	2				
	Treatment	4TL	2R	1TL 2NTL 3R		1NTL 1O						5	3	5	1
	Post-test					3		3	9	75.0	1				
	Delayed-test	1				2		3	11	78.6	2				
11	Pre-test			5				5	0	0.0	1				
	Treatment	1R		1TL 1R 2O		1NTL						1	1	2	2
	Post-test			3	5			8	10	55.6	2				
	Delayed-test			1				10	5	33.3	1				
12	Pre-test	1						1	0	0.0	1				
	Treatment	1TL		3R 1O								1		3	1
	Post-test		1	2		4		7	1	12.5	3				
	Delayed-test		1			5		6	2	25.0	2				
13	Pre-test		1					1	0	0.0	1				
	Treatment	2R	1TL 1R 1O	1R 1O		1TL						2		4	2
	Post-test	2	1			1		4	8	66.7	3				
	Delayed-test							0	11	100.0	0				
14	Pre-test				1			1	0	0.0	1				
	Treatment			1R		1NTL								1	1
	Post-test			1	2			3	2	40.0	2				
	Delayed-test				2			2	3	60.0	1				

Note. TL = targetlike pattern, NTL = non-targetlike pattern, R = repetition of the initial non-targetlike utterance, O = others (i.e., other types of response move). Values before TL, NTL, R, or O indicates the frequency. Number of types of pattern excludes the category 'Others'

Appendix 5.R

*Response moves (MO, Repetition, Other type of response, and multiple responses)
and subsequent use (Experimental group)*

	Post-test				Delayed test			
	Non-use		Use (i.e. non-targetlike)		Non-use		Use (i.e. non-targetlike)	
	Number of instances	%	Number of instances	%	Number of instances	%	Number of instances	%
MO (n = 7)	5	71.4	2	28.6	6	85.7	1	14.3
MO only with TL (n = 5)	3	60.0	2	40.0	4	80.0	1	20.0
MO only with NTL (n= 2)	2	100.0	0	0.0	2	100.0	0	0.0
Repetition only (n = 9)	5	55.6	4	44.5	7	77.8	2	22.2
Other types of response (n = 1)	0	0.0	1	100.0	0	0.0	1	100.0
Multiple responses including MO (n = 8)	1	12.5	7	87.5	4	50.0	4	50.0

Note. MO, MO only with TL and MO only with NTL did not involve any other response move.