The Effectiveness of Recasts in L2 Question Development

Kathryn Nichols

A thesis submitted to

Auckland University of Technology

in partial fulfillment of the requirements for the degree of

Master of Arts in Applied Language Studies (MA)

2009

The School of Languages and Social Sciences

Primary Supervisor: Professor John Bitchener

TABLE OF CONTENTS

ATTESTATION OF AUTHORSHIP

I hereby declare that this submission is my own work and that, to the best of my knowledge, it contains no material written by another person, nor material which to a substantial extent has been accepted for the qualification of another degree or diploma of a university or other institution of higher learning, except where due acknowledgement has been made in the acknowledgements.

ACKNOWLEDGEMENTS

I am indebted to many whose warm support, cheerful encouragement, and genuine guidance have been deeply appreciated.

First, I would like to express my immense gratitude to Professor John Bitchener, my supervisor, who has patiently guided me throughout the research process, for his insightful comments, invaluable guidance, thoughtful suggestions and inspiration. Thanks also go to Murray Black (Head of the School of Mathematical Sciences, AUT) and Dr Peren Arin (Senior Lecturer, School of Economics, Massey University) for their invaluable advice and guidance with statistical analysis.

I am profoundly grateful to all those who participated; the wonderful students at AUT International House, who acted as subjects. I would also like to thank Michael Giacon, the Head of School and Karen Rutherford, the Deputy Head of School, who allowed me to conduct this research study at their school.

I would like to make special mention of my husband, Isa Genç, who has given me an unending supply of all sorts of help and whose support has made it possible for me to complete this study. I would also like to acknowledge my son, Emre, who was born in the middle of this exciting, post-graduate adventure and who gave me the motivation required, to get the job done.

Thanks also go to family and friends who are not mentioned here for their support and encouragement.

This research was conducted under ethics approval number 06/32 on 13th June, 2006.

ABSTRACT

This study investigated the effects of recasts, a frequently provided type of oral feedback in language teaching, on the production of English question forms in second language (L2) learning. Despite the large number of studies which have examined the effects of recasts, mixed views exist as to their usefulness as a corrective feedback technique. The present study has continued the investigation into whether their frequent use is justified and also if learners actually perceive recasts as being corrective in nature. The focus of this study is to explore developmental effects, both immediately following and delayed, of recasts while considering these two key aspects. The primary aim is to explore delayed benefits of recasts while the secondary aim is to investigate whether or not these learners perceived the feedback as a correct version of their incorrect question form. The study involved 34 English as a second language (ESL) learners from various first language (L1) backgrounds who were enrolled in a general English programme, at an Auckland language school. The participants completed demographic questionnaires and took part in task-based interaction sessions, with the members of the experimental group receiving input containing intensive recasts. The proportion and types of questions produced in the immediate and two delayed post-tests were compared and analysed to consider whether exposure to recasts was beneficial. The main findings suggest that for more advanced question types, interaction with intensive recasts may be more effective than interaction alone, for delayed L2 development, despite only weak support in the literature. On the other hand, similar positive effects on immediate question production are not strongly evident. Therefore, the results in this study add support to the claim that recasts may be beneficial for delayed second language acquisition (SLA), when the linguistic target is questions, even when these recasts do not seem effective immediately following exposure. Furthermore, awareness of negative evidence was indicated, which may have contributed, in part, to the positive delayed findings. Finally, these results suggest that recasts can be an effective pedagogical tool in the communicative L2 classroom when providing linguistic feedback in a way that does not unduly disrupt the flow of the interactive task. Practical suggestions for future research were also identified.

LIST OF TABLES

LIST OF FIGURES

ABBREVIATIONS

AUTEC	Auckland University of Technology Ethics Committee
AUT IH	Auckland University of Technology, International House
ESL	English as a second language
L1	first language
L2	second language
NS(s)	native speaker(s)
NNS(s)	non-native speaker(s)
Max	highest number
Min	smallest number
SLA	second language acquisition
Std. Dev.	Standard deviation
UG	universal grammar

CHAPTER 1

INTRODUCTION

1.1   Background of the Study

The importance of the role of interaction in second language acquisition (SLA) has become a popular topic over the past three decades. It was initially on account of Hatch (1978) and her two seminal papers, on the benefits of taking part in communicative interaction for L2 learners, which led to this recognition. Previously, it was believed that firstly, the linguistic aspects of the L2 had to be learned before the language could be used communicatively. This early research was soon followed by Long (1980, 1981, 1983) and his ground-breaking exploration into how conversational and linguistic modifications which occur naturally, during conversational interaction, can lead to language acquisition. In his updated Interaction Hypothesis, Long (1996) refers to this type of interaction as negotiation of meaning and he suggests that the input provided at this time, facilitates learning. Another prominent researcher in this field, Gass (1997), claims that interaction should not only be seen as an opportunity for learners to practise the second language (L2) but as "the means by which learning takes place" (p. 104). Since this early research into the benefits of interaction, there has been opposition to the value of negotiation of meaning for L2 learning (Skehan, 1998). However, negotiation and its emphasis on message comprehension have continued to generate and sustain a great deal of interest in the field of SLA.

Within Long's (1996) Interaction Hypothesis, he suggests that feedback, in particular, implicit negative feedback (i.e. feedback that does not explicitly point out the error to the learner), which occurs during interaction, can lead to L2 development. Furthermore, exposure to negative feedback may lead to opportunities to produce more target-like second language (L2) speech (Mackey & Philp, 1998). The role of feedback in L2 classrooms is based on the view that learners benefit from information about the success

of their attempts to communicate in the target language (Long, 1977). A number of types of negative feedback exist, however, it is recasts (a correct L2 version provided in response to a learner error), which seem to be the most frequently provided in the L2 classroom (Lyster, 1997). Many definitions of the term 'recast' exist in the literature and at times it can be problematic as to whether they refer to the same type of corrective feedback (feedback which offers a correction to a learner error). Therefore, a general definition is often adopted. The current study utilised Ellis and Sheen's (2006) definition, as it was appropriate to the aim of using recasts in a way that allows a focus on linguistic form, without unduly disrupting the conversational flow of the interaction:

> "A recast consists of the teacher's reformulation of all or part of a student's utterance that contains at least one error within the context of a communicative activity" (Ellis & Sheen, 2006, p. 580).

As the field of SLA research has continued to grow, both researchers and teachers have become more and more intrigued by the notion of which types of conversational interaction lead to what kind of development (Mackey & Philp, 1998). Recasts have been a popular topic of interactional work within the SLA literature for more than two decades (e.g. Long, 1996; Long, Inagaki & Ortega, 1998; Lyster & Ranta, 1997; Mackey & Philp, 1998) and make up a high percentage of the corrective feedback offered by teachers in many meaning-oriented language classrooms (Ellis, Basturkmen & Loewen, 2001; Loewen, 2002, 2004; Lyster & Ranta, 1997; Oliver, 2000; Panova & Lyster, 2002). Why recasts seem to occur so often in the L2 classroom may be partly because they are quick to provide and do not unduly interrupt the communicative flow in the classroom. In addition, Long and Robinson (1998) claim that recasts may draw learners' attention to the inconsistency between their incorrect utterances and the target language forms. Long and Robinson's suggestion, coupled with their frequent usage, would suggest that recasts could be an effective means of feedback. However, research findings indicate mixed views and debate continues as to whether their frequent use, as a popular type of negative feedback, is effective for SLA. The current study sought to add to the literature in this research field, with recasts and their effects, especially delayed ones, as the focus of this

investigation. The relationship between this type of teacher feedback and SLA is examined with the production of question forms being the linguistic target.

1.2   Aims of the Research

The primary aim of this study is to investigate the relationship between exposure to recasts and the production of questions in task-based interaction. Delayed benefits of exposure to recasts are the main focus of the current study. At one week and 12 weeks following treatment post-tests were undertaken to look at whether recasts can be effective, even though immediate benefits may not be apparent. Some researchers claim that an immediate learner response following a recast is not always necessary for SLA to occur (Gass & Varonis, 1994; Lyster & Izquierdo, 2009; Mackey, 1999; Mackey & Philp, 1998; McDonough & Mackey, 2006; Ohta, 2000) and that recasts may still have an effect on L2 development (Lightbown, 1998; Mackey; Mackey & Philp). However, despite these claims, more studies which explore both short term and longer term recast effects are required. The reason for this is that debate surrounding the effectiveness of recasts in the classroom continues, mainly due to the following two key issues: is the teacher's high usage of recasts in the classroom justified and are learners aware of the fact that recasts contain negative evidence (indicating what is grammatically incorrect in an L2) and therefore offer corrections to their linguistic error? The latter area of concern is related to the secondary goal of the study, namely, to investigate whether L2 learners perceive the corrective nature of recasts, offered when an error has been made. In other words, do they realise that the recasts contained negative evidence? An alternative claim is that they may believe that the recast is simply confirmation of meaning (e.g. Mackey, Gass & McDonough, 2000). The current study aims to provide empirical evidence to add further input to the literature, regarding these two central aspects of interaction and recasts.

This research was carried out at a university language school in Auckland, AUT University, International House (AUT IH), with 34 intermediate level ESL students from a range of ethnic backgrounds. The study utilised a quasi-experimental approach (when an independent variable is manipulated in order to explore the influence this might have on a dependent variable) and involved an experimental group and a control group. A pre-test/post-test design was adopted in order to collect data over a four month period (from pre-test to the final delayed post-test – 12 weeks following treatment). In addition, a student questionnaire was administered to collect demographic information about the participants. Lastly, a post-data collection email was sent to the experimental group participants, to investigate their perceptions of the recasts they were exposed to during the treatment sessions. In other words, did they realise that these recasts were providing corrective feedback in response to question formation errors?

This study contributes to an understanding of the role played by negative feedback, namely, recasts, in L2 classrooms with a primary focus on delayed benefits of this type of feedback for learners. While a number of studies related to the present research have been undertaken in the past (such as, Mackey, 1999; Mackey & Philp, 1998), most longitudinal investigations of negative feedback and interaction administer a final post-test at three to four weeks following the treatment sessions. Therefore, studies of longer duration are lacking in this field (Long, Inagaki & Ortega, 1998; McDonough, 2005). The main advantage of conducting research over a longer term, according to Mackey and Gass (2005) is that the researcher is able to achieve results which show "a wider snapshot of treatment effects" (p. 149). Consequently, due to the fact that, in reality, language learning happens over a long period of time, these kinds of results may be more meaningful than those from data collected over a shorter research duration. The findings of this study could therefore be of benefit in L2 instruction, if they can convince teachers of the importance of providing learners with recasts when an error occurs, even though an immediate response to the recast may not be apparent.

1.3   Organisation of the Study

This thesis consists of five chapters. Following this introduction, Chapter 2 reviews previous literature and research that relates to the research questions addressed in the current study. It also introduces fundamental theoretical claims which have promoted interaction within the L2 classroom, as well as discussing relevant aspects of meaning negotiation and recasts. Major findings from empirical research studies concerning interaction and the use of recasts as corrective feedback are reviewed and gaps in research to date are subsequently identified. Finally, the research questions are raised for investigation.

Chapter 3 describes the methodological approach adopted in the current study. This research was aiming to uncover causal links between variables in order to answer the research questions. Specifically, whether there is a link between the use of recasts in the SLA classroom and delayed development of question forms. In order to generate data which could be interpreted from a longitudinal perspective, delayed post-tests were utilised in the quasi-experimental study design. Justification for this approach is provided. The major research instruments – the student demographic survey, pre- and post-tests and a post-data collection email – are identified and the procedures followed in collecting and analysing data are stated.

Main findings from an analysis of the research data are presented and interpreted in Chapter 4. These results are based on quantitative research techniques. In addition, findings are presented visually in graphs and tables in order to complement the analysis. This chapter is structured around answering the three research questions with reference to previous, related research results.

Chapter 5 summarises the key findings and focuses on both pedagogical and research implications of the study. Furthermore, limitations are indicated and finally, recommendations for future research are discussed.

CHAPTER 2

A REVIEW OF THE LITERATURE

2.1 Introduction

In order to gain a perspective on how the role of recasts fits into a broader theory of SLA, it is important to first introduce the theoretical model within which this research lies. Many prominent theories of SLA exist, including universalist theories, based on Chomsky's Universal Grammar, sociolinguistic models and socio-cultural models, based on the work of Vygotsky. However, for the purpose of this research, it is a cognitive approach to SLA which is the main focus. Gass's (1997) psycholinguistically motivated, cognitive model of SLA endeavours to conceptualise what a learner does when moving from L2 input exposure to the production stage of output. In particular, it is how classroom instruction and interaction, especially the role of feedback, may impact on the SLA process, which is of interest here.

Many pedagogical approaches to classroom instruction exist. This research is concerned with a communicative approach, where an emphasis on focus on form techniques is important. Interaction in the classroom and particularly, negotiation of meaning, as originally proposed by Long (1996), are considered vital to the attainment of mutual understanding between interlocutors and consequently, SLA. In the process of negotiation for meaning, feedback can take many forms. Among the different types of feedback, this study is interested in negative feedback as an approach to error correction. With negative feedback, the error is not intentionally pointed out to the learner, but is corrected in an unobtrusive way, such as by offering a recast, so as to try and maintain the communicative nature of the interaction. Recasts and the role they may or may not play in SLA have been a popular topic of interactional work in the SLA literature for more than two decades. Recasts are still being investigated as mixed views exist

regarding their effectiveness in the L2 classroom. Research to date suggests that learners may not use recasts very often immediately following exposure. However, the use of recasts beyond this initial exposure is an area of research that is largely unexplored (with the exception of Mackey & Philp, 1998 and Oliver & Mackey, 2003). This study focuses on that area and explores the use of recasts through production in delayed post-tests.

In this chapter, recasts are discussed and references made to Long and his well known Interactionist theory as well as to Gass (1997) and her theoretical model of SLA. Within this model, Gass argues that when learners are involved in meaning negotiation, they are exposed to useable input and are then led to notice a gap between their current L2 production and target linguistic forms. This noticing can lead to restructuring of learners' own linguistic forms and in turn, SLA.

Despite the large amount of research which has investigated recasts, debate continues as to whether their use, as a popular type of negative feedback, is effective for SLA. Results have indicated both support for and opposition to, both immediate and more delayed effects on learning. However, research involving delayed post-testing of longer than a few weeks is in short supply in the literature. Therefore, although immediate effects will also be explored, the primary focus of this study is the longer term benefits of using recasts as a form of negative feedback, within interaction in the communicative second language classroom.

In order to gain an understanding about the way in which negative feedback and in particular, recasts, may be of benefit to SLA, it is important to first present some background information concerning the ways in which second languages are learned. Therefore, the following sections will introduce and discuss two main positions on how L2 learning occurs and the ways in which they are purported to influence language learning.

2.2   Second Language Learning – Nature vs. Nurture

In SLA literature two positions of thought on how learning occurs are commonly referred to as nature and nurture. The first position refers to the possibility that both L1 and L2 learners can learn a language by means of innate (natural and instinctive) knowledge and the second position claims that language development is influenced and formed by the environment, "that is the interactions in which learners engage" (Gass, 2005. p. 225). It is these interactions which involve the central focus of this study, in particular, corrective feedback and recasts, and therefore the nurture position is considered to be of the most relevance here. However, in order to be able to contrast the two positions, the nature (innatist) position will be looked at briefly first, before turning to the nurture (interactionist) position.

2.2.1   Nature – the Innatist Position

In L1 acquisition, there is a lot of support for the existence of an innate linguistic knowledge system, which allows children to develop the grammar and other language features they require. This is known as the innatist position. Chomsky (1959) was the first to suggest that all children are biologically programmed to learn their L1 with little importance being placed on the environment, except to provide the people who communicate with the child. For Chomsky, language acquisition occurs naturally and there is no need for children to be formally taught. Proponents of the innatist position, refer to a child's innate ability as Universal Grammar (UG). UG is believed to consist of a set of principles which are common to all languages (Lightbown & Spada, 2006). White (2005) states that without these specific innate linguistic principles learning would not be possible.

This brings us to the question of whether L2 learners are also able to access UG. The UG position for SLA is less clear than for L1 acquisition (Gass, 2005). However, as with L1

acquisition, supporters of UG claim that grammatical properties of the L2 (i.e.the linguistic input a learner is exposed to) will interact with the innate UG system an individual brings to the language learning situation, in order to produce the grammatical structures of the L2.

For UG theorists the heart of linguistic knowledge is innate and exists independently of environmental input. Supporters of this position have argued that negative evidence cannot play an important role in language acquisition, as errors are considered simply to be a result of L1 influence (Adjemian, 1976; Corder, 1967; Nemser, 1971; Selinker, 1972). Conversely, social theorists consider that learner language is formed by and develops from, the surrounding environment and that this environment plays a crucial role in the language learning process. This brings us to the second position of language learning; the nurture or interactionist position, which assigns a key role to negative feedback, with recasts being one of the most frequent types. Therefore, this position is considered of great relevance in the current study.

2.2.2   Nurture - the Interactionist Position

For social theorists the focus is on the role which environmental language (input, output, interaction and feedback) plays in promoting SLA. This interactionist perspective has led to extensive empirical work examining the relationship between environmental language and the learners' developing L2 knowledge (their interlanguage systems). The focus of the interactionist theory of language teaching and learning is to provide access to meaningful and comprehensible input through conversational interactions. Within these interactions, corrective feedback indicates to learners that their speech is different to the norms of the target language and consequently initiates a sequence of negotiation, in order to try and solve the communication problem. Supporters of the interactionist view believe that this negotiation provides vital input and facilitates learning. The interactionist position will be discussed in more detail throughout the following sections, in order to illustrate where recasts fit into the L2 learning process. These recasts are a commonly used negative

feedback technique within interaction between non-native speakers (NNSs) and native speakers (NSs), First of all, the three language learning conditions (input, output and feedback) considered vital for SLA are introduced, leading into a wider discussion of interaction and its role in this process.

2.3 Language Learning Conditions

Three conditions argued to be central to SLA are; access to comprehensible input, opportunities to produce output of the language and exposure to feedback on learners' performance, particularly negative feedback (Gass, Mackey & Pica, 1998).

When learning a language "…it is an incontrovertible fact that some sort of input is essential …; clearly languages are not and cannot be learned in a vacuum" (Gass, 1997, p. 86). Krashen (1985), a leading pioneer in the research field of linguistic input, proposed that it was access to comprehensible input which leads to language acquisition. In his input hypothesis he defined comprehensible input as containing linguistic forms slightly in advance of a learner's current interlanguage system. However, as well as exposure to input, learners must have the opportunities to use and produce the language in order to acquire it (Swain, 1985, 1995).

Language acquisition is a process and as learners experiment with speaking and writing the L2 themselves, they can learn from their own output, provided the third key language learning condition is offered. Corrective feedback can lead to development of learners' interlanguage systems. In particular, negative feedback can alert the learner to the fact that linguistic modification of their output is required. The role of feedback within SLA is central to the success of interaction. It is discussed more fully in section 2.4.2, following a look at previous interaction research which indicates where the current study fits into the wider field. Early interaction research and the role of conversation within SLA provide

background information relevant to this study and will be examined in the next section (2.4).

2.4   Early Research into Interaction

Wagner-Gough and Hatch (1975) were among the first second language researchers to consider the role of conversation in second language learning. SLA theory had not focused on the importance of interaction before this (Pica, 1994). Sociolinguists had previously looked at interaction to discover links with social roles and features of NS and NNS relationships. However, Hatch (1978) brought the issue of interaction in language learning to the forefront of SLA research with her seminal papers, suggesting that interaction could also be the starting point for examining linguistic and cognitive aspects of the learning process, not just the social ones. This view set off intense debate and discussion, so much so, that her papers are still widely referred to today (Pica, 1994). Hatch appealed to SLA researchers to consider how the learning of L2 structure grows out of communication, rather than previous assumptions that L2 learning leads to communicative use. Long (1980, 1981, 1983) soon followed with his pioneering work, which he initially called interactional modification (referring to the social discourse that NSs and NNSs undertake in order to avoid and repair comprehension difficulties in their conversations) before it became known as the Interaction Hypothesis.

2.4.1   The Interaction Hypothesis

Communication in any language as either an L1 or L2 requires interaction between people. The interaction approach was stimulated by this communicative pressure and has become a widely researched area of SLA studies. Long (1983, 1996) proposes that interaction facilitates language acquisition because of the conversational and linguistic modifications

that naturally occur during interaction. He suggests that this negotiation of meaning provides the input which learners require to learn a language successfully. This is a direct challenge to Krashen's (1977) Input Hypothesis which maintained that the sole requirement for SLA was exposure to comprehensible input. In his initial interaction hypothesis, Long emphasized that in addition to input, participation in interaction and meaning negotiation are essential for facilitating comprehension, which in turn fosters development (Ellis, 2008; Leeman, 2003).

Long's theory has continued to guide and develop interaction research to the present day. He proposed that with NS/NNS conversations a greater amount of interactional modification exists than in NS/NS conversations. That is to say, that the interactional structure itself is worthy of examination. Much of the body of research that has followed on from this early work by Hatch and Long has concentrated on 'negotiation' (renaming Long's original 'interactional modification'). This term refers to "the modification and restructuring of interaction that occurs when learners and their interlocutors anticipate, perceive, or experience difficulties in message comprehensibility" (Pica, 1994, p. 494).

Other ways of modifying or restructuring interaction in L2 learning and their possible role in the learning process have also been studied, such as interrupting the flow of conversation by a correction or changing to a new topic. However, despite opposition to the usefulness of negotiation of meaning for L2 learning (Skehan, 1998), it is negotiation and its emphasis on achieving message comprehension which has generated and maintained considerably more interest in the field of SLA (Pica, 1994).

Consequent research which has followed on from that of Hatch and Long, has explored the role in which negotiated interaction plays in SLA. Strong support exists for Hatch's and Long's proposal that conversation is the means by which learning takes place (Gass & Varonis, 1985, 1989; Pica & Doughty, 1985; Pica, Young & Doughty, 1987). The emphasis has been primarily on the role that negotiated interaction between native and non-native speakers and between two NNSs plays in the development of a second

language, rather than considering conversational interaction purely as an opportunity for practice of specific features of the L2.

This notion has since been expressed in Long's updated Interaction Hypothesis (1996) which stresses the importance of negotiation for meaning and suggests that implicit negative feedback, such as clarification requests and recasts, which occurs during negotiation may be facilitative of L2 development; "Negative feedback obtained during negotiation work or elsewhere may be facilitative of L2 development, at least for vocabulary, morphology, and language-specific syntax, and essential for learning certain specifiable L1-L2 contrasts" (Long, 1996, p. 414). Pica (1994) and Gass (1997) have made similar claims. In fact, Gass' theoretical model explaining the process of SLA also looks at L2 learning from the interaction perspective. Gass focuses on the various discourse moves involved in negotiation of meaning throughout her model. This model of SLA is based on Varonis and Gass (1985) and attempts to conceptualize what a learner does when moving from exposure to L2 input to the production stage of output. Following on from the Interaction Hypothesis (Long, 1996; Long & Robinson, 1998) Gass claims that in current SLA models, interaction is seen as not only an opportunity for learners to practice the L2 but as "the means by which learning takes place" (p. 104).

Research into the ways in which interaction and in particular, negative feedback, may benefit SLA is robust and some of this research is discussed in the following section, in order to indicate how the current study is placed in the field of interaction in L2 learning.

2.4.2   The Role of Feedback within Interaction

Past research into interaction has placed an emphasis on the effects of interaction on L2 development (Ellis & He, 1999; Ellis, Tanaka & Yamazaki, 1994; Gass & Varonis, 1994; Loschky, 1994; Mackey, 1999; Polio & Gass, 1998). Results have shown benefits for participation in interaction, in terms of lexical acquisition (Ellis & He; Ellis et al.;

Loschky), improved native speaker comprehension of consequent production (Gass & Varonis; Polio & Gass) and the use of forms linked with more advanced developmental stages (Mackey). These studies are important to this field of research, in that they provide evidence of a direct correlation between interaction and development, however they do not clarify how or which interactional features are favorable (Leeman, 2003). More recent thinking has regarded participation in interaction as being of benefit, because it can lead to the negotiation of meaning, provide exposure to negative feedback and offer opportunities for modified language production (e.g. Gass, 1997; Long, 1996; Pica, 1992, 1994; Swain, 1985, 1995).

This brings us to the relationship between feedback and interaction and how together, they may affect learning. The value of feedback is supported by a number of SLA researchers who suggest that as long as learners are developmentally ready to receive the negative feedback, they will utilize these opportunities to produce more target-like utterances (Mackey & Philp, 1998; Oliver, 1995). In his updated Interaction Hypothesis, Long (1996) emphasises that interaction between native speakers and non-native speakers such as, negotiation involving implicit negative feedback and the resulting modified output, could serve a role in providing the comprehensible input needed for successful L2 learning. A major reason for this is that such feedback may draw learners' attention to mismatches between input and output. Recasts are incorporated into this research, as negative feedback, in order to alert learners to their question errors.

Observational studies investigating interaction between NSs and NNSs or NNSs with other NNSs can offer some understanding as to the various ways in which feedback is provided, as well as how learners respond to different types of feedback. Evidence now exists that NNSs modify their output in response to various signals from their partners and that corrections are often incorporated into their subsequent turns (Gass & Varonis, 1989; Lyster, 1998; Lyster & Ranta, 1997; Mackey, Oliver & Leeman, 2003; Oliver, 1995, 2000; Pica, 1992, 1994, 1996; Pica, Holliday, Lewis & Morgenthaler, 1989). However, a lack of developmental measures in these studies, such as pre- and post-tests contribute to difficulty in firmly establishing a direct relationship between any specific discourse structure and

acquisition (Leeman, 2003). Furthermore, when interactional features are looked at only in terms of immediate NNS responses there is no opportunity to observe delayed effects of feedback (Gass, 1997; Mackey, 1999; Mackey & Philp, 1998).

In contrast to these observational studies, experimental research which focuses on specific interactional features and feedback types makes it possible to explore individual effects of distinct features, such as recasts and clarification requests and their impact on L2 development. A number of such studies have suggested that recasts can have beneficial effects on learner performance in successive language tasks (Doughty & Varela, 1998; Long, Inagaki & Ortega, 1998; Mackey & Philp, 1998). Despite the important contribution to SLA that these studies represent concerning benefits of interaction and feedback, they do not yet provide clear indications for the role of negative evidence (Leeman, 2003). The current study hypothesises that through exposure to negative evidence (recasts) a positive effect on question production would result.

Further support for learning arising from negotiated interaction is available. Positive effects for negotiated interaction on production have been suggested by a number of researchers (Ellis, Tanaka & Yamazaki, 1994; Gass & Varonis, 1989; Mackey, 1999; Polio & Gass, 1998; Swain & Lapkin, 1998). Modifications to speech, which occur during interaction, have been argued to provide learners with implicit feedback on their own interlanguage production at a crucial time for learning; when mismatches occur between their utterances and the correct L2 feature (Gass 1997; Long 1996; Pica 1994; Swain 1995). Recently, researchers have tried to make the link between learning outcomes and interaction more explicit and direct (Gass, 2005). However, problems exist with this task, as the total amount of input a learner has been exposed to, as well as the extent of interactions the learner has participated in, are difficult to observe. This was also an issue in the current study and is referred to in Chapter 4. As mentioned previously, feedback is offered when mismatches occur between an L2 learner's utterance and the target L2 grammar. This feedback on errors explores linguistic form and it is important to look at how a focus on this form may be possible within the communicative nature of the L2 classroom of today. Therefore, section 2.5 presents a discussion of focus on form and

interaction. Traditional pedagogical techniques of focus on form are contrasted with more current L2 methods of instruction, which promote the benefits of teaching linguistic form in an interactive way.

2.5   Focus on Form and Interaction

Traditionally, grammatical forms were presented and practiced in isolation, one at a time (Ellis, 2003, 2008). Formal instruction provided input and negotiated interaction played little part in the learning process. Errors generated through natural communication were not considered to show progress, but rather were thought of as bad habits, which needed to be prevented before they occurred (Ellis, 1985; Lightbown & Spada, 2006). L2 teaching methods today adopt quite a different approach. The focus is on communicative classrooms, with meaning-based instruction, involving conversational interaction and opportunities to develop communication skills. Errors are considered as progress as they can indicate a discrepancy between the L2 learner's speech and the target language. A question of interest within this meaning-oriented instruction method is how can learners' attention be focused on the form of the language? Furthermore, should grammar be formally taught at all (Ellis, 2006)?

The role which form-focussed instruction (attention to the formal aspects of language, such as grammar, spelling and intonation) should play when teaching a language has long been a major issue in SLA. Form-focussed instruction has been argued to only be beneficial for beginner learners (Krashen, 1985; Long, 1980). Supporters of this view contend that form-focussed instruction can only affect the acquisition of simple structures as explicit knowledge and not the learning of implicit knowledge (Krashen, 1981, 1982, 1993). However, this claim was refuted by Ellis (1985), who argues that research does not show support and that for more advanced learners, this type of instruction continues to be useful. In addition, he suggests that instruction focusing on meaning will assist language

acquisition but he claims that the rate at which a learner will develop can be greatly enhanced with form-focussed instruction.

Interactionist supporters suggest that during conversation, learners are required to negotiate meaning in order to arrive at a mutual understanding of the intended message. For Long (1996), this negotiation is considered particularly beneficial for achieving meaning and reaching mutual understanding. Gass (1997) expands this view by suggesting that this negotiation includes both form and meaning with attempts on both sides to solve communication difficulties, rather than simply taking part in a "free-flowing exchange of information" (p. 107). Gass' model accounts for negotiation of meaning, suggesting that interaction is composed of a trigger and a resolution.

The trigger consists of 'the stimulus' for the subsequent negotiation, while the resolution "is an attempt to resolve the perceived difficulty" (p. 112). Gass (1997) then divides the resolution into the indicator and the response, with the indicator alerting an interlocutor to the fact that a problem exists. A response will generally follow the indicator, when the first speaker can acknowledge and address the problem. Finally, a reaction to the response can be made. The following exert from Pica (1994) illustrates this model.

NNS:	The windows are crozed.	trigger
NS:	The windows have what?	indicator
NNS:	Closed.	response/trigger
NS:	Crossed? I'm not sure what you're saying here.	indicator
NNS:	Windows are closed.	response
NS:	Oh, the windows are closed, oh, OK, sorry.	reaction to response

Gass (1997) suggests that this type of interaction benefits learners by providing useable input and helping them to notice a gap between their L2 production and the target-like forms. Gass refers to this noticing process as apperception where "apperception is an internal cognitive act in which a linguistic form is related to some bit of existing knowledge (or gap in knowledge)" (p. 4). Gass argues that once learners have acquired

this information, they realize that a gap exists, which needs to be filled and can then begin restructuring linguistic forms in the interlanguage.

Once the trigger has occurred, which in the case of this study is an incorrectly produced question form, it was followed by a response move from the NS (the researcher) for learners in the experimental group. Many studies have examined what happens in this move. Lyster (1998) and Lyster and Ranta (1997) conducted a series of studies of primary-age school children in French immersion classes in Canada and after examining the main feedback moves possible after a linguistic error, Lyster proposed the following categories:

1. Explicit correction; the teacher points out what was incorrect and supplies the correct form
2. Recasts; the teacher implicitly reformulates the student's utterance, or part thereof
3. Elicitation; the teacher prompts the student to reformulate their utterance
4. Metalinguistic clues; the teacher makes comments, provides information or asks questions regarding how well-formed the student's utterance is
5. Clarification requests; the teacher uses phrases such as 'I don't understand'
6. Repetition; the teacher repeats the ill-formed utterance, adjusting intonation to emphasise the error

Lyster (1998) found that recasts made up more than half of all feedback moves, following student errors at 55%, followed by elicitations (14%), clarification requests (11%), metalinguistic clues (8%), explicit correction (7%) and repetition (5%).

Problems arise incidentally during language use in the classroom and when focus is applied to form at this time, connections between form and meaning can be facilitated (Doughty & Williams, 1998; Long, 1988, 1991; Long & Robinson, 1998). This timing is important as the learner wants to know if they have communicated the message successfully. Therefore, a repair on form should happen at the time when a mismatch between the learner's

interlanguage form and the target language form is most apparent, rather than postponing attention to form until the next language lesson; a more traditional approach. Furthermore, learners can benefit from, and may require, focus on form to overcome incorrect or incomplete knowledge of specific target language features (Lightbown & Pienemann, 1993; Lyster, 1998; White, Spada, Lightbown & Ranta, 1991). In the case of the current study for the experimental group, the focus on form response move following an incorrect question form was provided in the form of a recast.

This brings us to the role that recasts may play in the L2 learning process. The following sections of this chapter will discuss the literature on recasts and how this type of corrective feedback may or may not affect second language learning. The language learning conditions discussed above and their role within form-focussed learning will be further explored with regards to corrective feedback and more specifically, recasts.

2.6   Recasts Defined

Recasts are one type of corrective feedback and are considered by some researchers to be implicit (implied although not directly expressed) negative feedback, providing negative evidence to the learner. Long (1996) and Long and Robinson (1998) view recasts as implicit and claim that because of this implicitness they aid acquisition. However, Nicholas, Lightbown and Spada (2001) point out that while it is clear that recasts provide positive evidence (what is grammatically acceptable in an L2) it is less clear as to whether or not they also expose learners to negative evidence, as learners may be unaware of the corrective intention of the recast. This view is shared by other researchers and will be discussed more fully in a later section (2.7.1).

Recasts are intended to provide a focus on linguistic form while not interrupting the flow of conversation in meaning-based interaction. They differ from other types of focus-on-form procedures which are explicit, such as overt error correction where the error is

pointed out to the learner before being corrected. Gass (2003) refers to recasts as "those instances in which an interlocutor rephrases an incorrect utterance with a corrected version, while maintaining the integrity of the original meaning" (p. 239).

Long (2006) offered the following recast definition:

> A corrective recast may be defined as a reformulation of all or part of a learner's immediately preceding utterance in which one or more nontarget-like (lexical, grammatical, etc.) items is/are replaced by the corresponding target language form(s), and where, throughout the exchange, the focus of the interlocutors is on *meaning,* not language as object. (Long, 2006, p. 77)

What is interesting about Long's (2006) definition is that to qualify as a recast, the reformulation of the learner's nontarget-like utterance must occur in the context of message-centered communication. According to Long, a recast motivated by a grammatical error rather than a message breakdown would not be considered as a recast, as it would not represent an attempt to solve a communication problem.

Many definitions of the term 'recast' have been offered and it is a point of contention that because various definitions exist, it can be doubtful as to whether they all refer to the same type of feedback. For this reason it can also be difficult to compare the results of L2 studies of recasts, as on many occasions these studies are not actually looking at the same thing. An attempt to overcome this confusion is that researchers will often operate with a general definition of recasts (e.g. Braidi, 2002; Long, 1996, 2006; Lyster & Ranta, 1997; Sheen, 2006). In the current study, the purpose of using a recast is to allow a focus-on-form without overtly interrupting the communicative flow of the task or conversation. With this in mind, Long's (2006) definition above does not fit the context of this study and therefore, perhaps the best definition to use is a general one. Therefore, this study will operate using Ellis and Sheen's (2006) definition;

> "A recast consists of the teacher's reformulation of all or part of a student's utterance that contains at least one error within the context of a communicative activity" (Ellis & Sheen, 2006, p. 580).

This type of approach recognises that the recast needs to occur within the context of an effort to communicate (as opposed to some focus on forms activities) but does not attempt to establish whether the recast is pedagogically or communicatively motivated.

2.6.1   Types of Recasts

As well as being difficult to define, recasts are not simple concepts to classify. In fact, Ellis and Sheen (2006) describe them as "chameleonlike" (p. 579). This is because many types of recasts exist. Some researchers have made an attempt to subcategorize them in relation to their differentiating characteristics (Braidi, 2002; Loewen & Philp, 2006; Philp, 2003; Sheen, 2006) whereas others have not (Lyster & Ranta, 1997; Mackey, 1999; Mackey & Philp, 1998).

Recasts are not pre-emptive (occur before a learner error). They serve as a reaction and are given in response to an utterance initiated by a learner within a communicative context. The interlocutor, when providing a recast following a learner's nontarget-like form, may be simply correcting the form of the utterance or attempting to clarify the meaning as well as the form. In either case the learner is exposed to a target language sample of what the interlocutor believes to be the intended message. The literature on recasts is abounding with a host of terms to describe the different kinds of recasts found to occur. Farrar (1992) in a study of child language acquisition suggested that they can be corrective or non-corrective. Corrective recasts were distinguished as those that correct a target error and non-corrective ones were those which provide a model of a target form. In SLA research, Lyster and Ranta (1997) utilised the same terms but defined a non-corrective recast as a reformulation of a learner's error-free utterances.

A further potentially important distinction is that between full and partial recasts. Full recasts comprise a reformulation of the whole erroneous utterance. An example of which

comes from Mackey and Philp's (1998) study of negotiated interaction on the production and development of question forms in ESL;

    NNS: what are they (.) what do they do in your picture?
    NS: what are they doing in my picture?   (p. 342)

In contrast to full recasts, partial recasts only involve part of the utterance containing the error. Furthermore, recasts can differ as to whether they are simple or complex (Nelson, Denninger, Bonvillian, Kaplan & Blake, 1983), which depends on whether the changes to the learner's incorrect utterance are minimal or substantial and on the nature of the change. The change may require a substitution of the flawed form, an addition, a deletion, or a reordering of the initial utterance. Finally, intensive recasts, those which are focused repeatedly on the same linguistic feature seem to be effective for acquisition, especially if these are enhanced in some way, such as accompanied by means of emphatic stress as in the case of the studies of Doughty and Varela (1998) and Han (2002). The current study provides full, corrective recasts, which are intensive, as they focus repeatedly on errors related to question formation. These recasts may be simple or complex depending on the number of changes the learner's original utterance requires.

This brings us to the way in which this research will fit into the existing body of work already undertaken in the field. This issue will be discussed below after first looking at some previous studies, which have investigated the role of recasts in SLA.

2.7   The Role of Recasts in SLA

Children seem to acquire their L1 without explicit instruction. A question of special interest in SLA is whether L2 learners can do the same. As the field of SLA research has continued to grow, both researchers and teachers have become more and more intrigued by the notion of which types of conversational interaction lead to what sort of development (Mackey & Philp, 1998). The benefits of interaction in SLA may be due to

the implicit negative feedback that occurs during negotiation of meaning, once a communication breakdown has occurred (Long, 1996). It is proposed that this corrective feedback contains data for language learning.

Recasts, being a frequent type of corrective feedback play a central role in claims about interaction-driven L2 development (Carpenter, Jeon, MacGregor & Mackey, 2006). Researchers were initially interested in investigating whether recasts may be effective in language learning. A large number of empirical studies in both experimental and classroom contexts have suggested that recasts can facilitate short-term L2 learning (Ayoun, 2001; Braidi, 2002; Doughty & Varela, 1998; Han, 2002; Havranek, 2002; Iwashita, 2003; Leeman, 2003; Long, Inagaki & Ortega, 1998; Mackey & Philp, 1998; Muranoi, 2000; Oliver, 1995, 2000; Philp, 2003). In addition, it has been suggested that exposure to recasts may lead to a more delayed effect on development (Ellis, Loewen & Erlam, 2006; Gass, 1997; Gass & Varonis, 1994; Lightbown, 1998; Mackey & Philp, 1998; Mackey, 1999). Recasts and delayed outcomes has become a hotly debated topic in the literature and will be discussed in more detail in a later section (2.9).

Following on from research which looked at whether recasts may be of benefit in SLA, researchers became interested in taking a more detailed look at how recasts impact L2 development, focusing on factors such as:

- how learners distinguish recasts from non-corrective repetitions in the discourse (Lyster, 1998; Lyster & Ranta, 1997; Panova & Lyster, 2002)
- various linguistic features (Han, 2002; Mackey, Gass & McDonough, 2000; Philp, 2003)
- the variable effects of developmental factors in recognising recasts (Mackey, 1999; Mackey, Philp, Egi, Fujii & Tatsumi, 2002; Philp, 2003)
- different instructional contexts and settings (Iwashita, 2003; Oliver & Mackey, 2003; Pica, 2002; Sheen, 2004)

This study aims to expand on research which has been conducted on linguistic features, by investigating the effects of recasts on question formation, as well as looking at developmental readiness as an important factor.

Despite the vast interest in recasts in L2 learning and the large number of experimental and observational studies undertaken to date, two issues remain unresolved. The first is that it is still not possible to draw conclusions on whether or not recasts' contribution to language development is as beneficial as their frequent classroom usage would suggest (Lyster, 1998; Lyster & Ranta, 1997; Nicholas, Lightbown & Spada, 2001; Panova & Lyster, 2002). This issue is referred to throughout this chapter. The second area of debate is that it also continues to be problematic as to whether recasts provide negative evidence or not. Due to the fact that this is an important issue in the current research, the following section is devoted to a discussion of whether or not recasts provide negative evidence. In addition, the necessity of being aware of the negative evidence aspect of recasts is also explored.

2.7.1 Do Recasts Provide Negative Evidence as well as Positive Evidence?

As previously discussed, learners require access to linguistic input in order to learn a language. This input needs to include information about what is possible in the L2 (positive evidence) as well as what is not possible (negative evidence) (White, 1991). Currently, SLA research is concerned with how learners process input and in so doing, extract language forms and match them to meaning and function (Doughty, 2005). As discussed in section 2.5, communicative teaching methods adhere to the notion that meaning is of central importance. Therefore, it is important to understand how a focus on form can be achieved while still maintaining meaning. With explicit instruction, grammar rules are explained, errors are pointed out to learners and then corrected. However, communicative teaching methods prefer a more implicit approach to instruction, which

makes no overt reference to rules or forms and aims to unobtrusively offer corrective feedback within meaningful contexts (Lightbown & Spada, 2006).

Implicit instruction providing negative evidence may be one way to focus on a specific linguistic feature while not unduly interrupting the flow of communication during an interactive task. Ellis (2003) suggests that task-based language teaching incorporating implicit negative feedback could provide instruction while maintaining a primary focus on meaning. This notion reflects Long's (1996) updated interaction hypothesis which stresses the importance of negotiation for meaning and suggests that negative feedback which occurs during negotiation may be facilitative of L2 development. Such feedback is purported to draw learners' attention to mismatches between target and nontargetlike L2 forms. If learners are focused on both form and message the negative feedback contained within recasts could provide a catalyst for change (Long, 1996; Long & Robinson, 1998; Oliver, 1995; Saxton, 1997).

The effectiveness of recasts as negative evidence has been the focus of a growing body of SLA research over the last two decades. It has been suggested that if a learner has no conscious perception that a recast is intended to be corrective, then recasts should be viewed as only providing positive evidence (Ellis & Sheen, 2006). Conversely, recognition of their corrective nature indicates a source of negative evidence. Research findings to date have been mixed and the question of how effective recasts actually are on acquisition has been raised.

Ongoing issues within this field of research include whether or not classroom learners perceive recasts as negative evidence or simply as input which offers an alternative way to say the same thing as the learner's incomplete utterance (Allwright, 1975; Fanselow, 1977; Lyster, 1998; Lyster & Ranta, 1997; Panova & Lyster, 2002). Lyster and colleagues in their often referred to research of French immersion programmes for young learners in Canada, question how learners distinguish recasts from non-corrective repetitions, if they do so at all. A further area of contention is whether recasts are simply treated as confirmation of meaning without awareness that they are grammatically

different to the learner's initial utterance and are offering feedback on form (Carroll, 1997; Lyster, 1998; Lyster & Ranta, 1997; Mackey, Gass & McDonough, 2000).

This ambiguity raises the issue of whether or not it is in fact necessary for learners to be aware of the recast as negative feedback when they are exposed to it, for there to be a beneficial effect on the developing L2 system (Gass & Varonis, 1994; Long, 2006; Mackey, 1999, 2006; Mackey & Philp, 1998; Mackey, Gass & McDonough, 2000; Ohta, 2000; Oliver, 1995, 2000). Further research in the area of uptake and SLA is required in order to investigate whether negative evidence could be beneficial without learners seeming to interpret it as corrective feedback. The current study incorporated a post-data collection question to the students about this issue, in order to be able to contribute to the literature. Findings are presented and discussed in Chapter 4 (section 4.5 and 4.5.1). This interpretation of recasts as well as the issue of uptake following recasts is considered and discussed in more detail in the following section as well as later in this chapter (2.7.3).

2.7.2   Do Learners Notice the Corrective Nature of Recasts?

Long (2006) claims that reactive recasts, being implicit and incidental, have several potential advantages from a psycholinguistic perspective over the same information offered as pre-emptive positive evidence, such as models. Firstly, recasts convey important information to the learner about the target language in context, when both the attention of the interlocutor and the learner are jointly focused on the target language at hand. Furthermore, this happens at a time when the learner already has prior comprehension of at least some of the message, thereby facilitating form-function matching. Long argues that learners have a desire for their message to be understood and "so will probably be motivated and attending, conditions likely to facilitate noticing of any new linguistic information in the input" (2006, p. 78). Secondly, as the interlocutor's response is a reformulation of the learner's own, then the learner will already understand all or part of it and this will free up attentional resources to concentrate on the form of the

response and again to form-function mapping. Thirdly, the dependence of recasts on non-target-like learner utterances means that the incorrect and correct utterances are juxtaposed. As Saxton (1997) stresses, a corrective recast simultaneously provides information about what is and is not acceptable in the target language, compared to positive evidence which only provides instances about what is acceptable. Saxton's Direct Contrast Hypothesis (1997) states quite simply that corrective recasts (negative evidence) are more effective than models (positive evidence) because they can inform the learner of two important points; firstly that the form modelled in the recast is grammatically correct and secondly, even more importantly, that their own form is ungrammatical. Models can only fulfil the first part of this task. However, it is important to consider whether L2 learners are aware of this distinction.

As Saxton (1997) and Long (2006) point out, corrective recasts are said to promote noticing of erroneous forms within a meaning-based context. A number of researchers have suggested that learners only benefit from comprehensible input if they notice the grammatical forms contained within it (Corder, 1967; Ellis, 1994; Gass, 1997, 2003; Long, 1996; Schmidt 1993, 2001; Swain, 1995, 2001). Schmidt (1990, 1993, 1994) claims in his Noticing Hypothesis that for there to be potential for learning, a learner must notice something about the input because intake (linguistic input which becomes incorporated into the learner's interlanguage system) is conditional upon noticing. A large number of researchers support the view that negative feedback, such as recasts, may help draw learners' attention to the language forms that they have produced and help them to detect gaps in their L2 development (Gass, 1997, 2003; Long, 1996, 2006; Pica, 1994; Swain & Lapkin, 1995). Perception of these gaps may lead to grammar restructuring (Gass & Varonis, 1994).

On the other hand, both Krashen (1985, 1989) and VanPatten (1988) take an opposing view to Schmidt (1990, 1993, 1994) and others, by denying any beneficial effects of focus on form techniques, at least in the early stages of learning. Krashen discounts any need for noticing of mismatched linguistic forms, as he believes that with rich enough comprehensible input the target-like forms will be available to the learner without any

specific focus on form. Debate also continues as to whether noticing of target-like forms requires immediate uptake of the correct form in the learner's consequent response, in order to be considered beneficial. In L2 recast studies, a number of measures of evidence of effectiveness have been proposed, such as, changes in the learner's interlanguage following a period of corrective recasts (Doughty & Varela, 1998); immediate repetition or repair (Lyster & Ranta, 1997); private speech (Ohta, 2000); immediate accurate recall of recasts (Philp, 2003) and perception of recasts as corrective feedback in stimulated recall (Mackey, Gass & McDonough, 2000).

Lyster and Ranta (1997) in their well-known study of four French immersion classes in Canada, define 'uptake' as the learner's utterance immediately following the teacher's feedback and when the student is clearly reacting to the teacher's intention to draw attention to some aspect of their original utterance. In this and further studies (Lyster, 1998; Panova & Lyster, 2002) Lyster and colleagues doubt that L2 learners notice the corrective nature of the recasts, due to the fact that uptake rates are low compared with other types of feedback, such as clarification requests. On the other hand, other researchers found a fairly high level of uptake following recasts (Ellis, Basturkmen & Loewen, 2001; Oliver, 2000) and suggest that this modified output, considered necessary for language acquisition (Swain, 1985, 1995) can aid language development (Ellis et al.).

This study hypothesises that learners can benefit from recasts despite the absence of immediate uptake, as described by Lyster and Ranta (1997). Support exists within the literature as a number of researchers suggest that the presence of some kind of overt learner response, such as uptake, following a recast may not be necessary for proof of noticing and potential for language development (Ohta, 2000; Lyster & Izquierado, 2009; Mackey, 1999; Mackey & Philp, 1998; McDonough & Mackey, 2006). Mackey and Philp go so far as to suggest that a delayed effect on learners' interlanguage systems could still be possible even without any obvious immediate uptake of the negative feedback. Long (2006) cautions that immediate uptake cannot automatically be equated with learning and as discussed in section 2.7.1 many researchers argue that even if learners fail to identify feedback as corrective, that does not mean that this feedback

cannot be beneficial for language learning. This leads us onto the question of which factors (besides noticing and uptake) may affect the benefits of using recasts. The following sections introduce a number of these factors and point out which ones may have affected the potential benefits of the recasts offered in the current study.

2.7.3 Factors which may affect the Benefits of Recasts

Results in L1 acquisition have helped to inspire a large quantity of recent L2 research which has focused on the role of implicit negative feedback, and in particular, the value of "so-called corrective recasts" (Long, 2006, p. 75). In meaning-oriented language classrooms, recasts make up a large percentage of the corrective feedback offered by teachers (Donato, 1994; Doughty, 1994; Ellis, Basturkmen & Loewen, 2001; Loewen, 2002, 2004; Lyster & Ranta, 1997; Oliver, 2000; Panova & Lyster, 2002; Seedhouse, 1997). However, despite the large amount of research which has been undertaken to date, relatively little is known about exactly how recasts lead to language acquisition (Carpenter, Jeon, McGregor & Mackey, 2006). Why is it that some learners seem to be more receptive to recasts than others and why do some structures appear to be more amenable to recasts than others? What leads to the differential effects of recasts? Answers to these questions are not easy to find, however, the success of recasts on acquisition may be affected because of the following variables:

- Opportunities for feedback and uptake may vary according to context (Loewen, 2002; Sheen, 2004). In the current study, opportunities for recasts varied according to the number of questions produced within the tasks, by individual students.
- Learners' perceptions might be influenced by the linguistic target of the feedback (Mackey, Gass & McDonough, 2000). In the current study, the linguistic target was the formation of questions, which occur frequently in the L2 and the students already had knowledge of them.

- The developmental level of the learner in relation to the linguistic target of the recast has also been shown to be of importance (Han, 2002; Mackey, 1999; Philp, 2003).

These variables and how they relate to the use of recasts in the current study will now be discussed in more detail.

2.7.3.1 Instructional Context/Setting

As discussed earlier (section 2.5), many researchers have strongly suggested that learners of an L2 seem to benefit from some kind of focus on linguistic form in classrooms, as well as meaning-focused communication (Doughty & Williams, 1998; Ellis, 2001; Long, 1991, 1996; Long & Robinson, 1998; Norris & Ortega, 2000; Spada, 1997). This focus on form is often associated with the incidental feedback provided by teachers during communicative activities (Ellis; Ellis, Basturkmen & Loewen, 2001; Long, 1996). A number of classroom studies have suggested that recasts tend to be the most frequent type of feedback type in communicative classrooms (Doughty, 1994; Ellis et al., Loewen, 2002, 2004; Lyster & Ranta, 1997; Oliver, 2000; Panova & Lyster, 2002; Seedhouse, 1997). Loewen and Philp (2006) suggest that this may be the case "because they are a non-threatening, unobtrusive means of supplying correction and do not interrupt the flow of the interaction" (p. 548). However, findings as to their effectiveness in both observational and experimental studies vary and one reason for this may be the instructional settings where research has taken place.

Research into negative feedback and recasts has been undertaken in a number of different contexts and results may differ between studies due to the context of the interaction and the type of instructional setting. Contexts in classroom observation studies have differed in various ways, such as whether the L2 instruction is content based (Lyster & Ranta, 1997) or communicative (Ellis, Basturkmen & Loewen, 2001; Oliver, 2000; Panova & Lyster, 2002), the age of the learners, children in the Lyster and Ranta and Oliver studies

and adults in the Ellis et al. and Panova and Lyster research, and finally whether the environment is immersion classes (Lyster & Ranta) or ESL (Ellis et al.; Oliver; Panova & Lyster). Furthermore, the level of the learners within the various contexts may affect results, due to motivation and cognitive ability. Ellis et al. suggest that their fee-paying adult learners were more motivated to study and committed to developing their English than the Canadian Anglophone young learners in Lyster and Ranta's study. These children may have been less concerned about making progress as they were simply attending school and studying their regular programme. Furthermore, generally, adult ESL learners respond to recasts providing uptake, more often than children which may be a result of being more cognitively able to attend to form. Another possible explanation is due to the nature of the instruction. Adult learners may have been induced to attend more to form (Ellis et al.), or they may have become aware of the target structures under investigation due to cognitive ability (Ohta, 2000).

Experimental research conducted in laboratory settings has generally shown more positive outcomes for recast usage than observational studies (Ayoun, 2001; Han, 2002; Iwashita, 2003; Leeman, 2003; Long, Inagaki & Ortega, 2001; Mackey, 1999; Mackey & Philp, 1998; Mackey, Gass & McDonough, 2000). However, Han advises caution due to the fact that it would be difficult to replicate experimental studies in real classrooms, which may explain why corrective feedback studies in real classrooms seldom generate positive findings (Lyster, 1998; Lyster & Ranta, 1997). The sample size of a study is also significant as small samples (Han; Mackey, Gass & McDonough) may not generate the same kind of results as larger samples or less controlled settings, such as classroom contexts or naturalistic settings. The sample in the current study was relatively small and the research was conducted in a controlled setting, so caution must be taken when generalising results to classroom or naturalistic settings.

While the current study was experimental, rather than observational, many of the points raised above relate to the sample used. These participants were fee-paying, adult L2 learners, taking part in meaning-focused communicative activities. It is therefore, considered that they would be motivated to develop their English and complete the tasks

to the best of their ability. Therefore, the focus-on-form technique used in the treatment and testing sessions, namely providing recasts in response to learner error, may have been of benefit. However, on the other hand, as suggested by Ohta (2000), due to the nature of the tasks the learners may have become aware of the target structure, which could have affected their performance. The linguistic target of recast studies and in particular, question formation, which was the focus in this research, will be discussed further in the following section.

2.7.3.2 Linguistic Target of the Feedback

One reason why research results may be difficult to implicate to other similar studies could be due to the targeted linguistic form under investigation in each study. Some researchers have suggested that the limitations of recasts may be differentially beneficial depending on the target form (Ellis & Sheen, 2006; Long, 1996; Mackey, Perdue & McDonough, 2000). Loewen and Philp (2006) comment that only a few experimental studies to date have included pre- and post-test measures and among these, recasts have shown to be effective for some particular forms but not for others (Iwashita, 1999; Long, Inagaki & Ortega, 1998; Mackey & Philp, 1998; Mackey, Philp, Egi, Fujii & Tatsumi, 2002; Ono & Witzel, 2003; Ortega, 1999). Mackey et al. suggest that studies which have looked at learners' perceptions of recasts as corrective feedback indicate differences caused by linguistic focus. As mentioned earlier (section 2.7.1) this study included a post data collection opportunity for participants to comment on whether learners had perceived recasts as corrective in nature. The results are presented and discussed in Chapter 4 (section 4.5 and 4.5.1).

A number of studies have explored the development of question forms (Mackey, 1999; Mackey & Philp, 1998; McDonough, 2005; McDonough & Mackey, 2006, 2008; White, Spada, Lightbown & Ranta, 1991) and exposure to negative feedback, with favourable results. There appears to be a general view that interaction, along with the provision of

negative feedback and opportunities to produce modified output can benefit interlanguage development in regards to question forms. Despite the fact that in Mackey and Philp's study, recasts were not usually repeated and rarely elicited modification by learners, they were still considered as beneficial for acquisition and may in fact lead to a more delayed effect on development (Mackey & Philp, 1998). Support for delayed outcomes of recasts (Ellis, Loewen & Erlam, 2006; Gass, 1997; Gass & Varonis, 1994; Lightbown, 1998; Mackey, 1999; McDonough & Mackey, 2006) will be discussed in more detail in a later section (2.8).

This brings us to issues regarding the developmental level of the learner and findings in recast studies. The extent to which exposure to recasts in the current study provided benefit to the learners in regards to their developmental level is explored in the next section.

2.7.3.3   Developmental Level of the Learner

Pienemann's (1989) teachability hypothesis states that learners must be ready to acquire language in order for acquisition to take place. Furthermore, learners may not notice input which is beyond their level of development. The underlying theory reflects the view that learners possess internal processing mechanisms which restrict them from advancing to the next stage in their interlanguage development, until they are ready to do so (Lightbown & Spada, 2006). Krashen (1985, 1989), on the other hand, opposes Pienemann's view and believes that with rich enough comprehensible input, the linguistic features will become available to the learner without any specific focus on form required. However, both Pienenmann and Krashen seem to agree that learners themselves will process and use for acquisition only those forms which are developmentally appropriate. With regards to negative feedback, some researchers believe that as long as learners are developmentally ready to receive the feedback then they will utilise subsequent

opportunities to produce more target-like utterances (Long, 1991, 1996; Mackey & Philp, 1998; Oliver, 1995).

Two experimental studies have compared the effects of various interaction opportunities on L2 development of question forms (Mackey, 1995; Mackey & Philp, 1998). The learners were considered to be at different stages of development. The treatment conditions of these two studies provided learners with exposure to a variety of question forms and not just those that were considered to be developmentally next. Both studies found support for the teachability hypothesis, suggesting that there may be a prime time in which recasts are effective in facilitating particular L2 structures, such as question forms. In one of the studies (Mackey & Philp) the learners who were considered ready, were reported to benefit more in the interaction with recasts condition, than the other learners, who were considered unready. Even though the current study did not separate learners into those considered ready or unready, developmental level was an important consideration, when looking at the results. Were these participants developmentally ready to benefit from the recasts for higher question stage levels or was this input beyond their current level of acquisition (Philp, 2003) and therefore, unnoticeable? Developmental readiness is discussed in relation to the findings in Chapter 4.

A number of researchers have suggested that developmental readiness may play a significant role for learners in recognising and accessing the corrective potential of recasts in SLA. Han (2002) suggested that those linguistic forms which are familiar to learners, that is to say the ones that they have some knowledge of and are "in the process of developing an ability to use properly in real operating conditions" (p. 568) may be more susceptible to utilising negative feedback. In other words, the learners are at a developmental stage ready to benefit from the negative and positive evidence provided by the recasts. The role of the recasts in Han's study was not to teach a new form but to increase the learners' awareness with regards to how to use the form on that particular task. This was similar to the role of recasts in the current study, as question formation was not new to the participants.

As discussed earlier (section 2.7.2), when learners notice the difference between their own erroneous utterance and their interlocutor's target-like one, then it is more likely that the correct form will be learned. Some argue that with the perception of this difference and the resolution of the mismatch of forms, the learner's interlanguage system can develop (Ellis, 1991; Gass, 1991, 1997; Gass & Varonis, 1994; Long, 1996). Recasts could speed up this process and if the learner is focused on both meaning and form at the same time, then this type of feedback may present a catalyst for change (Long; Long & Robinson, 1998; Mackey & Philp; Oliver, 1995; Saxton, 1997). However, Philp (2003) points out that if a learner is not developmentally ready to acquire the target feature, then they may not notice the input beyond their current level of development. Many researchers suggest that prior knowledge of the linguistic form may determine readiness to acquire and noticing of this form (Gass, 1997; Han, 2002; Philp; Saxton, 1997; Schmidt & Frota, 1986). Philp also suggests that more advanced learners may benefit from the increasing automaticity that comes with repeated practice. This may allow these learners to focus more of their attentional resources on the higher levels of the L2. Lower level learners may find the discrepancy between the recast and their own attempts too great, meaning that the recasts may be less accessible to them.

The next section looks at how the effectiveness of recasts is measured and whether these various measures may affect the value of research findings. Limitations of each method and the way in which recasts in the current study were assessed are also discussed.

2.7.3.4   Measures of the Effectiveness of Recasts

Various measures of the effectiveness of recasts in classroom studies to date have been utilised. The following measures were summarised by Loewen & Philp (2006) and include:

- uptake; the learner's immediate response to the recast (Lyster, 1998; Lyster & Ranta, 1997; Panova & Lyster, 2002)

- the use of uptake charts following a lesson (Nabei, 2002)

- unprompted production of the recast form obtained in task-based interaction (Egi, 2004; Lyster, 2004; Mackey & Philp, 1998);

- elicited production of the targeted form in custom-made individualised tests (Egi; Loewen, 2002, 2005; Nabei & Swain, 2001)

- use of tests which aim to access learners' underlying knowledge of a form, such as, grammaticality judgment tests (Ellis, Loewen & Erlam, 2006; Long, Inagaki & Ortega, 1998; Lyster, 2004; Nabei, 2002).

Limitations exist for each of the above measures. These include, the need for preselecting target forms in order to administer grammaticality judgement tests and for pre- and post-tests to be effective, as incidental focus on form is unpredictable (Loewen & Philp, 2006); uptake charts which require learners to record the linguistic forms noticed during the lesson do not measure actual use and it is difficult for the learners to articulate what was noticed; tailor-made tests, where forms which arose and received corrective feedback for each student during interaction are assessed, allowing for the testing of incidental focus on form are extremely time-consuming (Egi, 2004; Loewen, 2002, 2005; Nabei & Swain, 2001; Swain & Lapkin, 1998, 2001). Furthermore, these measures cannot provide information about learners' previous knowledge as pre-tests are generally not applied and therefore differentiation between the acquisition of new knowledge and the consolidation of suppressed knowledge is not possible. Finally, the measure of successful uptake is noted as being particularly problematic when investigating recasts (Loewen & Philp, 2006). Long (2006) points out that uptake does not indicate whether it is the learner's previous knowledge which has been stimulated by the corrective feedback or the acquisition of new knowledge. Furthermore, as mentioned previously, the noticing of recasts is not necessarily dependent on uptake (Mackey & Philp, 1998).

Further explanations as to why recasts can result in negligible uptake include a small sample size, such as Mackey, Gass and McDonough (2000) with only 17 learners and

possible problems for learners in perceiving the linguistic focus of the feedback, due to the communicative orientation of the classroom (Lyster & Ranta, 1997). Furthermore, learners may be unaware of the corrective intention of the feedback (Mackey et al.). Lyster (2004) concluded in his study, comparing the effects of recasts and prompts on young L2 learners' rule-based systems, that prompts encourage the use of more targetlike forms and lead to greater production of these forms in subsequent language processing, than simply hearing recasts. Lyster has related this finding to the ambiguity of recasts and the potential difficulties involved with noticing morphosyntactic errors. Potential ambiguity of recasts and whether they can still be beneficial to language development even when they are not necessarily perceived as an error correction technique is an important consideration in this study and will be discussed in the following section.

2.7.3.5   Potential Ambiguity of Recasts

Nicholas, Lightbown and Spada (2001) point out that feelings are mixed, regarding whether recasts comprise negative evidence, as learners may not realize their intended purpose of error correction, as discussed in section 2.7.1. A number of interaction researchers have claimed that recasts might be ambiguous from the learners' perspective (Allwright, 1975; Fanselow, 1977; Lyster, 1998; Lyster & Ranta, 1997; Panova & Lyster, 2002). In other words, due to the nature of the interaction (recasts often occur together with non-corrective repetitions and without explicit cues to identify their corrective purpose) the negative feedback provided in recasts, may be hidden and "learners may fail to see the difference between their erroneous utterances and the corrections supplied by their interlocutors in recasts" (Carpenter, Jeon, MacGregor & Mackey, 2006, p. 210). Lyster (1998) argued that this is particularly the case in a classroom setting where the primary focus is on meaning.

As previously discussed (section 2.7.1) the debate surrounding the potential ambiguity of recasts is closely connected to whether or not they are seen as providing positive or

negative evidence. Lyster (1998) argued that due to the fact that the formulating, corrective function of recasts is not transparent, their value towards acquisition is diminished. He adds that the problem is intensified due to the fact that native speakers will often repeat a learner's correct utterances as well, so that the learners can never be sure if their interlocutors are supporting their efforts to converse, by repeating what they have said or are in fact providing them with negative evidence. Learners do often fail to notice the corrective nature of recasts (Mackey, Gass & McDonough, 2000) and recasts have been said to be "among the least clear and direct forms of negative feedback" (Han, 2002, p. 550). However, despite these views, a good deal of empirical research has suggested that recasts can aid L2 learning and should not be discarded due to potential ambiguity and lack of uptake in certain contexts (Carpenter, Jeon, Macgregor & Mackey, 2006; Long, 2006). Therefore, as the literature offers robust support for their usage in the classroom as a focus-on-form technique, it is believed that this study can add to the body of research work investigating recasts, in spite of objections concerning their possible ambiguity.

2.8   The Focus of this Study

Longitudinal studies are needed to explore more fully, the delayed effects of negative feedback and interaction for all linguistic features (Long, Inagaki & Ortega, 1998; McDonough, 2005). Furthermore, it may be that an immediate learner response to a recast is not always necessary for this form of negative feedback to be beneficial for language acquisition (Gass & Varonis, 1994; Mackey, 1999, 2006; Mackey & Philp, 1998; McDonough & Mackey, 2006; Ohta, 2000). Some researchers suggest that despite the absence of immediate uptake, recasts may still have an effect on L2 development in the long term (Lightbown, 1998; Mackey, 1999; Mackey & Philp). Due to this fact, the literature suggests that further longitudinal studies are required to investigate effects over time (Gass, 1997; Gass & Varonis, 1994; Mackey, 1999; Ohta, 2000; Pica, 1994).

Ellis, Loewen and Erlam (2006), claim that positive effects of corrective feedback on learning of implicit knowledge exist. In their study, these effects were more obvious in a delayed post-test, which reflects Mackey (1999). Pica (1994) suggests that more experimental studies comparing effects on L2 learning outcomes of interaction with negotiation and also interaction without negotiation are needed. Pica expands on this view by proposing that these studies should use tasks designed to tap into grammatical modifications and be followed up with long term, focused experimental studies.

Taking into account research that has already been undertaken in the L2 field regarding recasts and their potential for language learning, as well as those aspects which require further investigation, this study explores the benefits on acquisition of question forms through the use of intensive recasts, both immediately following treatment and in the longer term. Tasks designed to elicit meaning-focused language, through communicative activity are adopted for this study. Only errors relating to question formation are recast for the participants in the experimental group. Retention from the immediate post-test is examined to indicate short-term effects of recasting. However, the main focus is on investigating delayed effects, as this appears to be an area of research largely unexplored (Mackey & Philp, 1998). In addition to an immediate post-test, two delayed post-tests, at one week and three months following treatment, are administered. Finally, after the delayed testing, an email was sent to the participants in the experimental group to explore their perception of the negative evidence contained within the recasts, to which they were exposed.

This study aims to answer the following research questions.

- Can the incorporation of recasts in task-based interaction benefit the immediate production of question formation in ESL?
- Can the incorporation of recasts in task-based interaction benefit the production and development of question formation over time (at one week and twelve weeks following treatment) in ESL?
- Do the participants in the experimental group perceive the recasts as containing negative evidence?

CHAPTER 3

RESEARCH METHODOLOGY

3.1   Introduction

This chapter introduces, and includes a discussion of, the methodological approach and research design, considered as best suited to provide an answer to the following research questions:

1.   Can the incorporation of recasts in task-based interaction benefit the immediate production of question formation in ESL?
2.   Can the incorporation of recasts in task-based interaction benefit the production and development of question formation over time (at one week and twelve weeks following treatment) in ESL?
3.   Do the participants in the experimental group perceive the recasts as containing negative evidence?

First, a quasi-experimental approach (where all features are the same as a full experimental design except for random participant assignment) with a pre-test/post-test design, is proposed in order to arrive at answers to the above research questions. An overview of the research design then follows, describing the key methods employed; that is to say, participant questionnaires, pre- and post-tests and task-based treatment sessions. Given the importance of design and validity when selecting research instruments, justification of each method used is provided. The specific data collection procedure and an outline of the data analysis methods used are described in the subsequent sections. In addition, ethical concerns concerning the research process are explained. The chapter concludes with a brief summary of the previous sections.

3.2 Methodological Approach

Regarding the various research methods available, it is important to acknowledge that "different methods are appropriate for different situations" (Patton, 1990, p. 39). Therefore, designing a study appropriate for a specific situation is largely determined by the purpose of the study, the questions under investigation, and the sources available. Taking these points into consideration as well as being aware of the limitations associated with every research method, a quasi-experimental approach to data collection with a pre- and post-test design was chosen as the most suitable for this study.

3.2.1 Quasi-Experimental Design

A quasi-experimental approach or "intervention research" (Dornyei, 2007, p. 119) is a type of empirical research where independent variables are manipulated in order to discover the influence this might have on one or more dependent variables. Dornyei suggests that this type of study is able to establish an explicit cause and effect relationship between variables. The relationship becomes apparent due to the fact that two groups of participants are similar in every respect except for the fact that one group (experimental) receives a special treatment, whereas the other group (control) does not. In quasi-experimental designs, certain deliberately manipulated processes occur in a firmly controlled environment, in which only the target variables are varied while others remain the same (Johnson & Christensen, 2004). The current study was a typical quasi-experimental design involving an experimental group and a control group. The control group provided a "baseline for comparison" (Dornyei, p. 116). As is the case with many research studies in applied linguistics, this research was aiming to uncover causal links between variables in order to answer the research questions. Specifically, whether there

was a link between the use of recasts in the SLA classroom and development of question forms.

In a true experimental study, participants are randomly assigned to either the experimental or control group. This measure is undertaken in order to make the two groups as similar as possible before applying any treatment. However, in this quasi-experimental study, even though random assignment was employed at first, there was also some intervention in the compilation of the student pairings. This step was necessary, due to the possibility that certain learner characteristics could have an impact on the target dependent variable. In other words, individual difference variables – namely, gender and nationality – were considered and measures were taken to avoid different gender and different nationality pairings wherever possible. How this was achieved is explained further in section 3.3 below.

3.2.2   Quantitative Research

In the field of SLA research, both qualitative and quantitative research is commonly carried out. Dornyei (2007) defines these two types of research as:

- "Quantitative research involves data collection procedures that result primarily in open-ended, non-numerical data which is then analysed primarily by statistical methods" (p. 24).

- "Qualitative research involves data collection procedures that result primarily in numerical data which is then analysed primarily by non-statistical methods" (p. 24).

This study utilised a quantitative research approach. Dornyei (2007) points out, that quantitative researchers are more interested in the common features of their participants rather than the individuals themselves. Therefore, as this study is mainly concerned with the relationships between variables which capture these common features – exposure to

recasts and the production of question forms - a quantitative approach to data collection was adopted in order to answer the research questions. Furthermore, quantitative research methods were appropriate as this study involved controlled measurement and examined numerical group scores and performance.

3.2.3   Longitudinal Research

Menard (2002) defines a longitudinal investigation as being research in which:

- data is collected for two or more distinct time periods
- the subjects or cases analysed, are the same or comparable from one period to the next
- the analysis involves some comparison of data between periods which is used to describe patterns of change and explain causal relationships

A vast amount of applied linguistics research examines developmental aspects of the learner. This is a reason why longitudinal research ("the ongoing examination of people or phenomena over time" (Dornyei, 2007. p. 78)) is regarded as one of the most significant research directions in this field (Abbuhl & Mackey, 2008). Therefore, it is somewhat of a surprise that so many researchers seem to shy away from long term studies with relatively few being reported in the SLA literature (Dornyei). Reasons for the tendency to focus on cross-sectional research (an analysis of the target phenomenon at a single point in time) could be due to the difficulties involved in conducting long term studies. Some of these potential problem areas are pointed out by Mackey and Gass (2005) as follows:

- Participant mortality, where the longer the study takes the more likely it is that participants will drop out.

- Participant conditioning, where regular treatment and test sessions could alter the participants' behaviour and responses.
- Participant maturation, where over a longer data collection period, some subjects, especially children may change in one way or another in addition to language development.
- The length of time required before meaningful results are produced.

However, despite these difficulties in implementing a longitudinal study, recent SLA research has seen an increase in the number of studies measuring language development by sampling over time (Dornyei, 2007). The main advantage of longitudinal research is that "one gets a wider snapshot of treatment effects" (Mackey & Gass, 2005, p. 149). Longitudinal investigations generally seem to focus on specific instructional techniques, such as corrective feedback. Ortega and Iberri-Shea (2005) point out that the quasi-experimental design, typically adopted in these studies has shown a marked improvement in recent years, with longer treatment periods and immediate as well as delayed post-tests.

In their recent overview of longitudinal research in SLA, Ortega and Iberri-Shea (2005) argue that due to the fact that language learning happens over time, a full longitudinal framework is required in order to produce meaningful L2 research results. They concluded that, "Ultimately, it is through cumulative longitudinal findings that the SLA research community would be able to contribute meaningful characterizations of the gradual process of attaining advanced second language and literacy competencies across various contexts" (p. 28).

The current study explored the role between recasts and L2 development patterns by conducting a pre-test and then post-tests both immediately following treatment as well as over time. It has been suggested that studies which utilise pre/post-test measures provide more direct evidence of the possible benefits of interactional feedback on SLA (Nassaji, 2009). Regarding the timing of delayed post-tests, due to the nature of L2 learning where learners generally enrol in language programmes of fairly short duration, 'over time'

often means a period of a few weeks. Ortega and Iberri-Shea (2005) observe that an increasingly common trend in L2 instruction research is to administer delayed post-tests just one month after instruction. Ishida (2004) investigated recasts on Japanese L2 development utilising post-tests up to seven weeks after treatment. Other studies of recasts have administered delayed post-tests one week and one month (Mackey, 1999) or one week and three weeks later (Mackey & Philp, 1998). However, this study considered that 'over time' needed to involve a more substantial period, in order to add weight to the results and effects of the treatment and therefore post-tests immediately, one week and three months following the treatment were administered.

3.3   Participants

Hatch and Lazaraton (1991) argue that regarding statistical concerns, a sample needs to contain 30 or more people in order to have a normal distribution. This is a basic requirement in quantitative research (Dornyei, 2007). Keeping this in mind as well as taking student numbers at AUT IH into consideration, a sample size of 34 participants was selected for this study. It was considered that adequate data could be collected from this number of participants to indicate relationships between the use of recasts in interactive tasks and question production, if they exist. These participants were all adult ESL learners, who were enrolled in a general English programme at AUT IH in Auckland. These learners were classified by their language programme as being at an intermediate English level with seven learners at the lower Intermediate 1 level, fourteen at Intermediate 2 and the remaining thirteen learners at the slightly higher Intermediate 3 level. It was assumed that the students were all at correct proficiency levels, as a strict proficiency test is administered to all new students at AUT IH to determine this, as well as ongoing monthly testing to monitor these levels. Besides proficiency level, the only other criterion for eligibility in this study was that their remaining period of enrolment at IH be at least three months from the onset of data collection. Participation was voluntary.

The data collection process involved two groups at two different times, the first including 26 learners and the second 8 learners. Two groups were necessary due to the lower number of intermediate level students enrolled at AUT IH during the first phase of data collection (commencing March 2007). The researcher began data collection with the second smaller group of volunteers in July 2007. Every aspect of the testing and treatment sessions was identical for both groups.

These 34 learners (17 males and 17 females) came from a range of ethnic backgrounds. The ethnicities were Chinese (8), Japanese (3), Korean (17), Thai (2), European (1) and Arab (3). Five participants were 18 to 19 years old, 18 were between 20 and 24 years old, nine were between 25 and 29 years old and the remaining two were over 30 years old. The average length of time spent studying English in their home country was 8.7 years with the average starting age being 11.5. The average length of time which the participants had already spent studying English in New Zealand before the start of the data collection, was 14.5 weeks with the longest length of time being 104 weeks (one participant) and the shortest only one week (for four participants). Among these 34 students, all but one listed English study as the main reason for being in New Zealand. 19 out of the 34 participants claimed that their reason for studying English is to improve future career opportunities, whereas 10 claimed that the main reason for studying English is to be able to enter a university in an English speaking country. The remaining five have other reasons for studying English in New Zealand. Finally, 20 out of 34 students made the decision themselves to come to NZ whereas the other 14 had the decision made for them by others. The above participant information is presented in tables and graphs in Chapter 4 (section 4.2)

3.3.1 Selection Procedure

Students at the Intermediate 1, 2 and 3 levels who were planning to continue their study at this school for at least a further three months were invited to participate in the research.

The classroom teacher for each Intermediate 1, 2 and 3 level class invited their students to attend a briefing at a designated time, regarding the research and the process that would be involved. This meeting occurred in a central location (the student lounge), 10 minutes before the end of morning class. It was explained that attendance was completely voluntary and those who did not wish to go, could remain in the classroom until the end of class, with the teacher. During the briefing, one of the Academic Leaders at International House distributed a copy of the Participant Information Sheet (Appendix A) to every student. She also verbally described each stage of the research and what the participants would be expected to do during the study, emphasising that participation was voluntary and any data collected and results found would be completely confidential. Students were able to ask questions at this time. They then had 24 hours to consider the invitation before being given an opportunity to sign a Consent Form (Appendix B). Signed consent forms were returned to the same Academic Leader, without the researcher or any classroom teachers being present in the room.

As mentioned in section 3.2.1, in order to reduce feelings of inhibition which could negatively affect the data collection process, it was an aim of this study that where possible, pairs would be made up of same-gender and same-nationality students. Therefore, initially, students were randomly assigned to one of the two groups and then within these groups, were selected for pairs keeping these aims in mind. Due to the number of females and males within each group, as well as the various nationalities involved, it was not possible to have perfect matching throughout, but the majority of pairs consisted of same-gender and same-nationality students. Furthermore, in almost every case, pairs consisted of students at the same stage of development as indicated by their class level e.g. Intermediate 1, 2 or 3.

3.4　　　　Data Collection Instruments

A variety of data collection instruments were used throughout the data collection process and these are discussed below.

3.4.1　　　　Questionnaires

A Student Questionnaire (Appendix C) was administered to each participant before the pre-test. This was a written questionnaire containing 15 simple questions requiring one or two word or multi-choice answers. Seliger and Shohamy (1989) point out that in SLA research, questionnaires are recommended to collect data on phenomena, which are not easily observed, such as attitudes and motivation. The majority of questions were closed (where the researcher determines the possible answers) as opposed to open-ended questions (which allow participants to answer in any way they determine). "Closed-item questions typically involve a greater uniformity of measurement and therefore greater reliability" (Mackey & Gass, 2005, p. 93). As suggested by Mackey and Gass, in order to maximise the effectiveness of the questionnaire, the researcher aimed to utilise a simple, uncluttered format with unambiguous, answerable questions. Simple language is particularly important when the participants are reading and completing the questionnaire in their L2. Benefits of using a questionnaire include the provision of answers to questions in a systematic and disciplined way, relative ease of construction, extreme versatility and the ability to gather a large amount of information quickly and in a readily usable form (Dornyei, 2007).

The purpose of this questionnaire was to determine individual participant characteristics including; ethnic background, age, reasons for studying ESL (individual learning differences), the age at which the learners started studying English and the number of

weeks spent studying in New Zealand. This type of information is not typically available from production data alone. The reason why this questionnaire was used was to explore whether the results, generated in this study, could be influenced by individual participant differences or demographic patterns in areas such as, gender, age or ethnicity. The purpose of the questionnaire was explained fully to the students verbally and on the participation information sheet as well as on the questionnaire itself. The participants also had many opportunities to ask questions regarding the content of the questionnaire.

In addition to the initial questionnaire, following the data collection process all of the participants in the experimental group were sent an email. This email asked them whether or not they had been aware of the intended purpose of the corrective feedback, which they had received in the treatment sessions. In other words, had they realised that the recasts given by the researcher were actually correct versions of their incomplete question forms?

3.4.2 Tasks

The objective of a task is to develop linguistic skills by eliciting meaning-focused language through communicative activity, resulting in incidental language acquisition (Ellis, 2003). However, they also require clear outcomes which are "what the learners arrive at when they have completed the task" (Ellis, p. 8). These proposed outcomes need to be achieved by using language in such a way as to promote learning. Ellis claims that this is what assesses the real success of a task. An example is when a task is intended to result in a pair of learners spotting the differences between two pictures. By simply showing each other their pictures, their language skills will obviously not benefit from the task at all, even though ultimately they have achieved a successful outcome.

As this study is interaction-based, the goal was to use carefully planned tasks to involve the learners in conversational interactions. All treatment and test tasks were two-way

communicative tasks, or in other words, tasks in which vital information is held by two or more parties and must be successfully exchanged in order to complete the task (Mackey & Gass, 2005). The crucial task requirement for the purposes of this study was the extent to which they enabled participants to produce question forms. All tasks involved were chosen to ensure elicitation of the targeted forms. Furthermore, only the participants initiated questions, therefore avoiding any modelling of these forms from the researcher. These interactive tasks were familiar to the students as they were used to performing similar ones during their general English lessons. The vocabulary involved in each task was considered appropriate for the proficiency level of the students and any potentially challenging words were pre-taught using vocabulary cards with pictures of the items. When considering the picture difference tasks to use in the test and treatment sessions, guidelines set out by Mackey and Gass (2005) were followed in the selection process. These guidelines included using pictures that:

- could be easily described but still required negotiation
- would elicit the targeted linguistic structure
- would provide sufficient opportunities for feedback and output, based on the research questions

A final piece of advice given by Mackey and Gass is to carefully pilot test the tasks. As with every step of this study, the tasks were piloted in order to confirm that they would be successful in eliciting the target form (question forms) and that opportunities to provide the intended feedback (corrective recasts) existed.

3.4.2.1   Question Formation

Second language acquisition research has extensively explored the development of question forms (Mackey, 1999; Mackey & Philp, 1998; McDonough, 2005; McDonough

& Mackey, 2006; Spada & Lightbown, 1993; White, Spada, Lightbown & Ranta, 1991) and exposure to negative feedback, with generally encouraging results. The main focus in these studies is on whether or not interaction, the provision of negative feedback and opportunities to produce modified output can benefit interlanguage development, particularly with regard to question forms. Despite the fact that in their well-known study, recasts were not often repeated and rarely elicited modification by learners, Mackey and Philp concluded that recasts can still be considered beneficial for acquisition of questions. In the current study, question forms were chosen as the measure of development (the dependent variable), because previous research has shown that they can be readily elicited (Mackey, 1995, 1999, 2006; Spada & Lightbown, 1993), that different question forms are available at all levels of learning and because question forms fall into the category of complex linguistic structures, which some researchers have suggested may be affected by interaction (Mackey; Spada & Lightbown). Furthermore, the elicitation of questions can be fairly easily controlled through the selection of well-designed and implemented tasks, such as picture differences. Finally, empirical research for the stages of acquisition of question formation is relatively sound (Mackey & Philp; Pienemann, 1998; Pienemann & Johnston, 1987; Pienemann, Johnston & Brindley, 1988; Pienemann & Mackey, 1993; Spada & Lightbown).

As discussed above, in this study, all question forms targeted in treatment and testing were part of the developmental sequence for question formation in ESL identified by Pienemann and Johnston (1987) and illustrated in Table 3.1 below. This sequence was adapted by Spada and Lightbown (1993) for their study of the effects of instruction on question formation and used by Mackey (1995) and Mackey and Philp (1998). The table below displays the version adapted by Mackey and Philp. Development was determined by improvement within this sequence, with a higher proportion of later stage questions – stage 4 and 5 – produced in the post-tests, compared with the pre-test scores. Sustained development involved increases of these more advanced question forms in all of the post-tests, particularly the delayed ones, at one week and 12 weeks following treatment. These findings would increase the likelihood that development had occurred over the longer term.

51

Table 3.1: Examples of question forms and developmental stages

Developmental Stage	Example
Stage 2 SVO? *Canonical word order with question intonation*	It's a monster? Your cat is black? You have a cat? I draw a house here?
Stage 3 Fronting: Wh/Do/Q-word *Direct questions with main verbs and some form of fronting*	Where the cats are? What the cat doing in your picture? Do you have an animal? Does in this picture there is a cat?
Stage 4 Pseudo Inversion: Y/N, Cop. *In Y/N Questions an auxiliary or modal is in sentence-initial position.* *In Wh-questions the copula and the subject change positions.*	(Y/N) Have you got a dog? (Y/N) Have you drawn the cat? (Cop) Where is the cat in your picture?
Stage 5 Do/Aux 2nd Q-word->Aux/modal->subj (main verb, etc) *Auxiliary verbs and modals are placed in second position to Wh-Q's (& Q-words) and before subject (applies only in main clauses/direct Q's).*	Why (Q) have (Aux) you (sub) left home? What do you have? Where does your cat sit? What have you got in your picture?
Stage 6 Canc Inv, Neg Q, Tag Q *Cancel Inv: Wh-Q inversions are not present in relative clauses.* *NegQ: A negated form of Do/Aux is placed before the subject.* *Tag Q: An Aux verb and pronoun are attached to endof main clause.*	(Can Inv) Can you see what the time is? (Can Inv) Can you tell me where the cat is? (Neg Q) Doesn't your cat look black? (Neg Q) Haven't you seen a dog? (Tag Q) It's on the wall, isn't it?

Note. This table is based on Pienemann and Johnston (1987) and Pienemann, Johnston and Brindley (1988). Stage 1 consists of single word words such as "why?' and "cat?" As all participants in the study were beyond this level, stage 1 question examples are not included in this table.

3.5   Data collection procedure

A quasi-experimental methodological approach was utilised for this research incorporating a pre-test, three treatment sessions, one immediate and two delayed post-tests (one week and 12 weeks later). The 34 participants were randomly assigned to either the experimental or control groups. Within their respective groups, the participants were then divided into pairs with a partner of the same ethnic background and gender whenever possible. The majority of pairs remained the same for all of the treatment sessions in the data collection process. However, due to some participants dropping out of the study, for various reasons beyond the control of the researcher, a small number of new pairings were necessary for post-testing. Picture difference tasks were used to elicit various question forms from the participants. During the treatment sessions, intensive recasts were provided by the researcher in response to incorrect questions produced by members of the experimental group. It was intended that these recasts would not unduly interrupt the flow of conversation between the students and the researcher. Therefore, the type of recasts used was implicit as opposed to explicit where the error is pointed out and a recast offered. Participants in the control group were not exposed to any intentional recasts. Participants took part in one pre-test, three treatment sessions, one immediate post-test and one or two delayed post-tests (two for students still at AUT IH three months later) lasting approximately 10 ~ 15 minutes each. Each test and treatment session took place before morning class, at lunchtime or immediately following the end of afternoon class. The participants chose the times which suited them best. Data was analysed using measures of frequency to look for development of question formation over time and results were presented in tables and graphs. The individual steps in the data collection process are discussed in more detail in the following sections.

3.5.1   Pre- and Post-Test Sessions

The pre-test, as with every test and treatment session, was administered to one pair of participants at a time. It consisted of a picture difference interactive task (see Appendix F for task material). Each pair was given clear written instructions appropriate to their developmental level (See Appendix E). The researcher supervised all of the tests and the participants had the opportunity at any time to ask questions. However, the researcher's role was a passive one during the actual task. Any potentially challenging vocabulary was pre-taught using vocabulary picture cards. The test-room was selected for its quiet location. A barrier was placed on the table between the two participants in each pair so that neither of them could see the other's picture. Furthermore, the written instructions clearly stated that they must not, at any time, show their picture to each other. This is crucial to the success of the experiment (Mackey & Gass, 2005). Each session was recorded on an audio tape recorder. The participants asked each other questions in order to discover a pre-determined number of differences between the two pictures. The data collected in the pre-tests was analysed to determine the current developmental level of question formation ability exhibited by the learners.

One immediate post-test following straight after the completion of the third treatment session was held for all participants and was conducted in exactly the same way as the pre-test (see Appendix G for task material). One week later, the first delayed post-test took place (see Appendix H for task material) and for those students who were still enrolled in AUT IH 12 weeks following the last treatment session, a second and final delayed post-test was administered (see Appendix I for task material). Both delayed post-tests followed the same procedure as the earlier pre- and immediate post-test. As recommended by Mackey and Gass (2005), the delayed post-tests were undertaken, in order to clarify whether the effects of this particular treatment could be considered long-lasting.

3.5.2   Treatment Sessions

The treatment sessions took place over the three weeks following the pre-test. One session per week was held for each pair. A new information gap task was utilized in each treatment session. During the first treatment session the two participants were holding the same picture and the researcher was holding the corresponding one. The participants asked questions to elicit the differences. They alternated turns. The researcher did not ask questions. For the pairs in the experimental group the researcher fully recast any non-target-like question forms as often as possible. As is often the case with task-based teaching, the focus was on meaning rather than form and the aim of this treatment was to retain the central meaning of the task at all times. Therefore, if at any time a recast would seem to interrupt the flow of meaningful interaction, then it was not offered. The control group performed the same task in the same way. However, they did not receive any intentional recasts. The researcher simply provided answers to their questions regarding the picture differences, with no corrective feedback on linguistic errors. The researcher concentrated on the content of their questions rather than form or structure. It was assumed that the participants would be comfortable taking part in this type of close proximity task, with both other NNSs and a NS, due to the teaching techniques they are exposed to in the regular ESL classroom. This did, in fact, seem to be the case and no uncertainty or nervousness was detected due to unfamiliarity of the task type.

The second treatment session also utilised an information gap task but rather than a picture, one student held a set of key words from which they had to form questions in relation to either a tourist attraction or a computer café's details, such as opening times and cost. Their partner held the answers. Once again in the same way as for the first treatment session, the participants in the experimental group had their ill-formed questions recast by the researcher, whereas those in the control group did not.

The third and final treatment session was another picture difference task which was conducted in exactly the same way as the first treatment.

3.6   Data Validity and Reliability

Two of the most important issues which need addressing in any study are the validity of the data and its reliability. Data validity and reliability contribute greatly to the credibility of the study design, data collection procedure and data analysis process. According to Seliger and Shohamy (1989), reliability and validity are "the two most important criteria for assuring the quality of the data collection procedures" (p. 184).

3.6.1   Data Reliability

Reliability provides information about the consistency and accuracy of the data collection procedure (Seliger & Shohamy, 1989). To achieve reliability, in the present study, the instruments and data collection procedures were thoroughly tested in a pilot study with a small sample of L2 students comparable to the sample population of the actual study. The aim of a pilot study, as suggested by Glesne and Peshkin (1992), is to learn about the research process and to get a general sense of the research setting. By undertaking a pilot study, the researcher was able to revise and where necessary, modify the instruments on the basis of new information. This early, critical step in the multi-dimensional research process is "an important means of assessing the feasibility and usefulness of the data collection methods and making any necessary revisions before they are used with the actual research participants" (Mackey & Gass, 2005, p. 43).

The following aspects of the study were comprehensively checked in the pilot study and where necessary, improvements made:

- The pre-data collection stage, that is distribution and explanation of the information sheet and consent forms, as well as responding to questions offered by the participants, regarding the research.

- The amount of time required to answer the questions in the questionnaire as well as to complete the testing and treatment sessions.

- The wording, relevancy and clarity of the questions as well as the comprehensibility of the instructions in the questionnaire.

- The appropriateness of the location and effectiveness of the barrier between the participants.

- The quality of the audio-recording equipment.

- The physical position of the researcher in the room, during the testing and treatment sessions, as well as the role of the researcher at these times and how much intervention seemed appropriate while minimising task interruption.

- Checking for any feelings of inhibition or unwillingness to interact with either the other participant or the researcher.

- The content and language involved in the tasks and level of difficulty.

- The amount of time involved in transcribing the recordings, so that this could be factored in for the main study.

Regarding the level of difficulty of the tasks, following the pilot test, the vocabulary in the pre-test was confirmed to be at an appropriate level for these students. The pre-test vocabulary level also matched that used in the post-tests. This measure was necessary in order to enhance reliability of any apparent improvement of question formation, following the treatment sessions. It is important, in a pre-test/post-test experimental or quasi-experimental design, that the pre-test be comparable in difficulty to the post-test and once this is determined to be the case then researchers can ascertain the effects of treatment immediately (Mackey & Gass, 2005). Furthermore, each pair was given clear written instructions appropriate to their developmental level in order to reduce any confusion as to what they were expected to do.

Other measures taken in order to enhance reliability include, the researcher being the only person to administer the treatment and testing sessions in order to minimize interviewer effects (Patton, 1990) as well as, all questionnaire responses being coded and all audio-recorded test transcripts being transcribed verbatim by the researcher.

3.6.2　　　　Data Validity

Data validity is an estimate of the extent to which a study or a set of instruments measures what it claims to measure (Newman & Benz, 1998; Seliger & Shohamy, 1989). Mackey and Gass (2005) stress that after spending a great deal of time and effort in designing a study, researchers want to then make sure that the results of the study are valid. That is, they need to be meaningful and have significance not only for the sample population, which was tested but also to a broader, relevant population.

3.6.2.1　　　　Internal Validity

Dornyei (2007) stresses that a full "pre-test-post-test control-group design" (p. 120) is one of the best ways to control for various threats to the internal validity (the extent to which the differences that have been found for the dependent variable can be directly related to the independent variable) of the experiment. The following measures were taken in order to minimise threats to internal validity:

- Proficiency level of participants; a strict proficiency test is administered to all new students at IH to determine their individual level of proficiency, as well as ongoing monthly testing to monitor these levels.
- Inhibition due to pairings of participants; the majority of pairs consisted of same-gender and same-nationality students.
- Participant mortality; at the beginning of the data collection process, enrolment information showed that all participants would remain at IH for at least three months. However, due to factors beyond the researcher's control, a number of students did withdraw from the school for various personal reasons.
- Participant deception; the participants were aware that this study was concerned with teacher feedback but not specifically, recasts. It would not have been

appropriate to inform them of this low-level deception at the beginning, as data and findings could have been compromised as a result. However, all participants were given a Post Data-Collection Participant Information Sheet (Appendix D) at the conclusion of the data collection phase, to explain what the deception was and why it had been necessary.

- Instrumentation effects; the pilot test showed that the questionnaire instrument had strong content validity (the extent of representativeness of the measurement regarding the phenomenon about which we want information). In addition, the tasks used in treatment and testing, were shown to have face validity (refers to the familiarity of the instrument) as the participants were used to performing these type of information gap tasks from classroom activities. An anticipated advantage of this validity was that as Mackey and Gass (2005) suggest, when participants perceive the research treatments and tests as having a connection with other language learning activities, they are far more likely to take the experiment seriously.

- Test effects; following the pilot test and feedback from the participants, it was shown that the pre- and post-tests were equivalent, in terms of vocabulary and language difficulty.

- Instructions and questions; the pilot test indicated that the instructions and questions on the questionnaire as well as the instructions for the testing and treatment sessions were clear and appropriate for the developmental level of the participants.

A study must be conducted with special attention given to internal validity, as this is a prerequisite to external validity (Mackey & Gass, 2005).

3.6.2.2 External Validity

In this study, the researcher was concerned with the extent to which the findings generated could be relevant, not only to the research population, but also to the wider population of second language learners. This generalizability of results is referred to as external validity. In order to try and minimise threats to external validity, the researcher utilised convenience sampling.

Convenience sampling (the selection of individuals, who happen to be available) is widely used in L2 research (Mackey & Gass, 2005). The selected sample is the basis of generalizability of results (Mackey & Gass) and ideally the most representative type of sample is one where each individual who could be selected to participate has the same chance of selection as any other. However, in L2 research, due to the focus of the study, the research questions under investigation and the available participants at the time of data collection, convenience sampling is far more common. In this study, delayed effects of instruction were the focus and the required developmental level of the participants was early intermediate (as determined by their general English classroom). Therefore, students at either an elementary or more advanced level, as well as those intermediate students enrolled in this language school for less than three months, at the start of data collection, were not considered appropriate candidates and were consequently excluded from the sample. Due to these selection criteria, it is argued that this sample was representative, insofar as lower intermediate level students are concerned, and of a sufficient size for results to be considered generalizable to a wider population of private language school students at the same level of language development.

A further point regarding the size of the sample is that in L2 studies, small groups of students are often appropriate, compared with general education research where larger pools of participants can usually be accessed. In the case of smaller sample size, the data analysis methods must take the numbers into account (Mackey & Gass, 2005). Fraenkel and Wallen (2003) suggest that for experimental and quasi-experimental studies, 15 to 30

participants per group can be considered a minimum guideline. However, this number depends on how tightly controlled the experiment is.

Finally, Dornyei (2007) points out that a quasi-experimental methodological approach can leave a study more vulnerable to validity threats compared with a full experimental design. However, just because a threat is possible, it does not necessarily mean that it will occur (Johnson & Christensen, 2004). All care was taken in the design process to reduce this risk in order to produce valid and credible results.

3.7   Data Analysis

Data analysis methods employed in this study were determined by the research questions guiding the study and the data collection methods used. Firstly, the summary statistics from the student questionnaire were calculated and the findings presented in tables and graphs in order to gain an understanding of participant characteristics, such as age, gender and the amount of time previously spent studying English. This information was then used to interpret whether the two groups within this study shared a similar composition or not.

Secondly, all the test and treatment sessions were audio-recorded and pre- and post-test sessions were transcribed by the researcher. Descriptive statistics were used to provide a picture of the data, allowing the researcher to gain a better overall understanding of the data set. Data from the tests was numerically coded by the researcher and then these codes were used in the data analysis process. This data was made up of all questions produced during the tests and these question types were identified and coded according to developmental stage, as presented by Pienemann and Johnston (1987) (see section 3.4.2.1 and Table 3.1 for examples of stages in Pienemann and Johnston's developmental sequence for questions). Below is a sample of data produced in an immediate post-test. Table 3.2 follows the data and displays the numerical codes assigned to the different

question forms included in the sample. This data comes from a pair of students (participants 27 and 28) in the experimental group. The questions are underlined. R indicates the researcher:

28 <u>Does he eat some food?</u>

27 No, no he's only waiting for the bus. Different.

28 <u>What is he doing in front of the bakers?</u>

27 They are talking each other and maybe she is holding something.

28 Same in my picture.

R Good. Next one.

27 In front of the parking sign, there is a left bicycle. <u>Can you see that?</u>

R Can you see what? (*prompting*)

27 <u>Can you see the bicycle that has been left on the road?</u>

28 A little bit far from parking sign. <u>Direction?</u>

R Ask again. (*B misunderstood the question*)

27 <u>Is the bicycle broken down on the road?</u>

28 No, he is riding it.

Table 3.2: Questions and numerical codes used for data analysis

Question	Question stage* (numerical code)
Direction?	2
Does he eat some food?	3
Can you see that?	4
Is the bicycle broken down on the road?	4
Can you see the bicycle that has been left on the road?	4
What is he doing in front of the bakers?	5

* Pienemann and Johnston's (1987) developmental sequence for question formation in ESL

2 = SVO? E.g. *It's a monster?*
3 = Fronting: Wh/Do/Q-word. E.g. *Where the cats are?*
4 = Psuedo Inversion: Y/N, Copula. E.g. (Y/N) *Have you got a dog?* (Cop) *Where is the cat in your picture?*
5 = Q-word->Aux/modal->subj (main verb etc). E.g. *Where* (Q) *does* (Aux) *your cat* (subj) *sit?*
6 = Cancel Inversion, Neg Q, Tag Q. E.g. (Can Inv) *Can you see what the time is?* (Neg Q) *Haven't you seen a dog?*

Please refer to this chapter, section 3.4.2.1 for an outline of Pienemann and Johnston's (1987) table.

This numerically coded data was then analysed using measures of frequency to indicate how often questions are formed and to look for patterns of sustained question development. Measures of frequency are often adopted in second language studies (Mackey & Gass, 2005) and were chosen here, because they indicate how often a certain behaviour occurs and then are based on counting the number of these occurrences (Seliger & Shohamy, 1989). In this study, these frequencies indicated how often questions are formed. Basic tabulations of frequency and graphs were used to make descriptive comparisons using percentage scores. Due to the nature of the interactive tasks used, where the participants were free to ask as many questions as they needed to, in order to complete the task at hand, the total number of questions produced varied between learners. Therefore, in order to make descriptive comparisons between the categories over the two groups, it was decided that the proportion of questions produced at each question development stage would be presented as a percentage of the total.

As previously discussed, the question forms targeted in treatment and testing were coded according to the developmental sequence for question formation in ESL identified by Pienemann and Johnston (1987). The frequency of these forms, provided data which when measured against this sequence was used to make comparisons between pre- and post-test scores. This study involves the analysis of nominal, categorical data. An example of how this nominal data was produced is as follows. The number of times a question at a particular question stage was correctly produced is calculated. For example, participant X produces four stage 2 questions, seven stage 3 questions, six stage 4 questions, eleven stage 5 questions and no stage 6 questions in test Y. The nominal data for this participant is presented in Table 3.3 below.

Table 3.3: An example of nominal data production

Development stage*	Number of questions produced
2 (e.g. It's a van?)	4
3 (e.g. Where the man is?)	7
4 (e.g. Where is the ball?)	6
5 (e.g. What does the man do?)	11
6	0

* Pienemann and Johnston's (1987) developmental sequence for question formation in ESL

2 = SVO? E.g. *It's a monster?*
3 = Fronting: Wh/Do/Q-word. E.g. *Where the cats are?*
4 = Psuedo Inversion: Y/N, Copula. E.g. (Y/N) *Have you got a dog?* **(Cop)** *Where is the cat in your picture?*
5 = Q-word->Aux/modal->subj (main verb etc). E.g. *Where* **(Q)** *does* **(Aux)** *your cat* **(subj)** *sit*?
6 = Cancel Inversion, Neg Q, Tag Q. E.g. (Can Inv) *Can you see what the time is?* **(Neg Q)** *Haven't you seen a dog?*

Please refer to this chapter, section 3.4.2.1 for a full copy of Pienemann and Johnston's (1987) table.

The nominal data displayed in Table 3.3 above is added to data from the rest of the participants in the group – either the experimental or control – and then the total for each question category within each group is calculated. These totals were used in statistical tests. As this study used nominal data, nonparametric tests for comparing category scores from the two groups were used for inferential statistical testing. Nonparametric tests are generally used with frequency data (Mackey & Gass, 2005). These tests were used to determine the effect of the independent variable (recasts) on the dependent variable (question formation) and to investigate whether there was a statistically significant relationship between the two variables.

The most commonly used nonparametric test in applied linguistics is the Chi-square test (Dornyei, 2007). It is often used with categorical (i.e. nominal) data (Mackey & Gass, 2005). Chi-square testing was carried out on the data collected from the pre- and post-tests. The purpose of calculating the chi-square value was to see whether a statistically significant relationship between the use of recasts and development of question forms existed. These results, if proven to be statistically significant, can then be used to suggest generalizations concerning benefits of using this type of corrective feedback to the wider population, which in the case of this study is other lower intermediate L2 learners studying ESL at language schools.

3.8   Ethical Issues

Social research concerns people's lives and therefore must certainly involve ethical issues (Dornyei, 2007). In accordance with the ethical guidelines issued by Auckland University of Technology Ethics Committee (AUTEC), privacy and confidentiality were respected throughout the research process. During both the participant selection and the data collection process the researcher was aware that as with all research conducted in New Zealand, the social and cultural sensitivity of all participants must be respected. It is possible that students from these various ethnic backgrounds would not have had any previous experience of taking part in research. For this reason, the researcher and classroom teachers took great care and consideration in fully explaining every step of the research process. It was also stressed that should any students decide they do not wish to take part in the research, neither their test mark nor class result would be affected. Participants were told that the activities they would be performing during the study are ones which are familiar to them from their regular class curricula. Furthermore, any sensitive steps in the research process (such as audio-taping) were discussed with the participants beforehand and written permission received before any data collection took place. All participants were assigned a number to ensure their identities remained confidential. Furthermore, they were assured that no identifying information would be included in the study.

During the data collection process the researcher was not the regular classroom teacher for any of the students involved in this study. This measure was taken to reduce the possibility of participants perceiving themselves in a dependent relationship with the researcher. However, as it was possible that the researcher had taught some of the students previously, measures were taken to try and avoid any discomfort to the participants in this situation. These included providing all students with the contact details of the Project Supervisor, who was available for consultation throughout the study. They also had the other teachers and staff in the school to discuss the research procedure with, if necessary. Furthermore, the session explaining the Participant

Information Sheet (Appendix A) and the research process was overseen by one of the Academic Leaders at AUT IH. Those students who volunteered for the study could also ask questions of the Academic Leader regarding the research, at all times. They also returned signed consent forms to this person. This step was taken in order to avoid any perception on behalf of the students of losing face, in front of the researcher or their teacher, if they decided not to participate in the study. Neither their regular classroom teacher nor the researcher was present in the room at this time. It was explained to everybody that participation was completely voluntary and that whatever they decided to do would not in anyway affect their class grade and test result. Furthermore, the participants were informed that even if they began the research, they could withdraw at anytime, with no questions asked. In addition, they had 24 hours to consider the invitation to participate before making their decision.

Another example of showing sensitivity to the individuals involved in the research was that during the treatment and testing sessions, participants were paired with someone of their own gender and ethnic background, wherever possible in order to further reduce potential anxiety.

One final ethical concern is that as this study involved low level deception, all participants were given a Post Data-Collection Participant Information Sheet (Appendix D) to explain what it was and why it was necessary, at the conclusion of the data collection phase. It would not have been appropriate to inform them of the deception beforehand as biased data and findings could have been the result. They were, however, aware from the beginning that this study was concerned with teacher feedback but not specifically, recasts.

3.9　　　**Summary**

This chapter has outlined in detail the research design and described the research procedure used in this study. After taking into consideration, available resources and time and location constraints associated with the nature of conducting research at a private language school a quantitative quasi-experimental design approach to data collection with a pre- and post-test design was chosen as the most suitable for this study. This methodological approach was adopted in order to present a comprehensive analysis of the research questions as well as an attempt to achieve validity and reliability in the data collection procedure. Data was gathered by initially pre-testing participants and then following three treatment sessions, conducting post-tests two or three times, with the final test taking place twelve weeks after the last treatment. The objective of delayed testing, in this way, was to determine whether any long-term effects of the treatment had occurred. In addition to test data, demographic information was also collected through short-answer questionnaires which were analysed quantitatively, in order to explore relationships between language performance and individual participant characteristics. Finally, a post-test email was sent to all the participants in the experimental group to find out whether or not they had been aware of the corrective nature of the feedback they received on their question forms. Data validity and reliability were achieved through pilot testing every aspect of the data collection process. Adjustments were made where necessary. Finally, effort was made to ensure the amalgamation of ethical considerations into the research process.

CHAPTER 4

PRESENTATION AND DISCUSSION OF RESEARCH OUTCOMES

4.1 Introduction

This chapter presents the results from the analysis of the data and the research questions posed in Chapter 2 are restated and addressed. Quantitative results from the data, collected by means of a student questionnaire, the pre- and post-test tape-scripts and the post-data-collection email are examined and key findings are presented in tables and graphs. The results relating to the three research questions are presented in turn, followed by a discussion of their significance, in relation to the wider literature presented in Chapter 2.

Before looking at the findings, which relate directly to the research questions, the results from the student questionnaire are presented. These findings were gathered and analysed in order to ensure that the two groups within the study – experimental and control – were evenly composed. If these two groups had displayed different demographic features, then the statistical results created from the test data could have been biased as a result and the desired outcome of generating reliable and valid findings would have been compromised.

4.2 Demographic Features

Participant summary statistics are presented in this section. The information was gathered by use of the Student Questionnaire (Appendix C) which was administered before the pre-test. The purpose of including this questionnaire in the data collection process was to find out individual participant characteristics; gender, ethnic background, age, reasons for

studying ESL, the age at which the learners started studying English and the number of weeks spent studying in New Zealand prior to the start of the research. This was a written questionnaire containing 15 questions requiring one or two word or multi-choice answers. These summary statistics are presented in pie charts below.

GENDER

Females 50%

Males 50%
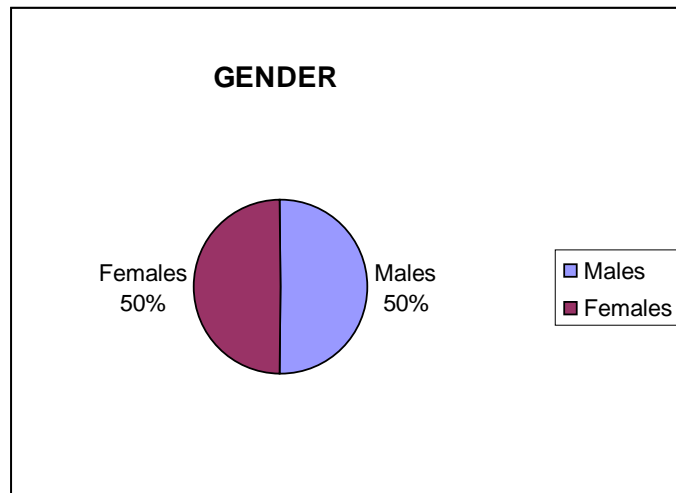
□ Males
■ Females

Figure 4.1: Gender of participants

The sample includes an equal number of male and female students (17 each).

AGE

6%   15%

26%

53%

□ 18 ~19
■ 20 ~24
□ 25 ~29
□ > 30
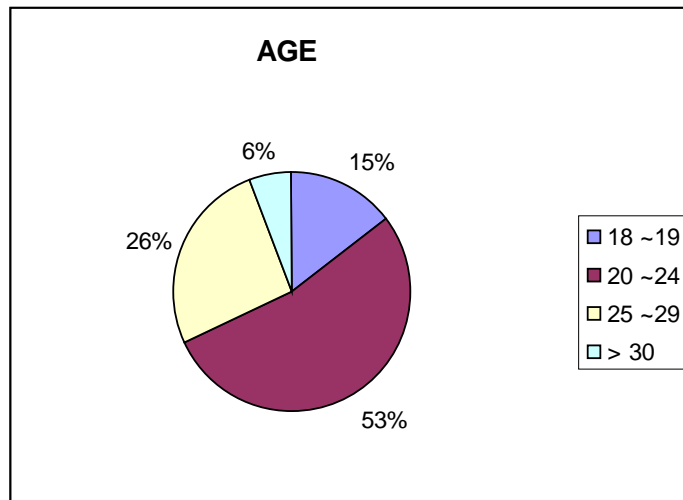
Figure 4.2: Age of participants

Of these 34 students, five of them (15%) are in the age group 18 to19, eighteen of them (53%) are in the age group 20 to 24, nine of them (26%) are in the age group 25 to 29, and finally, two of them (6%) are aged over 30.
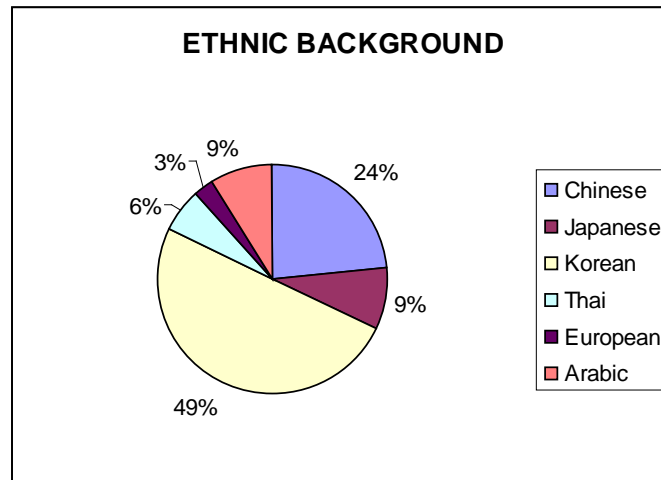
Figure 4.3: Ethnic background of all participants

When we look at the ethnic distribution of the students in the sample, we see that, eight of them (24%) are Chinese, three of them (9%) are Japanese, two of them (6%) are Thai, 17 of them (49%) are Korean, one of them (3%) is European, and finally, three of them (9%) are Saudi Arabian.
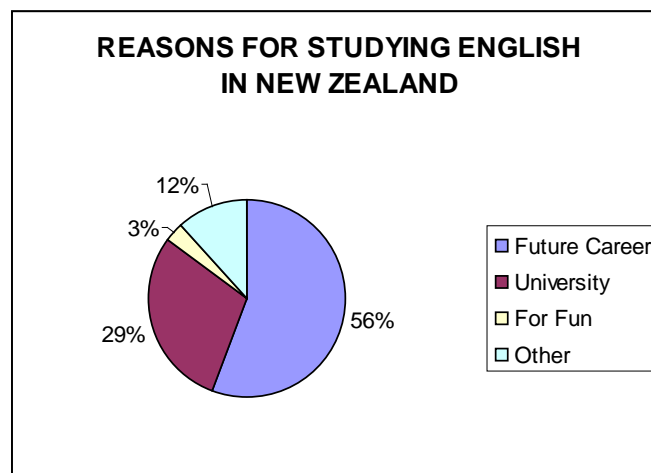
Figure 4.4: Reasons for studying English in New Zealand

Among these 34 students, all but one of them listed English study as the primary reason for being in NZ. 20 out of 34 students (59%) made the decision themselves to come to NZ and 19 out of 34 (56%) claimed that the reason for studying English is to improve future career opportunities, whereas 10 of them (29%) claimed that the main reason for studying English is to be able to enter a university in an English speaking country. Finally, four of them (12%) have other reasons for studying English in New Zealand while one of them (3%) is here for enjoyment. These results are presented in Table 4.1.

Table 4.1: Other individual learning statistics

Variable	Mean	Std. Dev.	Min	Max
Length of time spent in NZ (in weeks)	14.9697	19.96948	1	104
Age first started English study	11.58824	3.124853	5	18
Number of years of English study to date	8.705882	3.451213	2	17

The mean value for the time already spent in NZ is 14.9697 weeks. The mean value for the age when the students started to learn English is 11 years old and the mean value for the number of years which the students have been learning English to date is 8.7.

The results of the student questionnaire, which have been presented above, were interpreted by the researcher as showing that for most aspects, there is diversity and variation in this data set. The participants were randomly assigned to either the experimental or control group and looking at each of these two groups in turn, we can see that they were created with a similar distribution.

Experimental group characteristics
- eight females and 10 males

- five Chinese, three Japanese, eight Korean, one Saudi Arabian and one Taiwanese

- three 18 to19, eight between 20 to 24, six between 25 to 29 and one over 35 years old

- nine had spent two months or less in New Zealand and nine had spent more than two months in New Zealand prior to the start of the research

- nine had studied English for eight years or less and nine had studied for more than eight years prior to the start of the research

- four were from Intermediate 1, eight were from Intermediate 2 and six were from Intermediate 3 general English levels

Control group characteristics
- nine females and seven males

- two Chinese, two Japanese, nine Korean, two Saudi Arabian and one Tahitian

- three 18 to19, 10 between 20 to 24, two between 25 to 29 and one over 35 years old

- eight had spent two months or less in New Zealand and eight had spent more than two months prior to the start of the research

- nine had studied English for eight years or less and seven had studied for more than eight years prior to the start of the research

- three were from Intermediate 1, six were from Intermediate 2 and seven were from Intermediate 3 general English levels

As can be seen from the information above, the experimental and control groups make up a cross-sectional sample (representative of the general population). The majority of participants in both groups came from Korea (eight in the experimental group and nine in the control group) and the remaining learners were made up of four other nationalities in both cases. With regards to age, more than half of the participants in both groups were between 20-24 years old (eight in the experimental and 10 in the control) with one learner in each group who was over 35. The next section will look at the results relating to research question one.

4.3   Research Question One: Recasts and Immediate Benefits for SLA

This question examined whether the incorporation of recasts in task-based interaction can benefit the production and development of question formation immediately after the treatment, for L2 learners. First of all, the audio-recorded data from these two tests for both the experimental and the control groups was transcribed by the researcher and was then numerically coded. The data was coded by allocating the number of the appropriate stage in Pienemann and Johnston's (1987) developmental sequence for question formation in ESL, to each correctly formed question. This developmental sequence is illustrated, including examples of questions from each stage, in section 3.4.2.1 and 3.7 in Chapter 3. The totals of correctly formed questions at each stage were then added together.

As an important first step in the data analysis process, the pre-test scores for the two groups were tested to confirm that they were not statistically significantly different. This step was necessary in order to ensure that the experimental and control groups were independent. In other words, if pre-test scores (the number of correctly formed questions at each developmental stage) were found to be statistically significantly different, then the internal validity of the results from the three post-tests would have been threatened. Chi-square testing was used to determine group independence. Chi-square analyses must be calculated from raw frequencies (as opposed to relative frequencies or percentages) (Mackey & Gass, 2005). The raw data for both groups for this chi-square analysis is presented in Table 4.2.

Table 4.2: Correctly formed questions for the experimental and control groups

in the pre-test

Developmental stage *	2	3	4	5	6	Total
Experimental group pre-test	56	83	70	14	0	**223**
Control group pre-test	36	67	66	22	3	**194**

* Pienemann and Johnston's (1987) developmental sequence for question formation in ESL

2 = SVO? E.g. *It's a monster?*
3 = Fronting: Wh/Do/Q-word. E.g. *Where the cats are?*
4 = Psuedo Inversion: Y/N, Copula. E.g. (Y/N) *Have you got a dog?* **(Cop)** *Where is the cat in your picture?*
5 = Q-word->Aux/modal->subj (main verb etc). E.g. *Where* **(Q)** *does* **(Aux)** *your cat* **(subj)** *sit*?
6 = Cancel Inversion, Neg Q, Tag Q. E.g. (Can Inv) *Can you see what the time is?* **(Neg Q)** *Haven't you seen a dog?*

Please refer to Chapter 3, section 3.4.2.1 for an outline of Pienemann and Johnston's (1987) table.

The chi-square result for these pre-test scores was calculated with a level of significance (α) of 0.05 or 5%, which is generally the value used in L2 research (Mackey & Gass, 2005), and 4 degrees of freedom (*v*) by using a 2 x 5 table (2 rows and 5 columns). The critical chi-square statistic (χ^2_{crit}) was 9.488, meaning that if the resulting chi-square value is less than 9.488, then we can conclude that the two groups are independent. The result did in fact show that this was the case and therefore the post-test results could be interpreted as reliable ($\chi^2 = 9.11 < 9.488$). In order to answer research question one, the pre-test scores were compared with the immediate post-test scores. First of all, descriptive statistical tests were undertaken. Descriptive statistics are presented in order to provide a summary of the data and to obtain a better overall understanding of the data set (Mackey & Gass, 2005). As this study is concerned with frequencies, these statistics are presented as percentages in order to provide a clear overview or picture of the data and to show changes clearly. The percentages of correctly formed questions at each stage, in both the pre- and immediate post-tests are presented in Tables 4.3 and 4.4 below.

In Table 4.3 below, the proportion of correctly produced questions at each developmental stage, which were calculated from the raw frequency data, for the experimental group is presented. Results are shown for the pre- and immediate post-test.

Table 4.3: Correctly formed questions for the experimental group in the
pre-test and immediate post-test

Developmental stage *	2	3	4	5	6
Pre-test	25.1%	37.2%	31.4%	6.3%	0%
Immediate Post-test	11.7%	26.4%	48.1%	13%	0.8%

* Pienemann and Johnston's (1987) developmental sequence for question formation in ESL

2 = SVO? E.g. *It's a monster?*
3 = Fronting: Wh/Do/Q-word. E.g. *Where the cats are?*
4 = Psuedo Inversion: Y/N, Copula. E.g. (Y/N) *Have you got a dog?* (Cop) *Where is the cat in your picture?*
5 = Q-word->Aux/modal->subj (main verb etc). E.g. *Where* (Q) *does* (Aux) *your cat* (subj) *sit*?
6 = Cancel Inversion, Neg Q, Tag Q. E.g. (Can Inv) *Can you see what the time is?* (Neg Q) *Haven't you seen a dog?*

Please refer to Chapter 3, section 3.4.2.1 for an outline of Pienemann and Johnston's (1987) table.

In Table 4.4 below, the proportion of correctly produced questions at each developmental stage, which were calculated from the raw frequency data for the control group are presented. Results are shown for the pre- and immediate post-test.

Table 4.4: Correctly formed questions for the control group in the
pre-test and immediate post-test

Developmental stage *	2	3	4	5	6
Pre-test	18.6%	34.5%	34%	11.4%	1.5%
Immediate Post-test	9.6%	32.5%	52.2%	5.7%	0 %

* Pienemann and Johnston's (1987) developmental sequence for question formation in ESL

2 = SVO? E.g. *It's a monster?*
3 = Fronting: Wh/Do/Q-word. E.g. *Where the cats are?*
4 = Psuedo Inversion: Y/N, Copula. E.g. (Y/N) *Have you got a dog?* (Cop) *Where is the cat in your picture?*
5 = Q-word->Aux/modal->subj (main verb etc). E.g. *Where* (Q) *does* (Aux) *your cat* (subj) *sit*?
6 = Cancel Inversion, Neg Q, Tag Q. E.g. (Can Inv) *Can you see what the time is?* (Neg Q) *Haven't you seen a dog?*

Please refer to Chapter 3, section 3.4.2.1 for an outline of Pienemann and Johnston's (1987) table.

From Table 4.3 above, we can see a noticeable observed difference for stage 2 and 3 questions following the treatment sessions. With exposure to recasts, it was expected that the proportion of lower stage questions would decline in contrast to increases for higher stages. For the experimental group this pattern is illustrated in the results in Table 4.3, as

the proportion of both stage 4 and 5 questions showed an increase in the immediate post-test. The results for the control group (Table 4.4) differed to those of the experimental group in the fact that stage 5 questions decreased in the immediate post-test. These findings are displayed in the bar charts below (Figures 4.5 and 4.6).

4.3.1 Proportion of Questions Produced in the Immediate Post-Test

For these intermediate level L2 learners, it is stage 4 and 5 questions which are of special interest. These two stages indicate more advanced structures, particularly stage 5 which was considered at a slightly more developed level for these learners, before treatment.
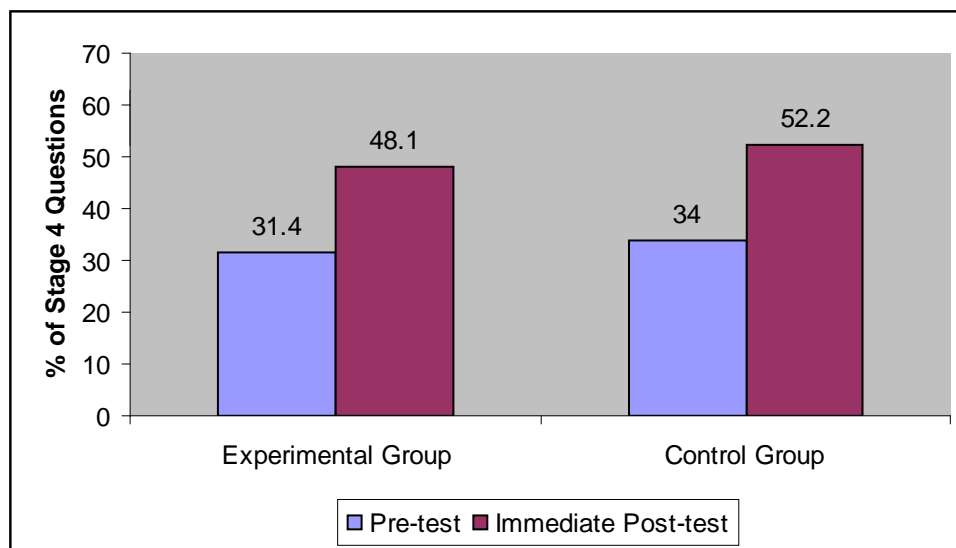

Figure 4.5: Correctly formed stage 4 questions in the pre-test and immediate post-test
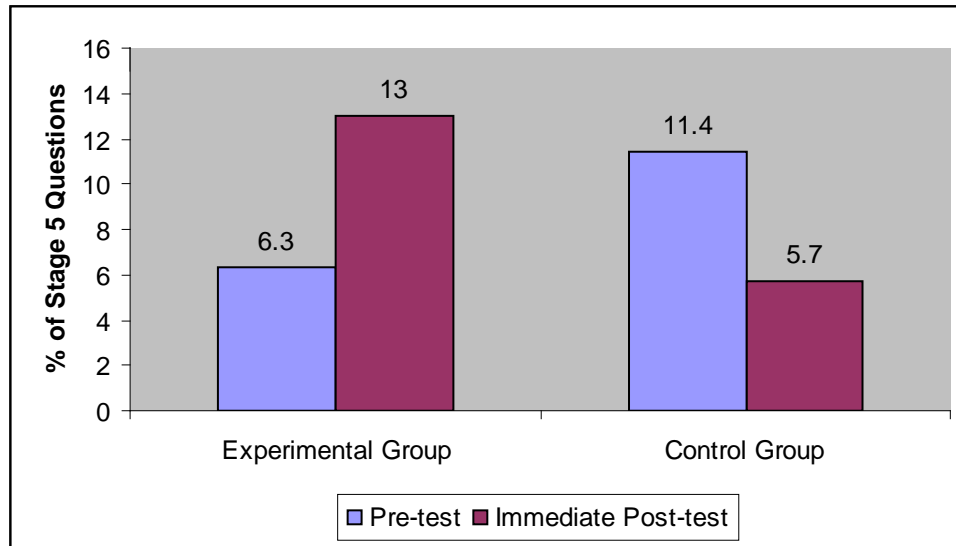
Figure 4.6: Correctly formed stage 5 questions in the pre-test and immediate post-test

Stage 4 questions went up from 31.4% in the pre-test to 48.1% immediately following treatment, so an increase of 16.7%. For stage 5 questions there was an increase of 6.7% between the pre-test and the immediate post-test (from 6.3% to 13%). These stage increases were expected and indicate that recasts may have been effective, for these learners, at least in the short term. On the other hand, when we look at the results for the control group (Table 4.4), there was an 18.2% increase in the percentage of correctly formed stage 4 questions only (from 34% to 52.2%). Below are some examples of stage 4 and 5 questions produced by participants from the experimental group in the immediate post-test. All of these four learners increased the proportion of stage 5 questions they produced correctly following treatment. The numbers (9, 10, 19 and 20) indicate the participant number, within the study.

Example 1:   9. Can you see the person in the newsagents? (*stage 4)* What is he doing? (*stage 5)*
          10. He is reading a newspaper. Can you see the scooter? (*stage 4)*

Example 2:   19. What is the man with backpack doing? (*stage 5)*
          20. He is waiting for bus.

Regarding stage 6 questions, both groups produced very few or none in either test. The experimental group produced 0% in the pre-test and 0.8% in the immediate post-test, whereas the control group produced 1.5% in the pre-test but 0% in the immediate post-

test. The stage 6 questions which were produced are interpreted as being outliers in the data, as they seem to be atypical of the rest of the dataset (Mackey & Gass, 2005). That is to say that they are not representative of the data elicited from the group as a whole. Due to these results making up such a small percentage of the total amount of questions produced, a decision was made not to further examine the contribution to development, of stage 6 questions.

4.3.2 Chi-Square Testing for Overall Significance in the Immediate Post-Test

Following descriptive statistics, inferential tests were also performed to analyse the statistical significance of the data. The reason for performing these inferential statistical tests was to investigate the extent to which the findings may be generalizable beyond the results of the current study. That is to say, to make inferences from the sample to the L2 learning population at large. Chi-square testing was chosen in order to determine whether a relationship may exist between the two categorical variables (question formation and exposure to intensive recasting). In other words, can a conclusion be made statistically, that the use of recasts in the classroom may benefit the production of question forms for the L2 learner?

With a level of significance (α) of 0.05 or 5% and 3 degrees of freedom (v) chi-square results for the total number of correctly formed questions in the immediate post-test were calculated. The critical chi-square statistic (χ^2_{crit}) was 7.815 by using a 2 x 4 table (stage 6 question data was excluded in all post-tests, therefore there were only 4 columns of scores). This means that if the resulting chi-square statistic has a value greater than 7.815, then we can conclude that in the short term, a relationship between exposure to recasts and question formation in the general population may be real. In other words, a statistically significant relationship exists. The raw frequency data for both groups is presented in Table 4.5 and 4.6.

Table 4.5: Correctly formed questions for the experimental group in the

pre- and immediate post-test

Developmental stage *	2	3	4	5	Total
Pre-test	56	83	70	14	**223**
Immediate post-test	28	63	115	31	**237**

* Pienemann and Johnston's (1987) developmental sequence for question formation in ESL

2 = SVO? E.g. *It's a monster?*
3 = Fronting: Wh/Do/Q-word. E.g. *Where the cats are?*
4 = Psuedo Inversion: Y/N, Copula. E.g. (Y/N) *Have you got a dog?* (Cop) *Where is the cat in your picture?*
5 = Q-word->Aux/modal->subj (main verb etc). E.g. *Where* (Q) *does* (Aux) *your cat* (subj) *sit*?
6 = Cancel Inversion, Neg Q, Tag Q. E.g. (Can Inv) *Can you see what the time is?* (Neg Q) *Haven't you seen a dog?*

Please refer to Chapter 3, section 3.4.2.1 for an outline of Pienemann and Johnston's (1987) table.

Table 4.6: Correctly formed questions for the control group in the

pre- and immediate post-test

Developmental stage *	2	3	4	5	Total
Pre-test	36	67	66	22	**194**
Immediate post-test	22	74	119	13	**228**

* Pienemann and Johnston's (1987) developmental sequence for question formation in ESL

2 = SVO? E.g. *It's a monster?*
3 = Fronting: Wh/Do/Q-word. E.g. *Where the cats are?*
4 = Psuedo Inversion: Y/N, Copula. E.g. (Y/N) *Have you got a dog?* (Cop) *Where is the cat in your picture?*
5 = Q-word->Aux/modal->subj (main verb etc). E.g. *Where* (Q) *does* (Aux) *your cat* (subj) *sit*?
6 = Cancel Inversion, Neg Q, Tag Q. E.g. (Can Inv) *Can you see what the time is?* (Neg Q) *Haven't you seen a dog?*

Please refer to Chapter 3, section 3.4.2.1 for an outline of Pienemann and Johnston's (1987) table.

For the experimental group, the chi-square statistic calculated from immediate post-test results was $\chi^2 = 29.05 > 7.815$. This result indicates a strong level of statistical significance for the use of recasts, at least immediately following treatment, at the 5% level of significance. However, the control group also showed a statistically significant result at 5% in the immediate post-test with a chi-square statistic of $\chi^2 = 18.03 > 7.815$. The fact that both groups showed statistically significant results overall, suggests that recasts may not make a difference to the performance of learners in

the production of correctly formed questions. The following section will discuss the significance of these results more fully.

4.3.3   Discussion of Results for Research Question One

Research question one investigated whether L2 learners who receive intensive recasts within interactive tasks show an immediate development of question formation. As discussed in Chapter 2, a large number of studies have found evidence to support the incorporation of recasts immediately following treatment (e.g. Iwashita, 2003; Leeman, 2003; Long, Inagaki & Ortega, 1998; Mackey & Philp, 1998; Philp, 2003). However, despite the findings of the studies mentioned here, the debate continues in the L2 learning field as to whether recasts are beneficial as a focus on form method of instruction. Two key issues remain unresolved. The first is that it is still not possible to draw conclusions on whether or not recasts' contribution to language development is as beneficial as their frequent classroom usage would suggest (Lyster, 1998; Lyster & Ranta, 1997; Nicholas, Lightbown & Spada, 2001; Panova & Lyster, 2002). The second is that it also continues to be problematic as to whether L2 learners perceive recasts as negative evidence or simply as input offering an alternative way to say the same thing as the learner's incorrect utterance (Allwright, 1975; Fanselow, 1977; Lyster, 1998; Lyster & Ranta, 1997; Panova & Lyster, 2002). This may have been the case in the current study following the immediate post-test, where even though the participants who had received the recasts showed improvement, those in the control group also improved, albeit to a lesser extent but not significantly so.

The proportion of stage 2 and 3 questions showed a large decrease in the immediate post-test for the experimental group. This result had been expected as a trade-off for the anticipated positive effects on the higher stage 4 and 5 questions following exposure to recasts. While this turned out to be the case and there was an increase in both stage 4 and 5 question production for the experimental group, the control group also showed an

increase in stage 4 questions. Therefore, except for stage 5 questions, overall the two groups seemed to follow a similar pattern of change immediately following treatment. It is interesting to note, however, that the participants in the experimental group showed improvement for stage 5 questions, whereas those in the control group did not. This may be because recasts have been shown to be beneficial for more developmentally advanced structures (Mackey & Philp, 1998). Even with this being the case, chi-square testing failed to show that the use of recasts along with interaction is of more benefit to L2 learners than interaction opportunities alone. Therefore, these results may indicate that the existence of immediate benefits for recasts still needs to be debated and requires further research, which could incorporate another form of corrective feedback, such as clarification or elicitation techniques, as well as recasts to compare the benefits of the different types.

The methodological approach of this study was modelled on that of Mackey and Philp (1998). In their experimental study examining the effects of negotiated interaction on the production and development of question forms, they found that the use of intensive recasts seemed to be of benefit for higher level learners in both immediate and short term interlanguage development. It was therefore expected in this study, that the results of the immediate post-test would also show improvement for the participants who were exposed to recasts, but not for those who took part in interaction alone (the control group). This expectation was due to the intermediate general English proficiency level of all the students. According to the data, this expectation seemed to have been met for the higher stage 5 questions for members of the experimental group, however, this was not the case for stage 4, as both groups showed improvement in the production of these question types. One possible reason for this result may be that the control group also showed improvement, due to the fact that they had taken part in the interaction, albeit without any exposure to recasts. In the L2 literature, there exists a large amount of support for development following on from interaction (Ellis & He, 1999; Ellis, Tanaka & Yamazaki, 1994; Gass & Varonis, 1994; Loschky, 1994; Mackey, 1999; Polio & Gass, 1998). The emphasis has been primarily on the role that negotiated interaction plays in the development of a second language, rather than considering conversational interaction

purely as an opportunity for practice of specific features of the L2. The research design used in the current study meant that all participants, those in the experimental group and the control group, took part in all three treatment sessions. Even though the control group participants were not offered any recasts at all when they produced grammatically incorrect questions, negotiation was obviously a part of the process required for each of the tasks, in order for the treatment sessions to be successfully completed. It may be that the interaction involved (asking and answering questions) during these interactive tasks, contributed to some improvement, across both groups immediately after the treatment.

Another possible reason for the improvement shown by the control group is that the tasks involved in the treatment sessions were all based on asking questions. This fact may have alerted the participants of the control group to become more aware of how they formed these questions. Consequently, as the immediate post-test followed straight after the final treatment session, this thought could still have been fresh in their minds, which may have influenced the improved results.

Finally, one reason for the lack of stage 6 question production may in fact be due to the higher developmental level of these types of questions, which could have been too advanced for the participants in the current study.

4.4 Research Question Two: Recasts and Longer Term Benefits for SLA

Although, this research question is similar to number one discussed above, question two focused on the delayed effects of recasts. That is to say, this question looked at whether recasts can be beneficial for L2 learners, beyond the immediate post-test. The data collected from the two delayed post-tests (one week and 12 weeks later) was analysed, using descriptive and inferential statistics. Key results are presented in tables and graphs. Following this section, will be a discussion of the significance of these results.

4.4.1   Proportion of Questions Produced in the Post-Test (One Week Later)

In Table 4.7 below, the percentages of correctly produced questions at each developmental stage are presented. These results were calculated from the raw frequency data, for the experimental group. Results are shown for the pre- and first post-test, one week later.

Table 4.7: Correctly formed questions for the experimental group in the pre-test and post-test (one week later)

Developmental stage *	2	3	4	5	6
Pre-test	25.1%	37.2%	31.4%	6.3%	0%
One Week Post-test	6.9%	32.8%	49.5%	10.8%	0%

* Pienemann and Johnston's (1987) developmental sequence for question formation in ESL

2 = SVO? E.g. *It's a monster?*
3 = Fronting: Wh/Do/Q-word. E.g. *Where the cats are?*
4 = Psuedo Inversion: Y/N, Copula. E.g. (Y/N) *Have you got a dog?* **(Cop)** *Where is the cat in your picture?*
5 = Q-word->Aux/modal->subj (main verb etc). E.g. *Where* **(Q)** *does* **(Aux)** *your cat* **(subj)** *sit?*
6 = Cancel Inversion, Neg Q, Tag Q. E.g. (Can Inv) *Can you see what the time is?* **(Neg Q)** *Haven't you seen a dog?*

Please refer to Chapter 3, section 3.4.2.1 for an outline of Pienemann and Johnston's (1987) table.

In Table 4.8 below, we present the percentages of correctly produced questions at each developmental stage, which were calculated from the raw frequency data, for the control group. Results are shown for the pre- and first post-test, one week later.

Table 4.8: Correctly formed questions for the control group in the pre-test and post-test (one week later)

Developmental stage *	2	3	4	5	6
Pre-test	18.6%	34.5%	34%	11.4%	1.5%
One Week Post-test	14.5%	37.1%	42.5%	5.4%	0.5 %

* Pienemann and Johnston's (1987) developmental sequence for question formation in ESL

2 = SVO? E.g. *It's a monster?*
3 = Fronting: Wh/Do/Q-word. E.g. *Where the cats are?*
4 = Psuedo Inversion: Y/N, Copula. E.g. (Y/N) *Have you got a dog?* **(Cop)** *Where is the cat in your picture?*
5 = Q-word->Aux/modal->subj (main verb etc). E.g. *Where* **(Q)** *does* **(Aux)** *your cat* **(subj)** *sit***?**
6 = Cancel Inversion, Neg Q, Tag Q. E.g. (Can Inv) *Can you see what the time is?* **(Neg Q)** *Haven't you seen a dog?*

Please refer to Chapter 3, section 3.4.2.1 for an outline of Pienemann and Johnston's (1987) table.

Interestingly, these results reflect those of the immediate post-test data for the experimental group, with the lower stage 2 questions showing a decrease and higher stage question changes (stage 4 and 5) showing increases. However, it is of interest to note that results for the control group show that the proportion of stage 3 questions increased slightly by 2.6% in this post-test. Regarding the higher 4 and 5 stages the control group only improved in stage 4. These results are illustrated in the bar charts below (figures 4.7 and 4.8).
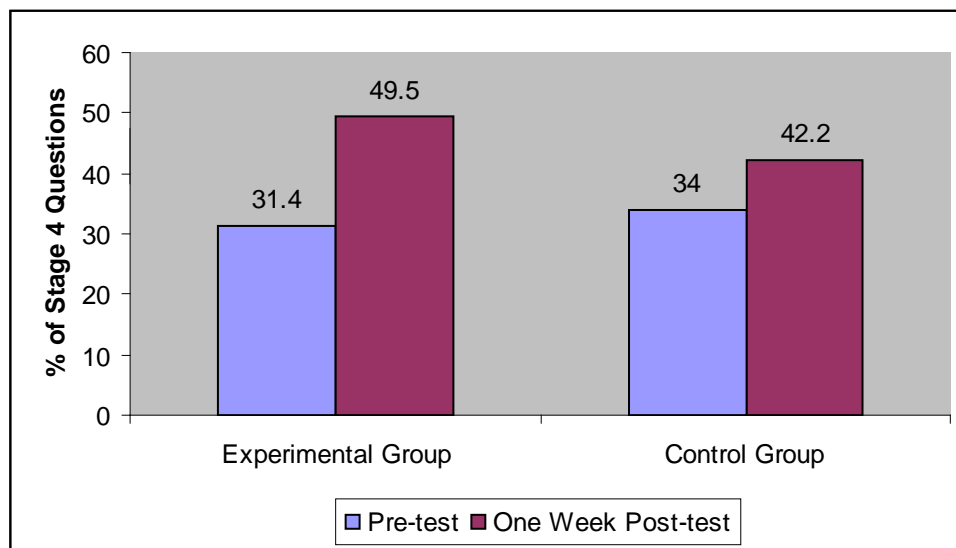
Figure 4.7: Correctly formed stage 4 questions in the pre-test and post-test (one week later)
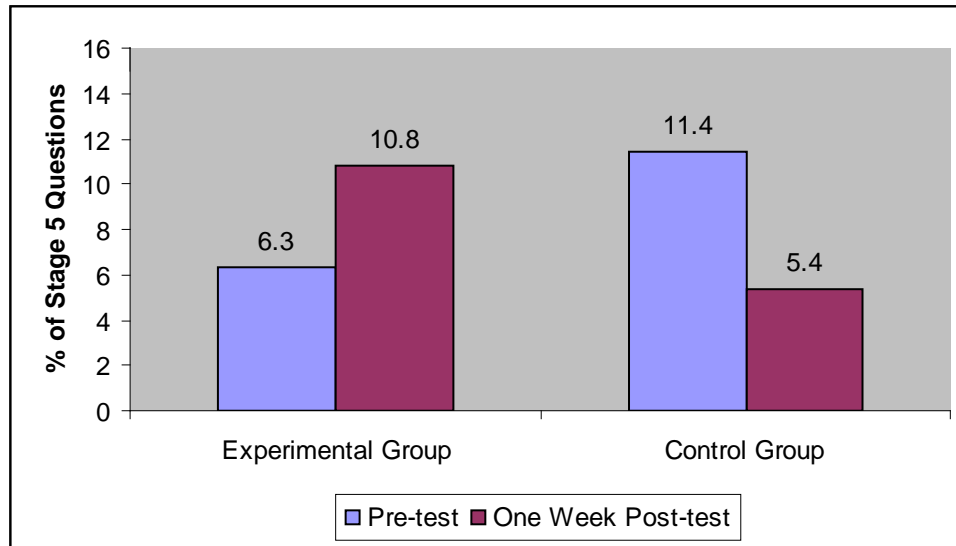
Figure 4.8: Correctly formed stage 5 questions in the pre-test and post-test (one week later)

For the experimental group (Table 4.7), stage 4 questions went from 31.4% in the pre-test to 49.5% one week following treatment, resulting in an increase of 18.1%. For stage 5 questions there was an increase of 4.5% between the pre-test and the post-test one week later (from 6.3% to 10.8%). Regarding the results for the control group (Table 4.8), there was an increase (8.2%) in the quantity of correctly formed stage 4 questions (from 34% to 42.2%) but a drop of 6% for stage 5 (from 11.4% to 5.4%).

4.4.2   **Proportion of Questions Produced in the Delayed Post-Test (12 Weeks Later)**

In Table 4.9 below, the proportion of correctly produced questions at each developmental stage are presented. These results were calculated from the raw frequency data, for the experimental group. Results are shown for the pre- and second delayed post-test (12 weeks later).

Table 4.9: Correctly formed questions for the experimental group in the

pre-test and post-test (12 weeks later)

Developmental stage *	2	3	4	5	6
Pre-test	25.1%	37.2%	31.4%	6.3%	0%
12 Week Post-test	10.6%	22.4%	50.9%	16.1%	0%

* Pienemann and Johnston's (1987) developmental sequence for question formation in ESL

2 = SVO? E.g. *It's a monster?*
**3 = Fronting: Wh/Do/Q-word. E.g.** *Where the cats are?*
4 = Psuedo Inversion: Y/N, Copula. E.g. (Y/N) *Have you got a dog?* **(Cop)** *Where is the cat in your picture?*
5 = Q-word->Aux/modal->subj (main verb etc). E.g. *Where* **(Q)** *does* **(Aux)** *your cat* **(subj)** *sit***?**
6 = Cancel Inversion, Neg Q, Tag Q. E.g. (Can Inv) *Can you see what the time is?* **(Neg Q)** *Haven't you seen a dog?*

Please refer to Chapter 3, section 3.4.2.1 for an outline of Pienemann and Johnston's (1987) table.

In Table 4.10 below, the proportion of correctly produced questions at each developmental stage, which were calculated from the raw frequency data, for the control group are presented. Results are shown for the pre- and second delayed post-test (12 weeks later).

Table 4.10: Correctly formed questions for the control group in the

pre-test and post-test (12 weeks later)

Developmental stage *	2	3	4	5	6
Pre-test	10.2%	38.1%	37.5%	12.5%	1.7%
12 Week Post-test	15.5%	31%	42.6%	10.9%	0 %

* Pienemann and Johnston's (1987) developmental sequence for question formation in ESL

2 = SVO? E.g. *It's a monster?*
3 = Fronting: Wh/Do/Q-word. E.g. *Where the cats are?*
4 = Psuedo Inversion: Y/N, Copula. E.g. (Y/N) *Have you got a dog?* **(Cop)** *Where is the cat in your picture?*
5 = Q-word->Aux/modal->subj (main verb etc). E.g. *Where* **(Q)** *does* **(Aux)** *your cat* **(subj)** *sit***?**
6 = Cancel Inversion, Neg Q, Tag Q. E.g. (Can Inv) *Can you see what the time is?* **(Neg Q)** *Haven't you seen a dog?*

Please refer to Chapter 3, section 3.4.2.1 for an outline of Pienemann and Johnston's (1987) table.

The delayed post-test (12 weeks later) results show the same pattern of change as the earlier delayed post-test outcomes even though there was a time difference of 11 weeks between these two tests. The experimental group displayed a sustained increase in the advanced stage 4 and 5 questions, whereas the control group showed an increase only for stage 4. These results are presented in figures 4.9 and 4.10 below.
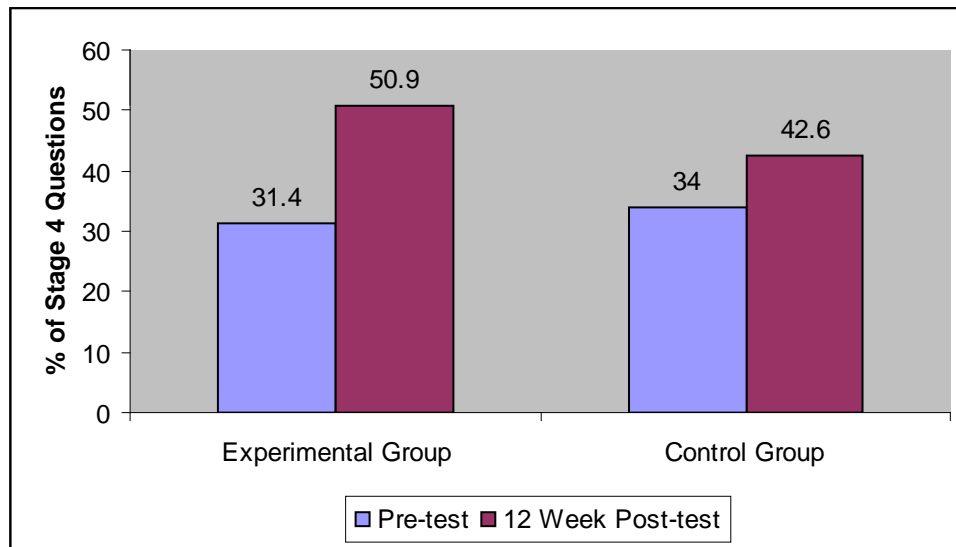
Figure 4.9: Correctly formed stage 4 questions in the pre-test and post-test (12 weeks later)
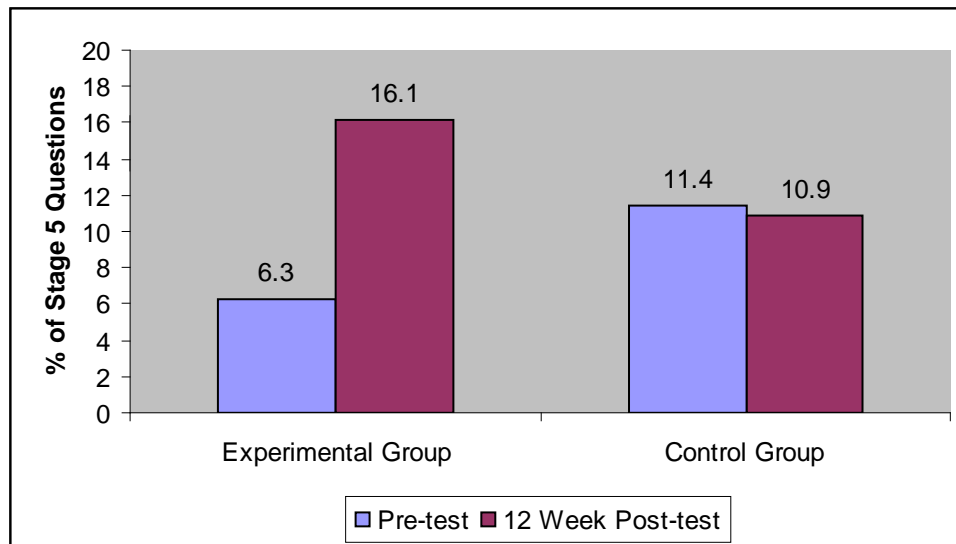
Figure 4.10: Correctly formed stage 5 questions in the pre-test and post-test (12 weeks later)

The percentages reported for the experimental group show an increase of 19.5% at stage 4 (from 31.4% to 50.9%) and an increase of 9.8% at stage 5 (from 6.3% to 16.1%) in the

12 week post-test. This pattern seems to indicate a delayed benefit in the use of recasts as a form of negative feedback. While the control group also showed an increase for stage 4 questions, it was a much smaller change than that of the experimental group at only 8.6% (from 34% to 42.6%). Stage 5 questions resulted in a small decrease in the 12 week post-test for the control group (from 11.4% to 10.9%).

4.4.3   Chi-Square Testing for Overall Significance in Delayed Testing

With a level of significance (α) of 0.05 or 5%, chi-square results for the total number of correctly formed questions in the two delayed post-tests were calculated. The degree of freedom (*v*) was 3 and the critical chi-square statistic (χ^2_{crit}) was 7.815 by using a 2 x 4 table. As with chi-square analyses for research question one, if the resulting chi-square statistic (χ^2) has a value greater than 7.815, then we can conclude that a relationship between exposure to recasts and question formation in the general population could exist. In other words, a statistically significant relationship is present.

The nominal, raw data for the delayed post-tests of the two groups is presented in Table 4.11 and 4.12 and the resulting chi-square test statistics for both the experimental and control groups in the delayed post-tests, one-week and twelve weeks later, are displayed in Table 4.13 below.

Table 4.11: Correctly formed questions for the experimental group in the
pre- and delayed post-tests

Developmental stage *	2	3	4	5	Total
Pre-test	56	83	70	14	**223**
Post-test (one week)	14	67	101	22	**204**
Post-test (12 week)	17	36	82	26	**161**

* Pienemann and Johnston's (1987) developmental sequence for question formation in ESL

88

2 = SVO? E.g. *It's a monster?*
3 = Fronting: Wh/Do/Q-word. E.g. *Where the cats are?*
**4 = Psuedo Inversion: Y/N, Copula. E.g. (Y/N)** *Have you got a dog?* **(Cop)** *Where is the cat in your picture?*
5 = Q-word->Aux/modal->subj (main verb etc). E.g. *Where* **(Q)** *does* **(Aux)** *your cat* **(subj)** *sit?*
**6 = Cancel Inversion, Neg Q, Tag Q. E.g. (Can Inv)** *Can you see what the time is?* **(Neg Q)** *Haven't you seen a dog?*

Please refer to Chapter 3, section 3.4.2.1 for an outline of Pienemann and Johnston's (1987) table.

Table 4.12: Correctly formed questions for the control group in the

pre- and delayed post-tests

Developmental stage *	2	3	4	5	Total
Pre-test	36	67	66	22	**194**
Post-test (one week)	27	69	79	10	**185**
Post-test (12 week)	20	40	55	14	**129**

* Pienemann and Johnston's (1987) developmental sequence for question formation in ESL

2 = SVO? E.g. *It's a monster?*
3 = Fronting: Wh/Do/Q-word. E.g. *Where the cats are?*
4 = Psuedo Inversion: Y/N, Copula. E.g. (Y/N) *Have you got a dog?* **(Cop)** *Where is the cat in your picture?*
5 = Q-word->Aux/modal->subj (main verb etc). E.g. *Where* **(Q)** *does* **(Aux)** *your cat* **(subj)** *sit?*
6 = Cancel Inversion, Neg Q, Tag Q. E.g. (Can Inv) *Can you see what the time is?* **(Neg Q)** *Haven't you seen a dog?*

Please refer to Chapter 3, section 3.4.2.1for an outline of Pienemann and Johnston's (1987) table.

Table 4.13: Chi-square test statistics from the delayed post-tests;

one week and 12 weeks later*

	One Week	12 Week
Experimental Group	$\chi^2 = 33.4$	$\chi^2 = 34.6$
Control Group	$\chi^2 = 5.81$	$\chi^2 = 2.18$

*$\chi^2_{crit} = 7.815$

It is apparent from the results in table 4.13 above that the experimental group chi-square test statistics were far greater than the critical chi-square value of 7.815 in both of the delayed post-tests ($\chi^2 = 33.4$ and $\chi^2 = 34.6$ respectively). These results illustrate a strong statistical significance in the delayed post-tests for the use of recasts on the production of questions. This positive result appears to be strengthened when we look at the chi-square test statistics for the control group in delayed testing. Both tests failed to produce a

statistically significant result ($\chi^2 = 5.81$ in the post-test, one week later and $\chi^2 = 2.18$, 12 weeks later). These statistical test results suggest a sound argument for the use of recasts in L2 learning environments, especially for delayed benefits. We will now discuss the significance of these results.

4.4.4   Discussion of Results for Research Question Two

In addition to immediate effects, the current study also investigated delayed effects of recasts at two different times following the treatment sessions; after one week and after 12 weeks. It was a primary goal of this study to contribute to the wider literature which focuses on delayed benefits of negative feedback techniques, such as recasts. While recasts, as a frequent type of conversational interaction have been a popular topic of research in the SLA field for more than two decades, it has been suggested that a greater number of longitudinal studies are needed, to explore the delayed effects of negative feedback and interaction (McDonough, 2005). Loewen and Philp (2006) agree that further research is required that investigates "the effectiveness of recasts in terms of learners' subsequent use of the linguistic items rather than through uptake alone" (2006, p. 542).

The results above for research question two show an increase in the proportion of stage 4 and 5 questions, correctly produced by the experimental group in both delayed post-tests (one week and 12 weeks later). These results indicate delayed benefits for learners who have been exposed to recasts and add support to previous research, which has suggested that the use of recasts as a form of negative feedback can be effective (Ellis, Loewen & Erlam, 2006; Gass, 1997; Lightbown, 1998; Mackey, 1999). The linguistic target of this research was the production of questions and a common theme in studies looking at question formation is that negative feedback may benefit more advanced linguistic structures (Mackey, 1999; Mackey & Philp, 1998; McDonough, 2005). The results here add support to the claim that development may not be only seen as an immediate effect of

interaction (Mackey, 1999), as the increase in developmentally more advanced structures for the experimental group was sustained in delayed post-tests. In fact, the largest improvement in the higher level stage 5 questions took place in the second delayed post-test (12 weeks later). This was 9.8% compared with 6.7% and 4.5% in the immediate and first delayed post-test (one week later) respectively. Furthermore, it is suggested that a learner needs to be ready to acquire language in order for acquisition to take place (Pienemann, 1989) and we will now discuss this theory in relation to research question two.

These results indicate that the intermediate level learners involved in this study may have been developmentally ready to benefit from the focus on form instruction they were exposed to in their treatment sessions, namely intensive recasts of all incorrect questions. If this was the case, then these students may have received the negative feedback at a prime time, to then be able to utilise subsequent opportunities to produce more target-like utterances (in the post-tests). Developmental readiness may have played a significant role for the experimental group in recognising and accessing the corrective potential of the recasts (Han, 2002) and consequently could have contributed to delayed improvement.

Regarding the statistical significance of the findings, in contrast to the results for research question one, chi-square testing of the data from the delayed post-tests strongly indicated that the use of recasts along with interaction is of more benefit than taking part in interaction alone. The data analysis for the experimental group came up with chi-square statistics of $\chi^2 = 33.4$ from the first delayed post-test (one week later) and $\chi^2 = 34.6$ from the second delayed post-test (12 weeks later). Keeping in mind that the critical chi-square value was $\chi^2 = 7.815$, these results were compelling. On the other hand, the chi-square results for the control group were $\chi^2 = 5.81$ and $\chi^2 = 2.18$ in the delayed post-tests respectively. These findings add support to the interaction hypothesis claim that recasts as a type of negative evidence can be effective in promoting SLA. The results of the current study indicate that the positive benefits of interaction can be enhanced when intensive recasts are provided (Mackey & Philp, 1998).

All participants in this study took part in task-based interaction. Iwashita (2003) suggested that this type of interaction can make the language forms essential or at least useful to the task, as well as be facilitative of learning. The results discussed in this section lend support to this argument, as producing a large number of comprehensible questions was necessary for the participants, in order to complete their information gap tasks. Furthermore, by taking part in task-based interaction, the participants from the experimental group were offered a recast of their non-target-like questions. This negotiated interaction, centred on completing a task, may have contributed to increased confidence or familiarity with using these question types (Mackey, 2006; Mackey & Philp, 1998) for the members of the experimental group and consequently may have led to the improvement shown in the findings. However, it is important to be aware of the fact that the learners continued with their classroom lessons during the time between the immediate and delayed post-tests as well as having many opportunities to hear the target language outside of the classroom. This subsequent exposure may have contributed to the delayed pattern of improvement displayed by the experimental group (Mackey, 1999). On the other hand, even though members of the control group also had an equal amount of language instruction and outside contact, over the same time period, the findings show that they failed to make the same level of improvement in their production of questions. This is an important point as the opportunity for additional input and practice was the same for both groups, yet the delayed test results were different.

Keeping this in mind, the improvement from the experimental group may, at least in part, be due to the recasts which they were exposed to during the three treatment sessions. Future studies could take the amount and type of further instruction into account and factor this information in, when delayed post-test results are analysed. Furthermore, participants could be asked to keep a diary record of the amount of outside contact they experience during the period of time between the final treatment session and last delayed post-test.

As mentioned earlier, the current study was modelled partially on that of Mackey and Philp (1998). In their study, modified output or uptake was one focus, where the

immediate response of the learner following the recast was recorded and analysed. Some researchers believe that the amount of uptake which results from being exposed to recasts may not be a reliable measure of success and that the noticing of recasts is not dependent on uptake (Loewen & Philp, 2006) nor that uptake necessarily indicates that acquisition has taken place (Ellis, Basturkmen & Loewen, 2001). Despite the suggestion that recasts may not be useful when it comes to correcting student errors, due to low levels of uptake and repair (Lyster & Ranta, 1997; Lyster & Panova, 2002), the current study lends support to the argument that immediate uptake is not always necessary for acquisition to take place (Mackey & Philp; McDonough, 2005). It is interesting to note that, in the current study, the students in the experimental group rarely repeated the recast, yet improvement was noticeable in the delayed post-tests. Further research which investigates delayed effects and modified output, at the same time, could be useful in delving further into this issue.

The answers provided here to research question two suggest that the use of recasts in the L2 classroom as incidental focus on form techniques could be productive for learners. Other feedback forms have their benefits (see Lyster, 1998, 2004), such as clarification requests, which some researchers suggest, have a higher rate of uptake than recasts (Lyster, 1998; Panova & Lyster, 2002). However, due to the fact that recasts maintain the focus on meaning and are time-saving, non threatening to the learner and less intrusive to the flow of interaction (Loewen & Philp, 2006), they certainly seem to be a useful pedagogical tool in the modern day communicative L2 classroom setting.

4.5   Research Question Three: Perception of Negative Evidence in Recasts

Research question three was a secondary question which was concerned with whether or not participants in the experimental group were aware of the fact that the recasts offered in response to their question form error, in fact, contained corrective feedback. Following the second and final delayed post-test (12 weeks later) an email was sent to the 18

participants in the experimental group. The question was asked regarding whether or not these learners were aware that the researcher had been providing a grammatical correction following their incorrectly formed question, or was this feedback regarded as simply a repetition of or an alternative way to form their question? Twelve out of the 18 students in this group sent a response to the email. It is thought that the email addresses provided on the initial consent forms were incorrect or no longer in operation for the students who did not respond. The answers were as follows:

Table 4.14: Students' perceptions of negative evidence

I was <u>aware</u> that the researcher was offering negative evidence in response to my error.	8
I was <u>unaware</u> that the researcher was offering negative evidence in response to my error.	4

As can be seen from the data presented in Table 4.14 above, two-thirds of the participants who responded to the question, regarding awareness of negative evidence answered that they had been aware of the corrective nature of the researcher's feedback (8 out of 12 responses). On the other hand, one third said that they had been unaware of this negative evidence (4 out of 12).

4.5.1   Discussion of Results for Research Question Three

This question examined the perceptions of the participants in the experimental group, regarding the corrective nature of the recasts. In other words, were they aware of the negative evidence involved in the researcher's feedback? A suggestion in the literature is that recasts may help draw learners' attention to the language forms that they have produced and help them to identify gaps in their L2 development (Gass, 2005; Long, 2006). In this study, if the learners in the experimental group did perceive these gaps,

then grammar restructuring and improvement in question production may have been achievable as a result. It is, however, important to note, as in the Mackey and Philp (1998) study that improvement was only considered in relation to question formation and that no claims about the overall development of participants' target language grammar are made.

It has been suggested that learners must notice and recognise the fact that they have made an error and also that negative feedback is contained within the recasts, in order to address their errors (Gass, 2005; Long, 1996, 2006). The improvement that the experimental group displayed in the delayed post-tests when compared to the control group may be due to the fact that in some way they noticed the corrective intention of the feedback. In addition, the responses from the post-data collection email indicate that this did, in fact, occur. However, the timing of the interaction could also have played an important role in this awareness. When the attention of both the learners and the native speaker is jointly focussed on the target language, then negative feedback can be beneficial (Long, 2006). During all three of the treatment sessions, recasts were provided incidentally, directly after an incorrect question was produced. The learners may have been more aware, whether consciously or not, of the error, due to this timing. Furthermore, in order to complete the task at hand, the learners would probably have been motivated and wanting to be understood, which are also conditions likely to assist noticing of the language forms being used (Long, 2006).

On the other hand, it has been suggested that recasts may be seen as input offering an alternative way to say the same thing as the incomplete form produced by the learner (Panova & Lyster, 2002). Lyster (1998) argued that this could be particularly true in a classroom setting where the primary focus is on meaning. However, it is important to note that the learner type was different in the Lyster (1998) study. In his study they were young L2 learners who were potentially less cognitively able to perceive what is unacceptable in the L2 through recasts, compared with the mature, adult L2 learners in the current study, who are thought to also possess a higher proficiency of English.

Despite the suggestion that recasts may be considered by learners as an alternative way of uttering the same thing, the results from the delayed post-tests as well as the positive responses for the majority of those who replied to the email suggest support for the use of recasts as a beneficial feedback technique when L2 learners are developmentally ready.

4.6   Summary

In this chapter, three research questions relating to immediate and delayed benefits of using recasts in the L2 classroom, as well as perceptions of negative evidence were addressed. For the primary research questions (one and two) descriptive and inferential statistical analyses were carried out to investigate development in question formation following exposure to negative evidence, namely recasts, provided as a corrective feedback technique. For the secondary research question three, a post-data-collection email was sent to the participants in the experimental group to ask them whether they had perceived the recasts as containing negative evidence. The findings were discussed with reference to each of the research questions and also considered in relation to relevant previous studies. Support was found for the following claims in the literature:

- The use of recasts as a form of negative feedback can be effective for SLA.

- Development may not be only seen as an immediate effect of interaction as delayed benefits were apparent.

- Negotiated interaction, centered on completing a task, can increase confidence and familiarity when forming questions.

- Negative feedback, such as recasts, may benefit more advanced structures, when the linguistic focus is on question formation.

- The use of recasts along with interaction is of more benefit than taking part in interaction alone.

- Developmental readiness may play a key role in recognising and accessing the corrective potential of the recasts.

- When the learners are aware of the negative evidence contained in corrective feedback, development can occur.

- Recasts may help draw learners' attention to the language forms that they have produced and help them to identify gaps in their L2 development.

The results here indicate support for the two main issues relating to the use of recasts as negative evidence in the L2 field, which were discussed in Chapter 2 as well as in the discussion sections above. First of all, the results of this study, signifying delayed benefits of recasts contributing to SLA, suggest that the high frequency of recast use in the L2 classroom may in fact be justified. Secondly, the fact that the majority of participants from the experimental group indicated that they had perceived the recasts as providing negative evidence, suggests that recasts as a form of corrective feedback can be effective for L2 learning as long as the learners are developmentally ready to receive it.

On the other hand, it is possible that due to all the testing and treatment sessions involving the necessity to ask a large amount of questions, this may have alerted the learners to the fact that the focus was on question formation. Furthermore, outside exposure and continued instruction in the classroom, between the treatment and final delayed post-test (12 weeks later) may also have contributed to delayed benefits. However, it is important to note that any such benefits would also have been available for the participants in the control group, who did not perform as well as the experimental group in delayed testing.

Overall, these findings indicate support for continued use of recasts as a useful corrective feedback technique in the L2 classroom. However, it is important to remember that improvement was examined in relation to question formation and therefore, no claims can be made about development of the participants' L2 grammar as a whole.

CHAPTER 5
CONCLUSION

5.1   Introduction

This chapter firstly presents a review of the aims and methodological approach of the study. Secondly, the main findings from the research are summarised, followed by a suggestion of the significance of these findings for theory and research development, as well as pedagogical implications for L2 teachers and institutions. Recommendations for future research in this field are proposed. Lastly, the limitations of the study are considered and the chapter concludes with an overview of the preceding sections.

5.2   Aims of the Research

The primary aim of this study was to investigate the relationship between exposure to recasts and the production of questions in task-based interaction; and in particular, whether delayed benefits (one week and 12 weeks following treatment) exist, even when immediate effects are not apparent. The secondary aim was to investigate whether L2 learners perceive the corrective nature of recasts, offered when an error has been made. Alternatively, do they think that recasts act simply as confirmation of meaning? The study focused on answering the following three research questions:

- Can the incorporation of recasts in task-based interaction benefit the immediate production of question formation in ESL?
- Can the incorporation of recasts in task-based interaction benefit the production and development of question formation over time (at one week and twelve weeks following treatment) in ESL?

- Do the participants in the experimental group perceive the recasts as containing negative evidence?

5.3   Methodological Approach

This research was carried out at a university language center in Auckland (AUT IH), with intermediate level students from a range of ethnic backgrounds. The study involved an experimental group and a control group. A pre-test/post-test methodological approach was adopted in order to collect data at four different time periods over a four month period (from pre-test to the final delayed post-test – 12 weeks following treatment). In addition, a student questionnaire was administered to collect individual characteristics of the participants in order to explore whether results of the study could be influenced by demographic features. Lastly, a post data-collection email was sent to the experimental group participants to examine their perceptions of the recasts they were exposed to during the treatment sessions. In other words, were they aware that these recasts were intended to offer corrective feedback in response to question formation errors?

5.4   Summary of Key Findings

A preliminary finding from looking at the descriptive statistics was that, for more developmentally advanced structures (question stage 5), recasts seemed to benefit learners immediately following treatment. However, when inferential statistical tests were performed, these results did not prove to be statistically significant. In other words, chi-square testing failed to show that the use of recasts along with interaction is of more benefit to SLA than interaction opportunities alone. While previous studies have demonstrated support for taking part in interaction and exposure to recasts in this immediate turn, the findings of the current study do not add support to the literature.

In spite of a lack of significance for immediate benefits, a key finding of this study is the effectiveness of recasts on delayed L2 production. The two delayed post-tests produced statistically significant results favouring recasts along with interaction. That is to say, the experimental group showed improvement in more developmentally advanced question stages, in the longer term, whereas the control group did not. These statistical results show support for the use of recasts in L2 learning environments because even though immediate benefits were not evident, delayed effects were significant.

A secondary finding was that contrary to much of the previous research, more than half of the participants in the experimental group indicated that they had, in fact, been aware of the corrective nature of the recasts. These findings suggest that recasts can be beneficial in L2 instruction, at least when questions are the linguistic form under investigation.

5.5   Significance of Findings for Theory and Research Development

The results summarised above add support to the claims that recasts, as a corrective feedback technique used in response to student error, can be beneficial for delayed learning. It is believed that these positive results came about in relation to the corrective feedback which the participants were exposed to, rather than because of any other factor, such as further instruction or outside language exposure. The reason for this belief is that these other conditions were the same for the members of the control group, who did not produce similarly strong results. The only difference between the two groups was the provision of recasts (experimental group) or absence of recasts (control group).

The findings also support previous suggestions that immediate uptake may not be necessary for learning to take place (Mackey & Philp, 1998; McDonough, 2005; McDonough & Mackey, 2006). Higher level stage 5 questions showed sustained development in both delayed post-tests and, in fact, the largest improvement took place in

the final post-test, twelve weeks after treatment. However, a student needs to be ready in order to acquire the linguistic form (Pienemann, 1989) and if the level of input is too advanced, this acquisition will not occur. This may have been the case for stage 6 questions, as there was a lack of production of these question forms in the tests. Therefore, it is believed that this type of question was still too advanced for these learners, or it may have not been available in the input in a significant enough amount to affect the output. Furthermore, as Philp (2003) suggests, recasts which differ greatly from the learners' original attempt are difficult to comprehend and the correct L2 version is less likely to be repeated. It is a consideration of this research that for these intermediate level students, stage 6 questions were beyond their current developmental level and therefore, not possible to acquire at this time.

5.6   Pedagogical Implications

Recasts are generally implicit in nature and do not unduly interrupt the communicative flow in the classroom. They are also time-saving and non-threatening to the learner (Loewen & Philp, 2006) and therefore seem to be a useful way of providing feedback in the modern L2 learning situation. Furthermore, even though doubt has been cast on the effectiveness of recasts on correcting student errors, the results presented here do not support this claim and, in fact, add weight to the suggestion that recasts should be an integral part of a teacher's feedback approach. The fact that many of the learners in the experimental group indicated awareness of the corrective nature of the recasts, suggests that as long as the students are developmentally ready, then the use of recasts to provide negative evidence is worthwhile. In addition, the timing of the interaction, when both the attention of the learner and the NS are on the target language at the same time may further add to the benefits of the negative feedback. It is important to note, however, as mentioned earlier, that the current study was looking at recasts in relation to producing questions and therefore caution should be applied when generalizing these results to other linguistic forms (McDonough & Mackey, 2008).

Lastly, a suggestion which could be of benefit to teachers is that they may like to explain to their students that during the course, recasts which they provide will contain target-like corrections and therefore when they are noticed in later communicative interactions, the learners will be more aware of what they are communicating.

5.7   Limitations of the Study

Regarding the lack of significance of results in the immediate post-test, this finding may have been due, at least in part, to the interaction opportunities which the participants in the control group were exposed to. Despite the fact that they were not offered any recasts in response to their errors, in order to achieve task outcome, negotiation of meaning was involved, which required asking and answering a number of questions. The learners in the control group may have been alerted to the fact that question formation was the key target feature and therefore they may have become more aware of the way they were producing this particular linguistic form. Furthermore, the literature suggests that the negotiation involved in task-based interaction can contribute to increased confidence or familiarity with the language forms as well as facilitate learning (Iwashita, 2003, Mackey & Philp, 1998). This claim along with the students' motivation to successfully complete the task could account, at least in part, for the results.

Another potential limitation of this research regards the length and frequency of the treatment sessions. This study incorporated three sessions, one per week, lasting 10 – 15 minutes each. Longer treatment sessions with less of a gap in-between may have been more effective (Mackey, 2006; McDonough & Mackey, 2008; Nassaji, 2009). In the Mackey and Philp (1998) study, participants took part in five initial sessions, on consecutive afternoons, lasting 15-25 minutes each. These sessions were then followed up with three more 15 – 20 minute sessions spread over three weeks. Overall, the Mackey and Philp study design incorporated treatments which were closer together and in a greater number than the ones in this study. However, due to the following reasons, the

design of the data-collection process in the current study was considered to be the most suitable:

- the number of participants involved
- the researcher being the sole administrator of testing and treatment sessions, owing to other teachers at IH being too busy to assist, and
- the limited time outside of classes, during which students could participate

Future research could attempt to replicate Mackey and Philps' study more closely in order to see whether long-term retention would be further enhanced.

5.8   Recommendations for Future Research

While this study adds support to the claim that recasts can benefit SLA, the linguistic form under investigation here was question formation. This was a form which all students already had some knowledge of and therefore future research is required which focuses on new linguistic forms in order to see whether similar findings would result (Ellis, Loewen & Erlam, 2006).

Furthermore, future studies could explore in more detail whether L2 learners involved in the research are aware of the fact that the recasts offered are in fact corrective, in nature. In the current study, after the entire data collection process was completed, an email was sent to all the participants in the experimental group to ask them this question and results of this step were presented and discussed in the previous chapter (section 4.5 and 4.5.1). However, as this email was sent at a later date, following the completion of the treatment sessions and tests, the participants may have found it difficult to remember the type of feedback which they had been exposed to by this time. For that reason, a more detailed questionnaire or interview with each participant immediately following the final post-test could be administered in future research to investigate this issue in more depth.

Another area which could add to the literature on the effectiveness of recasts on SLA is experimental research where emphasis or a sign is given to the students to indicate that an error has occurred, before the recast is offered. Doughty and Varela (1998) repeated the error as a way of incorporating an attention getting phase into the interaction. This does, of course, increase the explicitness of the feedback and add potential disruption to the communicative nature of the task. On the other hand, it could make the corrective feedback more noticeable to the learners and therefore, lead to greater effectiveness. The recasts in the current study were designed to be of an implicit nature and so minimal disruption was an aim of this study. However, more explicit studies could compare results to explore whether an attention-getting phase adds more benefit to the findings. Alternatively, a single study design which includes both types of corrective feedback – implicit and explicit – to investigate whether there is a difference, could generate useful findings.

As discussed in Chapter 2, many types of corrective feedback exist. Debate continues in SLA research as to whether feedback which gives the learners the answer, such as using recasts, is as effective as those types of feedback which require the students to produce the answer, such as clarification requests and elicitation (Lyster, 1998; Lyster & Ranta, 1997). However, as recasts were the only feedback technique under investigation in the current study, future task-based research involving adult L2 learners could look at other types which require the participants to come up with the answer themselves, as well as exposure to implicit recasts in order to compare findings.

Another important issue concerns outside L2 exposure and subsequent classroom instruction. In this study, all participants continued with regular lessons in between treatment and testing sessions. Furthermore, as they were living in an English speaking country during the research process, they were exposed to an unlimited amount of English outside of the classroom as well. Future studies could consider this extra input when interpreting results. Possible ways to do this include: the participants keeping a record of how often they encounter the L2 outside of the research or conducting a number of short interview sessions throughout the research process which could help the

researcher to gain an understanding of the amount of extra L2 contact involved. However, it is important to remember that the participants in the control group were also exposed to the same conditions as those in the experimental group, but without the same positive results, therefore, these factors may not have had significant impact on findings.

Regarding stage 6 questions, future research could utilise tasks which allow more opportunities for production of this type of question as the tasks in the current study may not have included enough of these kinds of chances. However, as mentioned above, another reason for the lack of stage 6 questions may have been due to the developmental level of the participants.

This section has discussed recommendations of the researcher in regards to future research involving recasts and SLA. The following and final section of this chapter is a summary of the research and its key findings.

5.9   Final Comment

The primary purpose of this study was to investigate the relationship between recasts, as a frequent corrective feedback technique in the L2 classroom and question production. The most important finding from this research was that despite the absence of significant results for the use of recasts for question formation in the immediate post-test, delayed benefits were evident and this adds support to the literature regarding the use of recasts in the L2 classroom. One of the key issues in the debate which surrounds recast use is whether high pedagogical usage is worthwhile. The study provided valuable empirical evidence that the common use of recasts as a corrective feedback technique in the L2 classroom seems to be justified. Findings also suggest that as long as the learners are developmentally ready to benefit from this type of negative feedback, then using recasts can be an effective tool in the classroom. Another positive finding from the research was that overall, learners seem to be aware that recasts contain negative evidence and

therefore offer corrections to their linguistic error. This knowledge can lead to restructuring of their L2 grammar and ultimately, SLA.

REFERENCES

Abbuhl, R., & Mackey. A. (2008). Second language research methods. In K. King & N. Hornberger (Eds.), *Encyclopedia of Language and Education,* Vol. 10, Research Methods. New York: Springer.

Adjemian, C. (1976). On the nature of interlanguage systems. *Language Learning*, *26,* 297-320.

Allwright, R. (1975). Problems in the study of the language teacher's treatment of learner error. In M. Burt & H. Dulay (Eds.), *On TESOL '75: New directions in language learning, teaching, and bilingual education* (pp. 96-109). Washington, DC: TESOL.

Ayoun, D. (2001). The role of negative and positive feedback in the second language acquisition of the passé composé and imparfait. *Modern Language Journal, 85,* 226-243.

Braidi, S. M. (2002). Reexamining the role of recasts in native-speaker/nonnative-speaker interactions. *Language Learning, 52,* 1-42.

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

Carroll, S. (1997). The irrelevance of verbal feedback to language learning. In L. Eubank, L. Selinker, & M. Sharwood Smith (Eds.), *The current state of interlanguage* (pp. 73-88). Amsterdam: John Benjamins.

Chomsky, N. (1959) On certain formal properties of grammars. *Information and Control, 1*, 91-112.

Corder, P. (1967) The Significance of Learner's Errors. *IRAL, 5,* 161-170.

Donato, R. (1994). Collective scaffolding. In J. P. Lantolf & G. Appel (Eds.), *Vygotskian approaches to second language research* (pp. 33-56). Norwood, NJ: Ablex.

Dornyei, Z. (2007). *Research methods in applied linguistics: quantitative, qualitative, and mixed methodologies*. Oxford. Oxford University Press.

Doughty, C. (1994). Fine-tuning of feedback by competent speakers to language learners. In J. E. Alatis (Ed.), *Georgetown University roundtable on languages and linguistics* (pp. 96-108). Washington, DC: Georgetown University Press.

Doughty, C. (2005). Instructed SLA: Constraints, compensation, and enhancement. In C. Doughty & M. Long (Eds.), *The handbook of second language acquisition* (pp. 256-310). Oxford: Blackwell.

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). Cambridge: Cambridge University Press.

Doughty, C., & Williams, J. (Eds.). (1998). *Focus on form in classroom second language acquisition.* Cambridge: Cambridge University Press.

Egi, T. (2004). *Recasts, perceptions and L2 development.* Unpublished doctoral dissertation, Georgetown University, Washington, DC.

Ellis, R. (1985). *Understanding second language acquisition.* Oxford: Oxford University Press.

Ellis, R. (1991). The interaction hypothesis: A critical evaluation. In E. Sadtono (Ed.), *Language acquisition in the second/foreign language classroom* (pp. 179-211). Singapore: SEAMEO Regional Language Centre.

Ellis, R. (1994). Uptake as language awareness. *Language Awareness, 4,* 147-160.

Ellis, R. (2001). Investigating form-focused instruction. *Language Learning, 51(Suppl. 1),* 1-46.

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

Ellis, R. (2006). Researching the effects of form-focussed instruction on L2 acquisition. In K. Bardovi-Harlig & Z. Dornyei (Eds.), *Themes in SLA research: AILA review, Volume 19* (pp. 18-41).

Ellis, R. (2008). *The study of second language acquisition.* 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, 28,* 339-368.

Ellis, R., & Sheen, Y. (2006). Re-examining the role of recasts in L2 acquisition. *Studies in Second Language Acquisition, 28,* 575-600.

Ellis, R., Tanaka, Y., & Yamazaki, A. (1994). Classroom interaction, comprehension, and the acquisition of L2 word meanings. *Language Learning, 44,* 449-491.

Erlam, R (2006). Elicited imititation as a measure of L2 implicit knowledge: An empirical validation study. *Applied Linguistics 27,* 464-491.

Fanselow, J. (1977). The treatment of error in oral work. *Foreign Language Annals, 10,* 583-593.

Farrar, M. J. (1992). Negative evidence and grammatical morpheme acquisition. *Developmental Psychology, 28,* 90-98.

Fraenkel, J., & Wallen, N. (2003). *How to design and evaluate research in education.* (5th ed.). New York: McGraw-Hill.

Gass, S. M. (1991). Grammar instruction, selective attention and learning processes. In R. Phillipson, E. Kellerman, L. Selinker, M. Sharwood Smith, & M. Swain (Eds.), *Foreign/second language pedagogy research* (pp. 134-141). Clevedon, UK: Multilingual Matters.

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

Gass, S. M. (2005). Input and interaction. In C. J. Doughty & M. H. Long. (Eds.), *The handbook of second language acquisition* (pp. 224–255). Oxford: Blackwell.

Gass, S. M., & Mackey, A., & Pica, T. (1998). The role of input and interaction in second language acquisition: Introduction to the special issue. *The Modern Language Journal, 82,* 299-307.

Gass, S. M., & Varonis, E. (1994). Input, interaction and second language production. *Studies in Second Language Acquisition, 16,* 283–302.

Glesne, C., & Peshkin, A. (1992). *Becoming qualitative researchers: An introduction.* White Plains, NY: Longman.

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

Hatch, E. (1978). Acquisition of syntax in a second language. In J. Richards (Ed.), *Understanding second and foreign language learning* (pp. 34-70). Rowley, MA: Newbury House.

Hatch, E., & Lazaraton, A. (1991). *The research manual: Design and statistics for applied linguistics.* New York: Newbury House.

Havranek, G. (2002). When is corrective feedback most likely to succeed? *International Journal of Educational Research, 37,* 255-270.

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.

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

Iwashita, N. (2003). Negative feedback and positive evidence in task-based interaction: Differential effects on L2 development. *Studies in Second Language Acquisition, 25,* 1-36.

Johnson, R., & Christensen. L. (2004). *Education research: quantitative, qualitative, and mixed approaches.(* 2nd ed.). Boston: Allyn and Bacon.

Krashen, S. D. (1977). The monitor model for adult second language performance. In M. Burt, H. Dulay, & M. Finocchiaro (Eds.), *Viewpoints on English as a second language* (pp. 152-161). New York: Regents.

Krashen, S. (1981). *Second language acquisition and second language learning.* Oxford: Oxford University Press.

Krashen, S. D. (1982). *Principles and practices in second language acquisition.* London: Pergamon.

Krashen, S. D. (1985). *The Input Hypothesis: Issues and implications.* London: Longman.

Krashen**,** S. D. (1989). We acquire vocabulary and spelling by reading: Additional evidence for the input hypothesis. *The Modern Language Journal, 73,* 440-464.

Krashen, S. 1993. *The power of reading.* Englewood, CO: Libraries Unlimited.

Leeman, J. (2003). Recasts and second language development. *Studies in Second Language Acquisition, 25,* 37–63.

Lightbown, P. (1998). The importance of timing in focus on form. In C. Doughty & J. Williams (Eds.), *Focus on form in classroom second language acquisition* (pp. 177-196). New York: Cambridge University Press.

Lightbown, P. & Pienemann, M. (1993). Comments on Stephen D. Krashen's 'Teaching issues: Formal grammar instruction'. *TESOL Quarterly, 27,* 717-722.

Lightbown, P., & Spada, N. (2006). *How languages are learned.* Oxford: Oxford University Press.

Loewen, S. (2002). *The occurrence and effectiveness of incidental focus on form in meaning-focused ESL lessons.* Unpublished doctoral dissertation, The University of Auckland, New Zealand.

Loewen, S. (2004). Uptake in incidental focus on form in meaning-focused ESL lessons. *Language Learning, 54,* 153-187.

Loewen, S. (2005). Incidental focus on form and second language learning. *Studies in Second Language Acquisition, 27,* 361-386.

Loewen, S., & Philp, J. (2006). Recasts in the adult L2 classroom: characteristics, explicitness and effectiveness. *Modern Language Journal, 90*, 536-556.

Long, M. H. (1977). *Face to Face.* London: Evans Bros.

Long, M. H. (1980). *Input, interaction and second language acquisition.* Unpublished doctoral dissertation, University of California, Los Angeles.

Long, M. H. (1981). Questions in foreigner talk discourse. *Language Learning, 31,* 135-158.

Long, M. H. (1983). Linguistic and conversational adjustments to non-native speakers. *Studies in Second language Acquisition, 5,* 177-194.

Long, M. (1988). Instructed interlanguage development. In L. M. Beebe (Ed.), *Issues in second language acquisition: Multiple perspectives* (pp. 115-141). Cambridge, MA: Newbury House.

Long, M. H. (1991). Focus on form: A design feature in language teaching methodology. In K. de Bot, R. Ginsberg, & C. Kramsch (Eds.), *Foreign language research in cross-cultural perspective* (pp. 39-52). Amsterdam: John Benjamins.

Long, M. H. (1996). The role of the linguistic environment in second language acquisition. In W. C. Ritchie & T. K. Bhatia (Eds.), *Handbook of second language acquisition* (pp. 413-468). San Diego, CA: Academic Press.

Long, M. H. (2006). *Problems in SLA.* Mahwah, NJ: Erlbaum.

Long, M. H., Inagaki, S., & Ortega, L. (1998). The role of implicit negative feedback in SLA: Models and recasts in Japanese and Spanish. *Modern Language journal, 82,* 357-371.

Long, M., & Robinson, P. (1998). Focus on form: Theory, research and practice. In C. Doughty & J Williams (Eds.), *Focus on form in classroom second language acquisition* (pp. 15-41). Cambridge University Press.

Loschky, L. C. (1994). Comprehensible input and second language acquisition: What is the relationship? *Studies in Second Language Acquisition, 16,* 303-325.

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., & Izquierdo, J. (2009). Prompts versus recasts in dyadic interaction. *Language Learning, 59,* 453-498.

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, Australia.

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. (2006). Feedback, noticing and instructed second language learning. *Applied Linguistics, 27,* 405-430.

Mackey, A., & Gass, S. (2005). *Second language research: methodology and design.* New Jersey: Erlbaum.

Mackey, A., Gass, S., & McDonough, K. (2000). How do learners perceive interactional feedback? *Studies in Second Language Acquisition, 22,* 471-497.

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., Perdue, S., & McDonough, K. (2000). How do learners perceive interactional feedback? *Studies in Second Language Acquisition, 22,* 471-497.

Mackey, A., & Philp, J. (1998). Conversational interaction and second language development: Recasts, responses and red herrings. *The Modern Language Journal, 82,* 338-356.

Mackey, A., Philp, J., Egi, T., Fujii, A., & Tatsumi, T. (2002). Individual differences in working memory, noticing of interactional feedback and L2 development. In P. Robinson (Ed.), *Individual differences and instructed language learning* (pp. 181-210). Amsterdam/Philadelphia: John Benjamins.

McDonough, K. (2005). Identifying the contributions of negative feedback and learners' responses to L2 development. *Studies in Second Language Acquisition, 27,* 79–103.

McDonough, K., & Mackey, A. (2006). Responses to recasts: Repetitions, primed production and linguistic development. *Language Learning, 54,* 693-720.

McDonough, K., & Mackey, A. (2008). Syntactic priming and ESL question development. *Studies in Second Language Acquisition, 30,* 31-47.

Menard, S. (2002). *Longitudinal research.* (2nd ed.). Thousand Oaks, Calif.: Sage.

Muranoi, H. (2000). Focus on form through interaction enhancement: Integrating formal instruction into a communicative task in EFL classrooms. *Language Learning, 50,* 617-673.

Nabei, T. (2002). *Recasts in classroom interaction: A teacher's intention, learner's attention, and second language learning.* Unpublished doctoral dissertation, University of Toronto, Canada.

Nabei, T., & Swain, M. (2002). Learner awareness of recasts in classroom interaction: A case study of an adult EFL student's second language learning. *Language Awareness, 11,* 43-63.

Nassaji, H. (2009). Effects of recasts and elicitations in dyadic interaction and the role of feedback explicitness. *Language Learning, 59,* 411-452.

Nelson, K., Denninger, M., Bonvillian, J., Kaplan, B., & Baker, N. (1983). Maternal input adjustments and non-adjustments as related to children's advances and to language acquisition theories. In A. D. Pellegrini & T. D. Yawkey (Eds.), *The development of oral and written language in social contexts* (pp. 31-56). Norwood, NJ: Ablex.

Nemser, W. (1971). Approximative systems of foreign language learners. *International Review of Applied Linguistics*, *9,* 115-123.

Newman, I., & Benz, C. (1998). *Qualitative-quantitative research methodology: Exploring the interactive continuum.* Carbondale, IL: Southern Illinois University Press.

Nicholas, H., Lightbown, P., & Spada, N. (2001). Recasts as feedback to language learners. *Language Learning, 51,* 719-758.

Norris, J., & Ortega, L. (2000). Effectiveness of L2 instruction: A research synthesis and quantitative meta-analysis. *Language Learning, 50,* 417-528.

Ohta, A. (2000). Rethinking recasts: A learner-centered examination of corrective feedback in the Japanese classroom. In J. K. Hall & L. Verplaeste (Eds.), *The construction of second and foreign language learning through classroom interaction* (pp. 47-71). Mahwah, NJ: Erlbaum.

Oliver, R. (1995). Negative feedback in child NS-NSS conversation. *Studies in Second Language Acquisition, 17,* 459-481.

Oliver, R. (2000). Age differences in negotiation and feedback in classroom and pair work. *Language Learning, 50,* 119-151.

Oliver, R., & Mackey, A. (2003). Interactional context and feedback in child ESL classrooms. *Modern Language Journal, 87,* 519-533.

Ono, L., & Witzel, J. (2003). *Recasts, salience and morpheme acquisition.* Paper presented at the Second Language Research Forum, Tucson, AZ.

Ortega, L. (1999). Planning and focus on form in L2 oral performance. *Studies in Second Language Acquisition, 21,* 109-148.

Ortega, L., & Iberri-Shea. G. (2005). Longitudinal research in second language acquisition: Recent trends and future directions. *Annual Review of Applied Linguistics, 25,* 26-45.

Panova, I., & Lyster, R. (2002). Patterns of corrective feedback and uptake in an adult ESL classroom. *TESOL Quarterly, 36,* 573-595.

Patton, M. (1990). *Qualitative evaluation and research methods.* (2nd ed.). Newbury Park, CA: Sage Publications.

Philp, J. (2003). Constraints on "noticing the gap": Non-native speakers' noticing of recasts in NS-NNS interaction. *Studies in Second Language Acquisition, 25,* 99-126.

Pica, T. (1992). The textual outcomes of native speaker-non-native speaker negotiation: What do they reveal about second language learning? In C. Kramsch & S.

McConnell-Ginet (Eds.), *Text and context: Cross-disciplinary perspectives on language study* (pp. 198-237). Lexington, VA: D.C. Heath.

Pica, T. (1994). Research on negotiation: What does it reveal about second-language learning conditions, processes, and outcomes? *Language Learning, 44,* 493-527.

Pica, T. (1996). Do second language learners need negotiation? *International Review of Applied Linguistics in Language Teaching, 34,* 1-19.

Pica, T. (2002). Subject-matter content: How does it assist the interactional and linguistic needs of classroom language learners? *Modern Language Journal, 86,* 1-19.

Pica, T., & Doughty, C. (1985). Input and interaction in the communicative language classroom: a comparison of teacher-fronted and group activities. In S. Gass & C. Madden (Eds.), *Input in second language acquisition* (pp. 115-132). Rowley, MA: Newbury House.

Pica, T., Young, R., & Doughty, C. (1987). The impact of interaction on comprehension. *TESOL Quarterly, 21,* 737-758.

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.

Pienemann, M. (1989). Is language teachable? Psycholinguistic experiments and hypotheses. *Applied Linguistics, 10,* 52-79.

Pienemann, M. (1998). *Language Processing and Second Language Development: Processability theory*. Amsterdam: John Benjamins.

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, AMEP.

Pienemann, M., & Mackey, A. (1993). An empirical study of children's ESL development. In P. McKay (Ed.), *ESL development. Language and literacy in schools, Vol. 2* (pp. 115-259). Canberra: National Languages & Literacy Institute of Australia and Commonwealth of Australia.

Pienemann, M., Johnston, M., & Brindley, G. (1988). Constructing an acquisition-based procedure for second language assessment. *Studies in Second Language Acquisition, 10,* 217-243.

Polio, C., & Gass, S. M. (1998). The role of interaction in native speaker comprehension of nonnative speaker speech. *Modern Language Journal, 82,* 308-319.

Saxton, M. (1997). The contrast theory of negative input. *Journal of Child Language, 24,* 139-161.

Schmidt, R. (1990). The role of consciousness in second language learning. *Applied Linguistics, 11,* 129-158.

Schmidt, R. (1993). Awareness and second language acquisition. *Annual review of Applied Linguistics, 13,* 206-226.

Schmidt, R. (1994). Implicit learning and the cognitive unconscious: Of artificial grammars and SLA. In N. Ellis (Ed.), *Implicit and explicit learning of languages* (pp. 165-209). London: Academic Press.

Schmidt, R. (2001). Attention. In P. Robinson (Ed.), *Cognition and second language instruction* (pp. 3-32). Cambridge: Cambridge University 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.237-322). Rowley, MA: Newbury House.

Seedhouse, P. (1997). The case of the missing "no": The relationship between pedagogy and interaction. *Language Learning, 47,* 547-583.

Seliger, H.W., & Shohamy.E. (1989). *Second language research methods.* Oxford: Oxford University Press.

Selinker, L. (1972). Interlanguage, *IRAL, 10,* 209-231.

Sheen, Y. (2004). Corrective feedback and learner uptake in communicative classrooms across instructional settings. *Language Teaching Research, 8,* 263-300.

Sheen, Y. (2006). Exploring the relationship between characteristics of recasts and learner uptake, deconstructing recasts: Do they constitute an implicit or explicit corrective feedback strategy? *Language Teaching Research, 10,* 361-392.

Skehan, P. (1998). *A cognitive approach to language learning.* Oxford: Oxford University press.

Spada, N. (1997). Form-focussed instruction and second language acquisition: A review of classroom and laboratory research. *Language Teaching, 30,* 73-87.

Spada, N., & Lightbown, P. (1993). Instruction and the development of questions in L2 classrooms. *Studies in Second Language Acquisition, 15,* 205-224.

Swain, M. (1985). Communicative competence: Some roles of comprehensible input and comprehensible output in its development. In S. Gass & C. Madden (Eds.), *Input in second language acquisition* (pp. 235-253). Rowley, MA: Newbury house.

Swain, M. (1995). Three functions of output in second language learning. In G. Cook & B. Seidlhofer (Eds.), *Principles and practice in applied linguistics: Studies in honor of H. G. Widdowson* (pp. 125-144). Oxford: Oxford University Press.

Swain, M. (2001). Examining dialogue: Another approach to content specification and to validating inferences drawn from test scores. *Language Testing, 18,* 275-302.

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. *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 pedagogical tasks* (pp. 99-118). London: Longman.

VanPatten, B. (1989). Can learners attend to form and content while processing input? *Hispania, 72,* 409-417.

Varonis, E., & Gass, S. M. (1985). Non-native/non-native conversations: A model for negotiation of meaning. *Applied Linguistics, 6,* 71–90.

Wagner-Gough, J., & Hatch, E. (1975). The importance of input data in second language acquisition studies. *Language Learning, 25,* 297–308.

White, L. (1991). Adverb placement in second language acquisition: Some effects of positive and negative evidence in the classroom. *Second Language Research, 7,* 133-161.

White, L., Spada, N., Lightbown, P., & Ranta, L. (1991). Input enhancement and L2 question formation. *Applied Linguistics, 12,* 416-432.

Participant Information Sheet

AUT

UNIVERSITY

TE WĀNANGA ARONUI O TAMAKI MAKAU RAU

Date Information Sheet Produced:

February 6 2006

Project Title

The role of teacher feedback on English question formation over time.

Invitation

As a full-time student at IH, AUT you have been selected to take part in this research.

My name is Katy Nichols and I am an ESL lecturer at International House (IH), AUT. I am also studying for my Master's degree at AUT. For my study I will be researching learning a second language and the way you ask questions. I would like to invite you to take part in my research study.

What is the purpose of this research?

The purpose of this research is to look at language teaching today and see if it can be improved to help students learn English more successfully. This research may be published in language learning journals when it is finished.

How are people chosen to be asked to be part of this research?

All Intermediate 1 level students studying at IH who will continue their study for at least two more months will be invited to take part in this study. It is voluntary and you decide if you want to take part or not.

What happens in this research?

First, you will fill in a short questionnaire. This information will be used to see if the results of the study may be related to your background. After this there will be 3 or 4 short speaking tasks with one other student. There will also be one speaking task with a teacher. Each session will take about 10 minutes and you will be audio recorded. These recordings will be listened to and transcribed (written out). You will have the chance to check them to make sure that you are happy with the information in them. The researcher will email these to you when they are ready.

What are the benefits?

All students will benefit from the extra English speaking opportunities. Also, the results of this research could give us information that can help students to learn English more successfully.

How will my privacy be protected?

You will be given a number and your name will not be used on any documents or on the results. The Consent Forms, on which your name is recorded, will be stored securely at AUT.

What are the costs of participating in this research?

You will spend about 1 ~ 1.5 hours doing the speaking tasks. Each task will take around 10 minutes. These tasks will take place on a Friday afternoon or immediately after the end of morning or afternoon classes. The research will take place in the second half of this year. Taking part in this study will not affect your test mark or regular classes in any way.

What opportunity do I have to consider this invitation?

You have 24 hours to think about this invitation to take part. If you decide not to take part it doesn't matter at all. The information collected in this research will remain confidential and your name will not be anywhere on any data or results.

How do I agree to participate in this research?

After reading this information sheet, you will have the chance to ask any questions you want. Then you will have 24 hours to think about taking part in this research and if you are happy and want to take part, please sign the consent form and return it tomorrow.

Will I receive feedback on the results of this research?

The results of this research will be available to you after March, 2007. If you would like to receive a copy of the findings, please tick this ☐ box.

What do I do if I have concerns about this research?

Any concerns regarding the nature of this project should be notified in the first instance to the Project Supervisor, Associate Professor John Bitchener, *john.bitchener@aut.ac.nz*, ph: 921-9999 ext 7830.

Concerns regarding the conduct of the research should be notified to the Executive Secretary, AUTEC, Madeline Banda, *madeline.banda@aut.ac.nz*, 921 9999 ext 8044.

Who do I contact for further information about this research?

Researcher Contact Details:
Katy Nichols, *katy.nichols@aut.ac.nz*, 921-9999 ext 4713

Project Supervisor Contact Details:
Associate Professor John Bitchener, *john.bitchener@aut.ac.nz*, 921-9999 ext 7830.

APPENDIX B

<u>CONSENT TO PARTICIPATION IN RESEARCH</u>

Title of Project: **The role of teacher feedback on English question formation over time.**

Project Supervisor: **Associate Professor John Bitchener**

Researcher: **Katy Nichols**

- I have read and understood the information provided about this research project (Information Sheet dated February 6, 2006.)

- I have had an opportunity to ask questions and to have them answered.

- I understand that the teaching sessions will be audio-taped and transcribed.

- I understand that I may withdraw myself or any information that I have provided for this project at any time prior to completion of data collection, without being disadvantaged in any way.

- If I withdraw, I understand that all relevant tapes and transcripts, or parts thereof, will be destroyed.

- I agree to take part in this research.

- I wish to receive a copy of the report from the research: tick one: Yes O   No O

Participant signature: ....……………………..

Participant name: ………………………………………………………………….

Participant Contact Details (please write your email address here to receive a copy of your taped transcripts):

………………………………………………………………………………………..

………………………………………………………………………………………..

………………………………………………………………………………………..

………………………………………………………………………………………..

Date:

APPENDIX C

Student Questionnaire

Please complete the following information or circle the appropriate answer. This survey is confidential (i.e. nobody will know your answers except the researcher). The reason why this information is important is to compare it to the results of the research to see if there are any patterns connected to your background.
Please circle one answer or write a short answer in the space provided.

1.   Gender: Female / Male

2.   Age: 18 ~ 19   20 ~ 24   25 ~ 29   30 ~ 34   Over 35

3.   What's your nationality?   _____

4.   Which country were you born in?   _____

5.   Which country did you go to school in? _____

6.   When did you come to New Zealand?   ___ month(s) ___ week(s) ago

      (e.g. 1 month 2 weeks ago)

7.   How long have you been studying English in New Zealand?

      __ month(s) ___ week(s)

8.   How many more months/weeks are you going to study English in New Zealand from now?

      ___ month(s) ___ week(s)

9.   How old were you when you first started studying English? _____

10.   How many years have you been studying English now?   ____ years

11.   What did you do in your country before coming to New Zealand?

_____ (e.g. I was a student / working in a restaurant)

12.   What is your <u>main</u> reason for coming to New Zealand?

a.   To study English

b.   To make new friends

c.   To experience New Zealand lifestyle

d.   To travel

e.   Other _____

(Please write the reason here)

13.   Who made the decision about you coming to New Zealand?

a.   Myself

b.   My parents

c.   Another family member _____ (who?)

d.   Somebody else _____ (who?)

14.   Why are you studying English?

a.   To go to a university in New Zealand.

b.   To get a good job in the future.

c.   Because my parents wanted me to.

d.   For fun / as a hobby.

e.   Other _____

(Please write the reason here)

15.   How much do you enjoy learning English?

a. Very much   b. Its OK   c. Not really   d. Not at all

Participant Information Sheet

Dear Participant

Thank you very much for helping with this research.
Now that the data collection process is finished I want to tell you about the type of information this study wanted to find out. As you know the aim of the study was to look at teacher feedback and to see if one type of feedback from the teacher can help students to produce questions more correctly. During the teaching session of the research half of you received one type of teacher feedback and half of you didn't. The reason for this was to find out if this type of feedback helps students learn English better. This type of feedback is called giving 'recasts' which means that the teacher repeats incorrect language (in this case questions forms) in a correct way, so that the student can hear this. The researcher wants to know if hearing this correct version of your question will help you to produce better questions. Now, data collected from those of you who received recasts will be compared with the data from those of you who didn't receive these recasts in order to find any important results. You will be told about these results after the data analysis process is completed.

Thank you again

Katy Nichols

The Researcher

APPENDIX E

TEST TASK INSTRUCTIONS

Find the differences …

1. Sit at the table, opposite your partner.

2. Look at your picture. Your partner has a picture which is a little different to your picture. ***Don't look at your partner's picture.***

3. If you don't know how to say a word in English, you can ask Katy what it is. You can point to the picture and say 'What is this?'

4. Ask your partner a question about their picture to try and find a difference.

5. Listen to your partner's answer and if you have found something different, Katy will tell you.

6. Next, answer your partner's question.

7. Continue until you have found the total number of differences. Katy will tell you how many there are.

Good luck!

Appendix F
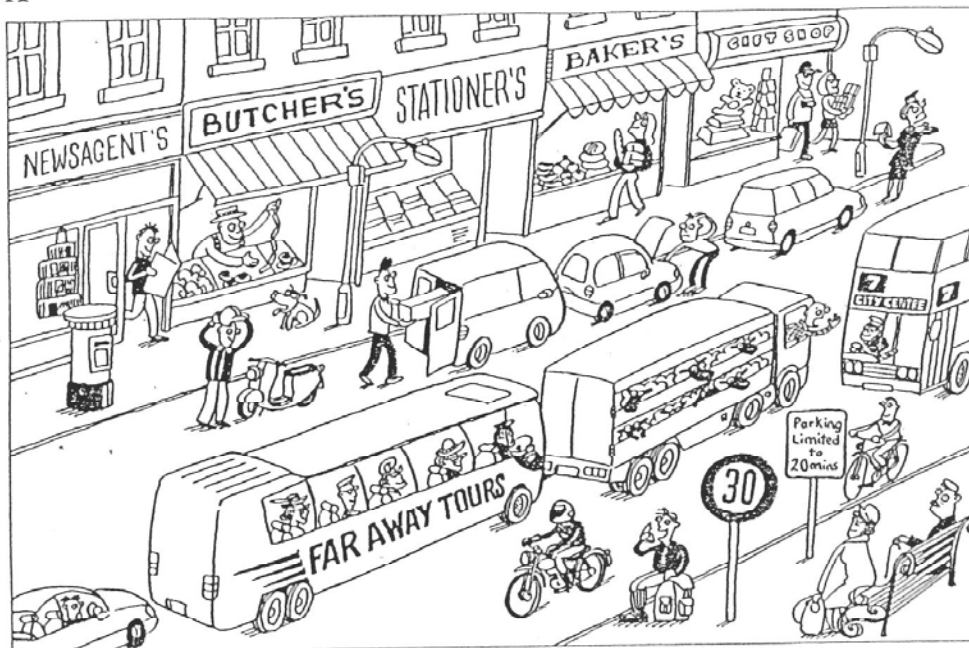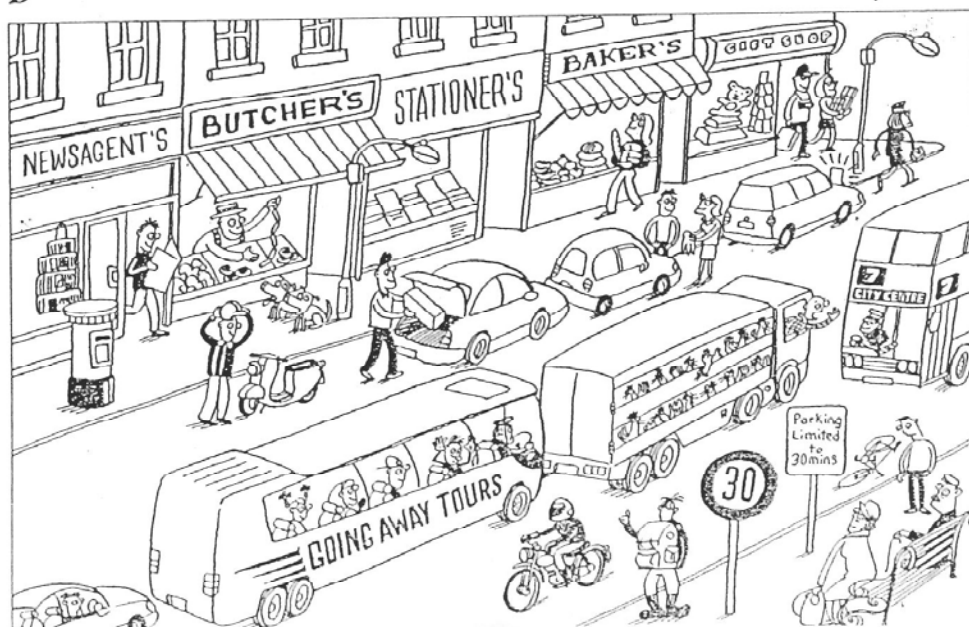
Pre-Test Task

Picture A

Picture B

Source: Watcyn-Jones, P., & Howard-Williams, D. (1995). *Grammar 1 Games and Activities: New Edition* (pp. 69-70). Essex: Penguin English.

Appendix G

Immediate Post-Test Task

A

B

Source: Oxenden, C., Latham-Koenig, C., & Hamilton, G. (1999). *English File Intermediate Teacher's Book* (p. 154). Oxford: Oxford University Press.

One-Week Post-Test Task

Picture B

Source: Kay, S., & The Lake School of English. (1997). *Reward Elementary Resource Pack* (11-15a/11-15b). Oxford: Macmillan Heinemann.

One-Week Post-Test Task

Picture A

Source: Kay, S., & The Lake School of English. (1997). *Reward Elementary Resource Pack* (11-15a/11-15b). Oxford: Macmillan Heinemann.

12-Week Post-Test Task

Park scene A

Park scene B

Source: Hadfield, J. (1999). *Beginners' Communication Games* (p. 97). Toronto: Addison Wesley Longman.