This is to certify that the

dissertation entitled

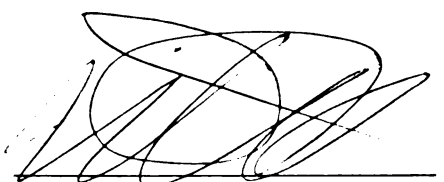
THE ROLE OF PRAGMATICS IN THE
ACQUISITION OF REFLEXIVE BINDING IN L2

presented by

Mahide Demirci

has been accepted towards fulfillment
of the requirements for

PhD degree in Linguistics



Major professor

Date 06/12/1997

THE ROLE OF PRAGMATICS IN THE ACQUISITION OF REFLEXIVE BINDING IN L2

By

Mahide Demirci

A DISSERTATION

Submitted to
Michigan State University
in partial fulfillment of the requirements
for the degree of

DOCTOR OF PHILOSOPHY

Department of Linguistics and Germanic, Slavic,
Asian and African Languages

1997

ABSTRACT

THE ROLE OF PRAGMATICS IN THE ACQUISITION OF REFLEXIVE BINDING IN L2

By

Mahide Demirci

This study investigates the effects of pragmatic constraints on the acquisition of the binding of English reflexives by adult Turkish L2 learners. Pragmatically biased and pragmatically neutral sentences are compared to determine whether pragmatic bias towards a non-local antecedent overrides the parameter setting of English and causes learners to choose as possible antecedents NPs outside the binding domain. The study uses an acceptability judgment test with 27 sentences divided into nine clause types, including neutral vs biased types, finite vs nonfinite types, and biclausal vs triclausal types. Group results indicate that pragmatically biased sentences compel the subjects to consider pragmatic information to the extent that it can override their syntactic intuitions. Acquisition theories need to account for the role that pragmatic information plays in the assignment of possible antecedents for reflexives. This study incorporates Huang's (1994) pragmatic theory of anaphora in which the interpretation of a reflexive is subject to the I-principle, a pragmatic strategy which finds an antecedent for the reflexive that gives the most informative, stereotypical interpretation in keeping with our knowledge about the world.

ACKNOWLEDGMENTS

I would like to thank my committee members, Sue Gass, Grover Hudson, and Alan Munn. I would especially like to thank the chairperson of my committee, Dennis Preston, who has been my professor, mentor, boss, and my friend for many years. I would not have been able to finish my program without his guidance and help. I am tremendously indebted to him as a scholar and as a human being.

My most sincere and endless thanks go to Brian Kleiner, who put almost as much effort into this dissertation as I did. He contributed as an editor, statistician, typist, and as a linguist. Most importantly, he tolerated all of my stressful nights and days involved in the production of this work. Finally, thanks to my daughter, Feride, for being there with me.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

ABL	ablative
CL	classifier
DAT	dative
EMP	emphatic particle
EXP	experiential aspect marker
FUT	future
GEN	genitive
INT	interrogative
MM	modifier marker
NOM	nominative
OBJ	objective
PASS	passive
PAST	past tense
PFV	perfective aspect marker
POSS	possessive
PL	plural
PP	pause particle
PROG	progressive
REFL	reflexive
SING	singular

INTRODUCTION

This study investigates the effects of pragmatic constraints on the acquisition of the binding of English reflexives by adult Turkish speakers learning English as a second language. It employs one of the subtheories of Universal Grammar (UG); the Binding Theory proposed and later modified by Chomsky (1981, 1986a, 1988, 1995), which deals with the referential properties of noun phrases. The data of this study are the results of an experiment which examined the interpretation of English reflexives by native speakers of Turkish at five different proficiency levels. The experiment involved an interpretation task which sought to uncover the subjects' judgments of coreference across a variety of both finite and nonfinite biclausal and triclausal sentences containing reflexives.

English and Turkish differ in the domains in which reflexives are bound — Turkish has the least restrictive setting (long-distance binding), whereas English has the most restrictive setting (local binding). This study explores whether and to what extent the acquisition of the locality requirement of English reflexives is influenced by pragmatic constraints. Pragmatically biased and pragmatically neutral sentences are compared in order to determine whether pragmatic bias towards a nonlocal NP overrides the setting for English reflexives and causes the learners to choose as possible antecedents NPs outside the local domain. In addition, the study attempts to observe developmental stages in the acquisition of English binding and to see whether the influence of pragmatic factors decreases over time with increasing proficiency. Finally, this study also tries to find out

how other factors such as nonfiniteness and level of embedding interfere with the acquisition of English binding.

The results of this study will demonstrate that pragmatics is the main factor which causes the learners at any proficiency level to violate the syntactic requirement of English binding, namely, the locality condition. Therefore, purely syntactic descriptions of the acquisition of English reflexives by Turkish learners appear to be inadequate. Acquisition theories need to account for the role that pragmatic information plays in the assignment of antecedents for English reflexives for Turkish second language (L2) learners. Therefore, this study attempts to incorporate the "Pragmatic Theory of Anaphora" developed by Huang (1994), which proposes that the interpretation of reflexives is subject to the I-principle (informativeness principle) for a set of "pragmatic" languages, such as Chinese, Korean, Japanese, and to this list we might add Turkish. What the I-principle does is to find an antecedent for the reflexive that gives the most informative, stereotypical interpretation in keeping with our knowledge about the world. We will see that for the Turkish learners there is no proficiency level difference with regard to their acquisition of English reflexive binding, and that the transfer of pragmatic principles from the first language (L1) is dominant and enduring.

Chapter 1

REFLEXIVES IN LINGUISTIC THEORY AND SLA RESEARCH

1.1. Introduction.

This chapter reviews a variety of views regarding the role of UG and L1 in SLA. In addition, it discusses Principle A of the Binding Theory of Chomsky (1981, 1986a) and two theoretical frameworks which treat reflexive binding in L2, namely, the GCP approach and LF movement. A review of these syntactic approaches to reflexive binding will set the stage for a different view in later chapters which incorporates the insight that pragmatics may play a significant role in L2 reflexive binding.

1.2. The acquisition of binding principles and the role of pragmatics.

This study accepts Huang's contention that the contribution of pragmatics to the binding of reflexives, at least with respect to pragmatic (discourse-oriented) languages (e.g, Chinese), is much more crucial than has commonly been believed (Huang, 1994). Although Chomsky (1982) argues that the development of grammar takes place independently of other kinds of social and cognitive interaction, this study reveals that pragmatics plays an important role in the interpretation of reflexives in L2 as well as in L1.

This study focuses on the effects of pragmatics on the acquisition of the locality requirement of English reflexives because, following Huang (1994), it is believed that

"...syntax and pragmatics are interconnected, though they are distinct levels and modes of explanation. Moreover, syntactic rules can and sometimes must make reference to pragmatic principles, contrary to the autonomy of syntax hypothesis (p.259)." In addition, Huang states that:

> ...the full understanding of the nature and ontogeny of knowledge of language appears to be partially dependent on a better understanding of the (creative) use of that knowledge. If these conclusions are correct, it seems then that a large portion of linguistic explanation currently sought in grammatical terms may need to be shifted to pragmatics... (p.261)

In this study, we will see that reference to pragmatic principles is necessary in order to account for the interlanguages of Turkish L2 learners, since reference to syntactic theory by itself appears to be inadequate in explaining the results. By considering pragmatics in the acquisition of L2 binding, this study attempts to develop a more interactive approach between "Internalized" and "Externalized" models of reflexive binding acquisition in L2 rather than employing the full use of a UG perspective which investigates only internalized models of language acquisition.

1.3. The role of UG in Second Language Acquisition (SLA)

1.3.1. UG and Principles and Parameters. According to Chomsky (1981, 1986a, 1988), Universal Grammar (UG) is a theory of grammar and language acquisition which consists of a system of conditions and rules that are properties of all human languages. Chomsky (1988) characterizes these properties as consisting of principles and parameters. The term

"principles" refers to abstract properties of grammar which apply to languages in general and underlie the grammatical rules of all languages. The term "parameters" refers to principles that vary in certain restricted ways from one language to another. In other words, principles determine what is given and parameters determine what must be learned (Flynn, 1991). Parameters have two or more settings, with different languages selecting different settings. Any language in which a particular principle operates must realize one of the parameter settings associated with that principle. Chomsky (1988) likens parameters to the array of switches found in a switch box. The learner's task is to use experience (evidence) to determine which position each switch must be in. Chomsky states that:

> Each permissible array of switch settings determines a particular language. Acquisition of a language is in part a process of setting the switches one way or another on the basis of the presented data, a process of fixing the values of the parameters. (p.63)

As it is seen, language acquisition within this framework means setting the values for parameters that are allowed to vary in UG, and setting them by default to an unmarked setting without the need of positive evidence. If the unmarked default setting is not correct, then it must be switched on the basis of language-particular evidence to the marked setting.

1.3.2. UG and SLA. UG, as proposed by Chomsky, does not make any claims about L2 acquisition. However, several researchers (Flynn, 1987, etc.) claim that if principles of UG are sufficient to explain how L1 acquisition occurs, then UG might in some way underlie L2 acquisition as well; this means that L2 learning might also be constrained by

principles and parameters of UG. According to Flynn (1987), if learning an L1 is a matter of setting parameters, then learning an L2 is a matter of resetting parameters. Flynn (1991) proposes two possibilities about a theory of UG in adult L2 acquisition. First, if UG doesn't hold in adult L2 acquisition, then this will suggest something about the nature of UG as a biological component of the human mind. For example, it may be subject to a critical period beyond which the principles cannot be activated. If this is true, then adult L2 acquisition would involve an inductive learning component. In other words, adults' learning of language will not be structure dependent, translating lexical items one by one, from L1 to L2, or else using relations of linear precedence in second language acquisition. The second possibility is that if L2 learning is structure dependent, and L2 acquisition follows deductively from a set of linguistic principles observed in first language acquisition, then this would suggest that the language faculty plays a role in SLA.

Unlike first language learners, L2 learners come to SLA with parameters already set to the L1 value. Parameter values in the L1 may match those for the L2 in some cases and not in other cases. According to Flynn (1987, 1991), when a parameter value in the L1 does not match that of the L2, the acquisition might be affected negatively, because in this case a new value must be assigned to the parameter to match the L2, and the acquisition would be disrupted when compared to the case in which the parameters are matched. If the L1 and L2 values match, acquisition might be affected positively, because L2 learners would be able to consult the L1 parameter value in guiding their acquisition of the L2.

1.3.2.1. *Access to UG and the role of L1 in SLA* . According to the principles and parameters framework, acquisition of particular aspects of L1 grammar takes place through an interaction between the evidence provided by the input, and a set of grammatical principles available to all learners, i.e., UG. However, in SLA, as several researchers have argued, other factors such as the L1 of the L2 learners may also play an important role in the setting of L2 parameters. The extent to which UG is available to L2 learners, and the extent to which the L1 parameter value is involved in the setting of the L2 parameter value have been debated for the last decade. There are several hypotheses concerning the role of UG and the role of the L1 parameter setting in SLA.

1.3.2.2. *Hypotheses about the role of UG and the role of L1 in SLA* . There are three main hypotheses for the role of UG in SLA (Sharwood Smith, 1990). Each of these hypotheses entail different views on the role of L1 in L2 acquisition:

i. The parasitic hypothesis (as advanced by Bley-Vroman, 1989; Schachter 1988, 1989, 1996; Clahsen and Muysken, 1986). This theory has also been called "No access" by Cook (1996), or "UG-is-dead" by White (1989). Sharwood Smith claims that this hypothesis is called parasitic because the interlanguage is thought to be parasitic on the L1 grammar. According to this approach, UG is no longer available in SLA, and does not help the learner in creating new areas of the interlanguage if they are not based on the L1. Therefore, L2 learners can only make use of general learning strategies to acquire the L2. Any features which appear to conform to UG in the interlanguage may be traced back to those of the L1 carried over into the L2, which Sharwood Smith calls "UG -by proxy."

According to the Parasitic Hypothesis, L1 parameter values do not play any role in L2 parameter setting, because L1 and L2 acquisition follow different procedures. Learners acquire the L2 through general problem solving procedures and cognitive strategies, unlike L1 learners who have access to UG.

Another variant of the parasitic hypothesis, called "partial access," proposed by Clahsen and Muysken (1989), claims that L2 grammars are constrained by principles of UG that hold for all languages but these grammars do not have access to parametric options. Therefore, learners can have access to linguistic principles only through the principles found in their L1 and cannot have access to the full range of parametric values of UG.

The "partial access" view entails that L2 learners can access only the principles of UG which are found in their L1. If L1 and L2 have different values for some parameter, the L2 value will not be accessible. Therefore, parameter resetting to the L2 value is not possible when L1 and L2 have different settings.

ii. "The Recreative Hypothesis" (Mazurkevich, 1984). This theory has also been called the "UG Hypothesis" by White (1989), the "Pure UG Hypothesis" by Cook (1996), and the "Complete Access Hypothesis" by Ellis (1996). This hypothesis makes the strongest possible claim for the operation of UG in L2, and very few researchers working on UG in L2 acquisition support it. According to this hypothesis, the L2 grammar is acquired in the same way as the L1 grammar is acquired. L2 learners can employ the principles of UG fully and set the parameters without any reference to their L1 values. According to Sharwood Smith (1990), this approach might be called the "Back-to-Square-

One" view, because the learners go back to the beginning and recreate the L2 grammar as if they were native learners of that language with the guidance of UG.

The Recreative Hypothesis entails that L1 parameter settings have no effect on the resetting of L2 parameters. The L2 data is able to trigger the appropriate parameter value without the help of L1. This means that L2 learners may set the parameters to their default values at the onset of L2 acquisition regardless of the L1 settings. A softer version of the Recreative Hypothesis would predict that the L1 mediates the construction of the L2 grammar in that when L1 and L2 have different parameter settings, L2 acquisition will be disrupted, because learners will need to assign a new value to match the L2 value. When L1 and L2 parameter settings match, the L2 acquisition will be facilitated because the learners can consult the L1 parameter value in guiding their acquisition of the L2 (Flynn, 1987, 1988, 1991, 1993). Briefly, according to this view, the L1 setting has an effect on the L2 setting, however, it is never actually adopted in the interlanguage grammar.

iii. "Reorganization Hypothesis," proposed by White (1985, 1988, 1989), Sharwood Smith (1988), and van Buren (1988). This has also been called the "Indirect Hypothesis" by White (1989). According to this hypothesis, UG is still active in SLA, but L2 learners initially assume the L1 value of the UG parameters. However, they are still able to eventually reset the parameters to L2 values. Sharwood Smith proposes three developmental phases within this approach:

i) initial application of L1 instantiations of UG, and

ii) a recreative application of UG in areas where L1 parameters are not relevant on the basis of positive evidence, and

iii) reorganizing and revising the L1 settings of phase (1).

The Reorganizational Hypothesis entails that learners start out with an L1 parameter value, but many abandon this L1 value eventually as their knowledge of L2 improves. In this case, the L1 data serve as input to the interlanguage grammar until L2 data force the learner to reanalyze the interlanguage grammar in the relevant respect. Otherwise, the L1 parameter will continue to be represented in the interlanguage grammar.

To summarize the various hypotheses, according to the Parasitic Hypothesis, UG is dead and there is no parameter resetting in L2. The Recreative Hypothesis claims that UG is fully active and parameter setting is natural in L2. The Reorganization Hypothesis allows parameter resetting through the help of the L1.

1.4. Binding theory

1.4.1. Introduction. This section describes the reasons for examining reflexives in L2, and discusses the theory of the binding of reflexives which this study takes as a starting point. I accept the assumptions about binding common to versions of the Government-Binding theory proposed by Chomsky (1981, 1986a). Furthermore, I assume that local or long-distance binding is a consequence of parametric variation. This study examines "I-Language" (internalized language), but observes how "E-Language" (externalized language) interacts with it.[1]

1.4.2. Why study reflexives? According to Huang (1994), in recent years anaphora has become a central topic in linguistic theory and has attracted the attention of philosophers, psychologists, cognitive scientists and artificial intelligence researchers. Huang states that

anaphora has:

> aroused this interest because, on the one hand, some aspects of anaphora have repeatedly been claimed by Chomsky (1981, 1982a, b, 1986a, 1988, 1991b) to reflect underlying principles of innate Universal Grammar (UG), the biologically determined endowment of the human mind; and on the other hand, anaphora has been shown to interact with various syntactic, semantic and pragmatic factors. (p.1)

Therefore, as Huang claims, the study of reflexives may test various hypotheses having to do with the relationships among syntax, semantics, pragmatics in linguistic theory and the application of this theory to acquisition.

According to Thomas (1991a, 1993), learners' interpretations of reflexives offer a good microcosm to test UG accounts of L2 acquisition for several reasons. First, in GB theory, Binding is considered to be a principle of UG, i.e., a part of the learner's innate endowment. Therefore, interpretation of reflexives within binding theory is assumed to be subject to formal syntactic constraints determined by UG and is not subject to learned knowledge. Hence, the study of reflexives by L2 learners can reveal the role of UG in SLA. Second, although the behavior of reflexives is constrained by UG, the properties of reflexives differ from language to language. This variation in the binding of reflexives offers the researcher a way of investigating the resetting of parameters by L2 learners as well as the effect of the L1 setting on the resetting to the L2 value. Third, as Thomas (1991a, 1993) also argues, L2 learners do not normally receive overt instruction about the rules governing reflexives, and therefore overt instruction cannot be counted as a source for hypotheses about the interpretation of reflexives.

1.4.3. Introduction to binding theory. This section will summarize the binding theory as proposed by Chomsky (1981, 1986a). Section 1.5.1. reviews the parameterized version of the binding theory, and the following section 1.5.2. reviews a nonparameterized version of the binding theory, namely movement in LF. We will see in chapter two how both the parameterized and nonparameterized version have been employed in studies of the L2 acquisition of reflexive binding.

1.4.4. Binding theory. Binding theory is a subtheory of GB theory which deals with the referential properties of noun phrases, namely anaphors (reflexives, reciprocals), pronominals and R-expressions. Each of these noun phrases is subject to separate principles of the binding theory; Principle A, Principle B, and Principle C.

Chomsky (1981, p.188) presents these conditions on binding in Government-Binding theory:

(1) The Binding Principle (Chomsky, 1981, p. 188):

A. An anaphor is bound in its governing category[2]
B. A pronominal is free in its governing category
C. An R-expression is free

The term "bound" means "co-indexed with a c-commanding antecedent," where "c-command" and "governing category" are defined in the following ways:[3]

(2) C-command (Chomsky, 1986b, p.8):

A c-commands B iff A does not dominate B and every G that dominates A dominates B.

(3) Governing Category (Chomsky, 1981, p.211):

B is a governing category for A if and only if B is the minimal category containing A, a governor for A and a SUBJECT accessible to A.

A 'SUBJECT' accessible to A is, very roughly, AGR (a set of agreement features of the tensed verb) and the subject of a nontensed clause, and/or the subject/possessive element of an NP. Therefore, the governing category is the minimal NP or S in which A appears.

This study deals with only Principle A of the binding theory, which involves the binding of reflexives by other NPs within the same clause. Principle A imposes a locality restriction on the binding of anaphors in that an anaphor must be bound in its local domain, called the governing category, by a c-commanding antecedent within this local domain. Consider the following example:

(4) John told Bill that $_a$[Sam$_i$ hates himself$_i$]

In accordance with Chomsky's binding Principle A, the anaphor *himself* in the example above is bound within the governing category labeled (a) by its antecedent *Sam*, because both *himself* and *Sam* are in the same minimal category, namely, the subordinate clause, which contains a subject. On the other hand, *himself* cannot be bound by John or Bill because *himself* and these NPs are not contained within the same clause. As it may be seen, the governing category in English is the smallest clause which contains the anaphor

and its antecedent.

However, there are some languages in which the governing categories for anaphors are larger than that of English; hence, languages vary as to what counts as a governing category. Some languages allow for a relatively small local domain and others allow the governing category to be as large as the entire sentence. Therefore, in some languages anaphors can take only local antecedents, and in others antecedents can be some distance away from anaphors — a phenomenon called long-distance reflexivization.[4]

An example of this latter possibility is Turkish; in the Turkish equivalent of the example sentence (4) above, the reflexive can refer to any NP within the sentence — namely, John, Bill or Sam, as illustrated in (5):

(5) Can_i Beyhana_j Salihin_k $\text{kendisinden}_{i/j/k}$ nefret ettigini soyledi.
Can Beyhan-DAT Salih-GEN self-POSS-ABL hate-NOM-POSS tell-PAST

$_a$[John_i told Bill_j that Sam_k hates $\text{himself}_{i/j/k}$]

Example (5) shows that the governing category (labeled (a)) in Turkish is much wider than that of English, since Turkish allows a reflexive to take its antecedent anywhere in the sentence.

1.5. Revisions of the Binding Theory.

Over the last decade, the existence of long-distance reflexivization in several languages, like Turkish, Japanese, Korean, etc., has called into question the definition of

locality, determined by Principle A, and hence the Binding Theory has been continually challenged. To address this problem, several researchers have attempted to revise the Binding Theory. Two of the influential revisions, which will be discussed at greater length within this chapter, are as follows:

1. Parameterization of the local domain as Governing Category Parameter, proposed by Manzini and Wexler (1987). This revision states that the local domain in which a reflexive is bound to its antecedent(s) constitutes a parameter equipped with five values.

2a. Movement at LF of the anaphor, initiated by Lebeaux (1983), and further developed by Chomsky (1986a), and Cole, Hermon and Sung (1990) among others. This revision claims that long-distance reflexivization is due to the properties of LF movement.

2b. Lexical analysis, initiated by Pica (1987). This revision states that the morphemic status of anaphors (compound or monomorphemic) determines movement at LF.

1.5.1. Parameterized approach to reflexive binding.

1.5.1.1. *Governing Category Parameter* . Manzini and Wexler (Manzini and Wexler (1987), Wexler and Manzini (1987)) propose to parameterize the notion of governing category and of proper antecedent in order to offer an account of the variation in binding domains across languages. They define the notion of governing category in terms of a set of 5 parametric values; as in (6):

(6) Governing Category Parameter (GCP) Wexler and Manzini (1987, p.53):

y is a governing category for x iff y is the minimal category which contains x and a governor for x and:

a. has a subject, or
b. has an INFL, or
c. has a TNS, or
d. has an indicative TNS, or
e. has a root TNS

It is claimed that a language can choose among these parametric options for its governing category. For instance, the value (a) holds for English reflexives, the value (b) for the Italian reflexive 'se,' the value (c) for the Icelandic pronoun 'ham,' the value (d) for the Icelandic reflexive 'sig,' and the value (e) for Japanese 'zibun.' The governing category parameter is a hierarchy that divides languages into types according to how close an anaphor and its antecedent in a sentence must be in order to be properly bound. As it might be seen, each of the governing categories defined by the governing category parameter values ranging from (a-e) is potentially bigger than the previous one, yet it includes the previous possibilities, a prospect called the "Subset Principle" by Wexler and Manzini; therefore, governing categories allow the antecedent of a reflexive to be further and further away from the reflexive. Each successive language in the hierarchy allows a greater number of possible antecedents than the previous one(s), as in Figure 1:

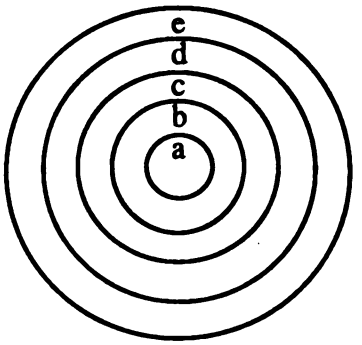


Figure 1: Subset relations among the values of the governing category parameter

Wexler and Manzini also propose the Lexical Parameterization hypothesis, which states that the values of the GCP are not specified for particular languages, but are associated with lexical items:

> (7) Lexical Parameterization hypothesis (Wexler and Manzini (1987, p.55)):
>
> Values of parameter are associated not with particular languages; but with particular items in a language.[5]

Turkish has two types of reflexive forms (Kornfilt, 1997). One of these reflexives ("kendi" — "self") is a morphologically simple form, whereas the other reflexive ("kendisi" — "self") is a morphologically complex form. The morphologically complex form ("kendisi") is associated with value (e) of the GCP, and it has long-distance binding. On the other hand, the morphologically simple reflexive ("kendi") appears to instantiate the same GCP value as the English reflexive does; hence, it is associated with value (a) of the GCP, and it has local binding.

1.5.1.2. *Proper Antecedent Parameter* . The other parameter relevant to the binding of reflexives, the Proper Antecedent Parameter (PAP), determines whether the reflexives must be bound by a subject NP or by any NP within the governing category:

(8) Proper Antecedent Parameter (Wexler and Manzini (1987, p.64)):

A proper antecedent for x is:
a) a subject y; or b) any element y whatsoever;

Languages like Japanese allow only subjects as antecedents of reflexives (value (a)), whereas languages like Turkish and English allow subjects and nonsubjects (value (b)). Consider the following example:

(9) Jane showed Mary a picture of herself

In this English example and its equivalent in Turkish, either "Jane" or "Mary" can serve as an antecedent of the reflexive, whereas in Japanese only "Jane" can.

1.5.1.3. *Markedness and the GCP* . Manzini and Wexler argue that the values of the GCP define a markedness hierarchy such that for anaphors the value (a) is unmarked and each subsequent value (b), (c), (d), and (e) is relatively more marked than the preceding value. Wexler and Manzini define markedness as follows:

(10) Markedness (Wexler and Manzini, 1987, p.430):[6]

Given the parameter P with values $P_1,..., P_n$, for every P_i and P_j,
$1 \leq i, j \leq n$ is less marked than P_j iff L (P_i) C L (P_j)

The notion of markedness within the UG framework can be defined as a measure of the amount of evidence a learner needs to arrive at the language-particular features of a grammar. The assumption based on the principles of UG is that the unmarked settings of parameters are those that the child learns from the least amount of positive evidence, whereas marked elements of UG need more evidence to be acquired than unmarked elements do.[7] Remember that the binding domains proposed by Wexler and Manzini are nested in a subset relation so that each successive domain includes the previous ones. They thus enter into a markedness hierarchy with each successive one requiring additional positive information for its selection as the proper one for the language. The parameter setting yielding the smallest subset language is the unmarked, initial assumption, and grammars yielding the superset are marked, adopted only if the relevant positive evidence occurs.

1.5.1.4. *Interaction of GCP and PAP*. Manzini and Wexler state that there is a correlation between the GCP and the PAP, such that languages that allow nonlocal binding require the antecedent to be a subject NP, whereas languages that require local binding allow the antecedent to be any NP. In other words, languages which permit long-distance binding cannot have both subject and nonsubject antecedents, and languages which permit only local binding can not have only subject antecedents.[8]

Wexler and Manzini propose a "Functional Principle" which prohibits both parameters of the GCP and PAP from being specified for the marked value simultaneously:

(11) Functional Principle (Manzini and Wexler (1987, p.437)):

An anaphor must be bound either in its unmarked governing category or by its unmarked proper antecedent

This principle prevents a language which has the marked value (e) for the GCP from instantiating the value (b) for the PAP. Eckman (1994, p.211) presents logically possible combinations of binding domains and antecedents:

(12)	Local	Local and nonlocal
Subject only	Type A (unattested)	Type B (Japanese, Korean)
Subjects and nonsubjects	Type C (English)	Type D (unattested)

It should be reminded that Turkish has the value (e) of the GCP, which allows both local and nonlocal antecedents, and has the value (b) of the PAP, which allows both subject and nonsubject antecedents. This means that Turkish falls into Type D, although Eckman as well as Wexler and Manzini claim that this type is unattested.

1.5.1.5. *The Subset Principle and the GCP* . Wexler and Manzini (1987) proposed the "Subset Principle" (SP) within a parametric theory of language acquisition to account for the learnability problem in cases where the parametric values, like those of the GCP, lie in a subset relation to one another. Wexler and Manzini (1987, p.414) state that the SP was developed as a modular theory of parameter setting. It is modular because markedness hierarchies are not built into UG but rather are derived from the interaction of UG and an

autonomous learning component. Hence, the SP is an independent learning principle that interacts with principles of UG during acquisition to guide the setting of UG parameters.

The subset relationship has a learnability side, the SP, and a linguistic side, the Subset Condition (SC). The SP and the SC are given as follows (Wexler and Manzini 1987, p.60):

(13) The Subset Condition:

For every parameter p and every two values i, J of p, the languages generated under the two values of the parameter are one a subset of the other, that is,

L (p(i)) L (p(J)) or L (p(J)) L (p(i)).

(14) The Subset Principle:

The learning function maps the input data to that value of a parameter which generates a language:
(a) compatible with the input data; and
(b) smallest among the languages compatible with the input data.

According to the SC, any two values of any parameter define languages which are a subset one of the other. According to the SP, on the basis of the input data, the learner selects the grammar that generates the smallest possible language that is compatible with the data. In order for the SP to apply, the languages generated by the various values of a particular parameter must lie in a subset relation to one another; that is, the SC must be satisfied. According to the SP, given two or more values of a parameter, the child will initially select the value which yields the smallest language.

1.5.1.6. *The learnability situation* . The learnability situation within the context of the GCP can be illustrated as follows; an English speaking child will initially select value (a) of the parameter, which generates the smallest language. Since the child will not come across any evidence which suggests that value (a) is not the correct one, the child will remain with the initially selected value (a), based on the assumption that the only type of data available to the child are positive. Negative data in the form of corrections or explanations are not available to the child. In the case of the Turkish speaking child, even if value (a) is selected first for the Turkish reflexive "kendisi," the child will abandon the value (a) for more inclusive parameter values when the child encounters additional positive evidence which shows that value (e) is correct for the Turkish reflexive "kendisi." The SP guides the child to choose the most restrictive setting available first, and to reject it only when positive data proves it to be wrong. In other words, the child will be guided by the SP to select value (a) first, and will move towards value (e) only in the presence of positive data which show that the smaller values are not correct.

Furthermore, a child learning English will require minimal positive evidence to set the GCP to the unmarked value (a), because this value defines the smallest language. On the other hand, a child learning Turkish will require relatively more positive evidence to arrive at value (e), because (e) is the most marked value which defines the largest language.

1.5.2. Non-parameterized approach to reflexive binding. A second revision of the standard Binding Theory is the non-parameterized approach (Lebeaux (1983), Pica (1987), Battistella (1989), Katada (1991), Reinhart and Reuland (1991), and Hestvik (1992))

which assumes that reflexives move in LF. The basic idea of the LF-movement approach is to keep the standard version of binding Condition A by proposing LF-movement for reflexives, thus attempting to reduce long-distance reflexivization to "a sequence of local dependencies" (Huang, 1994, p.101). According to the LF-movement approach, there are two types of LF-movement; 1) a local reflexive is adjoined to a position, such as VP, c-commanded by either a subject or an object, and 2) a long-distance reflexive is raised to a position, such as Infl, c-commanded only by a subject. Although the researchers have developed various versions of the proposal that reflexives undergo raising at LF, all of them claim that reflexives which are bound long-distance necessarily require subject antecedents, and that reflexives which can have subject or nonsubject antecedents must be locally bound.

According to Pica (1987), there is a relationship between binding domains and the morphological structure of reflexives. He proposes two types of reflexives; first, "compound reflexives," like English "himself," (which can be analyzed as [$_{np}$ [$_{spec}$ him] [$_{n'}$ self]]; that is, a maximal projection (or X^{max})), and they require local antecedents. The second type is "noncompound reflexives," like Japanese "zibun," which can be analyzed as [$_{np}$ [$_{n'}$ zibun]], an X^{o}, and allows binding across clause boundaries. According to Pica (1987, p.487), reflexives are referentially dependent expressions, and therefore they are unsaturated. Pica explains that reflexives are "unsaturated" in the sense that they are composed of a head noun ("self") and a specifier, but since a reflexive expresses "inalienable possession," it cannot be coindexed at S-structure with its specifier, and is thus

said to be an "open position." Since all unsaturated arguments (in an open position) at one level must find saturation at another level, Pica argues that this motivates the movement of reflexives at LF. Non-compound X^0 reflexives raise into INFL by head-to-head movement whereas compound X^{max} reflexives are maximal projections and they adjoin to maximal projections. Therefore, these reflexives adjoin to their immediate verb phrases and are bound at that adjunction place. Pica adds that after an X^0 reflexive is raised out of VP into INFL, the subject NP will c-command the reflexive at LF, but nonsubject NPs will not. Therefore, only subjects can be possible antecedents for X^0 reflexives. On the other hand, X^{max} reflexives are adjoined to VP, where they are still within the c-command domain of nonsubject NPs. Therefore, X^{max} reflexives may take either subject or nonsubject antecedents.

Battistella (1989) applies a "movement-to-INFL" at LF analysis to the Chinese reflexive "ziji" (an X^0 category). According to his analysis, "ziji" undergoes LF-movement first from the object position of the lowest clause to the INFL position of that clause, and then from the INFL position of the lowest clause to the INFL position of the intermediate clause, and finally from the INFL position of the intermediate clause to the INFL position of the matrix clause. In each case, movement leaves a trace. Long-distance binding of "ziji" is analyzed, therefore, as the result of cyclic head-to-head movement of "ziji" from INFL to INFL. Battistella argues that Chinese "ziji" is subject to binding Condition A to be bound to a subject in its governing category at LF.

Cole, Hermon and Sung (1990), like Battistella (1989), claim that long-distance long-distance reflexivization should be analyzed in terms of Infl-to-Comp-to-Infl movement at LF within Chomsky's (1986b) *Barriers* work. They claim that Infl in Chinese is lexical rather than functional. Therefore VP and CP in Chinese are not barriers for movement and accordingly "ziji" can undergo Infl-to-Comp-to-Infl at LF. Cole, Hermon and Sung (1990) try to find out how long-distance reflexivization is licensed in a language. Their answer is that long-distance reflexives can occur only in languages in which Infl is lexical and hence VP is not a barrier. On the other hand, long-distance reflexivization is forbidden in languages in which Infl is functional and VP is a barrier to movement.

In summary, the relationship between long-distance binding and subject orientation is the basis for the assumption of movement in LF. Pica (1987) indicates that the implicational relationship holds only in one way; long-distance binding entails subject antecedents, but not vice-versa. A nonsubject antecedent entails local binding, but not vice-versa. Consequently, anaphors which are subject to long-distance binding are subject-oriented, whereas anaphors which allow nonsubject antecedents must be locally bound.

1.5.3. Summary. The GCP, a parameterized approach to language acquisition, and LF movement, a nonparameterized approach to language acquisition, are the most common approaches to reflexive binding in SLA. The main difference between these two frameworks is that in LF movement, the L2 learners identify whether a reflexive moves locally or long-distance according to whether it is a compound (complex) or a noncompound (simple) reflexive, whereas in Manzini and Wexler's GCP framework, the

SP organizes learners' hypotheses about the properties of reflexives so that they will start with the unmarked setting of the binding parameter. On the other hand, these two theories appear to make a similar claim. According to LF movement, noncompound reflexives (like Turkish "kendisi"), which allow long-distance binding, can be bound by only subject antecedents, whereas compound reflexives (like English "himself"), which allow only local binding, can be bound by either subject or nonsubject antecedents. According to Manzini and Wexler, there is a correlation between the GCP and the PAP such that certain settings of the GCP can cooccur with only certain settings of the PAP. In other words, languages that have the unmarked setting of the GCP can have only the marked setting of the PAP, and conversely, languages that set the GCP for the marked value can have only the unmarked setting of the PAP. This means that, as LF movement also claims, languages cannot have a marked setting for both the GCP and PAP. According to Wexler and Manzini, long-distance binding and subject-only antecedents are the result of two different parameter settings, whereas at LF the connection between long-distance binding and a subject antecedent is a result of the movement of reflexives (Thomas, 1993, p.158).

A full analysis of the study of both parameterized and nonparameterized binding principles goes beyond the scope of this study. Furthermore, the experiment in this study was not designed to test the validity of the predictions of the GCP and LF movement. These theories describe L2 reflexives in interlanguage grammar only in terms of syntactic constraints. However, in this study, the interpretation of the anaphoric relations between antecedents and reflexives in L2 is explained in part in terms of a pragmatic theory. As a starting point, this study employs Principle A of the Binding Theory of Chomsky (1981,

1986a) on which the GCP framework is based, and also employs the parameterized approach in order to see how in L2 acquisition the knowledge of pragmatics interacts with syntactic knowledge, which is assumed to be autonomous and modular within the GB framework.

1.6. Plan of this work.

The body of this study will proceed as follows; chapter 2 will review and comment on previous studies on reflexive binding in SLA which have been carried out following Government-Binding theory and LF movement. Chapter 3 will discuss issues in SLA research methodology, especially approaches having to do with the elicitation of acceptability judgments. Chapter 4 will provide the data from an experimental study carried out in this work, including information about the subjects, method, aims and hypotheses. Chapter 5 will provide the results of the experiment along with discussion and conclusions.

Chapter 2

PREVIOUS STUDIES OF THE ACQUISITION OF L2 REFLEXIVES

2.1. Introduction.

Second language studies on reflexive binding have been conducted within several different binding frameworks, employing a variety of methodologies, and have examined different acquisition patterns. Most second language acquisition research on reflexive binding has been carried out within the framework of the GCP and the Proper Antecedent Parameter (PAP), proposed by Wexler and Manzini. Thus, it appears necessary first to review relevant previous studies which employ the framework of the GCP and PAP.

Research by Finer and Broselow (1986), Thomas (1989, 1991a), Finer (1991), Hirakawa (1990), Lakshmanan and Teranishi (1991), Eckman (1994), MacLaughlin (1995a) has examined whether adult learners can reset the GCP and PAP when the reflexive binding settings of their L1 differ from those of the L2. Research by Christie and Lantolf (1993), Thomas (1995), and White, Hirakawa and Kasaki (1995) has examined whether adult learners whose L1 allows only local binding can move reflexives in LF in an L2 which allows long-distance binding.

2.2. Studies within the parameterized approach.

2.2.1. Finer and Broselow (1986): "Second Language Acquisition of Reflexive Binding." Finer and Broselow's experiment investigates the acquisition of English reflexives by six adult native speakers of Korean, five intermediate level and one advanced, in order to find out whether adult learners rely on the subset principle in the setting of the parameters of L2. The experiment involved a picture identification task with sixteen complex sentences, both finite and nonfinite. The subjects were given two pictures for each stimulus sentence and asked to indicate which picture the sentence described. The following are examples of the stimulus sentences:

(15) i. Mr. Fat thinks that Mr. Thin will paint himself.

ii. Mr. Fat wants Mr. Thin to paint himself.

In terms of the Governing Category Parameter, English has the smallest value (a) at one end of the markedness hierarchy proposed by Wexler and Manzini (1987), whereas Korean has the largest value (e) at the other end. Therefore, in both (i) and (ii) only "Mr. Thin" can be the antecedent of "himself" in English, but in the Korean counterparts either "Mr. Fat" or "Mr. Thin" can be potential antecedents for the Korean reflexive.

The study suggested the following results:

i. Korean speakers would simply transfer the Korean parameter setting into English.

ii. The learners would start with the unmarked, subset value (a) for English. In this

case, the acquisition of the binding principles of L2 would be the same as that of the binding principles of L1.

iii. The interlanguage of the learners would exemplify the binding principles of neither English nor Korean:

(a) The grammar of learners might show the binding properties of no natural language.

(b) The grammar of learners might show one intermediate value between the two on the hierarchy.

The results of their experiment show that most of the learners' responses favored a local antecedent in finite clauses (92%), which indicates that learners appear to observe the SP. However, in nonfinite embedded clauses, nonlocal responses were quite frequent (38%), although a majority of the responses selected the local-only type. Finer and Broselow report that the Korean learners were making a distinction between the tensed and infinitival clauses although neither Korean nor English make the tensed/infinitival distinction for anaphoric binding. In terms of the hierarchy of binding domains proposed by Wexler and Manzini, the subjects were selecting a value for the GCP that was somewhere between the (a) value for English and the (e) value for Korean; they were possibly selecting the value (c) (e.g., for Russian). Thus, Finer and Broselow claim that the Korean learners have come up with a set of binding principles that are consistent with the parameters provided by UG, but inconsistent with either their first or their second language. Finer and Broselow argue that transfer appears to play no role in SLA; otherwise, if there were transfer, the subjects would have started with the Korean

parameter setting, and with exposure to negative evidence they would have gradually moved from the value (e) to (a). Finer and Broselow argue that their L2 learners didn't obey the subset principle either; otherwise, the learners would have started with the English setting because there would be no positive evidence from English to make the learners move to a more marked setting on the parameter. We would expect the learners to prefer the local interpretation of reflexives in all types of the English sentences.

Finer and Broselow bring another analysis to the data in which they suggest that the SP is fully operative. Besides the difference in binding domains, English and Korean differ in their choice of what can count as a proper antecedent for the reflexive (Proper Antecedent Parameter). In monoclausal sentences in Korean, only subjects can be antecedents, whereas in English both subjects and objects are potential antecedents. According to the Proper Antecedent Parameter, Korean is unmarked (smaller set) and English is marked (larger set). However, according to the GCP, English is unmarked (smaller set) and Korean is marked (larger set). Finer and Broselow assume that Korean learners are interpreting the subjects of infinitives as objects of monoclausal sentences and hence they are applying the Proper Antecedent Parameter to infinitives using the English setting. Therefore, the L2 learners assume that the choice of an antecedent within a nonfinite clause is free; that is, either the matrix clause subject NP or the subject of the nonfinite clause can be a potential antecedent of the reflexive. Consequently, the marked (larger) setting for the Proper Antecedent Parameter is operative in selecting the unmarked (smaller set) for the Governing Category Parameter.

2.2.2. Finer (1991): "Binding parameters in second language acquisition."

Finer (1991) conducted a further study to verify the results found in the previous study (Finer and Broselow (1986)). The method was similar; a picture identification task was employed. The subject population was enlarged to include 37 native speakers of Japanese (value (e) on the hierarchy); 30 native speakers of Korean (value (e) on the hierarchy) and 30 native speakers of Hindi (value (c) on the hierarchy). Subjects simultaneously heard and read 24 sentences involving reflexives; 13 sentences with either object control or exceptional case marking; such as "Mr. Fat expects Mr. Thin to paint himself"; nine sentences with tensed complements, such as "Mr. Fat thinks that Mr. Thin will paint himself," and two sentences with "subject control verbs," such as "Mr. Fat promises Mr. Thin to paint himself."

Finer claims that what he found in the replication for the Korean and Japanese speakers confirmed the general pattern shown by the Koreans of the initial experiment. That is, reflexives in tensed clauses were more likely to be bound locally (97% of the time for Korean subjects, and 91% for Japanese subjects) than reflexives inside infinitives (88% for Korean subjects, and 76% for Japanese subjects). While the speakers of Japanese and those of Korean made a distinction between clause types, the speakers of Hindi did not. Their English reflexives were bound locally in both types of examples, to an overwhelming degree.

Finer claims that his subjects were moving along the hierarchy of the GCP to a more restrictive setting. Their interlanguages at this point were compatible with UG based principles of binding; if there was movement along the hierarchy, it was movement towards

the target language that is consistent with a possible parametric option of the hierarchy. Finer's explanation entails that his subjects were not using the SP from scratch because that would have meant their adopting setting (a) of the GCP from the very beginning.

Finer, furthermore, uses the "Functional Principle" of Wexler and Manzini (1987) in order to explain his results. According to the Functional Principle, both parameters, namely the GCP and the PAP, are prohibited from being specified for the marked value simultaneously. Therefore, a language which has the value (e) for the GCP can not have the value (b) for the PAP. Japanese and Korean have a marked value for the GCP, and their anaphors are strongly subject-oriented, indicating the unmarked value of the PAP. English, however, has the least marked value (a) for the GCP, whereas it has the most marked value (b) for the PAP. Finer claims that in the absence of negative evidence, the Japanese and Korean learners selected an intermediate value of the GCP; a value between the native language and the target language. Finer argues that the recognition of a marked value for the PAP serves as a trigger because, in the context of the Functional Principle, these two parameters are basic linguistic knowledge. Therefore, the L2 learners will avoid a marked-marked pairing in their L2 mapping of the GCP and PAP to parametric values. The learners therefore recognize that a marked value on one parameter will require a retraction along the markedness scale on the other parameter.

2.2.3. Thomas (1989): "The interpretation of English reflexives by non-native speakers."

Thomas' study investigates the predictions of the transfer hypothesis and the Subset Principle in L2 reflexive binding. Her subjects consisted of 24 native Chinese speakers

and 29 native Spanish speakers learning English as a second language. She had 11 native English speakers as a control group. Her investigation is based on the fact that a reflexive in a biclausal finite clause needs a clause-mate antecedent, which Thomas calls the Clause-Mate Condition (CMC).

The experimental task was a 30 item multiple-choice questionnaire which included four types of sentences. In the questionnaire, a question following each sentence asked the subjects to indicate the antecedents of the reflexive by circling one of three multiple-choice answers. The following example illustrates this test format employed by Thomas (p.286):

(16) David could see that Bill was looking at himself in the mirror.

Who did Bill see in the mirror?

a. Bill
b. David
c. Either Bill or David

What Thomas calls Type I and Type II were biclausal finite sentences which checked the subjects' knowledge of the Clause-Mate Condition (sixteen tokens) while Type III and Type IV were monoclausal sentences which tested the subjects' knowledge of the "Subject Strategy" (fourteen tokens). The Subject Strategy refers to the notion that subject NPs are preferred to nonsubject NPs as antecedents in single clauses although both are possible antecedents. For example, in the sentence "John asked Bill to talk about himself," the Subject Strategy predicts that the subject "John" will be preferred to the nonsubject "Bill." We will examine only the part of her data obtained from testing the CMC that is

relevant to the concerns of this paper.

The experiment's results do not support the transfer hypothesis because the learners didn't simply carry the parameter settings of their L1 to their L2. The transfer hypothesis predicts that the learners whose L1 and L2 differ parametrically will have greater difficulty in establishing the correct L2 setting than learners whose L1 and L2 parameter values match. However, no statistically significant differences emerged between the Spanish and Chinese speaking groups although both languages have different values for the GCP for reflexives (Spanish has the value (a) for reflexives like English whereas Chinese has the value (e)). Sixty-nine percent of the Chinese responses and 60% of Spanish responses selected the local NP as antecedent in biclausal finite sentences. According to Thomas, the Subset Principle cannot explain the data related to the CMC because the subjects have set the GCP wider than is appropriate for English. If learners had reset the GCP back to its unmarked value, they would not have found any evidence to depart from the unmarked setting of their L2 (English).

Thomas (1989) also looked at the pragmatic influences on the interpretation of reflexives. Half of the biclausal sentences in the task were neutral (Type I sentences, 6 tokens) where both the local and nonlocal NPs were semantically and pragmatically likely antecedents for the reflexive. The other half of the biclausal sentences (Type II sentences, 10 tokens) were semantically and pragmatically biased in the direction of the nonlocal NP, the one not syntactically acceptable as an antecedent in English. Thomas' purpose in contrasting neutral and biased stimuli is to determine whether L2 learners' interpretation of reflexives are controlled by grammatical rules, or whether pragmatic considerations control

the choice of antecedent for anaphors. One of the pragmatically biased sentences in Thomas' study was (p. 287):

(17) Mary angrily told me that Susan has spilled a lot of paint on herself.

Type II sentences are designed to make the interpretation in which the reflexive is bound to the matrix clause subject (NP1) a more plausible reading than one in which the reflexive is bound to its clausemate (NP2).

Thomas' results showed that pragmatic favoring of the long-distance NP in biclausal sentences induces only about a 23% reduction in local binding among Chinese learners (49% selected the local NP, 34% the nonlocal NP) and Spanish learners (49% selected the local NP, 37% the nonlocal NP). According to Thomas, this means that the L2 learners resisted assigning a long-distance antecedent to a reflexive in pragmatically biased sentences. Thomas claims that pragmatic bias fails to induce L2 learners to allow long-distance binding. Therefore, it was concluded that the L2 learners' partiality for local antecedents is due to formal syntactic principles and not to pragmatic factors.

2.2.4. Thomas (1991a): "Universal Grammar and the interpretation of reflexives in a second language." Thomas' 1991 study tries to find out whether adult language learners have access to the principles and parameters of UG in SLA. Thomas tested the predictions of the transfer hypothesis and those of the SP for the interpretation of English reflexives by 132 native speakers of Japanese and Spanish, and for the interpretation of the Japanese

reflexive by 41 native speakers of English and Chinese. Within each language group, there were three proficiency levels.

The experimental task used in the study was a multiple-choice test which consisted of six types of sentences, three for learners of English and three for learners of Japanese, and each type was represented by three tokens. These sentence types were biclausal sentential complement (for learners of English and Japanese); biclausal relative clause (for learners of English); single clause: subject vs. nonsubject NPs (for learners of English); single clause: c-commanding vs. non c-commanding NPs (for learners of Japanese) and single clause: subject or nonsubject NPs (for learners of Japanese). We will only report the subjects' interpretations of biclausal sentential complement sentences and relative clauses which investigate GCP settings.

Results for the learners of English (Chinese and Spanish native speakers) show that in both biclausal sentences and relative clauses, most of the L2 learners bound reflexives only to local antecedents (70% to 91% of every group of L2 learners for biclausal sentences, and 60% to 96% of every group of L2 learners for relative clauses). A minority allowed either local or long-distance antecedents. According to Thomas, when all L2 learners' answers to these two types of sentence were pooled, the vast majority appeared to reset the GCP for English because 81% of the subjects consistently bound reflexives to local NPs.

When we turn to the interpretation of the data provided by learners of Japanese as a second language, Thomas reports that most of the responses from native speakers of English with low- and mid-level proficiency to biclausal Japanese sentences bound

reflexives locally. Thomas states that this may show that at these levels, learners do not have enough positive evidence to go to the larger setting for the Japanese reflexive. However, for the high-proficiency level learners with greater exposure to the L2, the responses which allowed either local or long distance binding increased to 30.8%, compared to 12.5% for low-level and 8.3% for mid-level learners. Half of the Chinese speakers (all subjects were at the same proficiency level) learning Japanese permitted long-distance anaphors exlusively, but the GCP has no value for such an interpretation of reflexives; in other words, there is no natural language which allows only long distance antecedents while excluding local ones. Furthermore, no Chinese learner consistently selected the Japanese reflexive to be bound by either a local or long-distance antecedent, although Chinese allows both local and long-distance antecedents as Japanese does.

Thomas claims that her data supports the principles and parameters of UG in determining the referential properties of reflexives in L2 and indicates that adult learners have direct access to UG. According to Thomas, the transfer hypothesis cannot be observed in the data because there is almost no difference between the native speakers of Spanish and Japanese in their interpretation of English reflexives. Furthermore, L2 learners are able to reset parameters successfully in their L2 because the vast majority of Japanese speakers even at lower levels of proficiency can bind English reflexives locally.

2.2.5. Hirakawa (1990): "A study of the L2 Acquisition of English reflexives."

Hirakawa's study examines how native speakers of Japanese acquire the syntactic properties of English reflexives in the context of the GCP and PAP. The study involved

four experimental groups of 65 native speakers of Japanese learning English as a second language at four proficiency levels; grades 10, 11, 12 and 13, and two control groups; one of 22 native Japanese speakers and another of 20 native English speakers. The subjects' knowledge of English reflexive binding was measured by a multiple choice grammaticality judgment test similar to Thomas' (1989). The test consisted of five sentence types; finite biclausal, finite triclausal, nonfinite biclausal, nonfinite triclausal, and single clauses. Each type was represented by five token sentences. The subjects were presented with a sentence containing the reflexive and were asked to indicate who the reflexive referred to by circling one of the given choices. The following example illustrates the test item from Hirakawa's study (p.70):

(18) John said that Bill hit himself

a. John
b. Bill
c. either John or Bill
d. someone else
e. don't know

Hirakawa concluded that there was no significant difference among the different proficiency levels in the interpretation of reflexives. Furthermore, the subjects did not follow the SP since they selected a non-local NP as the antecedent for the reflexives in several cases. She argues that the Japanese L2 learners set the parameter to the largest value, allowing a non-local antecedent not only in the case of nonfinite embedded clauses (55.1% local antecedent, 36.5% local one) but also in the case of finite embedded clauses (77% local antecedent, 17.1% nonlocal one). Moreover, her results indicate that the learners did best in finite biclausal sentences followed by finite triclausal sentences. The

subjects performed the worst in nonfinite structures. The L2 learners were much more accurate in finite clause sentences than in nonfinite clause structures. The learners interpreted more non-local NPs as the antecedent of the reflexive in nonfinite clauses than in finite ones, which replicates Finer and Broselow's findings. Hirakawa concludes that the subjects were affected by the infinitival more than by the levels of embedding.

Hirakawa reports that the Japanese control group showed a definite preference for the nonlocal NP over the local one. Hirakawa adds that the Japanese learners of L2 English did not seem to have been affected directly by their L1 preference since their responses do not show the Japanese control group's pattern of choosing nonlocal antecedents. Therefore, the learners do not transfer the rules of reflexive binding from their L1 to their L2. Hirakawa claims that Japanese learners can eventually reset the GCP to the English value. She reports that there were ten subjects (out of sixty-one) who responded 100% correctly. These subjects prove that resetting of the parameter in L2 is possible. Hirakawa adds that there were also six subjects who responded almost perfectly, making just one error. According to Hirakawa, although Japanese learners may fail to choose the right setting for English reflexives at the initial stages of L2 acquisition, they are able to succeed eventually over time.

2.2.6. Teranishi and Lakshmanan (1991): "Preferences vs. grammaticality judgments: Some methodological issues concerning the Governing Category Parameter in second language acquisition." Teranishi and Lakshmanan attempt to find out whether native speakers of Japanese acquiring English as a second language initially select the value

consistent with the most restrictive grammar associated with their L2, or whether they transfer the superset value associated with Japanese "zibun" to their L2.

The study consisted of two groups; a control group which consisted of 11 native speakers of English and an experimental group which included 34 native speakers of Japanese learning English as a second language. Twenty-three of the 34 students in the experimental group were in four different proficiency levels; beginning, low-intermediate, upper-intermediate and advanced. The 11 remaining of the 34 students in the experimental group were taking English composition courses and also served as a control group for Japanese.

In order to increase the responses to "either the local or non-local NP," Teranishi and Lakshmanan used an acceptability task that measured subjects' intuitions about who the reflexive *cannot* refer to. Subjects were asked to consider each statement and then indicate whether they disagreed or agreed with the statement by circling the relevant option, for example:

(19) John said that Bill saw himself in the mirror.

1. "himself" cannot be John	agree	disagree
2. "himself" cannot be Bill	agree	disagree

In their task, Teranishi and Lakshmanan attempted to avoid a semantic bias toward one of the two antecedents (the local or the non-local NP) through the use of neutral verbs. Teranishi and Lakshmanan argue that in previous research, native speakers of Japanese strongly preferred a reading where "zibun" was bound to a non-local antecedent; however,

native speakers of Japanese in their study do not display a preference for the non-local antecedent. Instead, a majority of the subjects' responses in the Japanese test (57.9% of the responses) were consistent with the reading wherein *zibun* is bound either to the local or non-local NP.

The majority of the responses of the Japanese L2 learners in the experimental group chose the local NP as the antecedent for the reflexive (77.9%). According to Teranishi and Lakshmanan, this suggests that the L2 learners in their study, considered as a group, were not treating the English reflexive like Japanese "zibun." In terms of the different proficiency levels, the beginning and intermediate level students had a higher percentage of incorrect responses when compared to the learner at the higher levels (17.5% vs. 7%). Sixty-four percent of responses of Level I and Level II learners selected the local NP, and 83% of Level III and Level IV learners selected the local NP. The 11 subjects who were taking English composition courses (81% selected the local NP) didn't make any non-local type response. All of their incorrect responses were of the "either" type.

Out of the 34 L2 subjects, 14 chose local antecedents only for all the ten sentences used in the English L2 test. This suggests that these subjects have successfully fixed the value of the GCP for the English reflexive.

Thirty-three of the incorrect responses to the English L2 test made by the remaining twenty L2 learners were consistent with the "either" choice. This may suggest that the subjects had transferred the GCP value that is associated with "zibun" to the English reflexive.

According to the Teranishi and Lakshmanan's results, 14 successful L2 learners

reset the GCP value correctly for the English reflexive even though this necessitated their moving from a larger grammar to a smaller one. In order to explain this movement, Teranishi and Lakshmanan state that the highest number of incorrect responses in the L2 test were consistent with the "either" choice. It is possible that during a previous stage these 14 successful learners were operating with a nontarget grammar. Although the Subset Principle doesn't permit a move from a larger grammar to a smaller one, these 14 learners came to reset the GCP to the English value. Teranishi and Lakshmanan claim that the main reason for this successful movement from a larger setting to the smaller one might be that Japanese has both types of reflexives; the morphologically complex reflexives, like "zibun-zisin" and "kare-zisin," and simple forms, like "zibun." The complex forms are associated with the smallest value of the GCP (value a) and the simple forms with the largest value of the GCP (value e). The English reflexive "himself" is initially analyzed as the morphologically simple form "self" by Japanese L2 learners, and associated with the Japanese simple reflexive form "zibun," causing a transfer of the GCP value for "zibun" (value (e)) to the reflexive in the English interlanguage). Later, when the L2 learners realize that English reflexives are morphologically complex, they associate English reflexives with the Japanese complex reflexives, which, in Japanese, have the smallest value of the GCP, like English. Then the learners transfer the subset value from Japanese to their interlanguage, setting the right value for English. Teranishi and Lakshmanan suggest that through this process of "interlingual identification," the L2 learners can successfully move from a superset to a subset grammar.

2.2.7. Eckman (1994): "Local and long-distance anaphora in SLA." Eckman's study considers the issue of the binding of reflexives with respect to two parameters of UG, namely the GCP and the PAP, from the viewpoint of two target languages — English and Japanese. The subjects in this study were twenty four university-age native speakers of English learning Japanese as a second language and twenty five ESL learners from four language backgrounds; 13 Arabic, 4 Japanese, 5 Mandarin, and 3 Spanish, as well as two control groups of English and Japanese speakers. The experiment involved a picture identification task in which the subjects had to match a picture with a sentence containing a reflexive. The task included four sentence types; simple ("Mr. Small hit himself"), biclausal finite ("Mr. Small said that Mr. Big looked at himself"), biclausal nonfinite ("Mr. Big asked Mr. Small to hit himself"), and simple clause with two NPs (Mr. Big gave Mr. Small a portrait of himself"), where each type had four tokens. Since the aim of the study was to find out whether an individual's interlanguage grammar obeys the constraints of UG, Eckman looked at individual results rather than group results. Since there were four tokens of each sentence type, Eckman decided that the threshold for systematicity was 3 out of 4. If a subject didn't reach the criterial threshold of 3, the IL was deemed to be unsystematic with respect to that aspect of coreference assignment. His results showed that 17 out of the 25 ESL learners performed systematically with respect to the English binding domain. Of these 17 subjects, 16 bound the reflexives only locally, and one allowed both local and nonlocal binding. When we look at the English learners of Japanese we find that 7 learners allowed only local binding with only subject antecedents, and 2 learners allowed only nonlocal binding. Fourteen subjects bound the reflexive both locally and nonlocally.

Eckman examined whether the subjects' interlanguage grammars obeyed the settings of the GCP and PAP. His data show that the L2 learners of English preferred local binding with the subject-only antecedent rather than local binding with both subject and nonsubject antecedents in simple clauses with two NPs. This type of combination of local binding with subject-only antecedent is not allowed by the GCP and PAP, and therefore is presumably not attested. However, according to Eckman, the subjects' interlanguage grammars obey UG constraints compared to native speakers of English, because even native speakers of English show the same preference, i.e., subject-only antecedent in simple clauses with two NPs, although their underlying grammar has local binding with both subject and nonsubject antecedents. Therefore, it is natural for interlanguage grammars to show the same strategy found in target grammars. Eckman states that the L2 learners have built a less marked version of the target language by choosing only subject antecedents within the local domain (the value (a) of the GCP and the value (a) of the PAP) to bind reflexives, although the grammars of both the native language and target language contain the more marked instantiations of reflexive binding. (Japanese selects the value (e) of the GCP and the value (a) of the PAP, while English has the value (a) of the GCP and the value (b) of the PAP.)

2.2.8. MacLaughlin (1995a): "Universal grammar and L2 acquisition of reflexive binding: Some learners acquire a non-L1/non-Target system?" In this study, MacLaughlin reanalyzes the data reported in Hirakawa (1990) in terms of a view of individual subjects. She tries to show that some of Hirakawa's second language learners acquire a system of

reflexive binding that is different from both the native language and the target language. She also addresses Yuan's (1994) claim that transfer is the only developmental factor in the acquisition of English reflexives by non-native speakers.

MacLaughlin distinguishes three types of reflexives and three respective domains of reflexive binding:

a) Type 1 reflexives are local and require a local domain,

b) Type 2 reflexives are bound by an antecedent outside an infinitival clause but not outside a finite clause and

c) Type 3 reflexives can be bound by an antecedent outside a finite clause.

For example, the English reflexive "himself" is a Type 1 reflexive, the Russian reflexive "svoj" is a Type 2 reflexive, and Japanese reflexive "zibun" is a Type 3 reflexive.

It should be noted again that Hirakawa tested 65 native Japanese speakers learning English as a second language as well as testing both English and Japanese controls. The L2 learners belonged to 4 different schooling levels or grades. The test consisted of a multiple-choice grammaticality judgment test containing 5 sentence types with 5 tokens of each sentence type. Her results showed that the L2 learners, as a group, allowed long-distance binding in both finite and non-finite sentences. However, their percentage of long-distance binding was lower than what was found with the native Japanese controls. The L2 learners allowed more long-distance binding in infinitival than in finite sentences.

The individual analysis of MacLaughlin shows that 10 subjects responded 100% correctly and that these subjects acquired Type 1, local binding. According to

MacLaughlin, one subgroup of subjects showed a Type 3 binding system, as in their native language. Furthermore, as MacLaughlin pointed out, another subgroup seemed to have acquired a Type 2 system, i.e., some subjects allowed reflexive binding outside an infinitival clause while consistently disallowing reflexive binding outside a tensed clause. According to MacLaughlin, this system of Type 2 binding is not found in the subjects' native language. Therefore, it cannot simply be the result of transfer, which appears to contradict Yuan's claim that studies of L2 reflexive binding do not give evidence for parameter setting or access to UG. According to Yuan, the results of these studies indicate that L2 learners transfer parameter settings from the L1. Yuan points out that the first languages of the subjects used in the studies on reflexive binding such as Japanese, Chinese, and Korean contain both Type 1 (local) and Type 3 (long-distance) reflexives. Therefore, according to Yuan, learners who correctly acquire the local binding properties of English in these studies have simply transferred the Type 1 local reflexive from the L1, whereas learners who allow long-distance binding have transferred properties of the Type 3 long-distance reflexive. Yuan believes that the tensed/infinitival asymmetry (Type 2 binding according to MacLaughlin) is the result of a misanalysis of the infinitival structure, because learners are analyzing the biclausal infinitival structures as monoclausal structures in which both the matrix subject and the second noun phrase are possible antecedents. Consequently, Type 2 binding out of infinitival clauses is, in fact, local binding within this misanalyzed binding domain. Yuan (1994) proposes that L2 learners misanalyze infinitival structures as follows:

(20) a. L2 learner structure: [Mary asked Ann to introduce herself]$_{\text{local binding domain}}$

b. Target structure: [Mary asked Ann [PRO to introduce herself]]$_{\text{local domain}}$

According to MacLaughlin, there are several problems with Yuan's proposal. For instance, the source of the misanalysis is not clear at all. Moreover, there is ambiguity in the transfer source since the native language contains two types of reflexives to transfer; some learners may transfer the local reflexive, while others transfer the long-distance reflexive. MacLaughlin states that learners who misanalyze infinitivals should reject binding of pronouns within the misanalyzed local domain too. For example, in sentence (18a), if the reflexive is replaced by a pronoun (like "her"), learners shouldn't bind the pronoun to the matrix subject ("Mary") to avoid violating Principle B. However, MacLaughlin says that experimental data disproves this claim. MacLaughlin concludes that some second language learners of English exhibit a pattern of reflexive binding that is not found in the native language, nor in the target language, but is within the range of possibilities allowed by UG.

2.3. Studies within the non-parameterized approach.

2.3.1. Christie and Lantolf (1993): "Bind me up bind me down: A UG study of reflexives in the second language." Christie and Lantolf's study deals with the availability of UG to L2 learners within the context of Binding Theory by using the theoretical framework of "Head Movement Analysis" by Cole, Sung and Hermon (1990) which assumes reflexive binding properties to be syntactically related to each other; long-distance reflexives are

morphologically complex and they can only be bound by subject antecedents. Their subjects were 22 Chinese and 15 Spanish learners of English, and 24 English learners of Chinese, as well as 26 English learners of Spanish. Their data collection involved a "truth-value judgment task" where subjects were shown a picture that represented the interpretation of the reflexive used in the stimulus test sentence. The subjects had to decide whether the meaning of the sentence matched the picture. The study included four types of structures; the first type included both finite and nonfinite biclausal sentences to test the binding domain of English and Chinese reflexives (e.g., "Steven thought that Rita had hurt herself in a car accident"). The second type was monoclausal sentences which tested the subject and object orientation of English reflexives (e.g, "Tom told Sam a secret about himself"). The third type was also monoclausal sentences which tested the c-command requirement (e.g, "Mary's father pointed at himself in the mirror"). The fourth type was simple sentences as control items (e.g, "John threw himself in the river"). The first 3 types were represented by 8 sentences each, and the last type was represented by 4 sentences.

In terms of data analysis, Christie and Lantolf mainly concentrated on responses of individual learners and used a cluster analysis on the results which compared individuals instead of group responses and determined the degree of correlation among several variables in the individual grammars. The variables in their study were long-distance binding of reflexives and subject orientation, the binding domains of reflexives, and c-command. This cluster analysis reveals how far away a subject in a group is from the 100% correct point on all three variables.

Individual results revealed that most of the English learners of Chinese allowed

long-distance binding in their L2. However, most of these learners also allowed binding of Chinese reflexives to either a subject or object antecedent, which is ruled out by head-movement analysis and UG. Furthermore, several English learners of Chinese allowed only object binding while disallowing subject binding in long-distance binding. Christie and Lantolf conclude that there is little evidence that long-distance binding correlates with subject orientation in L2 learners' grammars. Therefore, according to Christie and Lantolf, their data do not support the raising of anaphors in LF. Moreover, c-command violation is quite high for both Chinese and Spanish learners of English, which violates the principles of UG. Christie and Lantolf present two sorts of interpretation for their results. The strong interpretation suggests that UG is not available to the L2 learners in their study (since the data in most cases show no clustering effects). The weaker interpretation claims that the conception of UG theory is inadequate as far as binding aspects are concerned. Hence, binding principles are independent phenomena and should be investigated independently of UG.

2.3.2. Thomas (1995): "Acquisition of the Japanese reflexive "zibun" and movement of anaphors in logical form." Thomas' study looked at the acquisition of the Japanese reflexive "zibun" by native speakers of English, Chinese, Korean, Thai, French, Spanish and German, with the assumption that reflexives move in LF. She tries to find out whether L2 learners know that Japanese allows long-distance binding and know that "zibun" must take a subject as its antecedent. Thomas assumes that if L2 learners are constrained by UG, and if they move anaphors in LF, then they should reject the binding of "zibun" by

long-distance nonsubject NPs. In order to address these issues, Thomas proposes that research must do the following:

a) It must analyze individual responses
b) It must investigate an L2 with long-distance binding
c) It must use a methodology that prevents the problem of preference of one antecedent over another when the sentence is ambiguous

Her experiment consisted of 58 learners of Japanese as a second language, thirty-four in a low proficiency group and 24 in a high proficiency group, and a control group of Japanese native speakers. About one-third of the 58 learners of Japanese were native speakers of languages with long-distance reflexives, such as Chinese, Korean and Thai, while 32 subjects were native English speakers and 6 were speakers of French, Spanish and German. The subjects were given a truth-value judgment task which included 16 illustrated 3 to 5 sentence stories in Japanese with both Japanese orthography and romanized text. The subjects were supposed to say 'yes' or 'no' according to whether the statement after each story made sense with reference to the story. There were four sentence types, with four tokens per type. These sentence types were; a) monoclausal sentences with binding to subjects, b) monoclausal with binding to objects, c) biclausal sentences with binding to long-distance subjects, and d) biclausal with binding to long-distance objects. Examples of the sentence types from Thomas (1995, p.221) are:

(21) a. A spoke with B about self's problems.
b. The mother introduced A to self's teacher.
c. A knows that B will use self's car.
d. C told B that A likes self's book

Type (a) investigates whether subjects allow coreference between "self" and a local subject, and type (b) looks at whether subjects allow coreference between "self" and a local nonsubject. Type (c) examines whether subjects allow coreference between "self" and a long-distance subject, and type (d) looks at coreference between "self" and a long-distance nonsubject.

Thomas reports that in Japanese, binding to a subject is grammatical in both monoclausal and biclausal sentences (as in (a) and (c) above), whereas binding to object (as in (b) and (d) above) is ungrammatical in both types of clause. Her overall results show that high-proficiency L2 learners behave more like native speakers than low-proficiency L2 learners. As Thomas reports, however, even high-proficiency learners admit less long-distance binding of "zibun" than native speakers do. Thomas concludes that few of the L2 learners have acquired the full native speaker grammar of "zibun."

Her results show that 35 out of 58 subjects could not bind reflexives to long-distance antecedents at all. Therefore, according to Thomas, these subjects appeared not to acquire "subject-orientation" of "zibun." Of the 23 subjects who allowed long-distance binding in biclausal sentences, 16 allowed binding to subjects and disallowed binding to nonsubjects. The other 7 learners allowed long-distance nonsubject binding of "zibun" and therefore their grammars were unaccounted for by UG according to the LF movement approach (long-distance reflexives can be bound only by subject antecedents). Twelve out of 16 learners rejected both local and long-distance nonsubject antecedents. The remaining 4 learners permitted nonsubject antecedents in monoclauses but not in biclausal sentences. Eleven high-proficiency learners consistently selected long-distance subject binding and

rejected long-distance nonsubject NPs, whereas the remaining 13 high-proficiency learners did not seem to acquire the properties of "zibun."

Most of the subjects who accepted binding to nonsubjects were in the low proficiency groups. Thomas claims that subjects at lower proficiency levels may think that "zibun" is a pronoun rather than a reflexive. According to Thomas, data from lower proficiency learners do not support movement in LF, while data from higher proficiency learners somewhat support LF movement. Thomas suggests that the results appear to suggest developmental stages in the acquisition of the properties of "zibun" because the L2 learners initially allow binding to non-subjects and later move towards binding to subjects only, although few high-proficiency learners arrived at the full native speaker grammar of "zibun."

Thomas also looked at whether L1 exposure to long-distance anaphors had any effects on the high proficiency subjects' acquisition of Japanese. About one-third of the 58 learners of Japanese were native speakers of languages with long-distance reflexives (such as Japanese, Chinese and Korean). Subjects with such L1 exposure showed greater accuracy in monoclausal sentences involving binding to objects. However, the L1 exposure to long-distance anaphors was not helpful to L2 learners with long-distance binding to subjects in biclausal sentences, since most of them were not willing to accept long-distance subject binding.

2.3.3. White, Hirakawa, and Kawasaki (1995): "Second language acquisition of long-distance reflexives: effects and noneffects of input manipulation." The aim of this study

was to find out whether teaching the long-distance properties of the Japanese reflexive "zibun" leads to acquisition of its subject-oriented status. White, Hirakawa, and Kawasaki assume that L2 learners would initially treat the Japanese reflexive as having the properties of the English reflexive; however, as a result of positive evidence via teaching, the L2 learners would learn the long-distance binding and subject orientation of long-distance reflexives.

In this study there were 13 learners of Japanese from 4 L1 backgrounds (French, English, Korean and Chinese) in an intermediate level course in Montreal. Subjects were instructed about the properties of "zibun" for four weeks. The test given to subjects involved a truth-value judgment task based on pictures. Each picture had a sentence beneath it. Subjects had to decide whether or not the sentence correctly described the picture by saying 'true' or 'false.' The task consisted of 38 items; five sentence types with 4 tokens of each type. There were 18 remaining distracters.

Subjects in this study were pre-tested on their knowledge of "zibun" before instruction began, and they were given two posttests at separate times after the instruction was completed. Subjects received no instruction on reflexives between the two posttests.

The results indicated that learners correctly accepted binding to local subjects as being possible in both monoclausal and biclausal sentences, and there was no difference in results between the pre-test and post-test. The learners rejected long-distance binding to a subject in the pre-test, whereas they increased their acceptance in the first post-test. The learners accurately rejected binding to objects in both monoclausal and biclausal sentences. There was no significant difference between the three tests with regard to these sentence

types.

The analysis of individual responses indicates that at the pre-test 8 out of 12 subjects totally rejected long-distance binding, whereas only one subject totally accepted it. At the first post-test, four subjects totally rejected long-distance binding, whereas 7 subjects totally accepted it. At the second post-test (taken by only 10 subjects), four subjects totally rejected and 6 subjects totally accepted long-distance binding. Therefore, all changes took place between the pre-test and the first post-test.

White et al also examined whether any subjects showed evidence of having a UG inconsistent grammar. The following illustrates UG sanctioned binding grammars and one non-sanctioned one (E):

(22) UG-santioned binding possibilities (White et al, 1995, p.17):

A. Local subjects, local objects e.g. English "himself," Japanese "kare-zisin"
B. Local subjects, no objects e.g. Japanese "zibun-zisin"
C. LD and local subjects, no objects e.g. Japanese "zibun," Chinese "ziji"
D. LD and local subjects, local objects, e.g., Serb-Croatian "sebe," Icelandic "sig"
E. *LD and local subjects, LD and local objects

For example, if a subject accepts local subjects while disallowing long-distance (LD) subjects and objects, he or she will have built a type B grammar. In terms of these possibilities, the results showed that in the pretest the majority of subjects had a consistent type B grammar. At the post-test, several subjects indicated a type C grammar, the appropriate grammar for "zibun." Only one L2 learner showed evidence of the UG-

inconsistent type E grammar. White et al report that according to the individual results, about half of the L2 learners acquired the long-distance properties of Japanese "zibun" while half continued to accept only local subject antecedents. They conclude that although the teaching was effective in causing half of the subjects to acquire long-distance binding in Japanese, the other half could not make use of the instruction. Therefore, White et al suggest that teaching the properties of Japanese reflexives was not totally successful.

2.4. Non-syntactic approach to reflexive binding.

2.4.1. Matsumura (1994): "Japanese learners' acquisition of the locality requirement of English reflexives: evidence for retreat from overgeneralization." Matsumura's study looks at the acquisition of the locality requirement of English reflexives by Japanese learners of English, focusing on the process of acquisition over time. Matsumura's subjects were 29 Japanese learners of English in Grade 10, thirty in Grade 12, fifty-one first year university students, and 15 English native speakers as a control group. Matsumura used four types of sentence; biclausal finite, biclausal nonfinite, triclausal with two embedded tensed clauses, and triclausal with one infinitive clause) and each type was represented by four tokens in a multiple-choice task design in which the subjects were asked to mark one of the choices. An example from Matsumura (p.27) is shown in (23):

(23) Bob told Steve not to blame himself.

1. Bob
2. Steve
3. either Bob or Steve
4. someone else
5. don't know

Matsumura reports both group and individual results. She states that learners at higher proficiency levels had higher mean scores than lower levels for every type of sentence although the higher level learners still scored lower than the control group. This means that the higher proficiency levels (university students) were much better at the interpretation of English reflexives than lower levels were. According to Matsumura, L2 learners come to prefer local responses as they gain more overall proficiency in English, approximating native speaker judgments.

As to the results for individuals, out of the 7 learners who consistently provided local-only responses, 6 were from the university-aged group, and of the 5 subjects who responded almost perfectly (only one error) to the test sentences, all belonged to the university-aged group. Matsumura argues that the learners' improvement in the acquisition of the locality requirement of English reflexives is a result of increasing proficiency. She attempts to explain the progress observed in Japanese learners' acquisition of English reflexives in terms of a non-syntactic aspect of language, i.e., "viewpoint in sentence processing," which is used by native Japanese speakers to interpret the Japanese reflexive "zibun." According to the "in-sentence view perspective" process, in Japanese the reflexive "zibun" can refer to any grammatical person; third-person, second-person and first-person. Therefore, "zibun" is simply used for expressing "self" for whomever is intended in a sentence. In other words, "zibun" represents the first person as expressed by any particular person in a sentence whose perspective or viewpoint is employed. An illustration from Matsumura (p.32) is given in (24):

(24) Akiko thinks that Hiromi blames herself

In the Japanese equivalent of (24), the reflexive "herself" ("zibun") can refer to either Akiko or Hiromi, causing ambiguity in interpretation. The reflexive "herself" ("zibun") may refer to Akiko when we take the viewpoint of Akiko, because in this case the first person from the viewpoint of Akiko is Akiko herself. On the other hand, "herself" ("zibun") can refer to Hiromi when we employ Hiromi's viewpoint because this time the first person for Hiromi is Hiromi herself. According to Matsumura, most of the time, Japanese speakers prefer binding "zibun" to the subject of the main clause because it may be natural for speakers first to adopt the viewpoint of the subject NP since this NP appears first in the sentence or because it is the topic of the sentence. The viewpoint of the local NP is adopted only if it carries some kind of pragmatic consideration. Matsumura states that the relative infrequency of binding "zibun" to local NPs is a result of the burden of restructuring the first natural view.

Matsumura states that non-target-like judgments (i.e., choosing nonlocal antecedents for English reflexives), observed in earlier stages of the acquisition of reflexives by Japanese learners, may be due to the transfer of the perspective "in-sentence view" which the learners are familiar with in their L1 processing. In other words, Japanese learners use the "zibun" type processing of their L1 at the beginning stages of L2 acquisition. Matsumura explains that morphologically complex reflexives, like English "himself" and other Japanese reflexives such as "zibun-zisin," imply the processer's view (i.e., the hearer's view) rather than the in-sentence view. The notion of the processer's

view requires local binding, whereas the notion of the in-sentence view allows long-distance binding. According to Matsumura, what Japanese learners have to do to bind English reflexives locally is to employ the processer's view, because English reflexives are subject to the processer's view. Learners are able to give up the in-sentence view and adopt the processer's view in local binding of English reflexives due to improvement in proficiency over time and certain kinds of sentences which function as indirect positive evidence such as:

(25) Jane$_i$ thinks that she$_i$ is a super woman.

In the sentence above, the processer's view is employed to establish coreference between "she" and "Jane." "Jane" is perceived as the third-person and is perceived as such in relation to the processers. Matsumura claims that the constant occurrence of sentences (such as (25) above) and the nonoccurrence of sentences with the in-sentence view help the L2 learners to employ the processer's view in processing English sentences.

2.5. Summary

Most of the L2 studies have tried to address the issue of whether L2 learners have access to the principles and parameters of UG within the context of Wexler and Manzini's GCP and the SP, whereas only a few recent studies have examined this issue from the LF movement perspective. Many studies reviewed above are concerned with the ability of speakers of languages which allow the marked value (e) of the GCP to acquire the binding

setting which allows the unmarked value (a) of the GCP. These studies include Finer and Broselow (1986), Finer (1991), Thomas (1989, 1991a)), Hirakawa (1990) Lakshmanan and Teranishi (1994), Eckman (1994), and MacLaughlin (1995a). Most of these studies investigate the acquisition of English reflexives by native speakers of either Japanese, Chinese, or Korean. In addition, some of the studies are concerned with the ability of speakers of languages which allow the unmarked value (a) to acquire the binding system which allows the same value (a) of the GCP. For example, Thomas (1989, 1991a) looks at the acquisition of English reflexives by native speakers of Spanish. Some of the studies observed how speakers of languages which allow the unmarked value (a) of the GCP acquire the binding setting of languages which allow the marked value (e) of the GCP. For example, Thomas' (1991a) study examined the acquisition of Japanese reflexives by native speakers of English. Finally, some the studies observed the acquisition of Japanese (the value (e) of the GCP) by native speakers of Chinese (the value (e) of the GCP).

2.6. Generalizations from previous studies.

The previous studies examined lead to the following generalizations about how L2 learners acquire English reflexives (leaving aside those studies which looked at the acquisition of reflexives of languages other than English):

a. The results do not present conclusive evidence that the SP guides L2 learners, because the L2 learners do not initially assume the unmarked value (a) of the GCP for reflexives.

MacLaughlin (1995b) discusses four possibilities relevant to whether the SP may

be involved in L2 acquisition:

1) The SP operates just as in L1. The L2 learner selects the smallest parameter setting until there is evidence that a larger setting is necessary. According to White (1989, p.149), this is equivalent to the pure UG hypothesis and therefore it should be called "the subset hypothesis" for SLA.

2) The SP does not operate in L2 acquisition. The L2 learners cannot adopt the smallest parameter setting. Instead, the learners are influenced by their L1s. White (1989, p.149) calls this "the transfer hypothesis."

3) Although L2 learners may transfer their L1 value in the earlier stages of acquisition, they may reset the parameter and eventually reach the correct L2 grammar. This means that the SP operates and guides the L2 learner through successively nested languages with appropriate evidence. In order to test this hypothesis, according to MacLaughlin, we should observe what sort of development takes place after transfer, and we should observe whether this development confirms the predictions of the SP. MacLaughlin (1995b, p.179) call this the "transfer + subset hypothesis."

4) In cases where some parameters of UG have more than two values, L2 learners might select a value which is found neither in the L2 nor in the L1. However, it would be a value permitted by UG; i.e., possible in any natural language (White (1989, p.149)).

In terms of these possibilities, the results of the previous studies do not support the subset hypothesis (i.e., that the SP operates in L2 as in L1).

However, according to MacLaughlin (1995b, p.179), it is unwarranted to conclude that the SP is not operative in L2 because the L2 data in this area cannot distinguish

between transfer and the development after transfer (possibilities (2) and (3)). Some L2 learners show a sequence of development from a superset governing category setting to a subset setting (resetting the parameter from the value (e) to (a)), as in the studies by Hirakawa (1990), Finer (1991), Lakshmanan and Teranishi (1994), Eckman (1994), and Matsumura (1994). Finer (1991) argues that Japanese and Korean L2 learners of English initially transfer their L1 value, and therefore they start their acquisition process with value (e) of the GCP, because his subjects adopted the values of "zibun" and "caki" as the initial setting. In this respect, Finer and Broselow (1986), Finer (1991), Hirakawa (1990), and MacLaughlin (1995a) have discovered that some learners in their studies set their GCP to a middle value, like the value (c), although English reflexives have the value (a) of the GCP, and the Japanese and Korean reflexives have the value (e). This may suggest that the acquisition of the properties of English reflexives is achieved by resetting parameters gradually from the value (e) to the value (a).

Lakshmanan and Teranishi (1994) propose another approach to account for their L2 learners' progression in the direction of the native speaker grammar. They claim that Japanese learners of English associate English reflexives with Japanese monomorphemic "zibun" at early stages of acquisition. Hence, they bind English reflexives long-distance due to this morphological misassociation. However, when they recognize that "himself" is a complex morpheme, they retreat to local binding. Lakshmanan and Teranishi call the successful resetting of the GCP value from the superset to the subset grammar "interlingual identification."

From a different perspective, resetting to the correct parameter value for L2 after

transfer from L1 should not be possible if L2 learners initially incorrectly select a superset value for English reflexives, as many L2 learners did in the studies reviewed. They should not be able to reset to the subset value because it is assumed that only positive evidence guides acquisition. In such cases where learners make such generalizations based on their L1, negative evidence might be necessary to reset the larger setting of the parameter to the smaller one in such circumstances. According to White (1989, p.168), the failure to reset parameters leads to "fossilization." In the previous studies reviewed, fossilization may appear as the use of superset structures from L1 when the subset forms in L2 are actually required. This means that many L2 learners in these studies should have ended up constructing fossilized grammars for reflexive binding. However, some learners were able to move from the superset value to the subset one without any negative evidence and consistently bind reflexives locally, such as 10 learners in Hirakawa's (1990) study and 14 learners in Lakshmanan's (1994) study.

Some L2 learners exhibit possibility (4) mentioned above. They consistently reject nonlocal binding in finite clauses but allow it in nonfinite clauses. This means that L2 learners pick up a value which is found neither in their L1 nor their L2, for example, Finer and Broselow (1986), Finer (1991), and MacLaughlin (1995a).

b. The second generalization to fall out of the studies reviewed above about how L2 learners acquire English reflexives is that a majority of L2 learners assign local antecedents to reflexives in finite clauses. Still a large percentage of L2 learners allow long-distance binding in both finite and nonfinite clauses, although the data show more instances of long-distance binding out of nonfinite clauses than out of finite clauses (Finer and Broselow

(1986), Hirakawa (1990), and Finer (1991)).

c. The third generalization is that the performance of L2 learners significantly differs from both target language structure and from performance of both L1 and L2 control groups (Finer and Broselow, 1986; Hirakawa, 1990; Finer, 1991; Matsumura, 1994).

d. Learners of different L1s do not necessarily differ in their interpretations of reflexives. For example, both Spanish and Chinese learners of English in Thomas' (1989, 1991a) studies performed the same with regard to reflexive binding in English, despite the different binding settings of these two languages.

e. The majority of results appear not to support transfer. In Thomas' 1989 and 1991a studies, Spanish learners of English allowed long-distance binding although Spanish has local binding. In Thomas (1991a), Chinese learners of Japanese allowed local binding although Chinese has non-local binding. In Finer and Broselow (1986) and Finer (1991), Japanese and Korean learners of English made a distinction between tensed and infinitival clauses although none of these languages make the tensed/infinitival distinction for reflexive binding.

f. Individual results show that some L2 learners are able to acquire the right local binding for English reflexives. For example, Hirakawa (1990) states that 10 of 65 Japanese learners performed 100 percent correctly. Eckman (1994) reports that 23 of 25 learners exhibited local binding. Lakshmanan and Terinishi (1994) found that 14 of 24 Japanese learners performed 100 percent correctly. However, many subjects exhibited nonlocal binding as found in their L1. In Hirakawa's study, 55 learners allowed long-distance binding in English. Other studies also report varying incidents of nonlocal binding

in individuals' interlanguage grammars.

g. UG appears to be available to L2 learners, and parameters can be reset in SLA (Finer and Broselow, 1986; Hirakawa, 1990; Thomas, 1989, 1991a, 1993; Teranishi and Lakshmanan, 1991; Finer, 1991; Matsumura, 1994).

2.7. Some remarks on previous studies.

The studies by Finer and Broselow (1986), Thomas (1989), Hirakawa (1990), Teranishi and Lakshmanan (1991), Finer (1991), Matsumura (1994), and Eckman (1994) show that the SP does not necessarily exist in the early stages of acquisition in L2. These studies reveal that the operation of the SP in L2 is dependent upon the developmental stages of acquisition and upon the amount and type of input given to the learner. For example, Matsumura's data suggest that within the grammar being acquired (English in this case) there will be some structures which will serve as indirect positive evidence to help the learners to select the correct setting. Consider the following example from Matsumura (p.36):

(26) Jane$_i$ thinks that she$_i$ is a super woman.

In this example, "Jane" is referred to by the third-person form "she." "Jane" stands as the third-person and is perceived as such in relation to the processer (the speaker or hearer). Matsumura says that sentences like this function as indirect positive evidence for the acquisition of the locality requirement of English reflexives even though reflexives do not

appear in them, and they help learners to adopt the processor's view in English sentence processing, which might be the key to local binding of reflexives.

The grammar offers some indirect positive evidence to the learner in later stages of acquisition, which makes the learner reanalyze his or her initial hypothesis. In this way, the learners will be able to move from a larger grammar to a narrower one although both the SP and LF movement do not accept such a transition because the SP assumes that the learner will always start with the smallest setting compatible with the data while LF movement assumes that the clusters of the properties of reflexives will be instantiated at the same time; for example, long-distance reflexives are morphologically complex and subject-oriented. The previous studies indicate that learners need time and enough positive evidence to reset the parameter to the correct value (where positive evidence involves naturally occuring language use in the learner's environment).

Finer and Broselow (1986) claim that L2 learners did not obey the SP nor did they transfer the value from Korean. Instead, their L2 learners initially selected an intermediate value of the GCP based on the fact that they treated finite and nonfinite clauses differently. Finer and Broselow's results might have been interpreted in a different way — their subjects could not select the correct setting for English reflexives in nonfinite clauses not because they selected an intermediate value of the GCP, but rather because they had difficulty in parsing nonfinite clauses, and therefore they had difficulty in interpreting the reflexives inside nonfinite clauses. As White (1989) says, when testing the interpretation of reflexives in complex sentences, one should be sure that subjects can handle the structures in which the reflexives investigated operate. Thus, if subjects have a lack of

sufficient knowledge of the nonfinite clauses in L2, they might violate the binding conditions. At the point where English infinitives are acquired, the learners will probably realize that binding of reflexives outside of infinitival clauses is impossible. Although other later studies were able to replicate Finer and Broselow's results (Hirakawa, 1990; Matsumura, 1994; Eckman, 1994), none of these employed the same explanation as Finer and Broselow with regard to the finite/nonfinite clause distinction. However, MacLaughlin (1995a) arrives at Finer and Broselow's explanation by reanalyzing Hirakawa's data for individual results. MacLaughlin reports that some of the learners did in fact acquire a system of reflexive binding that is different from the one found in their native language and the one in their second language.

Although Finer and Broselow used a very small number of subjects and a limited variety of sentence types compared to Hirakawa's large number of subjects and rich variety of syntactic structures, both of the studies obtained similar results; the finite/nonfinite distinction effects the results more than levels of embedding. It should be noted that compared with Finer and Broselow (1986), Finer (1991) finds a much smaller incidence of long-distance binding of reflexives out of nonfinite clauses (7% nonlocal NP for Koreans and 12% nonlocal NP for Japanese), yet Finer still claims that the L2 learners were making a distinction between nonfinite and finite clause types and picked an intermediate value. However, in order to claim that subjects pick an intermediate value of the GCP, one must test more complex structures, such as; [*Mary thinks that* [*Jane wanted Anne* [PRO_i *to love herself*]]]. If the L2 learners make a distinction between finite and non-finite clauses in this type of sentence, they should not select "Mary" as the antecedent of "herself" since

they are assumed to pick up an intermediate value. However, if they select the non-local NP "Mary" as the antecedent, then they are choosing not an intermediate value, but the largest value of the GCP as in their L1. As a result, by only using biclausal nonfinite sentences it would be merely speculation to claim that learners acquire a binding system different from their L1 and their L2.

Thomas (1989) rejects both the SP and the transfer hypothesis, whereas her 1991 data show evidence for the resetting of the GCP by L2 learners but continues to reject the transfer hypothesis. In her 1989 study, Thomas did not use different proficiency levels, although to use different levels is necessary in order to observe developmental stages and the initial state in the acquisition of reflexive binding. Another weakness of the study was that the subjects were coming from different proficiency levels (although they were not distinguished as such), and there was too much variation in terms of age and level of education. For example, the subjects' ages ranged from 16 to 48, and some were high school students, some were college students, some were graduate students and some were adult education students. Moreover, Thomas' study did not use either a Chinese or a Spanish control group, and so it was impossible to determine whether her L2 learners were behaving differently from an L1 control. Although Thomas (1991a) used different proficiency levels in her study, she did not explain her results in terms of these levels at all.

Teranishi and Lakshmanan (1991) found that their subjects progressed towards the native speaker grammar. Their subjects increased local interpretation of English reflexives and decreased nonlocal interpretations over time. Teranishi and Lakshmanan explained this progress as L1 transfer in the form of interlingual identification which caused the learners

to perceive English reflexives as simple forms. This type of explanation is in fact closer to a nonparametric approach such as LF movement than to a parameterized binding approach such as the GCP. It should be remembered that according to LF movement, the syntactic properties of anaphors follow from their lexical and morphological structure. According to Teranishi and Lakshmanan's explanation, at early stages of acquisition Japanese speaking learners of English bind English reflexives nonlocally due to morphological misclassification because they think that English reflexives are morphologically simple like Japanese reflexives. The learners can arrive at the correct L2 grammar when they make an accurate morphological analysis of English reflexives.

In order for Teranishi and Lakshmanan to make such a claim, namely "interlingual identification," they must test this by using a different methodological framework with an appropriate number of sentences and sentence types. Otherwise their claim can only be considered speculative. Their explanation has several problems. The first problem is that according to their explanation, L2 learners cannot reset the parameter to the L2 setting because L2 learners are able to transfer only the rules of their L1 reflexive binding system to their L2. Moreover, Japanese has two types of reflexives which can be transferred. Hence, the question is, what factors might influence the choice of a transfer source? Learners may differ as to which source they pick for transfer; some transfer the local reflexive while others transfer the long-distance reflexive. Their explanation does not account for why learners at first transfer the simple form into their interlanguage instead of complex forms.

Teranishi and Lakshmanan's study consists of only one type of sentence, and there

were ten test sentences representing this one type, which compromises the validity and reliability of the study. Subjects might have guessed the answer after several sentences since all the sentences were structurally similar. Another methodological problem with this study is that the same experimental group was also used as an L1 control group. Since the same test sentences were used for both the experimental and control groups, this may have biased the subjects.

In sum, previous research led to the following main generalizations about reflexive binding in L2; first, L2 learners appear to transfer their L1 value in the earlier stages of acquisition, but are able to reset the parameter and eventually reach the correct L2 grammar, thus exhibiting a sequence of development from a superset to a subset setting. Second, L2 learners seem to be affected by the finite/nonfinite distinction and are more successful in selecting the correct antecedent in finite clauses than in nonfinite ones.

Chapter 3

METHODOLOGY

3.1. Introduction.

This chapter deals with the issues of L2 learners' grammaticality judgments and their use in the investigation of interlanguage competence. First, this chapter defines and discusses the advantages and problems of grammaticality judgments (a term which includes "acceptability" judgments). Second, this chapter deals with the various techniques for eliciting linguistic judgments and their corresponding problems, as well as with the experimental designs used in studies on L2 acquisition of reflexive binding.

3.2. Grammaticality judgments and their role in SLA research.

Within Government-Binding theory, it is assumed that a linguistic theory must be composed of the underlying linguistic knowledge of native speakers, and hence the validity of linguistic theories must be tested through metalinguistic judgments which act as indicators of speakers' knowledge of the grammatical structures:

> To be sure, the judgments of native speakers will always provide relevant evidence for the study of language, just as perceptual judgments will always provide relevant evidence for the study of human vision.... If a theory of language failed to account for these judgments, it would plainly be a failure. We might, in fact, conclude that it is not a theory of language but rather of something else. (Chomsky, 1986a, p.37)

Throughout the last thirty years, linguists have tested their syntactic theories by eliciting judgments or by using their intuitions about syntactic structures. After the 1980's, the main goal of SLA researchers has been to discover mental representations and processes that make acquisition possible. In order to study the mental structures and to test a specific hypothesis derived from a theory of UG, L2 researchers have borrowed experiment methods (such as grammaticality judgments) as well as linguistic theories from studies in first language acquisition. Gass (1983) points out that researchers have argued persuasively that learner-languages are subject to the constraints on natural languages, and therefore L2 learner-languages should be studied through linguistic intuitions:

> If we assume similarity to natural languages, we would further suppose that they could be investigated through the same methods as other types of natural languages, for which a chief methodological device is the use of intuitions of native speakers. (p.273)

Grammaticality judgments (a term often used interchangeably with 'intuitions'), as a part of metalinguistic ability, require L2 learners to decide which sentences are possible and which are not in the grammar of a particular language. It is impossible for L2 researchers to access the learners' interlanguage grammars directly. Therefore, the only way to do this is through learners' intuitions of grammaticality about the well-formedness of structures, because researchers assume that a sentence which is judged to be grammatical agrees with the learner's interlanguage grammar and reflects the development of interlanguage knowledge.

There is a variety of ways in which grammaticality judgments are employed in SLA

studies. Some researchers require learners to judge the well-formedness of certain sentences which might be grammatical and ungrammatical versions of a particular structure of interest to the researcher. Some researchers ask learners to correct sentences which are judged as incorrect by the learners. Other researchers ask learners to provide a grammatical description of the errors. Sentences used in these tasks are either designed by the researchers or actually produced by the learners.

3.3. Advantages of grammaticality judgments.

There are several advantages of grammaticality judgment tasks:

a. Some phenomena are not accessible to investigation in production data, because they occur either rarely or not at all. For example, without directly asking learners about their judgments on reflexive binding, there would be little hope of finding out what learners actually know, since in their speech reflexives are relatively difficult to capture, and even when they do occur, there is usually no overt indication of the coreference between the reflexive and its antecedent.

b. Learners avoid using structures that they find difficult, and therefore only grammaticality judgment tasks will reveal those structures which are problematic for learners.

c. Grammaticality judgment tasks allow researchers to test the sentences which violate universal constraints. Therefore, the subjects are required to consider whether a sentence which is impossible according to UG is also impossible in their interlanguage grammar (White, 1989). For example, in a sentence like "Mary's mother likes herself," if L2 learners bind the reflexive to "Mary" in a coreference judgment task, then they are violating

a universal constraint on reflexive binding (c-command requirement), and we are forced to accept that their interlanguage allows structures which should be impossible according to UG.

d. Grammaticality judgments serve as indicators of transfer which occurs at the level of abstract knowledge. Although it is easier to observe transfer at the performance level, without using grammaticality judgments it is harder to observe the effects of L1 on L2 structures which are related to the abstract principles and parameters of UG. For example, coreference judgment tasks used in the present study made possible the observation that some of the Turkish L2 learners used the Turkish binding setting for English reflexives; there was no way to make this observation at the performance level.

3.4. The problems with grammatical judgments.

Even though grammaticality judgments have been used intensively in L2 studies carried out within the UG framework, there are a number of problems which arise, mainly in three areas; validity, reliability, and indeterminacy of grammaticality judgments.

3.4.1. Validity. The issue of validity refers to the question of whether grammaticality judgments really measure the underlying knowledge of L2 learners, and whether they are free from performance constraints. With linguistic judgments, researchers aim to test grammatical competence. However, according to Gass (1994b), grammaticality judgments are not a direct reflection of competence because competence is an abstract entity. Such judgments only provide another sort of performance data which allow us to make

inferences about competence.

In addition, since our view of competence is dependent on performance data, extragrammatical factors may interfere with our conclusions:

i. Learners may be faced with the problem of parsability constraints. Learners may reject grammatical sentences which are complex and difficult to analyze. For example, in the present study learners may not be able to choose the right antecedent for a reflexive in triclausal sentences not because of a lack of knowledge of reflexive binding, but rather because of difficulties in analyzing multiply embedded constructions.

ii. Learners bring several test performance strategies to carry out the task. For example, they may guess if they are not sure, or they may try to establish a balance between the number of grammatical and ungrammatical sentences. Also they may lose patience if the test is too long and thus respond carelessly or not at all.

iii. The learners' judgments might be based on cognitive and pragmatic factors other than their competence. For example, when learners find test sentences semantically ambiguous, they tend to prefer the reading that represents the most common or pragmatically likely interpretation.

iv. The order of sentences also effects the learners' judgments. For example, a sentence might be judged as ungrammatical if it is placed after a set of clearly grammatical sentences. This indicates that judgments of sentences in isolation are very different from judgments of sentences among other sentences.

3.4.2. Reliability. The reliability problem refers to the issue of both intersubject consistency, i.e., agreement among judgments produced by different subjects with the same background (the same age, proficiency, years of instruction, L1, etc.), as well as intrasubject consistency, i.e., agreement among judgments produced by the same subject at different times. Learners who have identical backgrounds may reveal considerable variability in their grammaticality judgments, and a learner may judge the same structure as grammatical in one instance and deviant in another. Ellis (1991) states that:

> If it is accepted that a grammaticality task is a performance task, inter- and intralearner variation is neither surprising nor problematic. If it is claimed that a grammaticality judgment task reveals competence, and if the theory of competence is a probabilistic one, then it is necessary to demonstrate that the variability is systematic in some way. If the theory of competence is deterministic rather than probabilistic, then any variability in grammaticality judgments tasks is problematic; it can not be missed as the result of extraneous performance factors without challenging the whole theoretical rationale of metalingual judgments. (p.165)

According to Cowan and Hataca (1994), although it is not very easy to control for all factors which introduce variability, the reliability of studies can be increased by employing several standard procedures. For example, there should be at least three tokens for each sentence type, and there should be an equal number of distracter items. Because longer tests will increase reliability, there should be a total of sixty to seventy sentences in a test. Furthermore, by using large numbers of subjects it is possible to reduce variance within the same groups. Proficiency levels should be formed in a way which minimizes difference in background, which in return should minimize variability in judgments. Moreover, the use of a native speaking control group will show how flexible native

speakers are in their judgments, and this will help us in discovering the sources of variation of L2 learners.

3.4.3. Indeterminacy. Indeterminacy refers to the learners' incomplete knowledge or absence of knowledge of parts of the second language grammar (Gass, 1994b). In other words, indeterminacy means the indefiniteness of interlanguage rules which leads to the learner's inability to make a clear cut judgment of grammaticality (Sorace, 1988). According to Gass (1994b), it is clear that indeterminacy exists in interlanguage grammar:

> It is conceivable that it embraces an even greater range of data than for native speakers of a language. Because there is a wide range of data that are potentially indeterminate, it is particularly important that we have a principled basis for determining what is truly representative of a learner's knowledge and what is not. How can we know which data are good data, and which represent spurious exemplars of grammatical structures? (p. 305)

Although L2 researchers aim to have grammatical descriptions based on actual knowledge and not on guesswork, according to Gass (1994b) this does not appear to be an easy task to achieve:

> With grammatical judgments what we are asking learners to do is evaluate sentences of a language that they do not have total control over; many of the sentences being asked about are beyond the domain of their current knowledge. Thus, responses to these represent little more than guesses. What we want to know is which sentences actually represent those sentences that are part of a learner's grammatical knowledge and which ones do not. (p.306)

> The elicitation of valid and reliable data is a very important as well a complex task in

L2 studies (Sorace, 1988). The issues of validity and indeterminacy are correlated in interlanguage judgments because it is necessary to capture indeterminacy in order to obtain valid judgments. According to Sorace (1988), one of the factors which leads to indeterminacy is the openness of interlanguage grammars to other linguistic systems. This openness creates the conditions for the coexistence of more than one rule, either from different linguistic systems or from successive stages of interlanguage development, for the same aspect of grammar, which result in variability or indecisiveness in learners' judgments.

The second factor which, according to Sorace, contributes to interlanguage indeterminacy is the fact that SLA is shaped by two functionally different cognitive systems; UG, the innate language faculty, and a general problem solving faculty used in conceptual learning tasks. When the problem solving faculty is involved in acquisition, and interacts with or accompanies UG, interlanguages will be determined to a certain extent by the problem solving faculty; the result may be indeterminacy and partially account for learners' inability to produce clear grammatical judgments.

The third factor which generates indeterminacy in interlanguage grammars is the incompatibility of parameter settings of L2 learners where L1 and L2 have different settings for a given structure. In this case, the L1 setting may be transferred into the L2 setting.

3.4.3.1. *Types of indeterminacy* . Sorace (1988) claims that subsequent stages of interlanguage development are characterized by different kinds of indeterminacy.

i. The first sort, called "early indeterminacy," may be observed at initial stages of

acquisition because the interlanguage grammar lacks certain L2 syntactic rules, and hence is incomplete. For example, there is no need to test a principle like subjacency in a learner who has mastered only simple sentences, because subjacency is involved in the construction of complex sentences, but a test should be effective as soon as complex sentences and question formation are learned.

ii. "Intermediate indeterminacy" refers to a more advanced stage where indeterminacy is due to the learners' restructuring or reanalysis of knowledge within the interlanguage grammar. Sorace compares this restructuring of an L2 rule to U-shaped developmental patterns in language acquisition in which a rule becomes threatened after a period of stability, causing a process of restructuring that leads to temporary loss of determinacy of a structure.

iii. "Advanced indeterminacy" may be discovered in the very advanced stages of L2 acquisition in which there exists differences between the linguistic intuitions expressed by L2 learners and by native speakers. This type of indeterminacy might result from the failure to use available evidence and the difficulty of parameter resetting to the L2 value. Even though it is expected that at very advanced stages of acquisition nonnative speakers will have similar acceptability judgments to native speakers with respect to a given area of the L2 grammar, this is not necessarily the case because certain areas of L2 grammar may never become determinate in interlanguages. For example, Coppieters (1986) carried out research to find out competence differences between native speakers and near-native speakers of French in several areas of grammar; such as, contrast between two past tenses, between the two third person pronouns, and the use of adjectives before and after nouns.

The two groups studied by Coppieters seemed to be completely equivalent at the level of language use, and it was not possible to distinguish the near-native speakers (who had resided in France an average of 17 years) from native speakers. The results show that the interpretations of the near-native speakers of the grammatical forms tested in the study were quite different from the interpretations by native speakers. According to Coppieters, the non-native speakers in this study often could not produce clear judgments for some of the grammatical rules tested, and their preference for either one or the other form was unsystematic even though they were nearly indistinguishable in production from native speakers.

Again, following Sorace's (1988) discussion, let's consider the following case in order to consider how validity, reliability and indeterminacy may create problems for any grammaticality judgment test. Imagine a case in which learners are required to judge as "correct" versus "incorrect" a linguistic construction that is indeterminate in their interlanguage grammar. Any sentence which exemplifies that construction will be judged as either correct or incorrect by the learners although they do not have the knowledge of that construction in the interlanguage grammar. Therefore the learners' choices are completely random. If there are tokens of the same test sentence, the intuitions may be either consistent or inconsistent. If they are consistent, the researcher will think that the construction is acquired and stable. If they are inconsistent, then the researcher will not be able to decide whether this inconsistency is due to the random choice because the learners have not acquired the construction in question yet (in the earlier stages of acquisition), or whether it is due to the intermediate or advanced indeterminacy generated by reanalysis of

knowledge or unsuccessful parameter resetting (in the more advanced stages).

According to Sorace (1988), in order to improve the validity and reliability of grammaticality judgments, researchers should avoid requiring absolute judgments (either correct or incorrect) and three point rating scales (with a third category "not sure" or "don't know"). Preference tests are more reliable because they allow researchers to rule out random choice as a possible source of inconsistency and to assume that judgments are derived from actual states of interlanguage knowledge. Furthermore, the researcher should use a timed test in order to get immediate responses and to determine the difficulty of each sentence type where longer response time indicates greater difficulty (Sorace, 1988; Ellis, 1991; Gass, 1994b). Each sentence type in the task should be represented by at least two tokens, sentences should be randomized, and the subjects should be given different randomizations (Gass, 1994b). The task should control for sentence length as well as the syntactic, semantic and lexical complexity of the sentences. Subjects should be given clear instructions about what they are supposed to do in responding to the sentences. Subjects should have a short practice session by using different sentences than the ones in the task.

3.5. Judgment tasks and L2 studies of binding theory.

3.5.1. Introduction. Research in the area of reflexive binding in L2 has employed several varieties of comprehension or sentence interpretation tasks which ask subjects to interpret the meaning of test materials and are concerned with the competition of coreference.

3.5.2. Varieties of comprehension tasks.

3.5.2.1. *Picture identification task*. The first type is "picture identification" in which the subjects select one or more pictures out of a set of pictures which show each of the possible antecedents for a reflexive within a stimulus sentence. For example, Finer and Broselow (1986) gave the sentence "Mr. Fat asks Mr. Thin to paint himself" to L2 learners and presented them with two drawings, one of the local antecedent ("Mr. Thin"), and one of the nonlocal antecedent ("Mr. Fat"). The subjects had to choose the picture that the reflexive best described. Finer (1991) and Eckman (1994) also used this task.

3.5.2.2. *Multiple-choice task*. The second kind of comprehension task is the multiple-choice task, which has been the most common means of collecting data for reflexive binding in L2 (Thomas, 1989; Hirakawa, 1990; Thomas, 1991a; Thomas, 1993; Matsumura, 1994). This task asks subjects to indicate the antecedent of the reflexive in each sentence by circling one of the three or more multiple-choice answers. An example of the multiple-choice task is given below (Matsumura, 1994, p.28):

(27) Bob told Steve not to blame *himself*

1) Bob
2) Steve
3) either Bob or Steve
4) someone else
5) don't know

3.5.2.3. *Truth-value judgment task*. The third kind of comprehension task is the "truth-value judgment task," in which each potential interpretation of the reflexive is matched with

one picture or story. The subjects are not asked to choose among several possible pictures, but rather, for each picture or story they are required to give a "yes" or "no" response to a sentence matched with that picture or story. In this way, a particular interpretation is aimed to be forced on the subjects to prevent the problem of preference (Christie and Lantolf, 1993; Thomas, 1995; White et al, 1995). For example (White et al., 1995, p.4):

(28) Susan wanted a job in a hospital. A nurse interviewed Susan for the job. The nurse asked Susan about her experience, her education and whether she got on well with people.

The nurse asked Susan about herself. T F

According to this story, the statement ("the nurse asked Susan about herself") will be true (T).

3.5.2.4. *Acceptability interpretation task*. I call a fourth kind of comprehension task an "acceptability interpretation task," in which each sentence with a reflexive is followed by two or more statements concerning the interpretation of the reflexives in the sentence. Subjects are required to consider each statement separately and indicate whether they agree or disagree with the statement by circling the relevant option. The following example illustrates this type (Lakshmanan and Teranishi, 1994, p.195):

(29) John said that Bill saw himself in the mirror.

a) "Himself" cannot be John agree disagree
b) "Himself" cannot be Bill agree disagree

In a modified version of the same type of task, each test sentence is followed by a series of yes-no questions which question the subjects' interpretation of each possible antecedent of a reflexive (Berent (1994, p.21):

(30) Jack learned that Dan voted for himself.

a. Can "himself" = Jack	YES	NO
b. Can "himself" = Don	YES	NO
c. Can "himself" = another person	YES	NO

3.6. The problem of preference.

Although comprehension tasks are an obvious choice in the study of the interpretation of reflexives in L2 (Thomas, 1993), they have been faced with problems concerning the elicitation of judgments and with problems concerning task designs. Berent (1994) claims that the study of the acquisition of binding conditions is inherently difficult because:

> The study of binding theory and associated parameters is by nature more complicated than, say, the study of word order and associated parameters. Word order is a more salient feature of language than is the invisible association of two or nominal expressions through coreference. Given the formidable task of developing reliable measures for accessing coreference in emerging grammars, the results of many child language studies that address binding theory have been held suspect. (p.30)

The main problem which studies of binding principles in L2 has been faced with is related to the problem of validity — do these tasks reveal the L2 learners' preferred interpretations, or do they reveal their actual grammatical knowledge? This problem of

preference, a validity problem, of one interpretation over another may occur in ambiguous sentences in which there is more than one potential antecedent for the reflexives, as in (31):

(31) Mary showed Jane a picture of herself

Sentence (31) is ambiguous because there are two possible antecedents for the English reflexive "herself," a subject NP ("Mary") or a nonsubject NP ("Jane"). The results of previous studies indicate that native English speakers as well as L2 learners do not select both the subject NP and the nonsubject NP at the same time, but rather they prefer one over the other, even though both are theoretically possible antecedents for the reflexive. In other words, in such ambiguous sentences native speakers as well as L2 learners often prefer the subject NP as the antecedent and fail to recognize that the nonsubject NP can also be the antecedent of the reflexive. The problem then is whether this rejection of the nonsubject NP reflects a preference for the subject NPs, or the impossibility of nonsubject NPs as antecedents for English reflexives in monoclausal sentences. Consider Table 1 in order to see the preference for the subject NP over the object NP in English monoclausal sentences:

Table 1: Interpretation of *himself* in monoclausal sentences by native speakers of English, Japanese and Chinese

Study	L1	Subject	Object	Either
Hirakawa (1990)	Japanese	73.90%	20.30%	5.50%
Thomas (1989)	Chinese	59.70%	11.80%	28.60%
Thomas (1989)	English	72.82%	2.45%	24.73%

Now consider the following example:

(32) Bill thinks that John blames himself.

The same problem of preference is true of biclausal sentences as in (32) in which the ambiguity arises for native speakers of languages which allow long-distance binding, such as Japanese, Korean, Turkish, etc. In sentences like (32), both local and nonlocal subject NPs (e.g., "Bill" and "John") can serve as possible antecedents. Native speakers of these languages and even L2 learners fail to indicate both subject NPs as possible antecedents, and prefer just one of them. For example, Hirakawa (1990), Finer (1991), Thomas (1991a), and Eckman (1994) report that Japanese and Korean control groups in their studies showed a strong preference for nonlocal antecedents in their interpretation of Japanese "zibun" and Korean "caki," and failed to select the local NPs as antecedents. Table 2 illustrates the problem of preference in that native speakers of Japanese in two studies preferred the nonlocal NP as the antecedent of "zibun" (Lakshmanan and Teranishi, 1994, p.193):

Table 2: Interpretation of *Zibun* in Biclausal Tensed Sentences by Native Speakers of Japanese

Study	n	Nonlocal	Local	Either
Hirakawa (1990)	22	62.73%	26.36%	9.10%
Thomas (1990)	10	67.5%	2.50%	30.00%

According to Thomas (1991b), the preference for the nonlocal antecedent only by native speakers of Japanese, etc. violates the parameter setting of Japanese reflexives for the GCP, because "zibun" and other long-distance reflexives instantiate the superset value (e), which includes the subset values. A grammar which allows reflexives to have nonlocal antecedents while disallowing local antecedents is excluded by Wexler and Manzini's parameterized binding principles. Therefore, we would expect close to 100% or least a majority of native speakers of Japanese to allow either the local or nonlocal NP as antecedents. However, the previous studies do not support this expectation. The problem of preference is summarized by MacLaughlin (1995a):

> If subjects are really only reporting preferential interpretations, how do we interpret the data? Can we assume that a subject who only allows a reflexive to be locally bound actually has a grammar that disallows long-distance binding, or does the subject have a really strong preference for local binding? In other words, can a preference against long-distance binding be interpreted as a grammatical constraint that is part of the learner's interlanguage? (p.4)

3.6.1. Criticism of methodologies which cause the preference problem. Several researchers (Thomas, 1991b; Lakshmanan and Teranishi, 1994; White et al., 1995) argue that L2 coreference judgments reflect subjects' preferences rather than actual grammars. Lakshmanan and Teranishi (1994) criticize the multiple-choice format used in testing the L2 learners' interpretation of reflexives, because this format leads the learners to select one interpretation and to fail to see the other possible interpretations:

> The task appears to be methodologically flawed in that the subjects were required to identify who the reflexive could refer to, but not who the reflexive could not refer to. So it is possible that the subjects did not consider all possible options before arriving at their interpretations. In other words, a multiple-choice task may be tapping only subjects' preferences when interpreting reflexives and not their syntactic judgments. (p.192)

Lakshmanan and Teranishi (1994) also object to the picture identification task used by Finer and Broselow (1986), Finer (1991), and Eckman (1994), because when subjects are asked to identify the antecedents of a reflexive by identifying which of a series of pictures matches a particular sentence, they usually indicate only one antecedent even though others are theoretically possible:

> Subjects can be successful on this task by selecting the picture of what can be termed a reflexive action. In other words, simply by knowing that the sentence names the action of self-painting, for example, the subjects would be able to pick the right picture every time to match the sentence provided. Thus, their choice of correct picture may have been the result of adopting an ungrammatical strategy and may not have been based on their grammatical knowledge at all. (p.191)

White et al. (1995) also questioned whether the subjects' preference reflects a property of the grammar, or whether it relates to tasks such as multiple choice and picture identification tasks:

> It appears that such tasks may only elicit preferences on the part of the subject, preferences which may be strong enough to override the subject's perception of possible ambiguities. In other words, the fact that one interpretation is chosen doesn't necessarily mean that the other is excluded from the grammar. If so, it is a performance phenomena that we are dealing with and the question that arises is whether alternative methodologies can get around the problem.

3.6.2. Improving the experimental design. For Lakshmanan and Teranishi (1994) the preference for the subject NP over the nonsubject NP, and the preference for long-distance binding over local binding is indeed a performance phenomenon. According to Lakshmanan and Teranishi, the alternative methodology which can get around the problem is "a sentence interpretation test" in which each sentence was followed by two statements concerning the interpretation of the reflexive. Consider the example given by Lakshmanan and Teranishi (1994, p. 195):

(33) John said that Bill saw himself in the mirror

1. "Himself" cannot be John	agree	disagree
2. "Himself" cannot be Bill	agree	disagree

The subjects used in their study were instructed to think about each statement separately and then to circle the relevant option in each statement. The aim of the task was to require learners to think about what the reflexive cannot refer to as well as what it can refer to, in order to tap their syntactic judgments. Their results show that 57.9% of responses by the Japanese control group selected the "either" choice, while 20.6% of the responses selected the nonlocal NP, and 21.2% of responses selected the local NP as antecedent for the Japanese reflexive. Lakshmanan and Teranishi claim that their method is successful because a majority of responses selected the "either" choice; "Our results suggest that the test used in our study may have been more effective than the instrument used in previous studies in tapping the subjects' syntactic knowledge" (p.197). Although Lakshmanan and Teranishi are satisfied with their results and methods, it should be remembered that still

41.8% of responses by the Japanese control group chose only either the local NP or the nonlocal NP.

In their study of the acquisition of Turkish and Japanese reflexives by native speakers of English (1993), Demirci, Glew and Yamagata used the test format proposed by Teranishi and Lakshmanan in which subjects were asked to consider each NP and indicate whether or not it was impossible for the reflexive to refer to that NP. 42.5% of the responses of the Turkish control group selected the "either" choice in their L1, whereas 50% of the responses of the same Turkish control group selected only the local NP as the antecedent, excluding NP1 as a possible antecedent for the Turkish reflexive even though both NPs are syntactically possible antecedents for the Turkish reflexive. Although the study used the same interpretation task proposed by Teranishi and Lakshmanan and followed the same steps they employed, the study could not obtain the expected results from native speakers of Turkish in their L1 because less than half of the native speakers of Turkish allowed the "either" choice. This indicates that the methodology for these learners was not able to overcome the preference problem.

According to Berent (1994), the study of Lakshmanan and Teranishi carries a task design problem. This type of design in which a negative statement uses "cannot" followed by positive and negative agreement options places a heavy cognitive burden on subjects. Consider again the previous example (33). Note that agreeing requires a negative response to a negative statement. Therefore, it is not surprising, for Berent, to see that the subjects at the lower proficiency levels gave the highest percentage of incorrect responses. Berent's criticism is supported by the study by Demirci et al. in which the Turkish control group

complained about the burden of processing the negative statements in the test, even in their L1.

For White et al. (1995), the alternative methodology which will solve the problem of ambiguity is "truth-value judgment task":

> In order to overcome this problem, a number of acquisition researchers have turned to truth-value judgment tasks to bring certain interpretations to the attention of learners, thus enabling us to find out whether the learner accepts or rejects the interpretation being investigated. An additional advantage of those tasks is that the subject indirectly supplies a grammatically judgment without having to focus on the syntactic form of the sentence; indeed, the subject is not aware of making a judgment at all. (p.9)

Christie and Lantolf (1993) also use a "truth-value judgment" task in their study. They claim that in order to obtain more valid data about the underlying nature of interlanguage grammars, it was decided not to ask subjects what their interpretation of a given sentence was, but to present them with a certain interpretation and ask them to decide whether or not this given interpretation constituted an acceptable one.

Although the "truth-value judgment task" is designed to present learners with multiple interpretations and is used (with great expectations) by Christie and Lantolf (1993), White, Hirakawa and Kawasaki (1995), White et al (1995), and Thomas (1995), learners still appear not to be able to recognize alternative interpretations. Both Thomas (1995) and White, Hirakawa and Kawasaki (1995) report that the majority of Japanese native speakers in a picture truth value task accepted only nonlocal subjects as antecedents, even though in both studies a context for the local antecedent was supplied. White et al (1995) conclude that truth value judgment tasks could not overcome the problem of

preference in the case of potentially ambiguous sentences.

3.6.2.1. *Comparing two types of truth-value judgment tasks* . White, Bruhn-Garavito, Kawasaki, Pater, and Prevost (1995) attempt to explore the use of truth-value judgment tasks, comparing two different tasks of this type; one a story task and one a picture task. Subjects were Japanese speakers learning English as a second language and Quebecois speakers of English in Canada. There were four groups of subjects and two control groups depending on the type of test they took. For example, one group of Japanese speakers did the story task and the other Japanese group did the picture task. The study aimed to find out whether the preference for one NP over the other reflects a strong preference for this NP as an antecedent or the impossibility of the other NP as an antecedent. The test had three sentence types; monoclausal, biclausal finite, and biclausal nonfinite. With respect to the question of potential ambiguity, in English monoclausal sentences the antecedent of the reflexive can be either a subject or a nonsubject NP. In biclausal sentences in Japanese, either long-distance or local subjects can serve as the antecedents. In the truth-value judgment tasks they used, each potential interpretation of the reflexive is matched with one picture or story. Subjects were asked to accept or reject only one possibility. In the story task, which included 24 stories, each story was followed by a comment sentence and the subjects had to indicate whether the sentence was true or false according to the situation given in the story. For example, the possibility of long-distance binding is examined using stories like the following in which the context of the story allows learners to choose the nonlocal NP:

(34) Johnny and a little boy were playing with matches. Johnny lit a match and then dropped it on the little boy's leg. The little boy went screaming to his father and told him what had happened.

The little boy said Johnny burned himself. T F

In the picture task the context for the binding judgment is provided by a picture rather than a story. The subject looked at a picture with a sentence underneath it and indicated whether the sentence matched the picture by circling true or false. For local and long-distance binding, a biclausal sentence occurred twice in the test, once for local binding and once for nonlocal binding.

Their results show that there is a difference between the story and picture tasks in the case of monoclausal sentences involving binding to objects. Consistent acceptance of object antecedents are higher in the story task than the picture task. However, there were no differences across the two tasks for the other sentence types except for biclausal sentences with binding to a local antecedent where Japanese speaking subjects were more accurate in the picture task:

Table 3: Accuracy Scores based on mean scores (out of 4 total) (White et al)

	Mono-S (T)	Mono-O (T)	Bi-F-loc (T)	Bi-NF-LD (T)	Bi-F-LD (F)	Bi-NF-LD (F)	Total
J-stories	3.32	2.58	2.95	3.26	2.74	1.79	16.63
J-pictures	3.31	1.15	3.85	3.08	3.39	2.92	17.69
F-stories	3.68	3.09	3.41	3.50	2.68	2.5	18.86
F-pictures	3.4	1.47	3.6	3.47	3.27	3.07	18.27
C-stories	3.47	3.58	3.95	3.79	3.9	3.9	22.58
C-pictues	3.07	1.64	4	3.86	3.93	3.93	20.43

According to White et al, the story task is overall more successful in eliciting recognition of nonsubject antecedents in English monoclausal sentences than the picture task, although there does not appear to be any difference for biclausal sentences:

> In general, our results suggest that the truth value judgment tasks are not in fact able to totally overcome the problem of preferences in the case of potentially ambiguous sentences. This is especially true of truth value judgment tasks involving pictures. In conclusion, our results suggest that certain tasks can lead one to underestimate L2 learner's linguistic competence and that one must be cautious about making assumptions on the nature of the interlanguage grammar based on single tasks. (p.8)

Even if subjects receive training in detecting ambiguity (Thomas, 1995) and even if the researcher modifies the design, the preference problem persists. This may indicate that a preference for a certain antecedent can better be interpreted as a grammatical constraint rather than as merely a flaw of the experimental design. In other words, the notion of preference has been used by researchers as a convenient label for a still unexplained phenomenon in which one NP is preferred to another as an antecedent for a reflexive, even though both are theoretically possible. Such preference should be viewed as a phenomenon which itself results from the interlanguage grammar (in the broadest sense) and is not simply the byproduct of a flawed experimental design. For example, if a subject has 20 chances to select a local antecedent for a reflexive, but does so only 5 times, then it appears that the subject has a grammar that prefers nonlocal binding.

3.7. Individual results.

Up to this point we have discussed the issue of preference and its relation to the experimental designs used in L2 reflexive binding studies. Another issue in the methodology of reflexive binding studies is whether to consider group results or individual results. Most of the existing studies in this area reported their data as group means. However, according to Thomas (1995), in order to study the relationships between UG and L2 acquisition, one must analyze the responses of individuals instead of group responses, because group results may conceal important characteristics of individual grammars. The examination of individual responses enables the researcher to determine whether the group results are derived from identical behavior of all individuals or else from very distinct patterns of response from different individuals within the group. For example, suppose that a group of 15 subjects averaged 50% correct scores overall on some test. It might be the case that all of the 15 subjects scored exactly 50%. On the other hand, it might also be the case that 5 subjects scored 100%, 5 subjects scored 0%, and the other 5 scored 50%. In this latter case, the group result would not reveal the variation among the subjects. In this way, examination of individual results would present of clearer picture of the abilities of members within a group with regard to some feature.

Eckman (1994) also talks about the necessity of analyzing individual patterns and claims that data from individuals rather than group data are necessary to test whether interlanguage grammars obey the principles of UG, because it is impossible to determine whether the UG principles in question are being followed if the researcher cannot observe whether individual interlanguage grammars of each subject obey the constraints of UG:

> If a study reports that one of the L2 groups gave local responses on English reflexives in 80% of the cases as does Thomas (1991), this result does not tell us whether the principles of UG are being obeyed unless we know how those responses are being distributed across the subjects. Clearly, the case for UG can be made if all of the subjects scored 80%. If on the other hand, the 80% group results is a consequence of subjects' compiling many different scores, some of which indicate systematic adherence to UG principles and others that are either unsystematic or indicate nonadherence to UG, then the case for UG governing SLA is not at all clear. (p.213)

Hirakawa's study (1990) is a good example of how group results may conceal some characteristics of individual grammars. Hirakawa analyzed and reported her data in terms of group means. Later, MacLaughlin (1995a) reanalyzed Hirakawa's data for individual subjects and showed that it was impossible to ascertain the nature of the grammatical system that the learners have acquired by looking at group results. According to the group results reported by Hirakawa, 72.5% of responses by L2 learners selected a local NP in finite biclausal sentences while only 55% chose a local NP in nonfinite biclausal sentences. According to Hirakawa, these group results reveal that the L2 learners appear to be different from both the native English and the native Japanese controls and furthermore could not acquire the L2 binding setting for English reflexives. However, the individual results based on the same data indicate that ten subjects responded 100% correctly (5/5 consistent response) and 18 subjects responded 80% correctly (4/5 consistent response), and hence appeared to have acquired the native grammar with respect to English reflexive binding. Therefore, an examination of individual results proved that Hirakawa's interpretation based on group results was inaccurate.

3.8. Other problems with task design in studies of reflexive binding.

There are several other factors which might affect the subjects' responses in these tasks. For example, Thomas (1989) used a multiple-choice task where a test sentence was followed by a question and three choices, as in (p.286):

(35) David could see that Bill was looking at himself in the mirror

Who did Bill see in the mirror?

a. Bill
b. David
c. either Bill or David

Berent (1994) objects to this type of task design because it is not possible for subjects at lower levels to comprehend the question following the test sentence. Furthermore, the choice (c) will always be artificially low because it is last. In order for them to select choice (c), subjects have to rethink the options. Moreover, according to Berent, the disjunctions "either-or" are ambiguous because they can have an inclusive meaning or an exclusive meaning. Berent thinks that these factors weaken the reliability of a subject's response. Thomas (1993) mentions a problem of validity which any grammaticality judgment task can have. Any person for whom the task is too difficult or even a person with no knowledge at all of the language of the test can circle the answers randomly. Another mechanical problem Thomas has recognized is related to the picture identification task. Thomas wonders how Eckman's test sentence "Mr. Small found Mr. Big's comments on himself" could be portrayed visually so that subjects can readily distinguish this picture from a drawing of the competing interpretation ("Mr. Small found Mr. Big's comments on him").

3.9. Conclusion.

Research in L2 acquisition of reflexive binding has relied on comprehension or interpretation tasks as a source of information about the role of UG in L2 binding. As we have seen up to this point, these tasks have had two main problems. The first is data elicitation — do these tasks elicit information about L2 learners' knowledge of L2 reflexives or do they give information about preferences? The second is data analysis — should the grammars of individuals be studied or the grammars of groups? The second problem is easier to solve for researchers because any study can report both group and individual results, and hence it will be possible to find out what group results conceal about individual results.

Moreover, these tasks should be carefully designed to overcome the various problems discussed, in order to make claims about the nature of interlanguage grammars. If the problems mentioned earlier cannot be overcome, then there might be two options for researchers (Ellis, 1991). One is to abandon the use of these tasks until more is discovered about the type of data they elicit. The other is to work on the method by readministering different tasks, comparing the results obtained and reporting the changes in the judgments that occurred. In addition, researchers can obtain introspective information about the learners' thought processes in order to find out whether or not learners use their preference in responding to such sentences. All of the researchers mentioned in this chapter have expressed concern that the tasks used in their studies might tap the learners' preferred interpretation of the reflexives rather than their grammatical knowledge. Therefore, they are aware of the fact that judgment tasks used by them in L2 reflexive binding studies may

constitute performance data, and they are ready to investigate what kind of performance is involved. As long as they are aware of this, there is at least some promise that the elicitation methods will become more sophisticated in subsequent studies.

Chapter 4

TURKISH LEARNERS' KNOWLEDGE OF ENGLISH REFLEXIVE BINDING

4.1. Introduction.

This chapter describes an experiment which investigates the acquisition of English reflexive binding by Turkish adult second language learners of English with respect to a parameterized approach within the framework of UG. More specifically, the study tries to observe how crucial pragmatic factors are to the interpretation of reflexives in second language acquisition.

The binding theory on which the binding approaches such as the GCP and LF movement are based states only syntactic conditions under which an anaphor can be coindexed with a c-commanding antecedent within a local domain regardless of the nature of the antecedent. However, in several languages, like Turkish, Chinese, Japanese and Korean, the choice of antecedent among several possible ones is primarily of a pragmatic nature and the binding of reflexives in these languages may not be subject to the c-command condition at all.

4.2. The properties of Turkish reflexives.

This section examines the pragmatic and syntactic properties of long-distance anaphors in Turkish that distinguish these from local anaphors. According to Underhill (1976) and Kornfilt (1997), Turkish has two different reflexives; "kendi," which is the bare, noninflected form, and "kendisi," which is morphologically complex. Each of these reflexives has different binding properties. The simple form "kendi" is locally bound, whereas "kendisi" may be long-distance bound.

4.2.1. Local anaphors in Turkish. According to Kornfilt (1997) and Underhill (1976), the uninflected reflexive "kendi" is always locally bound within the minimal S, as in (36) (from Underhill (1976, p.356)):

36) Orhan$_i$, Mehmedin$_j$ kendine$*_{i/j}$ palto almasina
Orhan Mehmet-GEN self-DAT coat take-NOM-POSS-DAT

sevindi.
please-PASS-PAST

"Orhan$_i$ was pleased that Mehmet$_j$ bought a coat for himself$*_{i/j}$"

In (36), according to Underhill, "kendi" ("self") can only be coindexed with the local NP "Mehmet," and not with the nonlocal NP "Orhan." According to Kornfilt (1997), the noninflected "kendi" is always bound by a c-commanding local antecedent.

4.2.2. Long distance anaphors in Turkish. The following are the characteristics of the Turkish reflexive "kendisi" ("self"):

i. "Kendisi" can take more than one possible antecedent. Consider the following example from Underhill, 1976, p.356:

(37) Orhan$_i$, Mehmedin$_j$ kendisine$_{i/j}$ palto almasina
Orhan Mehmet-GEN self-POSS-DAT coat take-NOM-POSS-DAT

sevindi.
please-PASS-PAST

"Orhan$_i$ was pleased that Mehmet$_j$ bought a coat for himself$_{i/j}$"

Underhill (p.356) notes that in subordinate clauses, "kendisi" may refer to the subject of the subordinate clause or of the main sentence. So, for example, in (37), "kendisi" may refer to either the subject of the main sentence ("Orhan") or to the subject of the subordinate clause ("Mehmet").

ii. "Kendisi" can occur as the subject of the subordinate clause:

(38) Ahmet$_i$ kendisinin$_i$ cok akilli oldugunu saniyor.
Ahmet self-POSS-GEN so clever be-POSS think-PROG

"Ahmet$_i$ thinks that self$_i$ is so clever"
"Ahmet$_i$ thinks that he$_i$ is so clever"

iii. "Kendisi" can not take a first or second-person antecedent:

(39) Ben_i $Ahmetin_j$ $kendisinden*_{i/j}$ nefret ettigini biliyorum.
I Ahmet-GEN self-POSS-GEN-ABL hate-POSS know-PROG-1stSING

"I_i think that $Ahmet_j$ hates $self*_{i/j}$"
"I_i think that $Ahmet_j$ hates $himself*_{i/j}$"

iv. The binders of "kendisi" do not need to c-command the anaphor:

(40) $Alinin_i$ $babasi_j$ $kendisine_{i/j}$ palto alacak
Ali-GEN father-POSS self-POSS-DAT coat buy-FUT

"Ali's$_i$ father$_j$ will buy a coat for $self_{i/j}$"
"Ali's$_i$ father$_j$ will buy a coat for him_i/$himself_j$"

In (40), the potential antecedent "Ali" does not c-command the anaphor "kendisi."

v. "Kendisi" can have split antecedents:

(41) Ali_i $Ahmete_j$ sinavi $kendilerinin_{i/j}$ kazandigini soyledi.
Ali Ahmet-DAT exam-OBJ self-PL-GEN pass-PAST-POSS-OBJ tell-PAST

"Ali_i told $Ahmet_j$ that $self_{i+j}$ passed the exam"
"Ali_i told $Ahmet_j$ that $they_{i+j}$ passed the exam"

vi. "Kendisi" may not require an antecedent within the sentence but might be discourse bound:

(42) A: $Ahmet_i$ aradi mi?
Ahmet call-PAST INT

"Did $Ahmet_i$ call?"

B: Hayir, ama kendisi$_i$ geldi.
No but self-POSS come-PAST

"No, but self$_i$ came"
"No, but he$_i$ came"

It should be noted that although the traditional view appears to be that Turkish has two distinct reflexives, one local and the other nonlocal (Underhill, 1976), this view has been challenged by Kornfilt (1997), who argues that Turkish "kendisi" is actually a logophoric element rather than a genuine long-distance reflexive. Since "kendisi" is a logophoric element, it does not need to be c-commanded. Kornfilt claims that the use of "kendisi" is determined by discourse related factors rather than structural ones, unlike long-distance reflexives in languages like Chinese and Japanese.

Because "kendisi" has a wide range of possible antecedents, discourse/pragmatic factors will help in the choice of antecedent. Consider the following example:

(43) Kucuk kiz$_i$ annesinin$_j$ kendisine$_{i/j?}$ oyuncak almasina
Little girl mother-POSS-GEN self-POSS-DAT toy buy-POSS

cok sevindi.
very please-PASS-PAST

"The little girl$_i$ was pleased that her mother$_j$ bought herself$_{i/j?}$ a nice toy"

In (43), "the little girl" is the most plausible antecedent of the reflexive "kendisi," since our knowledge about the world tells us that mothers usually buy toys for their daughters and not for themselves (see section 4.5. for a more detailed description of what is mean by "pragmatic factors" or "pragmatic constraints.") But one could imagine other more marked

contexts in which the little girl's mother is the most plausible antecedent, for example, where the mother is an avid toy collector.

Kornfilt's position then, appears to be that Turkish has one true reflexive, namely the bare form "kendi," whereas the inflected form "kendisi" is actually not a genuine reflexive but rather a logophoric element. This view is attractive because it removes Turkish as an obstacle to an apparently uncontested universal generalization. In this work, however, I will assume the traditional view that Turkish maintains two distinct, genuine reflexives, one of which must be locally bound, and one which may be long-distance bound. However, I will make this assumption with the proviso that the complex anaphor "kendisi" functions more like a pronoun than a reflexive in several respects, as we have seen above (namely, "kendisi" shows agreement and also can have extrasentential antecedents). The one way in which "kendisi" functions more like a reflexive is that it may be locally bound, whereas pronouns can not be locally bound. Thus, it should be kept in mind that while the term "reflexive" will continue to be used to refer to "kendisi" in this work, the assumption that "kendisi" is a reflexive should be treated with some skepticism; "kendisi" is then, in many respects, as much like a pronoun as it is like a reflexive. Although this proviso, as well as Kornfilt's distinction between a genuine reflexive and an apparent one in Turkish, may allow for a modified explanation of how Turkish L2 learners acquire English reflexives, it will not considerably change the conclusions drawn in later chapters.

4.3. Binding domains in Turkish and English.

The Turkish reflexive "kendisi" falls into the largest setting of the GCP (value (e)), which is the most marked setting. It is a single morpheme, long-distance bound, and can be bound by both subject and nonsubject antecedents. On the other hand, English reflexives (e.g., "himself") have the smallest setting of the GCP, hence the unmarked setting. They are complex morphemes, are locally bound, and can be bound by either subject and non-subject antecedents. Consider the following example:

(44) John told Bill that Brian talked to Sam about himself

In the Turkish equivalent of this sentence, any NP (namely Sam, Brian, Bill and John) is a possible antecedent, whereas in English only Sam and Brian within the subordinate clause may be possible antecedents. The possible antecedents for Turkish and English are illustrated in (45) below:

(45)	a.	S_1 [	NP_i	V	NP_j	S_2 [	NP_k	V	NP_l	refl] S_2] S_1	
	b.		i		j		k		l	i/j/k/l	(Turkish)
	c.		i		j		k		l	*i/*j/k/l	(English)

As the illustration in (45) shows, English reflexives must be bound within the local domain, whereas Turkish reflexives permit long-distance binding. In other words, while English reflexives have to obey a clausemate condition, Turkish reflexives can be indefinitely far away from their antecedents.

4.4. The role of pragmatics in Turkish reflexive binding.

The role of pragmatics seems to be essential to the interpretation of reflexives in Turkish. In Turkish any noun phrase in any position with any gender can be a possible antecedent of "kendisi." Turkish speakers rely heavily on inference, context and knowledge about the world in order to choose among several possible antecedents. English speakers do not rely on pragmatic information nearly so much because reflexives are syntactically coindexed within the smallest domain, and hence the choice of antecedent is limited to local NPs. This study investigates the effect of pragmatic constraints on the interpretation of possible antecedents for English reflexives by Turkish L2 learners, because pragmatics plays a crucial role in Turkish reflexive binding whereas syntax plays the larger role in determining the antecedent in English reflexive binding. Therefore, this study aims to observe whether Turkish learners will transfer the pragmatic mode of interpretation of reflexives from their L1 to their L2. Furthermore, this study attempts to see whether with increasing amount of input over time, the L2 learners will be able to abandon the pragmatic mode of Turkish and adopt the syntactic mode of English.

4.5. The nature of "pragmatic constraints".

It is necessary to explain what is meant by "pragmatic constraints." Consider the antecedent of "himself" in the following sentence:

(46) The president ordered his bodyguard to protect himself during the speech

This pragmatically biased sentence leads to the interpretation that the reflexive is bound to the matrix clause subject. Bias in this sentence has to do with our knowledge about the way things are in the world, which creates a bias in favor of one NP over the other. In the sentence above, our knowledge that presidents are protected by bodyguards and not vice versa leads us to favor the president as antecedent. In other words, pragmatic bias refers to stereotypical relations obtained between referents or events. In the example above, if the hearer selects "the bodyguard" as the antecedent, he or she provides a statement that is, in Levinson's terms (1987) "informationally weaker" than knowledge of the world allows, and is consistent with the rules of English reflexive binding. If the hearer on the other hand selects "the president" as the antecedent, he makes the strongest statement consistent with what he knows, although this interpretation "breaks" the syntactic rule of English reflexive binding.

It is obvious that pragmatically biased sentences might induce L2 learners to consider pragmatic information to the extent that it could effect syntactic intuitions about the locality requirement of English reflexives. However, it is not clear how strong this pragmatic effect could be. This study compares pragmatically biased and pragmatically neutral sentences to examine the extent to which this pragmatic bias can effect syntactic intuitions. This study also examines the results of neutral (nonbiased) sentences in order to find out whether L2 learners have acquired the properties of English local binding, and finally this study looks at whether pragmatic constraints slow down the acquisition process.

4.6. Possible outcomes of the study.

This study will test the following possibilities:

4.6.1. The effect of pragmatic bias:

either a.: Turkish learners will interpret English reflexives according to the syntactic properties of English reflexives, i.e., they will be uninfluenced (at least as uninfluenced as native speakers) by the pragmatic bias of the test sentences. In this case, they will choose only a local NP as an antecedent. This will indicate that Turkish learners are able to make a transition from the predominantly pragmatic mode of their L1 to the syntactic mode of their L2.

or b.: Pragmatic information will override the parameter setting of English and cause learners to choose as a possible antecedent an NP which might be outside the binding domain. In this case, the L2 learners will select only the pragmatically biased NP. This will indicate that L2 learners transfer the pragmatic mode of reflexive binding from their L1 to their L2 in the interpretation of English reflexives.

4.6.2. The effect of syntactic constraints.

either a.: If the acquisition of binding principles in a second language proceeds as in the acquisition of the binding principles in a first language, then Turkish L2 learners of English will start with the smallest setting of the binding parameter (which is also the English setting). The L2 learners will remain at this value once and for all because they will not come across any positive evidence which could force them to change the binding setting for English (as the Subset Principle predicts). In this case, both the initial state and

the final state of the L2 learners will be the target language (English). Then we should expect very few instances of long-distance binding from Turkish learners because they will completely ignore the parameter setting of their L1.

or b.: If the acquisition of binding principles in a second language does not proceed as in the acquisition of the binding principles in a first language, then Turkish L2 learners of English will transfer their L1 binding setting to their L2. Hence, they will employ the most marked setting in their L2 reflexive binding system, although their L2 requires the most unmarked one. In this case, the Turkish learners will allow long-distance binding for English reflexives, and further they will not be able to reset the parameter value to the English setting over time. In this case, both the initial and final state of the L2 learners are Turkish.

or c.: Even though Turkish L2 learners will initially select the marked setting of their L1 for English reflexives, with increasing input over time they gradually will be able to reset the binding parameter setting to the L2 value. Here, "input" means exposure to English in the classroom setting, but where there is no overt instruction regarding the rules of reflexive binding in English. In this case, the initial state of the L2 learners is Turkish, but the final state is English. However, it should be remembered that this possibility is problematic for Manzini and Wexler's parameterized binding principles (and for the LF movement approach as well), because the initial state of the L2 learners in this case should have been that of the target language.

4.6.3. The effect of proficiency level.

either a.: There will be a tendency towards more accurate interpretation of reflexives with increasing proficiency, i.e., the higher the proficiency level, the fewer errors the L2 learners will make. In this case, Turkish L2 learners will move from the least restrictive setting to the most restrictive setting as they become more proficient in English. As it may be seen, this possibility is related to the possibility (c) above in which the acquisition of reflexive binding proceeds from the initial transfer of the L1 setting to the resetting to the L2 value. The following stages will then be observed:

1. Low-level learners may fail to set the correct value for English binding, because they transfer their L1 binding setting, and in addition, have not learned complex sentences where reflexives may be used. In this case, they will tend to choose nonlocal antecedents in both finite and nonfinite clauses. In addition, they will be most heavily influenced by pragmatic constraints.

2. Intermediate learners will not treat English like their L1 as low level learners do. In this case, they will show evidence for local binding, but will also allow non-local binding, and will be swayed to some degree by pragmatic constraints.

3. Advanced learners will set the correct binding setting for English regardless of sentence type or pragmatic bias. They will ignore pragmatic information and select antecedents for reflexives according to syntactic constraints.

or b.: There will not be any tendency towards more accurate interpretation of reflexives with increasing proficiency.

4.6.4. The effect of sentence type.

either a.: The learners will be affected by the levels of embedding, i.e., the learners, especially low-level learners, will select more nonlocal antecedents in triclausal sentences than biclausal ones. Previous research (e.g., Hirakawa, 1990; Matsumura, 1994) has shown that L2 learners tend to have problems with complex sentences. Therefore, they may have difficulty discovering the smallest clause for triclausal sentences which could contain possible antecedents.

or b.: The learners will be affected by nonfiniteness, i.e., they will interpret more nonlocal NPs as antecedents for reflexives in nonfinite clauses than finite ones because of the difficulty of parsing nonfinite sentences — nonfinite sentences lack both tense and agreement markers, and therefore it might be more difficult for L2 learners to locate the smallest clause which could contain the antecedent of the reflexive.

or c.: The complexity of structure of the different types of sentences will not have differential effects on choice of antecedents.

4.7. Subjects

4.7.1. Background information. The subject pool consisted of an experimental group of 170 native speakers of Turkish, a control group of 25 native speakers of English, and another control group of 25 native speakers of Turkish. The subjects in the experimental group and in the Turkish control group were full-time undergraduate students in an English department within the Faculty of Education at Uludag University in Turkey. In the department from which the subjects were drawn, English is the medium of education, and

the expected length of the program is five years. All of the subjects are expected to graduate as English teachers. In this study, there were 5 proficiency levels which were determined by the number of years the student had attended the university. These proficiency levels were prep year, 1st year, 2nd year, 3rd year, and 4th year. It is assumed that since nearly all of the subjects had started learning English at the same age (in middle school) and followed a strictly graded syllabus at the university, the grade of the learners generally corresponded to their level of proficiency. In addition, students who enter the department must take a proficiency exam and are subsequently placed either into the prep year or into the first year class. Further, students are required to follow a strict schedule of courses appropriate to their year, and cannot pass on to the next year without completing that schedule and passing those courses. Henceforth in this study, the grades from prep year to year four will be refered to as proficiency levels.

4.7.2. Proficiency levels.

i. Prep year. This group included 33 students, 19 females and 14 males. Their ages ranged from 18 to 22. Twenty-six students in this group started taking English classes in middle school, three of them started in high school, and four after finishing high school. Six of the students within this group began learning English in middle school where English was the medium of instruction. Two began to learn English in high school where English was the medium of instruction, and three learned English in both middle school and high school where English was used exclusively in the classroom.

ii. First year. This group included 34 students, 28 females and six males. Their

ages ranged from 18 to 22. Thirty-one learners started taking English classes in middle school, two started in high school and one after finishing high school. Thirteen students within this group began learning English in middle school where English was the medium of instruction. Five began to learn in high school where English was the medium of instruction, and eight learned English in both middle school and high school where English was the language used in the classroom.

iii. Second year. This group had 35 students, 19 females and 16 males. Their ages ranged from 19 to 25. Twenty-four learners started taking English classes in middle school, nine learners started in high school, and two after finishing high school. Five students within this group began learning English in middle school where English was the medium of instruction. Five began to learn in high school where English was the medium of instruction, and seven learned English in both middle school and high school where English was the medium of instruction.

iv. Third year. This group consisted of 35 students, 22 females and 13 males. Their ages ranged from 21 to 24. Thirty-four of them started taking English classes in middle school, and just one started after finishing high school. Three within this group began learning English in middle school where English was the medium of instruction. Four began to learn in high school where English was the medium of instruction, and seven learned English in both middle school and high school where English was the language used in the classroom.

v. Fourth year. This group consisted of 35 students, 28 females and seven males. Their ages ranged from 22 to 26. All thirty-five of them started taking English classes in

middle school. Seven within this group began learning English in middle school where English was the medium of instruction. Three began to learn in high school where English was the medium of instruction, and three learned English in both middle school and high school where English was the language used in the classroom.

Most the the subjects in this study did not have any usage of English outside of the classroom. A few subjects worked in tourist information centers in summers, a few of the subjects claimed to watch television channels broadcast in English, and a few said that they sometimes read English language newspapers. Very few of the subjects had visited an English-speaking country for vacation.

v. Turkish control group. This group consisted of 25 native speakers of Turkish of the same ages as those in the experimental group. The subjects of the control group were also students within the same English department. It should be noted that none of these subjects also belonged to the experimental group.

vi. English control group. This group consisted of 25 native speakers of English who were undergraduate students at Michigan State University. These subjects were drawn from the same humanities course, and were awarded extra credit for their participation.

4.8. Materials.

The experiment of this study was composed of an interpretation task in which each test sentence with a reflexive was followed by two or three statements (depending of the level of embedding) which questioned each NP of the sentence as a possible antecedent for

the reflexive by requiring an answer of "yes" or "no." The test format asked the subjects to consider each NP within the sentence separately and indicate whether or not it was possible for the reflexive to refer to that NP. Each test sentence contained at least two NPs with the same gender which agreed in number with the reflexive. The following example illustrates the test format used in the study:

(47) Ahmet said that Cem always criticized himself.

a. "himself" = Ahmet	YES	NO
b. "himself" = Cem	YES	NO

Subjects were asked to indicate their interpretation of the reflexive by circling either "yes" or "no" for each statement which followed the test sentence. Prior to the test, they were explicitly instructed not to assume that giving a "yes" response for one antecedent in one statement required a "no" response for another possible antecedent in the other statement. In the example above, there are two possible antecedents for the reflexive "himself"; "Ahmet" and "Cem." For the statement (a), a subject would circle either "yes" or "no" in considering "Ahmet" as an antecedent for the reflexive "himself." For the statement (b), a subject would circle either "yes" or "no" in considering "Cem" as a possible antecedent.

It was expected that this type of format would allow the subjects to check each possible antecedent separately and then to make a choice accordingly. Therefore, it would be possible for this study to discover whether the learners accepted or rejected each possible antecedent. It was hoped that this format would avoid the problem of limiting

subjects to just one choice of antecedent, because this study is interested in all the choices which the subjects' grammars allow.

To facilitate understanding among both high and low-level learners alike, the experimental materials employed simple vocabulary and conventional Turkish names whose genders are easily recognizable. Furthermore, prior to the test every subject was given a list of vocabulary items used in the test sentences along with their Turkish gloss. The subjects were requested to ask the meaning of any lexical item which they were not familiar with. They were read the instructions which were written on their test forms in both English and Turkish:

> Consider each sentence, and circle either YES or NO for each statement below it. Think about each statement separately, and don't assume that a YES answer for one statement means that the answer for the next statement must be NO.

The test sentences were mixed and randomized in three different formats in order to prevent the potential effects of the subjects' fatigue on the test results and in order to control for the effects of sentence order on subjects. The test was given to the subjects during their class period, and subjects responded to the test sentences at their own pace.

4.8.1. Test sentences. This study used nine types of sentences, divided between biclausal and triclausal sentences, finite and nonfinite sentences, and pragmatically biased and neutral sentences. Each type had three tokens, in all, 27 test sentences.

4.8.1.1. *Type I (NP1BNF)* . Type I sentences are biclausal, nonfinite, NP1 biased, in which NP1 refers to the subject of the matrix clause. The following example illustrates Type I sentences:

(48) The president ordered his bodyguard to protect himself during the speech

This pragmatically biased sentence is designed to lead to an interpretation which binds the reflexive to the matrix clause subject ("the president") rather than to the matrix clause object ("the bodyguard"). In this type of sentence the pragmatic bias is in the direction of a nonlocal NP which is not acceptable as an antecedent in English. The two candidate antecedents for the reflexive in the sentence above are the local NP ("the bodyguard"), which might be selected by L2 learners according to the syntactic constraints imposed by English reflexive binding, and the other is the nonlocal NP ("the president"), which might be chosen according to the pragmatic constraints imposed by Turkish reflexive binding. In selecting an antecedent (or antecedents) for the reflexives in Type I, the L2 learners will reveal how strongly they have constructed the binding setting for English in the face of the pressure of pragmatic constraints.

4.8.1.2. *Type II (NP1BF)* . Type II sentences are biclausal, finite, NP1 biased, in which NP1 refers to the subject of the matrix clause. (49) illustrates Type II sentences:

(49) The little girl was happy that her mother bought herself a nice toy

Type II sentences, like Type I sentences, are designed to find out if L2 learners identity reflexives with the syntactically potential antecedent which is a local subject NP, or with the pragmatically potential antecedent which is a nonlocal subject NP. If the pragmatic factors override the syntactic constraints on English binding, then the reflexive "herself" in the example above will be interpreted as "the little girl" rather than as "the mother." Choosing "the little girl" as the antecedent of the reflexive "herself" might be the correct interpretation of the reflexive for the L2 learners because such an interpretation clearly supports our real world knowledge that it is stereotypically mothers who buy toys for their children and that mothers do not ordinarily play with toys.

The only difference between Type I and Type II sentences is that the embedded clauses in Type I are nonfinite whereas those in Type II are finite. Comparing the results for these two types should reveal something about how subjects are effected by the finite/nonfinite distinction with regard to reflexive binding.

4.8.1.3. *Type III (NP2BNF)* . Type III sentences are biclausal, nonfinite, NP2 biased, in which NP2 refers to the object of the matrix clause (the local antecedent). The example below illustrates Type III sentences:

(50) The doctor asked the sick woman to take care of herself

This sentence is designed to impose an interpretation where the reflexive is bound to the local antecedent ("the sick woman"). In (50), the two candidate antecedents for the

reflexive are "the sick woman," which might be an antecedent due to both pragmatic and syntactic constraints, and the other is the nonlocal NP "the doctor," which might be an antecedent due to the reflexive binding of the subjects' L1 grammar.

4.8.1.4. *Type IV (NP2BF)*. Type IV sentences are biclausal, finite, NP2 biased. Consider the following example:

(51) Ahmet heard that the famous actor talked about himself on T.V.

In example (51) above, as in Type III sentences, the pragmatic bias favors the local NP "the famous actor" as the antecedent rather than the nonlocal NP "Ahmet." The only difference between Type III and Type IV sentences is that Type III are nonfinite (object control), whereas Type IV are finite (sentential).

Type III and Type IV sentences in which the local NP is pragmatically favored as an antecedent are designed to discover if the bias towards the local NP leads the L2 learners to increase their local binding. Furthermore, the comparison of these two types of sentence with Types I and II will indicate whether bias to NP1 and bias to NP2 are equally strong. The local NP (NP2) is the legal antecedent according to the grammar rules of English reflexive binding. Moreover, the same NP is also a likely antecedent according to the pragmatic rules of the subjects' L1 binding system. If the subjects tend to choose NP2 for NP2 biased sentences more than they choose NP1 for NP1 biased sentences, then we may claim that the syntactic rules of reflexive binding in English are interacting with pragmatics

in the interlanguage of L2 learners.

4.8.1.5. *Type V (BCNF)*. Type V sentences are biclausal, nonfinite, neutral structures. (52) illustrates Type V sentences:

(52) Cem asked Hakan to trust himself more

Neutral sentences are those in which either NP1 or NP2 are pragmatically possible antecedents for the reflexive. In neutral sentences, the learners' knowledge of reflexive binding in English should control their choice of an antecedent for the reflexive. If the binding condition for English is a part of the learners' grammars, then the learners will select only the local NP ("Hakan" in the example above), and not the nonlocal NP ("Cem").

4.8.1.6. *Type VI (BCF)*. Type VI sentences are biclausal, finite, and neutral. (53) gives an example of a Type VI sentence:

(53) Ali said that Salih criticized himself

There are two equally plausible antecedents for "himself" in (53), the local NP "Salih," and the nonlocal NP "Ali." The only difference between Type V and Type VI is that Type V sentences are nonfinite whereas Type VI sentences are finite.

4.8.1.7. *Type VII (TCNF)* . Type VII sentences are triclausal with two embedded nonfinite clauses. Like all of the triclausal sentence types in this study Type VII sentences are neutral:

(54) Ahmet asked Adnan to tell Cengiz to look at himself in the mirror

In (54) there are three candidate antecedents, "Ahmet," "Adnan," and "Cengiz."

4.8.1.8. *Type VIII (TC1NF)* . Type VIII sentences are triclausal with one nonfinite clause in the most embedded position and one finite clause before that:

(55) Fatih believed that Adnan asked Salih to respect himself more

4.8.1.9. *Type IX (TCF)* . Type IX sentences are triclausal with two embedded finite clauses:

(56) Ali remembers that Hakan said that Can hates himself

The triclausal sentences are designed to determine whether the level (and type) of embedding will influence the choice of antecedent.

4.8.2. Summary. In all of the biclausal sentence types reviewed so far, there are two candidate antecedents for a reflexive which is always in nonsubject position in either a finite

or nonfinite object complement clause. Both possible antecedents are either subject NPs or object NPs of the matrix clause, and they c-command and precede the reflexive. In the triclausal sentences, there are three candidate antecedents for a reflexive which is also always in nonsubject position. Furthermore, all three antecedents are either subject NPs or object NPs of the matrix clause, and c-command and precede the reflexive. In identifying the antecedents of the reflexives in these sentence types, the learners will reveal how they treat the binding domain for English reflexives, sometimes driven by pragmatic considerations, and sometimes driven by grammatical rules.

The division between pragmatically biased versus neutral sentences aims to show whether pragmatic constraints or syntactic ones are predominantly employed by L2 learners. The learners who have already acquired the locality conditions of English reflexives should resist pragmatic favoring of a nonlocal NP and bind reflexives exclusively to the local NP, even though the nonlocal NP is more pragmatically plausible as an antecedent.

The division between biclausal and triclausal sentences aims to reveal whether L2 learners are affected by the number of embeddings when they bind reflexives to an antecedent. It is obvious that the L2 learners may be faced with a problem of analyzing triclausal sentences and of determining clause boundaries, which will create a serious problem for L2 learners because they have to bind English reflexives with an antecedent within the smallest clause. That is, if learners cannot identify the boundaries of the smallest clause among three clauses, then they will be liable to (incorrectly) choose antecedents from larger domains. The results of previous studies showed that L2 learners permit more long-

distance binding in triclausal sentences than in biclausal ones (Hirakawa, 1990; Matsumura, 1994).

The division between finite and nonfinite sentences aims to discover if L2 learners will be able to correctly parse a nonfinite clause by determining what the matrix clause and matrix subject is, as well as what the subordinate clause and subordinate subject is. All of this information is necessary for L2 learners so that they will be able to find out what the local NP is in order to bind the reflexive with the subordinate subject of the nonfinite clause. L2 learners may misinterpret infinitivals as monoclausal sentences, thinking that the first NP is the subject and the second NP is the object — hence treating both NPs as local antecedents. The results of many previous studies indicate that L2 learners allow more long-distance binding across infinitival clauses than across finite clause boundaries (Finer and Broselow, 1986; Finer, 1991; Hirakawa, 1990; Eckman, 1994). Furthermore, according to the results of these studies, L2 learners may reset their L2 setting to an intermediate value of the GCP which allows long-distance binding across infinitival but not finite clauses.

In each of these English sentence types, the reflexive is bound only to the local NP, whereas in the Turkish equivalent of these sentence types, "kendisi" ("self") can be coreferential with any of the NPs appearing in the sentence. It should be remembered that Turkish does not make any finite/nonfinite distinction. Therefore, there were only four types of sentence to be answered by the Turkish control group. These are biclausal NP1 biased sentences, biclausal NP2 biased sentences, biclausal neutral sentences, and triclausal sentences. Each type was represented by six tokens, and all of the sentences were in

Turkish and corresponded to their English equivalents.

4.9. Analyzing data.

There are four possible outcomes for each token for biclausal sentences and eight possible outcomes for triclausal sentences. For example, the L2 learners may interpret the biclausal example below in such a way so as to establish coreference between the reflexive and any of (57a-d):

(57) Ali said that Salih criticized himself

1) himself = Ali	Yes	No
2) himself = Salih	Yes	No

a. Yes Yes =	Ali	Salih	(either NP1 or NP2)
b. Yes No =	Ali		(NP1 only)
c. No Yes =		Salih	(NP2 only)
d. No No =			(neither NP1 nor NP2)

L2 learners may interpret the triclausal example below in such a way so as to establish coreference between the reflexive and any of (58a-h):

(58) Ali remembers that Hakan said that Can hates himself

1) himself = Ali	YES	NO
2) himself = Hakan	YES	NO
3) himself = Can	YES	NO

a. Yes Yes Yes =	Ali	Hakan	Can	(either NP1, NP2 or NP3)
b. Yes Yes No =	Ali	Hakan		(NP1 or NP2 only)
c. No Yes Yes =		Hakan	Can	(NP2 or NP3 only)
d. Yes No Yes =	Ali		Can	(NP1 or NP3 only)
e. Yes No No =	Ali			(NP1 only)
f. No Yes No =		Hakan		(NP2 only)
g. No No Yes =			Can	(NP3 only)
h. No No No =				(neither NP1 nor NP2 nor NP3)

The incidence of each of the four possible responses for biclausal sentences (57a-d), and of the eight possible responses for triclausal sentences (58a-h) were tabulated for three tokens of each sentence type. For five experimental groups, out of a total of 4590 potential tokens (170 subjects x 9 sentences types x 3 tokens), there were ten missing tokens because ten test sentences were not assigned any interpretation. Out of a total of 675 potential tokens for the Turkish control group (25 subjects x 27 sentences (three biclausal types x 6 tokens and one triclausal type x 9 tokens)), there were 28 missing tokens, 27 of which were triclausal sentences. Out of a total of 675 potential tokens for the English control group (25 subjects x 9 sentence types x 3 tokens), there were not any missing tokens.

Chapter 5

RESULTS

5.1. Introduction.

This chapter presents both group and individual results. First, the group results are analyzed in the following way; the analysis counts the number of responses provided by the subjects in terms of whether they chose NP1, NP2, all, and neither for biclausal sentences, and NP1, NP2, NP3, all, NP1-NP2, NP1-NP3, NP2-NP3 for triclausal sentences. Then, the percentages for each possibility for each subject group and for each sentence type are calculated. Furthermore, the analysis counts the mean scores based on the number of correct responses (NP2 for biclausal sentences and NP3 for triclausal sentences) for each proficiency group and for each sentence type.

As a second step, the analysis of the individual results counts the number of L2 learners in each group who chose only the local antecedents for reflexives sentence by sentence, and examines the differences in the percentages of the learners between the groups according to sentence types.

5.2. Group results

5.2.1. Results of biclausal sentences.

5.2.1.1. *Proficiency levels* . Tables 4 and 5 present the results of the judgment task by

biclausal sentence types for all five proficiency levels as well for the native and Turkish control groups. Table 4 gives the results in raw numbers in terms of choice of antecedent for each biclausal sentence type. Table 5 reports the percentage of the responses for each proficiency level which establish coreference between the reflexive and the indicated possible antecedents (NP1, NP2, ALL, and Neither) for biclausal sentences. Proficiency levels are listed by row, and sentence types are listed by column:[1]

Table 4: Raw scores for all groups for biclausal sentence types

		NP1BNF*	NP1BF	NP2BNF	NP2BF	BCNF	BCF
NP1	prep (31)	75	87	10	9	18	18
	1st year (35)	86	89	7	12	26	23
	2nd year (36)	80	81	4	3	21	20
	3rd year (33)	91	90	9	7	21	36
	4th year (35)	86	92	3	8	16	17
	Native con. (25)	18	14	1	0	6	3
	Turkish con. (25)	141	-	20	-	55	-
NP2	prep (31)	15	4	78	76	65	65
	1st year (35)	16	13	92	91	73	77
	2nd year (36)	18	16	95	99	69	67
	3rd year (33)	8	8	90	89	74	63
	4th year (35)	15	11	85	93	75	74
	Native con. (25)	54	57	73	75	66	70
	Turkish con. (25)	4	-	114	-	68	-
ALL	prep (31)	3	1	4	8	10	10
	1st year (35)	2	2	5	2	6	5
	2nd year (36)	10	11	9	6	18	21
	3rd year (33)	0	1	0	3	4	0
	4th year (35)	4	2	12	4	13	14
	Native con. (25)	3	3	1	0	3	2
	Turkish con. (25)	5	-	15	-	27	-
Neither	prep (31)	0	0	1	0	0	0
	1st year (35)	4	1	1	1	0	0
	2nd year (36)	0	0	0	0	0	0
	3rd year (33)	0	0	0	0	0	0
	4th year (35)	0	0	0	0	0	0
	Native con. (25)	0	1	0	0	0	0
	Turkish con. (25)	0	-	0	-	0	-

* NP1BNF = NP1 (nonlocal NP) biased, nonfinite
NP1BF = NP1 biased, finite
NP2BNF = NP2 (local NP) biased, nonfinite
NP2BF = NP2 biased, finite
BCNF = Biclausal neutral, nonfinite
BCF = Biclausal neutral, finite

Table 5: Percentages for all groups for biclausal sentence types

		NP1BNF	NP1BF	NP2BNF	NP2BF	BCNF	BCF
NP1	prep (31)	80.65	94.57	10.75	9.68	19.35	19.35
	1st year (35)	82.69	84.76	6.67	12.38	24.76	21.90
	2nd year (36)	74.07	75.00	3.70	2.78	19.44	18.52
	3rd year (33)	91.92	90.91	9.09	7.07	21.21	36.36
	4th year (35)	81.90	87.62	3.81	7.62	15.38	16.19
	Native con. (25)	24.00	18.67	1.33	0.00	8.00	4.00
	Turkish con. (25)	94.00	-	13.42	-	36.67	-
NP2	prep (31)	16.13	4.35	85.71	81.72	69.89	69.89
	1st year (35)	15.38	12.38	87.62	84.76	69.52	73.33
	2nd year (36)	16.67	14.81	87.96	91.67	63.89	62.04
	3rd year (33)	8.08	8.08	90.91	89.90	74.75	63.64
	4th year (35)	14.29	10.48	84.76	88.57	72.12	70.48
	Native con. (25)	72.00	76.00	97.33	100.00	88.00	93.33
	Turkish con. (25)	2.67	-	76.51	-	45.33	-
ALL	prep (31)	3.22	1.09	4.30	8.60	10.75	10.75
	1st year (35)	1.92	1.90	4.76	1.90	5.71	4.76
	2nd year (36)	9.26	10.19	8.33	5.56	16.67	19.44
	3rd year (33)	0.00	1.01	0.00	3.03	4.04	0.00
	4th year (35)	3.81	1.90	11.43	3.81	12.50	13.33
	Native con. (25)	4.00	4.00	1.33	0.00	4.00	2.67
	Turkish con. (25)	3.33	-	10.01	-	18.00	-
Neither	prep (31)	0.00	0.00	1.08	0.00	0.00	0.00
	1st year (35)	0.00	0.95	0.95	0.95	0.00	0.00
	2nd year (36)	0.00	0.00	0.00	0.00	0.00	0.00
	3rd year (33)	0.00	0.00	0.00	0.00	0.00	0.00
	4th year (35)	0.00	0.00	0.00	0.00	0.00	0.00
	Native con. (25)	0.00	1.33	0.00	0.00	0.00	0.00
	Turkish con. (25)	0.00	-	0.00	-	0.00	-

Table 5 shows that comparing the performance of the proficiency levels from prep to year 4, there is little difference between the levels in terms of their choice of antecedent for each sentence type. For example, for NP1BNF sentences, 80.65% of responses by the prep year group and 81.90% of fourth year group selected the biased nonlocal NP. For NP2BNF sentences, 85.71% of responses by the prep year group and 84.76% of the fourth year group selected the biased local NP2. For BCNF sentences, 69.89% of responses by the prep year group and 72.12% of the responses by the fourth year group selected the local NP. The results show that third year subjects did the worst while the second year subjects did the best in resisting the bias towards the nonlocal NP. Furthermore, the greatest variety among levels appears to be in the 'ALL' choice for all sentence types. For example, for NP1BF sentences, 1.09% of responses by the prep year group were 'ALL', whereas 10.19% of responses by the second year group were 'ALL.' For BCF sentences, 19.44% of responses by the second year group were 'ALL,' whereas 0.0% of the third year group were 'ALL.' Although the 'ALL' choice represents the correct binding setting for Turkish (either NP1 or NP2), very few selected it. For example, for BCNF sentences, only 10% of the responses by all levels were 'ALL,' whereas 69.94% of overall responses by all levels were NP1, the correct setting for English.

Comparing the levels using a chi-square test for all the biclausal sentence types combined, there were no significant differences among the proficiency levels (chi-square = 3.088, $p < .05$, 4 df, $X^2 < 9.49$). These results appear to indicate that there is no developmental sequence for these learners in the acquisition of English reflexives.

After 5 years of input through formal college education, the more advanced subjects are at the same point as the prep year subjects in the acquisition of English reflexive binding. Figure 2 compares percentages of correct responses for the proficiency levels in a graph for each of the biclausal sentence types:

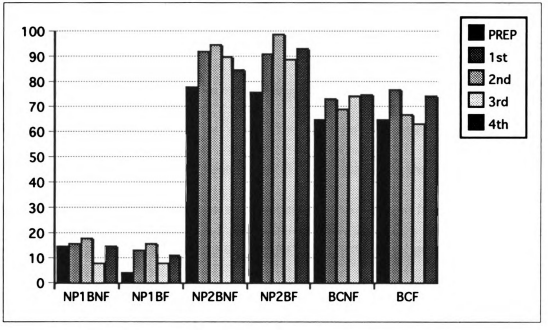


Figure 2: Percent of correct response by proficiency level for biclausal types

5.2.1.2. Turkish control group. With regard to the Turkish control group, Table 4 above shows that almost all responses (94%) of the Turkish control group selected the nonlocal antecedent (NP1) in NP1 pragmatically biased sentences. However, pragmatic favoring of the local NP (NP2) in NP2 biased sentences appears to have less of an effect on the choice of antecedent — 76.51% of the responses selected the local NP. It is interesting that 10.07% of the responses of the Turkish control group allowed

either NP1 or NP2 binding in the NP2 biased sentences, which is three times more than the percentage of 'ALL' responses in the NP1 biased sentences (3.3%). In neutral biclausal sentences, the responses of the Turkish control group were more evenly divided between the nonlocal NP (36.67%) and the local NP (45.33%), with a slightly greater tendency to select the local NP. Only 18% of responses were 'ALL' (either NP1 or NP2), which is the actual Turkish parameter setting.

5.2.1.3. *Native control group* . The native (English) control group did as one would expect, choosing the correct NP most of the time as the antecedent (91% in neutral sentences and 99% in NP2 biased sentences). It is still surprising that a substantial percentage of the responses of native speakers choose the nonlocal NP in NP1 biased sentences (21%). Figure 3 compares the results of the two control groups:

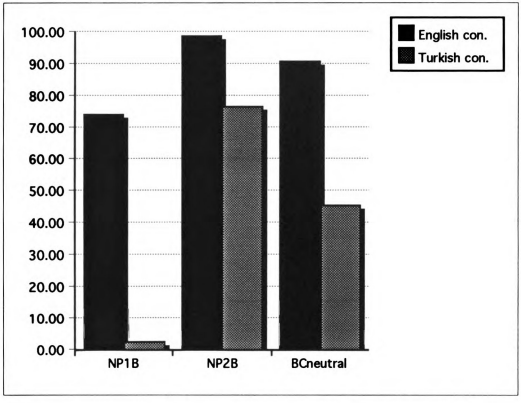


Figure 3: Percent of correct response for control groups for biclausal sentence types

5.2.1.4. *Mean scores* . Table 6 gives the mean scores (based on the number of correct responses) for each level and for each biclausal sentence type, as well as for neutral, NP1 biased, NP2 biased, and all the biclausal sentences combined:

Table 6: Mean scores for biclausal sentences

	BC n=18	neutral n=6	NP2B n=6	NP1B n=6	BCNF n=3
prep (31)	10.29	4.1	5.06	0.94	2.19
1st year (35)	9.94	4.29	5.14	0.69	2.06
2nd year (36)	9.94	3.56	5.42	0.94	1.81
3rd year (33)	9.85	4.03	5.45	0.42	2.18
4th year (35)	10.14	4.03	5.29	0.57	2
Native cont (25)	15.68	5.4	5.92	4.36	2.64
	NP1BF n=3	NP1BNF n=3	NP2BF n=3	NP2BNF n=3	BCF n=3
prep (31)	0.29	0.68	2.52	2.55	2.03
1st year (35)	0.34	0.34	2.54	2.6	2.2
2nd year (36)	0.44	0.47	2.78	2.64	1.81
3rd year (33)	0.18	0.24	2.7	2.73	1.88
4th year (35)	0.23	0.29	2.69	2.6	2.11
Native cont (25)	2.16	2.2	3	2.92	2.76

As Table 6 shows, there was almost no difference between levels in terms of choosing the correct antecedent (NP2); for example, for all biclausal sentences (18 total) the mean score for the prep year group was 10.29, while for the fourth year group the mean score was 10.14. On the other hand, the mean score for the native control group was 15.68. For all six neutral biclausal sentences, the mean score for the prep year group was 4.1, while it was 4.03 for the fourth year group. For NP2 biased sentences (n=6), the mean score for prep year was 5.06, while it was 5.29 for the fourth year group. As it is seen, this is the only sentence type where the learners come close to the native control group (mean score 5.92) in selecting the antecedent for reflexives. The biggest gap between learners and the native speakers was with NP1 biased sentences (ranging from .42 to .94 for the levels and 4.36 for the native control group). Nonetheless, mean

scores for the native control group show the effect of pragmatics; for neutral sentences this group had a mean score of 5.4, whereas for the NP1 biased sentences the mean score was 4.36.

5.2.1.5. *Combined group results* . Since there was not a significant difference between the levels (years prep through four) in the interpretations of reflexives for each biclausal sentence type, the results by proficiency levels were combined in order to examine the results for the finite/nonfinite distinction and pragmatically biased versus neutral sentences. Table 7 shows the combined results in raw numbers and percentages respectively:

Table 7: Combined results for biclausal sentences

		NP1B		NP2B		Neutral	
		nonfinite	finite	nonfinite	finite	nonfinite	finite
	raw						
NP1		418	439	33	39	102	114
NP2		72	52	440	448	356	346
ALL		19	17	30	23	51	50
Neither		4	1	2	1	0	0
	TOTAL	513	509	505	511	509	510
		NP1B		NP2B		Neutral	
		nonfinite	finite	nonfinite	finite	nonfinite	finite
	percent						
NP1		81.48	86.25	6.53	7.63	20.04	22.35
NP2		14.04	10.22	87.13	87.67	69.94	67.84
ALL		3.70	3.34	5.94	4.50	10.02	9.80
Neither		0.70	0.20	0.40	0.20	0.00	0.00

5.2.1.5.1. *Finite vs Nonfinite sentences* . It is apparent from the results shown in table 4 that there is very little difference in percentages between the finite sentences and their nonfinite counterparts. At most, comparing finite and nonfinite sentences, for any choice of antecedent (NP1, NP2 or ALL) there was no more than a 5 percentage point difference (81.48% for NP1 biased nonfinite and 86.25% for NP1 biased finite). Using the Z-test, there was no statistically significant difference between nonfinite and finite sentences ($Z = .8857$, $\alpha = .05$).[2] These results contradict the results of prior studies which revealed differences between finite and nonfinite sentences in the interpretation of reflexives by L2 learners (Finer and Broselow, 1986; Finer, 1991; Hirakawa, 1990; Eckman, 1994; MacLaughlin, 1995). Figure 4 represents the results comparing correct responses for finite and nonfinite sentences for all levels combined:

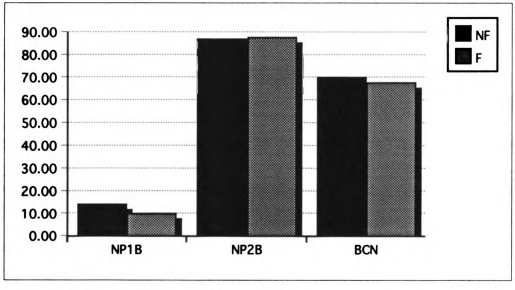


Figure 4: Finite vs nonfinite biclausal sentences

5.2.1.5.2. *Biased vs neutral sentences* . When we focus on the pragmatically biased versus neutral sentences, it is evident that when there is a pragmatic bias in favor of the nonlocal NP, then subjects are extremely likely to choose NP1 as the antecedent (84%). When there is pragmatic bias in favor of the local NP, then subjects are extremely likely to choose NP2 as the antecedent (87.4%). As it may be seen, the pragmatic bias is equally strong in both directions, that is towards the local and non-local NP. The results indicate that the subjects are heavily influenced by pragmatic bias. Finally, when there is no pragmatic bias in favor of either NP1 or NP2 (i.e., neutral), then subjects tend to choose NP2 (68.89%), although not as much as when the NP2 is biased. In neutral sentences, some responses selected the nonlocal NP (21.2%), and some selected 'ALL'(around 9.91%). Using the Z-test, there was a statistically significant difference between NP1 biased versus neutral sentences in terms of the number of correct responses ($Z = 51.61$, $\alpha = .05$), as well as between NP2 biased versus neutral sentences ($Z = 10.11$, $\alpha = .05$).

If we compare the levels (year prep though year 4) with the control groups, we find that there is a statistically significant difference between the levels and the native control group for all sentences combined ($Z = 17.30$, $\alpha = .05$), as well as between the levels and the Turkish control group ($Z = 7.198$, $\alpha = .05$). In addition, there is also a significant difference between the levels and the native control for NP1 biased sentences (Chi-square = 332.98, $p = .05$, 1 degree of freedom, $X^2 > 3.84$), for NP2

biased sentences (Chi-square = 16.74, p = .05, 1 degree of freedom, $X^2 > 3.84$), and for neutral sentences (Chi-square = 30.29, p = .05, 1 degree of freedom, $X^2 > 3.84$). Finally, there is a statistically significant difference between the levels and the Turkish control group for NP1 biased sentences (Chi-square = 10.65, p = .05, 1 degree of freedom, $X^2 > 3.84$), for NP2 biased sentences (Chi-square = 10.96, p = .05, 1 degree of freedom, $X^2 > 3.84$), and for neutral sentences (Chi-square = 17.57, p = .05, 1 degree of freedom, $X^2 > 3.84$).

These results indicate that the L2 learners do not behave like native speakers of English nor like Turkish speakers in the interpretation of reflexives. Turkish learners appear to have constructed grammars of English reflexives which carry little resemblance to the treatment of "kendisi" in their L1, and which aren't similar to the treatment of "himself" in their L2 either.

In sum, results of the biclausal sentences reveal that while there was no significant difference between the proficiency levels, there was a significant difference between the levels and the Turkish and English control groups for all sentence types. In addition, there was no significant difference between finite and nonfinite sentence types in terms of the selection of the correct antecedent. Finally, there was a significant difference between the pragmatically biased sentences and the neutral ones. On the other hand, there was no significant difference between the NP1 biased sentences and the NP2 biased sentences in terms of the selection of the biased NP.

5.2.2. Results of Triclausal sentences.

5.2.2.1. *Proficiency levels and control groups* . Tables 8 and 9 present the results of the judgment task by triclausal sentence types for all five proficiency levels as well for the English and Turkish control groups. Table 8 gives the results in raw numbers in terms of choice of antecedent for each triclausal sentence type. Table 9 reports the percentage of the responses for each proficiency level which establish coreference between the reflexive and the indicated possible antecedents (NP1, NP2, NP3, NP1-2, NP1-3, NP2-3, All, and Neither) for triclausal sentences. These results will be discussed within the following sections which look at the three triclausal sentence types. Proficiency levels are listed by row, and sentence types are listed by column:

Table 8: Raw scores for all groups for triclausal sentence types

		TCNF*	TC1NF	TCF			TCNF	TC1NF	TCF
NP1	prep (31)	23	4	7	NP1-3	prep (31)	7	5	4
	1st year (35)	26	10	9		1st year (35)	12	2	5
	2nd year (36)	14	3	7		2nd year (36)	16	2	5
	3rd year (33)	25	15	26		3rd year (33)	1	0	0
	4th year (35)	14	7	9		4th year (35)	10	3	7
	Native con. (25)	2	0	1		Native con. (25)	5	0	2
	Turkish con. (25)	43	-	-		Turkish con. (25)	12	-	-
NP2	prep (31)	1	32	12	NP2-3	prep (31)	1	4	9
	1st year (35)	6	40	18		1st year (35)	4	5	5
	2nd year (36)	5	35	20		2nd year (36)	3	13	9
	3rd year (33)	5	40	18		3rd year (33)	2	1	2
	4th year (35)	9	31	22		4th year (35)	9	13	8
	Native con. (25)	0	4	0		Native con. (25)	0	3	1
	Turkish con. (25)	37	-	-		Turkish con. (25)	19	-	-
NP3	prep (31)	50	34	55	ALL	prep (31)	7	7	4
	1st year (35)	53	41	62		1st year (35)	0	2	3
	2nd year (36)	63	44	56		2nd year (36)	5	6	8
	3rd year (33)	63	41	51		3rd year (33)	0	0	0
	4th year (35)	49	40	41		4th year (35)	10	6	11
	Native con. (25)	65	68	72		Native con. (25)	3	0	0
	Turkish con. (25)	65	-	-		Turkish con. (25)	12	-	-
NP1-2	prep (31)	2	5	2	Neither	prep (31)	0	0	0
	1st year (35)	4	4	1		1st year (35)	0	0	0
	2nd year (36)	2	5	3		2nd year (36)	0	0	0
	3rd year (33)	1	1	1		3rd year (33)	0	0	0
	4th year (35)	4	5	7		4th year (35)	0	0	0
	Native con. (25)	0	0	0		Native con. (25)	0	0	0
	Turkish con. (25)	10	-	-		Turkish con. (25)	0	-	-

*TCNF = Triclausal, nonfinite
TC1NF = Triclausal, one nonfinite
TCF = Triclausal, finite

Table 9: Percentages for all groups for triclausal sentence types

		TCNF	TC1NF	TCF			TCNF	TC1NF	TCF
NP1	prep (31)	25.27	4.40	7.53	NP1-3	prep (31)	7.69	5.49	4.30
	1st year (35)	24.76	9.62	8.74		1st year (35)	11.43	1.92	4.85
	2nd year (36)	12.96	2.78	6.48		2nd year (36)	14.81	1.85	4.63
	3rd year (33)	25.77	15.31	26.53		3rd year (33)	1.03	0.00	0.00
	4th year (35)	13.33	6.67	8.57		4th year (35)	9.52	2.58	6.67
	Native con. (25)	2.67	0.00	1.33		Native con. (25)	6.67	0.00	2.67
	Turkish con. (25)	21.71	-	-		Turkish con. (25)	6.06	-	-
NP2	prep (31)	1.09	35.16	12.90	NP2-3	prep (31)	1.09	4.40	9.68
	1st year (35)	5.71	38.46	17.48		1st year (35)	3.81	4.81	4.85
	2nd year (36)	4.63	32.41	18.52		2nd year (36)	2.78	12.04	8.33
	3rd year (33)	5.15	40.82	18.37		3rd year (33)	2.06	1.02	2.04
	4th year (35)	8.57	29.52	20.95		4th year (35)	8.57	12.38	7.62
	Native con. (25)	0.00	5.33	0.00		Native con. (25)	0.00	4.00	1.33
	Turkish con. (25)	18.69	-	-		Turkish con. (25)	9.60	-	-
NP3	prep (31)	54.95	37.36	59.14	ALL	prep (31)	7.69	7.69	4.30
	1st year (35)	50.48	39.42	60.19		1st year (35)	0.00	1.92	2.91
	2nd year (36)	58.33	40.74	51.85		2nd year (36)	4.63	5.56	7.41
	3rd year (33)	64.95	41.84	52.04		3rd year (33)	0.00	0.00	0.00
	4th year (35)	46.67	38.10	39.05		4th year (35)	9.52	5.71	10.48
	Native con. (25)	86.67	90.67	96.00		Native con. (25)	4.00	0.00	0.00
	Turkish con. (25)	32.83	-	-		Turkish con. (25)	6.06	-	-
NP1-2	prep (31)	2.20	5.49	2.15	Neither	prep (31)	0.00	0.00	0.00
	1st year (35)	3.81	3.85	0.10		1st year (35)	0.00	0.00	0.00
	2nd year (36)	1.85	4.63	2.78		2nd year (36)	0.00	0.00	0.00
	3rd year (33)	1.03	1.02	1.02		3rd year (33)	0.00	0.00	0.00
	4th year (35)	3.81	4.76	6.67		4th year (35)	0.00	0.00	0.00
	Native con. (25)	0.00	0.00	0.00		Native con. (25)	0.00	0.00	0.00
	Turkish con. (25)	5.05	-	-		Turkish con. (25)	0.00	-	-

5.2.2.1.1. Triclausal nonfinite. If we look at triclausal nonfinite sentences we find that for each level, the majority of responses selected NP3 (only) as the antecedent of the reflexive; for example, for prep year 54.95%; for 1st year 50.48%; for second year 58.33%; for third year 64.95%; and for fourth year 46.67% of responses. The second most selected antecedent for this sentence type and for each level was NP1 (only); for prep year 25.27%; for first year 27.76%; for second year 12.96%; for third year 25.77%; and for fourth year 13.33% of responses. As the results show, the fourth year group had the lowest percentage of correct responses (46.67%), which is unexpected because of their higher proficiency. The third year had the highest percentage of correct responses (64.95%). It appears that the distribution of responses for all the levels was similar in that NP3 was selected most frequently, then NP1, and then the rest of responses were fairly evenly distributed among the remaining choices. In addition, it appears that the responses were most evenly distributed across the remaining choices for the fourth year (e.g., NP2 — 8.57%, NP1-2 — 3.81%, NP1-3 — 9.52%, NP2-3 — 8.57%, All — 9.52%). Using a chi-square test, there was no significant difference between the levels in terms of the number of correct responses they made to triclausal nonfinite sentences (chi-square = 8.41, p =.05, 4 degrees of freedom, $X^2 < 9.49$).

For the triclausal nonfinite sentences, less than half of the responses by the Turkish control group selected NP3 (44.54%). 16.22% selected NP1 and the rest of the responses were fairly evenly distributed between the other choices, in the same way as for the five levels (e.g., NP2/12.16%, NP1-2/2.70%, NP1-3/9.46%, NP2-3/8.11%, All/6.76%). Only 6.76% of the responses by this group selected the 'ALL' choice,

which is the binding setting for Turkish reflexives. For triclausal nonfinite clauses, there appears to be a difference between the Turkish control group and the Turkish learners. Using a chi-square test there was a significant difference between the levels and the Turkish control (chi-square = 27.85, p =.05, 1df, $X^2 < 3.84$). The native control group correctly selected NP3 86.67% of the time. The rest of the responses were mostly divided between NP1-3 (6.67%) and All (4.00%). It is obvious that the native control group was far more successful than the levels in the choice of antecedent. Figure 5 illustrates the results for triclausal nonfinite sentences, with focus on the selection of NP3, NP1, and all other responses together:

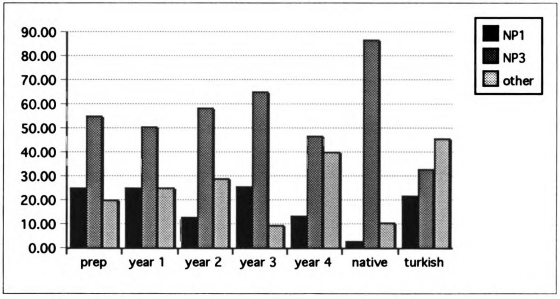


Figure 5: Levels and controls for triclausal nonfinite sentences

5.2.2.1.2. *Triclausal one nonfinite (TC1NF)*. If we look at triclausal one nonfinite sentences we find that for each level, the majority of responses correctly selected NP3 (only) as the antecedent of the reflexive. Clearly however, the percentage of correct scores for this sentence type was much lower for all the levels compared to the triclausal nonfinite sentences. For example, for prep year 37.36% selected NP3; for 1st year, 39.42%; for second year, 40.74%; for third year, 41.84%; and for fourth year, 38.10% of responses. The second most selected antecedent for this sentence type and for each level was NP2 (only); for prep year, 35.16%; for first year, 38.46%; for second year, 32.41%; for third year, 42.82%; and for fourth year, 29.52% of responses. As the results show, all levels produced similar results in terms of the number of correct responses. Using a chi-square test, there was no significant difference between the levels in terms of the number of correct responses they made to triclausal one nonfinite sentences (chi-square = .498, p=.05, 4 df, $X^2 < 9.49$). It appears that the distribution of responses for all the levels was similar in that NP3 was selected most frequently, then NP2, and then the rest of responses were fairly evenly distributed among the remaining choices.

For the triclausal one nonfinite sentences, Turkish control group selected NP2 a majority of the time (32.00%). A similar number selected NP1 (26.00%), and only 6.00% of the responses by this group selected the 'ALL' choice. It should be noted that only 16% of the Turkish control group responses selected NP3, which is the right setting for English. For triclausal one nonfinite clauses, there was no significant difference between the Turkish control and the Turkish learners (chi-square = 2.73,

p=.05, 1df, $X^2 < 3.84$). The native control group was highly successful for this sentence type, selecting the correct NP (3) 90.67% of the time. The gap between the native group and the levels is obvious. Figure 6 illustrates the results for triclausal one nonfinite sentences, with focus on the selection of NP3, NP2, and all other responses together:

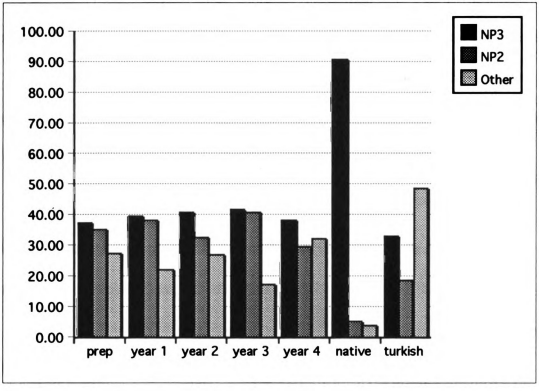


Figure 6: Levels and controls for triclausal one nonfinite sentences

5.2.2.1.3. *Triclausal finite* . As for the other two triclausal sentence types, the majority of responses for triclausal finite sentences selected NP3 (only) as the antecedent of the reflexive. For example, for prep year, 59.14%; for 1st year, 60.19%; for second year,

51.85%; for third year, 52.04%; and for fourth year, 39.05% of responses. The second most selected antecedent for this sentence type and for each level was NP2 (only) (similar to the result for TC1NF sentences); for prep year, 12.90%; for first year, 17.48%; for second year, 18.52%; for third year, 18.37%; and for fourth year, 20.95% of responses. As the results show, the fourth year group had the lowest percentage of correct responses (39.05%), while the first year had the highest (60.19%). It appears that the distribution of responses for all the levels was similar in that NP3 was selected most frequently, then NP2, and then the rest of responses were fairly evenly distributed among the remaining choices. This distribution is similar to that for TC1NF sentences. Using a chi-square test, there was a significant difference between the levels in terms of the number of correct responses they made to triclausal finite sentences (chi-square = 11.67, p =.05, 4df, $X^2 < 9.49$). This result is due in large part to the poor performance of the fourth year group compared to the other groups in terms of the correct choice of antecedent. As it may be seen, the other four levels had similar performances.

For the triclausal finite sentences, less than half of the responses by the Turkish control group selected NP3 (42.43%). 24.32% selected NP1 and the rest of the responses were fairly evenly distributed between the other choices, in the same way as for the five levels (e.g., NP2 — 16.22%, NP1-2 — 5.41%, NP1-3 — 2.00%, NP2-3 — 10.81%, ALL — 5.41%). Only 5.41% of the responses by this group selected the 'ALL' choice, which is the binding setting for Turkish reflexives. For triclausal finite clauses, there was no significant difference between the Turkish control and the fourth

year group (chi-square = .821, p =.05, 1df, X^2< 3.84). On the other hand, there was a significant different between the Turkish control and year prep (chi-square = 11.80, p=.05, 1df, X^2< 3.84), year one (chi-square = 14.33, p=.05, 1df, X^2< 3.84), year two (chi-square = 6.72, p =.05, 1df, X^2< 3.84), and year three (chi-square = 6.60, p =.05, 1df, X^2< 3.84). The native control group correctly selected NP3 96.00% of the time. It is evident that the native control group was far more successful that the levels or the Turkish control group in the choice of antecedent. Figure 7 illustrates the results for triclausal finite sentences:

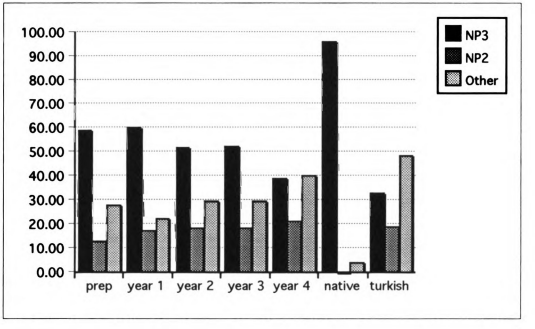


Figure 7: Levels and controls for triclausal finite sentences

5.2.2.1.4. Mean scores. Table 10 gives the mean scores for each level (based on the number of correct responses) and for each triclausal sentence type as well as for the three sentence types combined:

Table 10: Mean scores for triclausal sentences

	TC n=9	TCF n=3	TC1NF n=3	TCNF n=3
prep (31)	4.29	1.81	0.97	1.55
1st year (35)	4.54	1.82	1.23	1.51
2nd year (36)	4.42	1.53	1.14	1.77
3rd year (33)	4.38	1.52	1.18	1.91
4th year (35)	3.43	0.97	1.11	1.34
native cont (25)	8.69	2.84	2.8	2.92

5.2.2.2. Comparing sentence types. By conflating the four levels and comparing the three triclausal sentence types, we find that subjects appear to be more successful in selecting the correct antecedent for the triclausal finite and nonfinite types than for the triclausal one nonfinite type (TCNF = 54.94%, TCF = 52.27%, TC1NF = 39.53%). Table 11 gives the results for the three triclausal sentence types in both raw numbers and percentages:

Table 11: Combined results for triclausal sentences

		TCNF	TC1NF	TCF
	raw			
NP1		102	39	58
NP2		26	178	90
NP3		278	200	265
NP1-2		13	20	14
NP1-3		46	12	21
NP2-3		19	36	33
ALL		22	21	26
Neither		0	0	0
	TOTAL	506	506	507
		TCNF	TC1NF	TCF
	percent			
NP1		20.16	7.71	11.44
NP2		5.14	35.18	17.75
NP3		54.94	39.53	52.27
NP1-2		2.57	3.95	2.76
NP1-3		9.09	2.37	4.14
NP2-3		3.75	7.11	6.51
ALL		4.35	4.15	5.13

Figure 8 illustrates these results for the three sentence types:

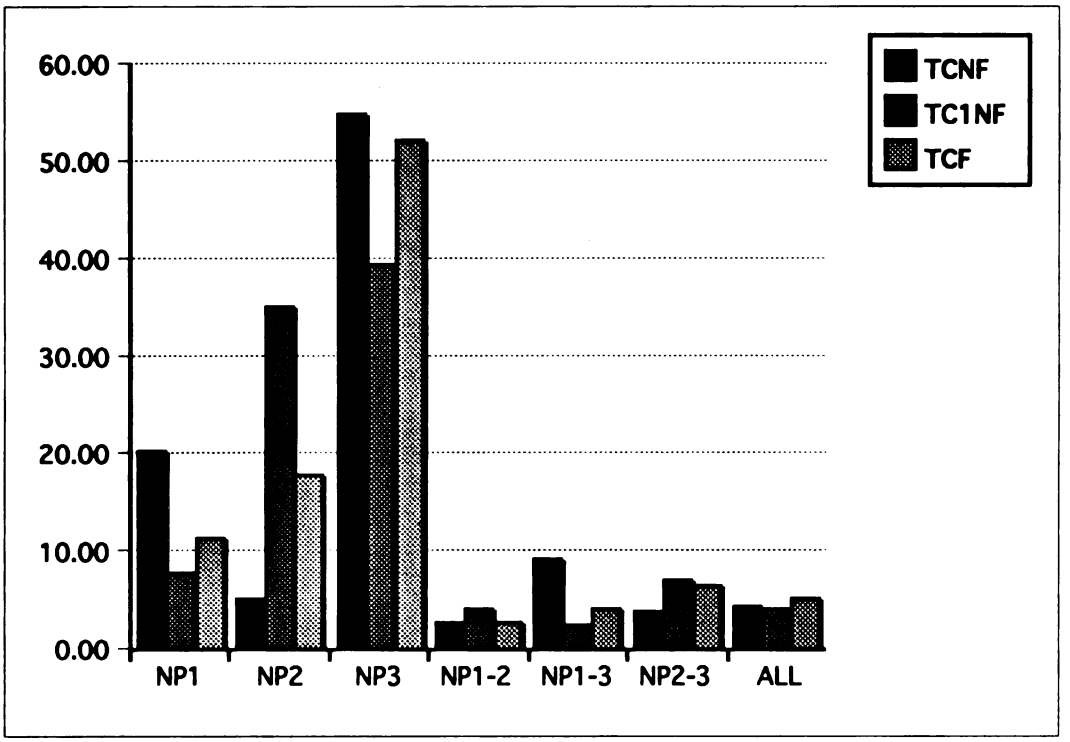


Figure 8: Responses for triclausal types with levels combined

There was no significant difference between TCNF and TCF in terms of correct choice of antecedent (chi-square = 2.24, p =.05, df1, $X^2 < 3.84$) for all proficiency levels combined. However, there is a significant difference between TCNF and TCF combined and TC1NF ($Z = 3.455$, $\alpha = .05$). Figure 9 illustrates the results for TCNF and TCF combined, with levels conflated (since there is no difference between levels for these two types:

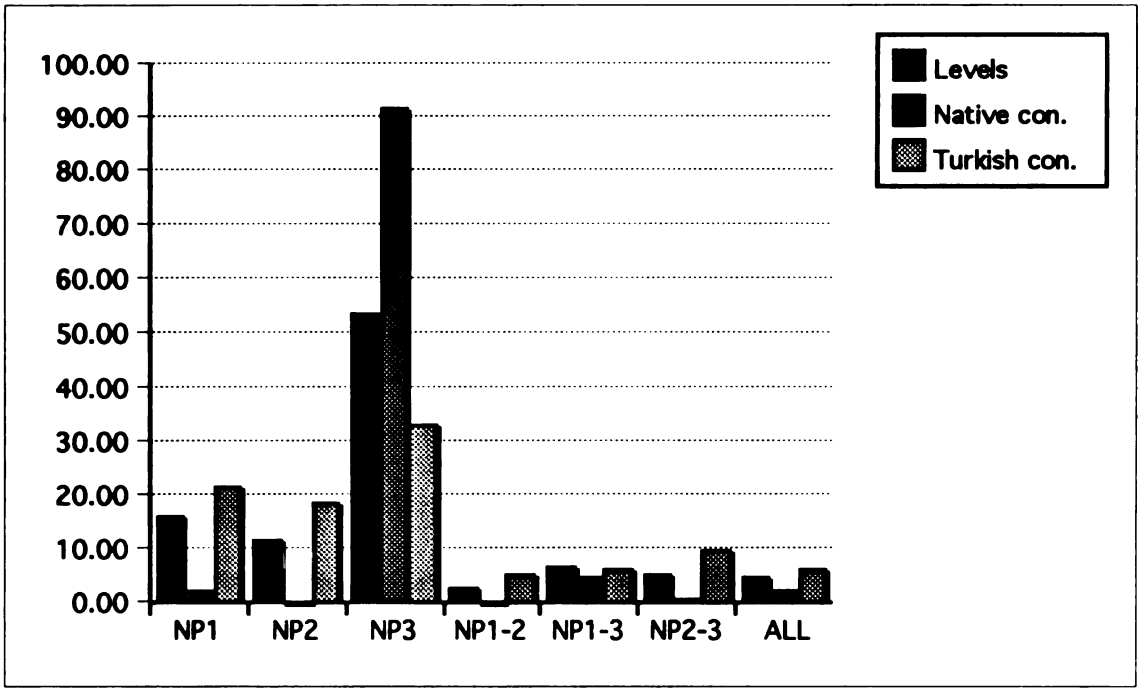


Figure 9: Results for TCNF and TCF combined.

5.3. Summary of group results.

a. Proficiency levels. There were no significant differences between the five proficiency levels for any sentence type, except for the triclausal sentences with one nonfinite clause.

b. Pragmatically biased sentences. In both NP1 and NP2 biased sentences, the L2 learners most of the time selected the NP which is pragmatically biased in both finite and nonfinite sentences. There was no difference between finite and nonfinite pragmatically biased sentences. The L2 learners treated neutral sentences significantly differently than the biased sentences. Finally, the L2 learners differed from both the native English and the native Turkish controls in responding to the pragmatically biased

sentences.

c. Neutral sentences. L2 learners allowed long-distance binding in both finite and nonfinite clauses. However, their percentage of long-distance binding for neutral sentences was lower overall than their percentage of local binding. Moreover, their percentage of long-distance was lower than that for the Turkish control, but higher than that of the native English control group. Thus, the L2 learners differed from both control groups. In addition, there was no significant difference between neutral finite and neutral nonfinite sentences in terms of the correct choice of antecedent. Finally, learners made significantly more nonlocal binding errors in triclausal sentences than they did in biclausal sentences ($Z = 7.079$, $\alpha = .05$), which means that the level of embedding was more influential than the finite/nonfinite distinction in selecting the correct NP.

d. Triclausal sentences. L2 learners allowed long-distance binding in the three triclausal sentence types. They scored worst on the triclausal sentences with one nonfinite clause, which was significantly different from the other two types. There was no significant difference between the triclausal finite and the triclausal nonfinite sentences. The L2 learners behaved differently from both the native control and the Turkish control group, with the exception of triclausal nonfinite sentences where the learners behaved the same as the Turkish control. Also, for triclausal finite sentences, both only fourth year learners did not behave significantly different than the Turkish control group.

5.4. Individual results.

As was discussed in chapter 3 on methodology, one important reason for examining individual results is that group results may be misleading. That is, it is important to show that it is not the case that group results are derived from some members getting everything right whereas others are making many errors. Examining individual results enables the researcher to confirm or disconfirm general conclusions drawn from group results and to provide evidence at the individual level that some UG principle being tested is adhered to.

5.4.1. Biclausal results for individuals. For the individual results, the response patterns of each individual were analyzed. First, each of the subjects' response patterns were evaluated with regard to each of the six sentence types separately. Since each type has three tokens, a subject could get 3/3 correct if he selected the local NP consistently for each type. After ascertaining subjects' response patterns on each sentence type, the results of finite and nonfinite were combined because subjects did not show any sign of tensed-infinitival asymmetry. As a result, the subjects' correct responses were evaluated for three types; biclausal neutral, NP1 biased, and NP2 biased. Subjects were evaluated twice, using two different criteria; 100% (or 6/6 correct responses), and 88.33% (or 5/6 correct responses). Following Eckman (1994) and MacLaughlin (1995), we can use these criteria to determine whether or not individuals have acquired the binding system of the target language, where one criterion is less strict than the other. Table 12 shows the individual results using the 100% criterion, and Table 13

show the results using the 88.33% criterion. In each table, the number of subjects from each group who satisfied the criterion is listed for biclausal neutral, NP1 biased and NP2 biased sentences. Table 12 is followed by a graph in Figure 10 which represents its results:

Table 12: Number of individuals with 6/6 correct (100% criterion) by sentence type:

Levels (n=170)	BCNeutral	NP2 biased	NP1 biased
prep (n=31)	6	16	2
year 1 (n=35)	12	21	1
year 2 (n=36)	6	24	2
year 3 (n=33)	5	19	0
year 4 (n=35)	7	19	0
TOTAL	**36 (21.2%)**	**99 (58%)**	**5 (3%)**

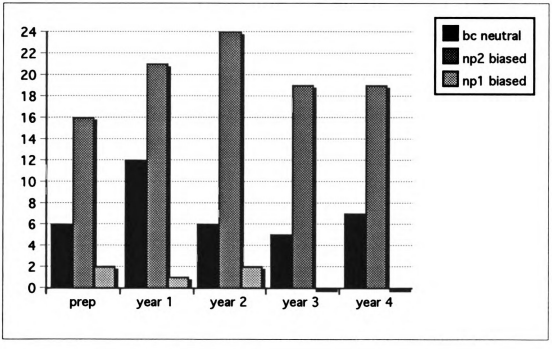


Figure 10: Number of individuals with 6/6 correct (100% criterion) by sentence type.

Table 13: Number of subjects with 5/6 correct (%83.33 criterion) by sentence type:

Levels (n=170)	BCNeutral	NP2 biased	NP1 biased
prep (n=31)	17	24	2
year 1 (n=35)	19	27	1
year 2 (n=36)	14	29	2
year 3 (n=33)	11	30	0
year 4 (n=35)	18	30	0
TOTAL	**79 (46%)**	**140 (82%)**	**5 (3%)**

5.4.1.1. *Neutral sentences* . As Table 12 shows, using the 100% criterion, only 36 subjects out of 170 chose NP2 on every occasion for biclausal neutral sentences. For this type (BCN) year 1 subjects were most successful because this group had the highest number of individuals (12) who scored 6/6 correct, while other groups had almost the same results in terms of the number of subjects scoring 6/6. The overall results for BCN sentences show that 36 subjects (20%) acquired a native like grammar for reflexive binding; in other words, 36 subjects were able to reset the parameter setting to the English value in terms of biclausal neutral sentences.

5.4.1.2. *NP2 biased sentences* . The analysis of individual response patterns for NP2 biased sentences using the 100% criterion indicates that 99 subjects (58%) chose NP2 on every occasion. The second year group had the most individuals (24) who scored perfectly on this sentence type, while other groups had a similar number of individuals scoring 6/6. It should be noted that for this type of sentence the number of subjects who selected NP2 as antecedent for all tokens doubled, tripled or even quadrupled for

each group compared to number of subjects who selected NP2 in all cases for biclausal neutral clauses. For example, 6 subjects from year 2 selected NP2 in all cases for neutral clauses, whereas 24 subjects from the same year selected NP2 in all cases for NP2 biased sentences. These results are important in showing how strong the pragmatic bias can be for individual learners in determining the antecedent of a reflexive.

5.4.1.3. *NP1 biased sentences* . The analysis of individual response patterns for NP1 biased sentences using the 100% criterion shows that there are only 5 subjects (3%) out of 170 who selected NP2 on every occasion. As table 12 shows, there was no individual from years 3 and 4 who selected NP2 on 6/6 occasions. It should be noted that 7 subjects from year 4 selected NP2 in all cases for neutral clauses, whereas none of these subjects were able to choose NP2 in every case in NP1 biased clauses. These results are crucial in showing how strong the pragmatic bias can be for learners in pushing them away from the correct antecedent of a reflexive in their L2.

5.4.1.4. *Tracking individuals across sentence types* . When we follow the response patterns for each subject across the three types, we find that only 3 out of the 6 subjects from prep year who scored 6/6 for neutral sentences also scored 6/6 for the NP2 biased sentences. None of these 3 subjects selected NP2 6/6 times for NP1 biased sentences. One subject from prep year who did score 6/6 for NP1 biased sentences was not among those 16 subjects from prep year who scored 6/6 for NP2 biased sentences, but that

same subject was among the 6 who scored perfectly for the neutral sentences.

Nine subjects out of 12 from year 1 who scored 6/6 for the neutral sentences also scored 6/6 for the NP2 biased sentences. One of the these 9 subjects scored 6/6 for NP1 biased sentences. Therefore, this individual scored 18 out of 18 for all biclausal sentences. Six subjects out of 6 for year 2 scored 6/6 for neutral clauses and also scored 6/6 for NP2 biased sentences, whereas none of these 6 subjects were 1 of the 2 subjects who scored 6/6 for NP1 biased. In fact, both of the subjects who scored 6/6 for NP1 biased sentences selected NP2 for neutral clauses 5/6 times, and therefore belong to the group of 14 from year 2 who scored at least 5/6 for neutral clauses. For year 3, 5 of the 5 subjects who scored 6/6 for neutral sentences also scored 6/6 for NP2 biased sentences. For year 4, 5 of the 7 subjects who scored perfectly on neutral sentences also scored perfectly on the NP2 biased sentences.

According to Table 13 (p. 158), which shows the results based on the 83.33% criterion, almost half of the subjects (79/170 or 46%) scored 5/6 correct for neutral clauses, whereas this number jumps to 140/170 or 82% for NP2 biased sentences. However, loosening the criterion to 83.33% does not have any effect on the results for the NP1 biased sentences — still only the same 5 subjects (3%) scored 5/6 correct. Loosening the criterion to 83.33% does not considerably change the results noted above using the 100% criterion. From prep year, of the 17 subjects who scored at least 5/6 for neutral sentences, 15 of them also scored at least 5/6 for NP2 biased sentences. From year 1, 18 out of 19 subjects who scored at least 5/6 for neutral sentences also scored 5/6 for NP2 biased sentences. From year 2, 14/14 subjects who scored at least 5/6 for

neutral sentences also scored at least 5/6 for NP2 biased sentences. From year 3, all 11 subjects who scored at least 5/6 for neutral sentences also scored at least 5/6 for NP2 biased sentences. Finally, from year 4, 16/18 who scored at least 5/5 for neutral sentences also scored at least 5/6 for NP2 biased sentences. Overall, 74 out of 79 subjects (93%) who scored at least 5/6 for neutral sentences also scored 5/6 for NP2 biased sentences.

The results of individual responses to NP2 biased and to NP1 biased sentences prove that pragmatics might be the most important factor in determining the antecedent for the Turkish L2 learners. Although 79 subjects appear to have reset the parameter to the English setting with the 83.33% criterion, only 5 subjects out of the 79 could resist the pragmatic pressure towards a nonlocal NP and thus stick with the English setting.

5.4.1.5. *Sample of individual results for biclausal sentences* . To better understand the distribution of responses for each individual across the 6 biclausal sentence types, consider the following Tables (14-18) which present a sample of 10 subjects from each proficiency level:

Table 14: Sample of 10 subjects from prep year.

	Subject	NP1BNF	NP1BF	NP2BNF	NP2BF	BCNF	BCF
prep	1	0	0	2	2	1	1
	2	0	0	3	3	3	3
	3	1	1	3	2	2	3
	4	0	0	3	3	3	2
	5	2	1	2	2	3	3
	6	0	0	3	3	2	3
	7	0	0	3	3	3	2
	8	1	0	3	3	2	2
	9	3	3	3	3	2	1
	10	1	0	1	1	1	0

Table 15: Sample of 10 subjects from year 1.

	Subject	NP1BNF	NP1BF	NP2BNF	NP2BF	BCNF	BCF
Year 1	1	2	2	3	3	3	3
	2	3	3	3	3	3	3
	3	1	1	0	0	0	0
	4	1	0	2	0	0	2
	5	1	0	3	3	2	3
	6	0	0	3	3	3	3
	7	0	0	3	3	2	2
	8	0	0	3	3	0	2
	9	0	0	3	3	2	2
	10	0	0	2	3	2	3

Table 16: Sample of 10 subjects from year 2.

	Subject	NP1BNF	NP1BF	NP2BNF	NP2BF	BCNF	BCF
Year 2	1	0	0	2	2	1	1
	2	1	0	3	3	1	0
	3	0	0	3	3	2	2
	4	2	0	3	3	3	2
	5	1	2	3	3	3	3
	6	0	0	3	3	3	3
	7	1	1	3	3	2	2
	8	2	2	3	3	3	3
	9	0	0	3	3	3	3
	10	0	0	3	2	2	2

Table 17: Sample of 10 subjects from year 3.

	Subject	NP1BNF	NP1BF	NP2BNF	NP2BF	BCNF	BCF
Year 3	1	0	0	1	0	1	0
	2	0	1	3	0	1	0
	3	0	0	2	1	3	1
	4	2	2	3	3	3	3
	5	0	0	3	3	2	3
	6	0	0	3	3	3	3
	7	0	0	3	3	3	3
	8	0	0	3	2	3	2
	9	1	0	3	3	3	3
	10	0	0	3	2	2	2

Table 18: Sample of 11 subjects from year 4.

	Subject	NP1BNF	NP1BF	NP2BNF	NP2BF	BCNF	BCF
Year 4	1	1	0	3	3	3	3
	2	0	0	3	3	3	2
	3	0	0	2	3	3	3
	4	1	0	2	3	0	1
	5	0	0	2	2	1	2
	6	0	0	1	2	0	0
	7	0	0	3	3	3	3
	8	0	0	3	3	1	3
	9	0	0	3	3	2	3
	10	0	0	2	3	2	2
	11	1	0	2	1	2	3

If we look at sample of individual results of ten subjects from each proficiency level in Tables 14-18, we can see that these results illustrate a regular distribution of correct responses with the most correct responses for the NP2 biased sentences, the fewest correct responses for the NP1 biased sentences, and the neutral sentences falling somewhere in between. Certainly some subjects are more successful in binding reflexives in their L2 than others, while very few subjects show irregular patterns in terms of the distribution of correct responses.

Consider subjects 1-8 for prep year in Table 14. The results for these subjects show a regular distribution with all subjects scoring best for NP2 biased sentences, worst for the NP1 biased sentences, and somewhere in between for the neutral sentences. Subjects 9 and 10 appear to be exceptions. Subject 9 is one of the few subjects from any group who scored 6/6 for NP1 biased sentences. Subject 10 had

great difficulty choosing the correct NP for any sentence type.

In Table 15, which gives a sample of results for year 1, we see the same regular pattern with subjects 5-10. On the other hand, subjects 1 and 2 from this group did very well, even for the NP1 biased sentences, whereas subjects 3 and 4 did very poorly on all biclausal sentence types.

Table 16 shows a sample of individual results for year 2 in which we again find the same regular distribution. One point to arise from these results is that although some subjects may score higher than others across the sentence types, the same pattern holds for each individual; scoring highest for the NP2 biased sentences, lowest for NP1 biased sentences, and in between for the neutral sentences. For example, compare subjects 1 and 7 from year 2 in Table 19:

Table 19: Comparison of two subjects from year 2.

Subject	NP1BNF	NP1BF	NP2BNF	NP2BF	BCNF	BCF
1	0	0	2	2	1	1
7	1	1	3	3	2	2

Here we see that although subject 1 consistently scored one less correct than subject 7 for each sentence type, both subjects scored worse for NP1 biased sentences, best for NP2 biased sentences and in between for neutral sentences.

Table 17 above shows a sample of results for individuals from year 3. Subjects

1 and 2 show an irregular distribution of correct responses; neither of these subjects were particularly successful in choosing the correct NP. Subject 4, on the other hand, was very successful, scoring 2/3 for both NP1 biased sentence types. All other subjects illustrate a regular distribution of correct responses.

For the sample in Table 18 for year 4, we find the same regular distribution for most subjects. Subject 4 is exceptional in that he or she scored as expected for the NP1 biased and NP2 biased sentences, but was very unsuccessful for the neutral sentences. Subject 6 scored poorly across all sentence types.

These samples show that overall there is a regular distribution of correct responses with most subjects scoring best for NP2 biased sentences, worst for NP1 biased sentences, and neutral sentences falling somewhere in between. Furthermore, this regular pattern holds across the different proficiency levels. There are, however, a small number of exceptional cases where the response patterns of individual subjects do not fit the overall pattern. In these cases, subjects may do exceptionally well for all sentence types, or exceptionally poorly, or else they may have some difficulty with a particular sentence type (e.g., NP2 biased or neutral), or else unusual success with some sentence type (e.g., NP1 biased).

5.4.2. Triclausal results for individuals. For the individual results for triclausal sentences, the response patterns of each individual were analyzed. First, each of the subjects' response patterns were evaluated with regard to each of the three sentence types separately (triclausal finite, triclausal one nonfinite and triclausal nonfinite).

Since each type has three tokens, a subject could get 3/3 correct if he or she selected the local NP consistently for each type. After ascertaining subjects' response patterns on each sentence type, the results of triclausal finite and nonfinite were combined because subjects did not show any sign of tensed-infinitival asymmetry. However, the triclausal one nonfinite sentence type was not combined with the others because there was a significant difference between this type and the other two with regard to the correct choice of antecedent. Finally, the results for all three types were combined. As a result, the subjects' correct responses were evaluated for five categories; triclausal finite, triclausal nonfinite, triclausal one nonfinite, triclausal finite and nonfinite combined, and all three triclausal types combined. For the three types separately, subjects were evaluated twice, using two different criteria; 100% (or 3/3 correct responses), and 66.67% (or 2/3 correct responses). Given the fact that 2/3 is the next available choice, the low 66.67% criterion was the only one possible. It is provided here for the sake of exposition. For the triclausal finite and nonfinite types combined the two criteria were 100% (6/6) and 83.33% (5/6). For all types combined the criteria were 100% (9/9) and 89% (8/9). In addition, a slightly lower criterion of 77% (7/9) was also applied to better illustrate the distribution of individual responses patterns for the triclausal sentences. Table 20 provides these results using the 100% criterion, and Table 21 gives the results using the other criteria just mentioned. Table 20 is followed by a graph in Figure 11 which represents its results:

Table 20: Individual results by level for triclausal types using 100% criteria.

levels	TCF 3/3	TCNF 3/3	TC1NF 3/3	TCF/TCNF 6/6	All 9/9
prep	8	5	2	2	1
year 1	8	5	4	2	2
year 2	10	12	7	6	4
year 3	4	8	3	2	2
year 4	2	4	3	0	0
Total	32 (18.82%)	34 (20%)	19 (11.18%)	12 (7.1%)	9 (5.3)

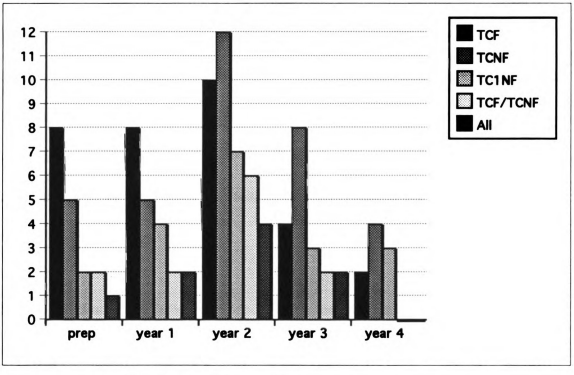


Figure 11: Individual results by level for triclausal types using 100% criteria.

Table 21: Individual results by level for triclausal types using lower criteria.

	TCF 2/3	TCNF 2/3	TC1NF 2/3	TCF/TCNF 5/6	All 8/9	All 7/9
prep	19	15	11	6	2	7
year 1	24	19	15	7	2	5
year 2	18	18	12	11	4	6
year 3	17	24	11	6	2	3
year 4	5	16	13	3	0	2
Total	83	92	62	33	10	23
percent	48.82%	54.12%	36.47%	19.41%	5.88%	13.53%

5.4.2.1. *Triclausal finite sentences*. If we look at the results of individual responses for triclausal finite sentences, we see that the year 4 subjects did the worst; only 2 out of 35 selected NP3 on each occasion. Year 3 subjects closely followed with only 4 out of 33 selecting NP3 on every occasion. However, subjects from other levels showed a better performance, such as 10/36 for year 2, and 8 subjects from both prep and year 1 selected NP3 in all three cases. Overall, 32 subjects (18.82%) out of 170 scored 3/3 correct for this sentence type.

5.4.2.2. *Triclausal nonfinite sentences*. The responses for triclausal nonfinite sentences show that the subjects from years 2, 3 and 4 performed better than they did for the triclausal finite clauses. Again, there were more subjects from year 2 (12 subjects) than any other group who scored 3/3 for the triclausal nonfinite sentences. Overall, 34 subjects out of 170 (20%) scored 3/3 for this sentence type, a very similar

result to the one for finite sentences.

5.4.2.3. *Triclausal one nonfinite sentences* . Responses for triclausal one nonfinite sentences reveal that far fewer subjects from each level scored 3/3 correct than for the other two sentence types. For example, only 2 subjects from prep year scored 3/3 for one nonfinite sentences compared to 8 for finite sentences. Again, 7 subjects from year 2 scored 3/3 correct on this sentence type, more than for any other year. Years 3 and 4 had only 3 subjects each who scored 3/3 correctly. Overall, only 19 out of 170 subjects (11.18%) scored 3.3 for this sentence type, much worse than for the other triclausal sentence types.

5.4.2.4. *TCNF and TCF combined* . If we look at individual results for the triclausal nonfinite and triclausal finite sentences combined, we see that as usual more subjects for year 2 (6 subjects) selected the correct NP on every occasion (6/6) than for any group, whereas year 4 had the fewest subjects (0). All other levels also had few subjects who scored 6/6 for this category. Overall, only 12 out of 170 subjects (7.1%) scored 6/6 for triclausal nonfinite and triclausal finite sentences combined.

5.4.2.5. *All triclausal types combined* . If we look at the individual results for all three triclausal sentence types together, we see that 9 subjects overall out of 170 (5.3%) selected NP3 for all 9 occasions; one subject from prep year, 2 subjects from year 1, 4 from year 2, 2 from year 3 and none from year 4. We may claim that these 9 subjects

have been able to reset the parameter setting to the English value for triclausal sentences. These 9 subjects have not been effected by the multiple levels of embedding and by nonfiniteness. They were able to parse complex sentences and analyze the reflexive which is buried inside the most embedded clause and were able to find the correct antecedent within the same clause as the reflexive. The performance of these subjects reveals that in their L2 they have acquired the complex sentences in which the target test item operated.

5.4.2.6. *Using the 2/3 criterion for triclausal types* . It is apparent from table 16 that using the very loose 2/3 criterion for triclausal finite sentences, year 4 subjects lagged far behind subjects from other groups with only 5 subjects to score at least 2/3. The difference between year 4 and year 1 which had 24 subjects who satisfied the 2/3 criterion is extraordinary. For triclausal nonfinite sentences the gap between year 4 and the other levels disappears; all groups had roughly the same number of subjects to satisfy the 66.67% criterion. For triclausal one nonfinite sentences, the overall number of subjects to satisfy the 66.67% criterion decreases to 62 subjects out of 170 (36.47%). The difference between the levels for this sentence type is minimal.

5.4.2.7. *Using the 5/6 criterion for TCF and TCNF combined* . If we look at the individual results in Table 21 for triclausal finite and triclausal nonfinite combined using the 83.33% criterion (5/6), we see that only 33 out of 170 (19.41%) selected the correct NP on at least 5 out of 6 occasions. Except for year 4 which had only 3 subjects

who scored at least 5 out 6, the number of individuals who scored at least 5 out of 6 correct appears to be similar for all levels. Again, year 2 boasted the most subjects (11) to satisfy the 83.33% criterion.

5.4.2.8. *Using the 8/9 criterion for all triclausal sentences combined*. Table 21 shows that for all the triclausal sentences combined using the 88.89% criterion (8/9), there is little difference from the 100% criterion — only 10 out of 170 (5.88%) subjects scored 8/9 for this category. If we consider a criterion of 77.78% (7/9), we see that the number of subjects satisfying this criterion increases to 23 out of 170 (13.53%). Here only 2 subjects from year 4 scored at least 7/9, whereas slightly more subjects from other levels satisfied this criterion (e.g., 7 for prep year, 6 for year 2).

5.4.2.9. *Sample of individual results for triclausal sentences*. Consider the following Table (22) which presents a sample of 10 subjects from each proficiency level for each of the three triclausal sentence types:

Table 22: Sample of 10 subjects from each level for triclausal types.

	subject	tcF	tc1NF	tcNF		subject	tcF	tc1NF	tcNF
prep	1	3	1	3	year 3	1	2	1	2
	2	3	0	2		2	2	1	1
	3	2	0	3		3	1	1	2
	4	3	0	1		4	2	1	2
	5	1	0	1		5	2	0	3
	6	0	0	0		6	2	2	2
	7	1	0	2		7	2	1	3
	8	3	0	3		8	2	1	2
	9	3	3	2		9	1	2	3
	10	2	2	2		10	3	3	3
year 1	1	3	3	3	year 4	1	0	0	0
	2	1	0	0		2	1	2	2
	3	0	0	1		3	1	0	1
	4	3	3	3		4	2	1	3
	5	2	0	2		5	2	1	2
	6	2	1	2		6	3	1	2
	7	3	1	2		7	1	1	1
	8	2	2	2		8	1	0	1
	9	1	0	1		9	1	3	1
	10	3	2	2		10	1	2	2
year 2	1	3	3	3					
	2	3	1	2					
	3	2	3	1					
	4	1	0	1					
	5	3	1	3					
	6	3	3	3					
	7	1	0	1					
	8	3	1	3					
	9	0	0	1					
	10	3	3	3					

Comparing the individual scores across the three triclausal sentence types, a pattern emerges, although not as distinctive as the one for the biclausal sentences. Here we find that for most subjects, scores for the triclausal finite and triclausal nonfinite sentences are similar and higher than the scores for the triclausal one nonfinite sentences. For example, subjects 1-5 for prep year all scored lower for the triclausal one nonfinite sentences compared to the other two types. Moreover, there are a few exceptions to this pattern. For example, subject 9 from year 4 and subject 3 from year 2 both scored higher for the triclausal one nonfinite sentences compared to the other two types.

As with the biclausal sentences, there are some subjects who deviate from this pattern by either scoring well for all sentence types, or by scoring poorly for all three triclausal sentence types. For example, subjects 1 and 4 from year 1, subjects 1 and 6 from year two, and subject 10 from year 3 all scored 3/3 for all three types. Subjects 5 and 6 from prep year, subjects 2, 3 and 9 from year 1, and subject 9 from year 2 all scored very poorly for the three types.

These samples also illustrate that there is more variation among subjects for triclausal sentences compared to biclausal sentences. That is, unlike for the biclausal sentences, the performance of subjects on the triclausal sentences is less predictable. Another observation is that there is no obvious difference in the response patterns of individuals from different proficiency levels.

5.4.3. Comparison of biclausal and triclausal sentences. Now consider Table 23 which shows the individual results for all biclausal sentence types combined as well as for all biclausal and triclausal sentence types combined using several different criteria:

Table 23: Individual results for biclausal and biclausal/triclausal combined.

levels	BC 18/18	BC 17/18	BC 16/18	BC 15/18	BC + TC 27/27	BC + TC 26/27	BC + TC 25/27	BC + TC 24/27	BC + TC 23/27
prep	0	0	0	2	0	0	0	0	0
year 1	1	1	2	2	1	1	2	2	2
year 2	0	2	3	5	0	2	2	3	4
year 3	0	0	1	1	0	0	1	1	1
year 4	0	0	0	0	0	0	0	0	0
Total	1	3	6	10	1	3	5	6	7

If we look at the individual results for all biclausal sentence types in Table 23, using the 100% criterion (18/18), we can see that only one subject (from year 1) out of 170 selected the correct NP in all 18 sentences. In fact, the same subject was the only one to select the correct NP in all 27 test sentences. Using a criterion of 17/18 (94.44%), only 3 subjects (2 of them from year 2) scored at least 17/18. In fact these 2 subjects from year 2 selected the correct NP in 26 out of 27 test sentences. Using a criterion of 16/18 (88.89%), only six subjects out of 170 satisfied this criterion for all biclausal sentences combined, and with 15/18 (83.33%) only 10 subjects overall satisfied this criterion. If we examine all triclausal and biclausal sentences combined, it is apparent from table 17 that

very few subjects (7) satisfy even an 85% criterion (23/27).

5.4.4. Comparison of biclausal neutral and triclausal finite and nonfinite sentences. If we compare the individual results of biclausal finite and nonfinite neutral sentences with triclausal finite and nonfinite sentences, we can conclude that more subjects performed well in biclausal sentences than in triclausal sentences. For example, six subjects from prep year scored 6/6 correct for biclausal neutral sentences, whereas only two subjects scored 6/6 for combined triclausal finite and nonfinite sentences. Twelve subjects from year 1 scored 6/6 correct for biclausal sentences, whereas only two (of them) scored 6/6 correct for the triclausal sentences. The same is true for other levels except for year 2 which had six subjects who scored 6/6 for biclausal neutral sentences, and six subjects who scored 6/6 correct for triclausal finite and nonfinite sentences combined. Two of the subjects who scored 6/6 for the triclausal sentences also scored 6/6 for the biclausal sentences, and the remaining four satisfied the 83.33% (5/6) criterion for the biclausal sentences. Overall, for all neutral biclausal and triclausal sentences types, the fewest number of subjects scored well for triclausal one nonfinite sentences. In sum, it can be claimed that the complexity of the sentence type has some effect on the performance of individual subjects at any proficiency level; the more embedding, the fewer subjects to satisfy the various criteria applied. It appears that more subjects have greater success in selecting the correct antecedent for (neutral) biclausal sentences in comparison to triclausal sentences.

5.4.5. Finite/nonfinite distinction in biclausal sentences. In order to find out whether individual subjects made a finite/nonfinite distinction in biclausal sentences, as the subjects did in the studies by Finer and Broselow (1986), Hirakawa 1990), Finer (1991), and MacLaughlin (1995), each subject in the study was examined. The criterion that was used here was that if there was more than a one point error difference between finite and nonfinite sentences of the same type (NP1 biased, NP2 biased, neutral), then the subjects were presumed to make a finite/nonfinite distinction. For example, if subject A scored 3/3 for NP2 biased finite sentences and 1/3 for NP2 biased nonfinite sentences, then that subject was considered to make a finite/nonfinite distinction.

Only 4 subjects from year 4 made such a distinction in biclausal neutral sentences; 3 of the subjects scored 1/3 correct for the nonfinite neutral sentences whereas they scored 3/3 for the finite neutral sentences. Another subject from year 4 scored 2/3 for the nonfinite neutral sentences, whereas he or she scored 0/3 for the finite ones. (That is, he or she did better in nonfinite clauses.) The same four individuals who made a distinction in biclausal neutral sentences did not make such a distinction in the NP1 biased and the NP2 biased sentences.

Only three subjects from year 3 made a finite/nonfinite distinction for neutral biclausal sentences, using the same criterion. All three scored 3/3 for nonfinite neutral sentences, but scored 1/3 for finite neutral sentences. For NP2 biased sentences, eight subjects from year 3 made a finite/nonfinite distinction. Seven of them scored 3/3 for the nonfinite sentences, whereas they scored 1/3 for finite sentences. (That is, they did better in nonfinite clauses.) The remaining subject scored 3/3 nonfinite sentences and 0/3 for

finite sentences. One of these eight subjects from year 3 also was 1 of the 3 subjects to make a finite/nonfinite distinction in the neutral sentences.

Only one subject from year 2 made a finite/nonfinite distinction for neutral biclausal sentences, and this subject was 1 of 2 subjects who made a distinction for NP2 biased sentences (scoring better for the finite sentences). Only 4 subjects from year 1 made a distinction in neutral biclausal sentences, scoring better in the finite sentences, whereas in NP2 biased sentences, 1 subject made a finite/nonfinite distinction, doing better in the nonfinite sentences. Only 2 subjects from prep year made a finite/nonfinite distinction for neutral sentences, scoring better in nonfinite sentences. One subject made a finite/nonfinite distinction for NP1 biased sentences, scoring better for nonfinite sentences.

In sum, these results show that with regard to the finite/nonfinite distinction at the individual level, using the criterion mentioned above, very few individuals made such a distinction, and those who did so made a distinction for only one type (neutral, NP1 biased or NP2 biased). This indicates that at the individual level a finite/nonfinite distinction is not being made by the learners.

5.4.6. Finite/nonfinite distinction in triclausal sentences. In order to find out whether individual subjects made a finite/nonfinite/one nonfinite distinction in triclausal sentences, each subject in the study was examined. Since there were three sentence types, three two-way comparisons were made. The criterion that was used here was that if there was more than a one point error difference between the types being compared,

then the subjects were presumed to make a distinction with regard to those types. For example if a subject scored 1/3 for triclausal finite sentences and 3/3 for triclausal nonfinite sentences, then that subject was deemed to make a distinction.

First let's compare the triclausal finite with the triclausal nonfinite sentences. According to our criterion, for year 4, only 4 out of 33 subjects made a distinction between these types. Three of them scored better on the nonfinite sentences and 1 scored better on the finite sentences. For year 3, only 4 subjects made a distinction, and all of them scored better on the triclausal nonfinite sentences. For year 2, five subjects made a distinction between triclausal finite and triclausal nonfinite; 4 of them scored better of nonfinite sentences and 1 scored better on finite sentences. For year 1, just 3 subjects made a distinction; 2 of them scored better for finite sentences, and 1 scored better for nonfinite sentences. Finally, for prep year, only 4 subjects made a distinction and all of them scored better for finite sentences. Overall, only 20 out of 170 subjects (11.76%) made a distinction between triclausal finite and triclausal nonfinite sentences using the criterion mentioned above. Twelve of the 20 subjects scored better on the nonfinite sentences.

Now compare triclausal finite and triclausal one nonfinite sentences using the same criterion. For year 4, five subjects made a distinction between triclausal finite and triclausal nonfinite sentences, and 4 out of those 5 scored better on the one nonfinite sentences. Of the year 3 subjects, there was only 1 subject to make a distinction, scoring better on the triclausal finite sentence. For year 2, there were 5 subjects to make a distinction; 3 of the 5 scored better on the finite sentences and 2 of them scored better on

the one nonfinite sentences. For year 1, there were 5 subjects to make a distinction, with 4 of them scoring better on the finite sentences. Finally, for prep year, there were 7 subjects to make a distinction, and all 7 scored better in the finite sentences. Overall, there were only 23 subjects out of 170 (13.53%) who made a distinction between triclausal finite and triclausal one nonfinite sentence given the criterion employed. Sixteen of the 23 subjects scored better on the triclausal finite sentences.

Last of all, compare the triclausal nonfinite sentences with the triclausal one nonfinite sentences. Using the same criterion, year 4 had 6 subjects to make a distinction and 4 of them scored better in the nonfinite sentences. For year 3, seven subjects made a distinction and all of them scored better on the nonfinite sentences. As for year 2, eleven subjects made a distinction; 9 of the 11 subjects scored better for the nonfinite sentences. For year 1, eight subjects made a distinction and 5 of them scored better on the nonfinite sentences. Finally, for prep year, 5 subjects made a distinction and all of them scored better in the nonfinite sentences. Overall, 37 subjects out of 170 (21.76%) made a distinction between triclausal nonfinite and triclausal one nonfinite sentences given the criterion mentioned earlier. Of those 37 subjects, 30 scored better on the triclausal nonfinite sentences.

In sum, given the criterion employed, the overall percentage of subjects who are making a distinction between the triclausal sentence types is fairly low. The fewest number of subjects (20/170 or 11.76%) made a distinction between triclausal finite and triclausal nonfinite sentences. Closely following were the 23 subjects out of 170 (13.53%) who made a distinction between triclausal finite and triclausal one nonfinite

sentences. The number of subjects to make a distinction between triclausal nonfinite sentences and triclausal one nonfinite sentences was larger but still rather low (37/170 or 21.76%).

Thus, it appears that individuals are not treating the triclausal sentences in very distinct ways, although when they do score better for one sentence type compared to another, they usually do so in a predictable way; 30 of the 37 subjects (82%) who made a distinction between triclausal nonfinite and triclausal one nonfinite sentences scored better on the nonfinite sentences. Sixteen of the 23 subjects (70%) who made a distinction between triclausal finite and triclausal one nonfinite sentences scored better on the finite sentences. Of the 20 subjects who made a distinction between triclausal finite and nonfinite sentences, 12 scored better on the nonfinite sentences. These results reflect the fact that overall subjects scored more poorly on the triclausal one nonfinite sentences compared to the other two types, and that the difference between the other two types was limited.

5.5. Summary of Individual results.

Some L2 learners have acquired a native-like binding system with regard to biclausal neutral sentences. Most of the learners have acquired a native-like binding system with regard to NP2 biased sentences. Extremely few learners have acquired a native-like binding system with regard to NP1 biased sentences. Very few learners acquired a native like binding system with regard to all biclausal sentences. Very few acquired a native-like grammar for triclausal sentences, and very few learners acquired a

native-like binding system for all sentences combined. With regard to the distribution of individual responses across the sentence types, we found that a large majority of subjects from all five proficiency levels showed a similar pattern in terms of the correct choice of antecedent. In addition, the results showed that individuals in general did not make a distinction between finite and nonfinite sentences, although for the triclausal sentences, subjects scored more poorly on the triclausal one nonfinite sentences. Finally, proficiency level did not appear to be an important factor in determining the number of subjects from each group who satisfied some criterion (such as 5/6, 6/6, 8/9, etc.).

5.6. Comparison of individual results with group results.

It was noted in an earlier chapter (pp.98-100) that any presentation of group results should be supplemented by individual results in order to get a clearer picture of the acquisition of particular constructions in a second language. This is because group results may be misleading in various ways and may conceal information about the interlanguage grammars of the individuals within the groups. So not only may individual results provide confirmation of group results, but they also lead to a rethinking of conclusions drawn only from group results. As will be shown below, this was the case with the data of this study. That is, for the most part the individual results are consistent with and lend credibility to the group results. In addition, the individual results in some aspects force us to rethink certain conclusions drawn only from the group results.

Both the group results and the individual results show the same distribution of correct responses. For example, for both sets of results, NP2 was selected most often in

NP2 biased sentences, NP2 was selected least often in NP1 biased sentences and NP2 was the preferred choice in the neutral sentences. In addition, we saw that for both sets of results, there was very little difference between the proficiency levels, and furthermore there was very little difference between finite and nonfinite sentences in terms of the correct choice of antecedent. For triclausal sentences, both the group results and the individual results showed that triclausal finite and triclausal nonfinite sentences were treated in a similar way, whereas subjects were less successful with the triclausal one nonfinite sentences.

On the other hand, it is also the case that the group results concealed certain significant information about the individuals within those groups. For example, the group results for the neutral biclausal sentences showed that 69% of responses correctly selected NP2, which indicates that the L2 learners were not able to reset the parameter to the English setting. However, individual results revealed that 34 subjects (or 20%), using the 100% criterion, and 79 subjects (46%), using the 83.33% criterion, were able to acquire the English binding system. The 69% overall group result for biclausal neutral sentences is the result of subjects compiling many different scores, some of which exhibit the complete acquisition English binding by some L2 learners, and others which show an unsystematic pattern of L2 reflexive binding. Therefore, the 69% group score for neutral biclausal sentences does not reveal anything about the 79 subjects (using the 83.33% criterion) who have apparently acquired the English binding system.

Group results for NP2 biased sentences show that 87.67% of responses by L2 learners selected NP2 as antecedent, while individual results show that 140 subjects

(82%) scored at least 5/6 correctly for these sentences (the 83.33% criterion). Thus, the group results are very similar to the individual results. This similarity might be due to the highly successful performance of most of the subjects in selecting the correct antecedent. Here the group results do not conceal individual differences since most subjects were very successful and very few subjects (17) scored 3/6 or lower for this sentence type.

Group responses for NP1 biased sentences show that 12.33% of responses selected NP2 while individual responses show that only 5 subjects (3%) correctly selected NP2 (using both the 100% and the 83.33% criteria). Here, group results appear much higher than individual results. This difference might be due to the extremely small number of subjects who correctly selected NP2 on at least 5/6 occasions.

Chapter 6

DISCUSSION

6.1. Main outcomes.

The following are the main outcomes of this study, the possibilities of which were forecast in chapter 4:

a. The effect of sentence type: The results did not show any significant differences between nonfinite and finite sentences for both triclausal and biclausal sentences, although subjects had greater difficulty with the triclausal sentences with one nonfinite clause. Furthermore, the results show that learners' interpretations of L2 reflexives were effected by the level of embedding in that they were more successful with the biclausal sentences than the triclausal sentences.

b. The effect of pragmatic bias: Turkish learners have interpreted English reflexives according to pragmatic information, which indicates that they have not been able to make a transition from the predominantly pragmatic mode of their L1 to only the syntactic mode of their L2.

c. The effect of syntactic constraints: Turkish L2 learners did not behave as native Turkish speakers nor as native English speakers do with regard to binding English reflexives. Turkish learners of English allowed long-distance binding, and apparently they

were not able to reset the parameter value to the English setting over time.

d. The effect of proficiency level; There was no tendency towards more accurate interpretation of reflexives with increasing proficiency. In fact, there were no significant differences between the levels in terms of the correct selection of antecedents for reflexives. Furthermore, almost all learners in the study were equally influenced by pragmatic constraints.

6.2. Explanation

6.2.1. The influence of finiteness or nonfiniteness.

The results of this study reveal that L2 learners were equally successful with nonfinite and finite structures, but had greater difficulty with triclausal than biclausal structures, and overall triclausal sentences with one nonfinite clause were the most difficult structures for L2 learners at every level to parse in order to bind reflexives with an antecedent. These results contradict Hirakawa's results which indicated that the finite/nonfinite distinction was more problematic for her L2 learners than the level of embedding. In fact, unlike the results of this study, almost all of the studies in this area have shown that L2 learners are more successful with binding in finite sentences than in nonfinite ones (Finer and Broselow, 1986; Hirakawa, 1990; Finer, 1991; Matsumura, 1993; Eckman, 1994). L2 learners in these studies might have had difficulty with nonfinite clauses due to the different kinds of verb morphology in finite and nonfinite clauses. An important morphological rule for verbs in English is that the verb in nonfinite clauses does not take any tense and agreement markers, unlike finite clauses. The L2 learners must

know the verb morphology sufficiently to distinguish the finite verb from the nonfinite verb in order to find the local antecedent for the reflexives. It appears that Turkish learners have acquired the English verb morphology because they were able to distinguish finite and nonfinite verbs in both biclausal and triclausal sentences.

6.2.2. The influence of levels of embedding.

The L2 learners in this study may have been less successful with triclausal sentences compared to biclausal sentences because for each sentence the parsing process is more complicated in that the learner must analyze two embedded clauses (instead of one) and there are three potential antecedents (instead of two) for the reflexive; this means that the subjects have to parse three clauses and use the knowledge of reflexive binding simultaneously in order to discover the right antecedent, which causes a heavy cognitive burden. The reason why subjects had such difficulty with the triclausal one nonfinite sentences might be that it is difficult for L2 learners to detect the border of the nonfinite clause which is buried inside the most embedded position in order to discover the local domain in which the subject NP, the antecedent for the reflexive, is located. However, since there are two finite clauses before the nonfinite clause, whose verb does not carry any tense and agreement markers, the subject NPs of the two finite clauses become more salient and recognizable than the subject NP of the nonfinite clause. Therefore, subjects were more likely to select the incorrect nonlocal NP as antecedent of the reflexive.

6.2.3. The influence of pragmatic bias.

The main outcome of this study was that Turkish learners at any proficiency level interpret English reflexives, at least for the pragmatically biased sentences, according to pragmatic strategies, rather than syntactic principles. Therefore, they almost always permitted a reflexive in a finite or nonfinite subordinate clause to be bound by an antecedent in a higher clause when that NP was pragmatically favored. In addition, the learners were significantly more likely to select the local NP when that NP was pragmatically favored compared to neutral sentences. This shows that Turkish learners were ready to violate the locality requirement of English binding in the face of pragmatic pressure. This result contradicts the results of Thomas (1989) which did not find any evidence that her subjects were affected by pragmatic bias.[1] Thomas explains her results by claiming that the locality requirement in English binding is a structural condition on the distribution of reflexives, and therefore is inhospitable to pragmatic pressure.

6.3. Huang's pragmatic theory of anaphora.

6.3.1. Introduction. It has been claimed that when a language selects a marked parameter value, this language has more than one NP which is syntactically eligible as the antecedent, and native speakers prefer one possible antecedent over the other one(s) (Thomas, 1991, 1993; Lakshmanan and Teranishi, 1994; White et al, 1995). Turkish has a marked parameter value for reflexive binding and thus has more than one eligible antecedent. If asked to select an antecedent for a reflexive, native Turkish speakers will usually select one or the other of the eligible NPs, but not both. Unlike other languages such as Chinese,

Korean and Japanese (Thomas, 1991b; Teranishi and Lakshmanan, 1991), Turkish native speakers do not habitually choose either the local NP or the nonlocal NP, but rather sometimes choose one and sometimes choose the other (Demirci et al, 1993). However, there does not appear to be any principled account for why one antecedent is selected over another in such languages. Thomas (1993) notes that:

> At present we have no theory of how a language's preferences for the interpretation of anaphors relate to the rest of its grammar.... we have no account of why Chinese speakers prefer local antecedents and Japanese speakers long-distance antecedents. I will assume that these preferences are pragmatic facts, and not attributable to the syntactic component of UG. (p.32)

Before we attempt to explain why Turkish learners are ready to succumb to pragmatic bias in their L2, we should initially examine how the antecedents of Turkish reflexives are selected from among several according to pragmatic constraints in their L1, since we will later claim that Turkish learners of English transfer from their L1 to their L2 the strategy of pragmatic interpretation for the selection of antecedents. This study employs the "pragmatic theory of anaphora" developed by Huang (1991, 1994) in order to explain how antecedents of reflexives are selected in Turkish.

6.3.2. Pragmatic vs syntactic languages.

According to Huang, the contribution of pragmatics to anaphora is much more fundamental than it has commonly been believed.[2] This is because syntax and pragmatics are interdependent in determining many processes of anaphora that are believed to occur within a grammar. Huang argues that pragmatics provides a set of complementary and

explanatory principles which constrains the interpretation of an utterance whose linguistic representation has already been antecedently cognisized. According to Huang, the extent to which syntax and pragmatics interact varies typologically. There is a class of languages such as Chinese, Japanese and Korean, where pragmatics plays a certain role in anaphora, which in European languages like English and French is played by syntax. In these pragmatic languages many of the constraints on grammatical processes, such as anaphora, are due to principles of language use rather than grammatical principles. Therefore, according to Huang, most of the linguistic explanations related to anaphora, which are currently done in syntactic terms, should be shifted to pragmatics. Huang states that a purely syntactic approach is not sufficient in specifying the set of possible antecedents for long-distance reflexives in pragmatic languages. According to Huang, since a syntactic approach has nothing to say about the selection of an antecedent for long-distance reflexives or the motivations between their use in languages like Chinese, Korean, etc., Huang develops a pragmatic theory of anaphora within the neo-Gricean framework of conversational implicature.[3]

6.3.3. Pragmatic theory of anaphora.

In Huang's theory, anaphora is largely determined by the systematic interaction of two neo-Gricean pragmatic principles, namely the M(anner)- and I(nformativeness)-principles, constrained by a world-knowledge based Disjoint Reference Presumption (DRP), information saliency and general consistency conditions on conversational implicature. These notions will be defined in what follows.

Huang provides the following interpretation principles for zero anaphors, pronouns and reflexives, and consistency constraints in his pragmatic theory of anaphora:

(58) A pragmatic theory of anaphora (Huang, 1994, p.144-145):

a. Interpretation principles
Assuming that a reflexive is necessarily referentially dependent, and a pronoun and a zero anaphor are optionally but preferably referentially dependent,

(i) the use of a zero anaphor will I-implicate a local coreferential interpretation;

(ii) the use of a pronoun will I-implicate a local coreferential interpretation, unless the pronoun is used where a zero anaphor could occur, in which case, the use of the pronoun will M-implicate the complement of the I-implicature associated with the use of the zero anaphor;

(iii) the use of a reflexive will I-implicate a local coreferential interpretation, unless the reflexive is used where a pronoun or a zero anaphor could occur, in which case, the use of the reflexive will M-implicate the complement of the I-implicature associated with the use of the pronoun or the zero anaphor, in terms of either reference or expectedness.

b. Consistency constraints
Any interpretation implicated by (a) is subject to the requirement of consistency with
(i) the DRP
(ii) information saliency, so that
(1) implicatures due to higher constructions may take precedence over implicatures due to lower constructions (i.e. 'matrix wins'), and
(2) implicatures to coreference may be preferred according to the saliency of antecedent in line with the following hierarchy: topic>subject>object, etc.; and
(iii) general implicature constraints, namely,
(1) background assumptions,
(2) meaning$_{-nn}$, and
(3) semantic entailments (e.g. referential dependence).

Under this theory, according to Huang, the interpretation of a reflexive is subject to the I-principle. All the I-principle does is to find an antecedent for the reflexive that gives the most informative, stereotypical interpretation in keeping with our knowledge about the world. The following sections discuss various features of Huang's theory.

world. The following sections discuss various features of Huang's theory.

6.3.3.1. *The semantics of reflexives* . The basic intuition Huang wants to capture with his theory is that in languages like Chinese, the preference for coreference is stronger than that for noncoreference. There is a distinction of referential dependence between reflexives and other anaphoric expressions due to semantics. Reflexives are marked and prolix because they appear to be semantically stronger than pronouns. This is because reflexives are necessarily referentially dependent while pronouns are not. Also, reflexives can be used where pronouns can occur (see footnote 4).

Huang tries to answer why a local coreferential interpretation is considered to be semantically more specific, thus informationally richer, and also why the general pattern of anaphora is assumed to be an instantiation of the interaction of the neo-Gricean pragmatic principles. Huang states that, according to Levinson (1987), the answers to these questions should be sought in the analysis of the notion of informativeness. First, at the level of reference, the smaller the number of possible state-descriptions which are compatible with a proposition, the more informative the proposition. Levinson (1987) states that "A proposition A is more informative than a proposition B if and only if the set of entailments of B is properly contained in the set of entailments of A." (Huang, 1994, p.121). Consider the following example:

(59) a. John loves himself

b. John loves him

In (59a), since a coreferential interpretation between "himself" and "John" reduces the number of possible entities referred to in the minimal domain of discourse, it is more domain-informative than a corresponding, noncoreferential interpretation as in (59b). In (59b), a noncoreferential interpretation (where "him" does not refer to "John") involves a larger set of possible referents and so is less domain-informative. Therefore, according to Huang, the preference for a local coreferential interpretation is the direct result of the I-principle.

6.3.3.2. *Neo-Gricean pragmatic principles* . Huang's theory relies on Levinson's (1987) reduction of the Gricean principles to three inferential strategies; these are the Q(uantity)-, I(nformativeness)-, and M(anner)- principles. These principles cover lexical NPs, pronouns, zero anaphors, and reflexives:

(60) Levinson's Q-, I-, and M- principles (Huang, 1991, p.305)

The Q-principle

Speaker's maxim:
Do not provide a statement that is informationally weaker than your knowledge of the world allows, unless providing a stronger statement would contravene the I-principle.

The I-principle

Speaker's maxim: the maxim of minimization
'Say as little as necessary', i.e., produce the minimal linguistic information sufficient to achieve your communicational ends,
(bearing the Q-principle in mind).
Recipient's corollary: the rule of enrichment
Amplify the informational content of the speaker's utterance,
by finding the most specific interpretation, up to what you judge to be the speaker's m-intended point. Specifically:
(i) Assume that stereotypical relations obtain between referents or events, unless: (1) that is inconsistent with what is taken for granted, (2) the speaker has broken the maxim of minimization by choosing a prolix

expression;
(ii) assume the existence or actuality of what a sentence is 'about' if that is consistent with what is taken for granted;
(iii) avoid interpretations that multiply entities referred to (assume referential parsimony); specifically, prefer coreferential readings of reduced NPs (pronouns or zeros).

The M-principle

Speaker's maxim:
Do not use a prolix, obscure or marked expression without reason.
Recipient's corollary:
If the speaker used a prolix or marked expression M, he did not mean the same as he would have had he used the unmarked expression U — specifically he was trying to avoid the stereotypical associations and I-implicatures of U.

The basic idea of the Q-principle is that the use of a semantically weaker expression in a set of contrastive semantic alternates Q-implicates the negation of the interpretation associated with the use of a semantically stronger one.[4] In other words, from the absence of an informationally stronger expression, one infers that the interpretation associated with the use of that stronger expression does not hold.

The I-principle operates in the opposite direction of the Q-principle. It induces from the use of a semantically general expression an inference that is semantically more specific, namely, an interpretation that is best in keeping with the most stereotypical and explanatory expectations.

The basic idea of the M-principle is that the use of a marked expression M-implicates the negation of the interpretation associated with the use of an unmarked expression in the same set. In other words, from the use of a marked expression, one infers that the stereotypical interpretation associated with the use of an unmarked expression does not obtain.

Huang adopts Levinson's (1987) proposal of an ordered set of precedents of these three pragmatic principles. The hierarchy of preference regarding the application of the Q-, M- and I- principles is; Q>M>I. This means that if an marked expression is used, then the M-implicature created will cancel and take precedence over any I-implicature. For example, if we say that "John killed the dog," the I-implicature of this will be that John killed the dog with that intention. However, if we say that "John caused the dog to die," the M-implicature (that John didn't intend to kill the dog) cancels and takes precedence over the I-implicature.

6.3.3.3. *Consistency constraints*. Huang incorporates "general consistency constraints" into the pragmatic apparatus by which the interpretations of anaphors are implicated (determined in large part by the systematic interaction of the M- and I- principles). The consistency constraints discussed by Huang which are relevant to the purposes of this paper are the DRP, and information saliency.

6.3.3.3.1. *Disjoint Reference Presumption* . Huang (1994) includes Farmer and Harnish's Disjoint Reference Presumption into his pragmatic theory of anaphora:

(60) Farmer and Harnish's Disjoint Reference Presumption (DRP) (p.87)

The arguments of a predicate are intended to be disjoint, unless marked otherwise.

This principle resembles Chomsky's condition B in the sense that the coarguments of a predicate tend to be disjoint in reference unless one of them is encoded by means of a

reflexive. However, according to Huang, unlike Chomsky's condition B, the DRP is of a pragmatic nature and describes a usage preference. The DRP is based on world-knowledge and on the fact that one entity acts upon another entity due to the way the world is stereotypically. According to Huang (1994, p.129), the advantage of attributing the DRP to world knowledge is that it will automatically prevent any inconsistent pragmatic implicature from arising or cancel it without violating the hierarchy proposed by Levinson (Q>M>I). Let's consider examples given by Huang in order to show how the DRP operates:

(61) a. Xiaoming changchang manyuan ø.
Xiaoming often blame
'Xiaoming$_1$ often blames (him$_2$/himself$_1$/me/you/us/them...).'

b. Xiaoming changchang manyuan ta.
Xiaoming often blame 3SG
'Xiaoming$_1$ often blames him$_2$/?himself$_1$."

c. Xiaoming changchang manyuan ziji.
Xiaoming often blame self
'Xiaoming$_1$ often blames himself$_1$.'

In (61a), the preferred, local coreferential interpretation of the zero-anaphor implicated by the I-principle will be cancelled by the DRP since the zero-anaphor is in the object position of a simple, transitive clause. Given this canceling by the DRP, the correct, noncoreferential interpretation is arrived at (where the zero-anaphor refers to something different from the local subject). The same holds true for the pronoun in (61b) in that the

DRP cancels the local coreferential interpretation. In (61c), on the other hand, the correct local coreferential interpretation is not blocked by the DRP since the DRP does not operate if a co-argument is a reflexive.

6.3.3.3.2. *Information saliency* : *"matrix wins"* Huang points out that one of the key distinguishing properties of conversational implicatures is "cancelability." This is where conversational implicatures evaporate in certain linguistic or non-linguistic contexts. Conversational implicatures may by cancelled in the face of background assumptions (world knowledge), $meaning_{nn}$, semantic entailments, context and priority pragmatic inferences. Consider Huang's example which involves the canceling of a conversational implicature based on world knowledge (where '+>' indicates conversational implicature, and '*+>' indicates its cancellation):

(62) a. John and Mary tested a TV set.
+> John and Mary tested a TV set together.
b. The American and the Russians tested an atom bomb in 1966.
*+> The American and the Russians tested an atom bomb together in 1966.

The implicature which arises in (62a) is that John and Mary tested a TV set together. However, given our knowledge about the world, this implicature is cancelled in (62b).

Huang then goes on to show that both M- and I- implicatures may be cancelled under certain conditions. Consider several of Huang's examples which show how an M-implicature may be cancelled:

(63) a. Xiaoming shuo ta xia ge yue jiehun.
Xiaoming say 3SG next CL month marry
'Xiaoming$_1$ says that he$_{1/2}$ will get married next month.'

b. Xiaoming shuo ziji xia ge yue jiehun.
Xiaoming say self next CL month marry
'Xiaoming says that he$_1$ will get married next month.'

Huang shows that in (63b) the M-implicature should arise that the reflexive "ziji" in the subject position is noncoreferential with the matrix subject. This is because the use of the more marked reflexive in a position where a less marked pronoun could occur (as in (63a)) M-implicates a contrast in reference so that the reflexive should be disjoint in reference with the matrix subject. However, this M-implicature does not obtain ("ziji" is coreferential with the matrix subject) because of the semantic constraint that reflexives must be referentially dependent.

Now Huang argues that there are instances where an M-implicature is not cancelled by background assumptions, nor by semantic entailments, but rather in some instances M-implicatures may be cancelled by I-implicatures, this in spite of the hierarchy that Q>M>I. To resolve this apparent contradiction, Huang appeals to the 'matrix wins' hypothesis proposed by Gazdar (1979) and Levinson (1983). This hypothesis states that implicatures due to higher constructions may cancel implicatures due to lower constructions. Consider the following examples given by Huang which illustrate how 'matrix wins' operates:

(64) a. Yang Daniang danxin nuer bu ken cihou ø.
Yang Grandma worry daughter not willing look after

'Grandma Yang$_1$ is worried that her daughter$_2$ is not willing to look after (her$_{1/3}$/herself$_2$/me/you/us/them...).'

b. Yang Daniang danxin nuer bu ken cihou ta.
Yang Grandma worry daughter not willing look after 3SG

'Grandma Yang$_1$ is worried that her daughter$_2$ is not willing to look after her$_{1/3}$.'

c. Yang Daniang danxin nuer bu ken cihou ziji.
Yang Daniang worry daughter not willing look after self

'Grandma Yang$_1$ is worried that her daughter$_2$ is not willing to look after her$_1$/herself$_2$.'

In (64a), the DRP blocks the coreferential interpretation between the zero anaphor and the subject of the embedded clause ("daughter") which is promoted by the I-principle. Thus, the zero anaphor is I-implicated to be coreferential with the matrix subject ("Grandma Yang"). In (64b), since the pronoun is used in the same position as the zero anaphor, this should M-implicate that the referent of the pronoun is different from the referent of the zero-anaphor. However, this interpretation (where the pronoun is M-implicated to be disjoint in reference with the matrix subject) is cancelled by the rival I-implicature that the pronoun is coreferential with the matrix subject due to the 'matrix wins' mechanism. According to Huang, this is because the matrix subject ("Grandma Yang") occupies the most prominent position in the sentence.

Now consider (64c) in which the reflexive "ziji" is coreferential with the matrix subject. Huang claims that in this sentence, "ziji" would first trigger an I-implicature which

would make "ziji" coreferential with the embedded subject. This interpretation is not prevented by the DRP (which does not operate with reflexives), nor by the semantics of "ziji." According to Huang, however, this interpretation is not the best one in that it does not make the most sense pragmatically given our knowledge about the world. The interpretation which best fits our knowledge about the world is the one in which the reflexive is coreferential with the matrix subject. Therefore the original interpretation will be cancelled in favor of the I-implicature where "ziji" is referentially dependent on the matrix subject.

The M-implicature which arises by the use of the reflexive in the same position where a pronoun could be used (as in 64b) loses out to this I-implicature because of the "matrix wins" mechanism. That is, despite the Q>M>I hierarchy, the M-implicature that the reflexive is disjoint with the matrix subject loses out to the I-implicature because the matrix subject occupies the most prominent position in the sentence.

In sum, according to Huang's pragmatic theory, an I-implicature may cancel both an M-implicature as well as another I-implicature due to the "matrix wins" mechanism as long as the interpretation which results is best in keeping with our knowledge about the world. We will see later that the pragmatically biased sentences in the data of this work lend themselves easily to such a framework. In particular, the NP1 pragmatically biased sentences in this study are similar to example (64c), and the interpretations of respondents find their explanation in Huang's pragmatic theory.

6.3.3.4. *Antecedent search procedure.* Huang proposes an antecedent search procedure for reflexives (p.178):

(66) An antecedent search procedure for reflexives.

In a structure of the sort [s_2 [s_1 R]], where R is a reflexive, R is interpreted as referentially dependent according to the following preference order:
(i) R is referentially dependent on the local subject; failing which:
(ii) R, especially when morphologically complex, is referentially dependent on the local object; failing which:
(iii) R, especially when morphologically complex, is referentially dependent on both the local subject and the local object (split antecedents); failing which:
(iv) (i)-(iii) is recursively applied to the next, higher clause, until an antecedent is found; failing which:
(v) find the nearest antecedent in the discourse, preferably a topic; failing which:
(vi) settle for an 'arbitrary' interpretation.

Hence, to find an antecedent for a reflexive, a local subject is in general preferred to a local object, especially when the reflexive is morphologically complex. A non-split antecedent will be preferred to a split one especially again when morphologically complex. If none of these NPs qualify as a possible antecedent, then a more remote clause will be checked and so on, recursively until the root clause is reached. Failure to find an intrasentential antecedent leads to the search for a previous discourse antecedent that is topic. If no antecedent is found, settle for an "arbitrary interpretation," which appears to be similar to the interpretation of generic pronouns.

In order to see how the antecedent search procedure works with reflexives, in what follows we will consider several examples relevant to the purposes of this study. These examples will show how an actual (preferred) antecedent is selected out of the set of possible antecedents, where the preferred antecedent means the interpretation that is the

most favored one out of a number of other possible interpretations. Now consider the first example (provided by Huang) in which there is only one possible antecedent within the sentence:

(67) Xiaoming changchang manyuan ziji.

Xiaoming often blame self

'Xiaoming$_i$ often blames self$_i$.'

In this sentence, the preference for a coreferential interpretation induced by the I-principle will be unblocked since the DRP is not in operation. Since the DRP is not in force here, 'self' is subject to the I-inferred coreferentiality with the local subject, and this interpretation is further reinforced by the semantics of 'self.' In addition, the M-principle is not at play here because of the semantics of the reflexive, namely that a reflexive is normally referentially dependent.

Consider another example given by Huang:

(68) Ziji de erzi de le jiang shi Lao Wang hen gaoxing.

self MM son win PFV prize make Lao Wang very happy

'That his$_1$ son had won a prize makes Wang$_1$ very happy.'

By the antecedent search procedure, since there is no subject binder in (68), the reflexive should be referentially dependent on the object ("Wang"). The following example

illustrates the split antecedent interpretation:

(69) Xiaoming gaosu Xiaohua Xiaoqiang bu xiangxin tamen ziji.

Xiaoming tell Xiaohua Xiaoqiang not trust 3PL self

'Xiaoming$_1$ tells Xiaohua$_2$ that Xiaoqiang$_3$ does not trust them$_{(1/2)}$.'

The next example illustrates a topic construction:

(70) Wo a jiushi qiang dui zhe ziji de xiongtang,

1SG PP EMP gun point EXP self MM chest

o ye bu pa.

EMP not afraid

'I, even if a gun is pointed at my chest, (I) won't be afraid.'

Since the comment clause does not contain a possible antecedent, the I-principle selects the topic as antecedent ("I"). Consider another example:

(71) Ziji pian ziji bu hao.

self deceive self not good

'To deceive oneself is not good.'

In (71) the first "ziji" is not bound at all and there is no apparent intrasentential antecedent

available, and so in this case "ziji" receives an arbitrary interpretation.

The following example given by Huang is most relevant to our purposes, and involves an instance in which there are two structurally possible antecedents:

(72) Lao Wang yiwei Lao Li bu zhidao ziji hui jisuanji.

Lao Wang think Lao Li not know self know computer

'Wang$_1$ thinks that Li$_2$ does not know that he$_{1/2}$ knows how to use a computer.'

In example (72), by the antecedent search procedure the reflexive "ziji" would first be interpreted to be bound (by the I-principle) to the subject of the lower clause ("Li"). However, according to Huang, this first inference would be rejected given our world knowledge — people usually know whether or not they can use a computer. This would lead to another operation of the I-principle such that the subject of the matrix clause would be implicated to be the antecedent of "ziji." Consider one more example which illustrates the same point:

(73) Xiaoming xihuan [[biaoyang ziji de] laoshi].

Xiaoming like praise self MM teacher

'Xiaoming$_1$ likes the teacher$_2$ who praises him$_1$/himself$_2$.'

In (73) the antecedent search procedure would lead to the I-implicated interpretation that the subject of the intermediate clause was the preferred antecedent. However, this

interpretation would be rejected given our background assumptions about the way things are in the world — teachers usually praise their students and not themselves. Thus, the original I-implicature would evaporate in favor of the more reasonable I-implicature where "ziji" is coreferential with the subject of the matrix clause.

Huang concludes his discussion of such examples (72 and 73) as follows; "All this must suffice to point to the conclusion that it is pragmatics that is responsible for determining the actual, preferred antecedent where there is more than one structurally possible antecedent" (p. 184).[5] The presentation of Huang's pragmatic theory of anaphora up to this point has been geared towards an account of sentences like those in (72) and (73), since these are the kinds of "NP1 biased" sentences in the study at hand.

In sum, Huang provides a pragmatic theory of anaphora which may account for how antecedents are selected in so-called pragmatic languages.[6] For all of the Chinese sentences which were translated into English above, the equivalent sentences in Turkish would involve the same I-principle and antecedent search procedure, and the same NPs would be chosen as the antecedents of the reflexives. This is because with respect to reflexive binding, Turkish belongs to the group of pragmatic languages proposed by Huang. Since reflexives occur very freely in Turkish and they may have more than one possible antecedent, it is claimed that the assignment of an actual antecedent for a reflexive is not restricted by grammatical rules, but by pragmatic inferences. Consequently, the I-principle developed by Huang is operative for Turkish speakers in selecting an actual antecedent among several possible ones in Turkish.

6.4. The transfer of Huang's pragmatic principles by Turkish L2 learners.

It will be argued in this section that the results of this study give evidence that Turkish L2 learners, regardless of their proficiency, appear to transfer pragmatic interpretive strategies for the binding of reflexives from their L1 to their L2. Consider the following examples from the study in order to see how Turkish learners employ the I-principle in the selection of an antecedent in English. It should be remembered that the learners' L2 (English) is a language whose binding system presumably operates on purely syntactic principles. The first example involves what has been called an NP1 biased sentence:

(74) The little girl was happy that her mother bought herself a nice toy

According to the antecedent search procedure, "herself" will create an I-implicature and therefore "herself" will be referentially dependent on the embedded subject ("mother"). Although this is a plausible interpretation, prevented neither by the semantics of "herself" nor by the DRP, it is not the best interpretation because it is not the most sensible interpretation according to our knowledge about world — i.e., mothers ordinarily buy toys for their children and not for themselves. Therefore, the initial interpretation is cancelled and the I-principle will then provide an interpretation that "herself" is dependent on the matrix clause subject by preference. The second interpretation fits our knowledge best. As we saw in the results, in such NP1 biased sentences, Turkish L2 learners almost always selected the nonlocal NP as the antecedent of the reflexive, and it seems plausible that this

results from the application of the I-principle as explained above.

Now consider the following NP2 biased sentence from the study:

(75) Ahmet heard that the famous actor talked about himself on T.V.

According to the antecedent search procedure, "himself" will create an I-implicature in (75) and thus will be referentially dependent on the embedded subject ("the famous actor"). Further, this interpretation is not contradicted by the semantics of "himself" nor by the DRP. Furthermore, it is an interpretation that is consistent with what we know about the way things are in the world. In addition, selecting "Ahmet" would be inconsistent with what is known about the world — it is probably doubtful that the famous actor would be talking about Ahmet on television (assuming Ahmet is an ordinary person). As we saw in the results, learners at all levels selected NP2 a large majority of the time in NP2 biased sentences, and it is claimed here that this results from the application of pragmatic strategies as set down by Huang.

Finally, consider a biclausal neutral sentence from the data:

(76) Ali said that Salih criticized himself

In (76), according to the antecedent search procedure, "himself" will create an I-implicature and thus will be referentially dependent on the embedded subject ("Salih"). Since both "Ali" and "Salih" are equally plausible interpretations given what we know about the world

(people can criticize both themselves or others), the first interpretation should hold. That is, there is no reason to cancel the initial I-implicature. The results of this study show that for neutral sentences, respondents are more likely than not to select the local NP as the preferred antecedent in neutral sentences (69%).

In the study, it was shown that for neutral sentences, the Turkish control group scored significantly differently than each of the experimental groups; it appears that for Turkish speakers, when two or more antecedents are structurally possible and where neither interpretation makes more sense of the way things are stereotypically in the world, then Turkish speakers prefer the local NP as antecedent, following Huang's antecedent search procedure (37% of responses selected NP1, 45% selected NP2, and 18% selected both). However, the results of the Turkish control group for neutral sentences do not provide very robust support for Huang's antecedent search procedure, because a large number of responses also selected NP1 for these neutral sentences (37%).

Despite this apparent problem for Huang's antecedent search procedure (an issue which is beyond the scope of this dissertation), the results of the Turkish control group with regard to the three biclausal sentence types help to shed light on an issue which is crucial for our purposes — namely, how can we determine whether the results for the experimental group point to the acquisition of syntactic knowledge of the L2 (Binding condition A for English), or on the other hand to complete transfer of pragmatic strategies from the L1(including the antecedent search procedure)? It appears that both of these alternative explanations account for the results.

When we consider the results for neutral sentences by the experimental group, we

found that the majority of responses correctly selected NP2 as antecedent (69%), and 79 out of 170 individuals scored at least 5/6 correct for these sentences. This latter result appears to indicate that almost half of the learners have acquired some knowledge of English reflexive binding, at least with regard to neutral sentences. On the other hand, this success is also predicted by Huang's antecedent search procedure in which local NPs are preferred to other possible antecedents where such an interpretation fits our world knowledge.

How then can we determine whether the respondents of this study have actually acquired syntactic knowledge of English or are merely applying the pragmatic strategies of their L1? This brings us back to the results of the Turkish control group. It could be argued that if the experimental group had not acquired any syntactic knowledge of English reflexive binding, then this group would have performed similarly to the Turkish control with regard to the neutral sentences. The results show that the experimental group was significantly more likely than the Turkish control to select NP2 as antecedent in neutral sentences. That there was a significant difference between these two groups for neutral sentences suggests that the subjects of the experimental group have begun to acquire syntactic knowledge and are not merely employing L1 pragmatic strategies.

In addition, comparing the two groups for the NP1 biased sentences, we find that the experimental group was more likely to select NP2 (12%) than the Turkish control group (2.6%). This significant difference (see section 5.2.1.5.2.) between the two groups suggests again that at least the experimental group has made some progress in the direction of acquiring some knowledge of English reflexive binding and are not simply transferring

pragmatic strategies for the neutral sentences. Thus it will be assumed in what follows that at least some of the subjects of this study acquired some syntactic knowledge of the target language, at least with respect to neutral sentences.

In order for Turkish learners to reset the binding parameter to the L2 value, they should be able to make a transition from a predominantly pragmatic mode to only a syntactic mode. However, the results indicate that they have not been able to do that so far. The evidence for this is that 79 out of 170 subjects were able to acquire the L2 binding system of English for neutral sentences, whereas only 5 out of these 79 subjects were able to ignore pragmatic constraints in NP1 biased sentences; when the sentence was biased in favor of the NP1, almost all subjects incorrectly selected NP1. Further, it must be assumed that those 79 subjects who appeared to have acquired the syntactic rules for English binding readily relinquished those syntactic rules in the face of pragmatic pressure. In addition, for NP2 biased sentences, we again find the pull of pragmatic pressure, since 140 of the 170 subjects selected NP2 at least 5 out of 6 times.

Moreover, the fact that 84% of responses selected NP1 in NP1 biased sentences, 87% of responses selected NP2 in NP2 biased sentences, and 69% of group responses went to NP2 in neutral sentences, indicates that L2 learners do not simply transfer the Turkish L1 setting, which permits both local and nonlocal NPs at any position within the sentence to be antecedents, into their L2. Rather, what they transfer is the strategy of pragmatic interpretation, namely the I-principle, from their L1 to their L2. This is because both the group and individual results for neutral sentences indicate that many Turkish learners show evidence of a syntactic interpretation of English reflexives, that is, the

locality requirement. However, that these 79 learners readily relinquish the syntactic interpretation in pragmatically biased sentences suggests that these learners are transferring something from their L1, and it is claimed here that this is the set of pragmatic principles and strategies which guide their interpretation of reflexives in their L1. For these subjects, for the pragmatically biased sentences, there arises a clash between the acquired syntactic constraints of the L2 binding system and the pragmatic principles transferred from their L1. The data give evidence that whenever such a clash arises, the Turkish L2 learners opt for the transferred pragmatic principles over the syntactic constraints.

In conclusion, it might be claimed that Turkish learners carry the I-principle, which operates on common perceptions about the world, from their L1 to their L2 in order to select the antecedent which fits their world knowledge in the best way. The operation of the I-principle by L2 learners in the L2 may lead to the suppression of their knowledge of L2 binding conditions. The pragmatic constraints governed by the I-principle seem to be so dominant over the syntactic ones that the employment of the pragmatic constraints does not stop with increasing proficiency and amount of input over time.

According to the results, 23 out of 35 subjects from the fourth year, after receiving 5 years of intensive English study in an English department, still were not able to select the correct NP as antecedent in NP1 biased sentences, while 30 subjects from the same year scored at least 5/6 correct for NP2 biased sentences and 18 subjects scored at least 5/6 correct for neutral sentences. Similarly, 16 out of 31 subjects from prep year, who were at the beginning of their five year program, could not select the correct NP in NP1 biased sentences, while 24 subjects from the same year scored at least 5/6 correct for NP2 biased

sentences, and 17 subjects scored at least 5/6 correct for neutral sentences. These results indicate that subjects at two levels, the most advanced and the least advanced, do not show any difference in their treatment of English reflexives in any of the three biclausal sentence types. If we focus on the NP1 biased sentences, we see that the L2 learners at these two levels with extremely different amounts of input and time are behaving in almost exactly the same way in using the I-principle to select the antecedent.

These results show that the pragmatic interpretive principles might have fossilized in the learners' interlanguage grammar. According to Selinker and Lakshmanan (1996), fossilization is the persistence of nontarget-like structures in the interlanguage of L2 learners, even for those who are at very advanced stages in their L2. White (1988, 1989, 1996) defines fossilization as the use of the L1 parameter setting when the L2 setting is needed. Evidence for this fossilization in the study is that no progress is made between prep year and year four in the acquisition of English reflexive binding. This fossilization seems to interfere with the development of the learners' acquisition of the English binding system, even when there is no pragmatic pressure towards any possible antecedent. For example, only half of the year 4 subjects were able to acquire the English binding system (using the 5/6 criterion) in neutral sentences, even though they were at the end of their formal instruction period. Similarly, only slightly more than half of the prep year subjects were able to acquire the English binding system at the beginning of their instruction program.

Both the group and individual results show that L2 learners were not able to set the parameter to the L2 value (there was a significant difference between the learners and the

English control). As the results indicate, the learners haven't revealed any development in the acquisition of English reflexive binding with increasing input and time. Their interlanguage grammar is different from their L1 system but also is different from the L2 system. We may claim that some Turkish learners do in fact reset the parameter to the L2 value, but in an unstable way, in that the learners are very susceptible to pragmatic pressure. The resetting is unstable because although 79 subjects selected the correct NP at least 5/6 times for neutral sentences, almost none of them selected the correct antecedent at least 5/6 times for the NP1 biased sentences.

As a result, the underlying knowledge of many of the Turkish L2 learners appears to be composed of two main competing properties: pragmatic constraints transferred from the L1, and syntactic constraints acquired from English. The syntactic constraints break down whenever faced with pragmatic constraints, which prevents learners from fully reconstructing the L2 grammar of reflexives.

Therefore, the issue which needs to be addressed is why the pragmatic constraints are so dominant for the Turkish learners. That is, what prevents the learners from fully acquiring the locality requirement for English reflexive binding? One explanation might be that the acquisition of the unmarked setting for English reflexive binding, which has a smaller domain and involves only syntactic processes, might be very difficult for L2 learners with the marked setting in their L1, which has a larger binding domain and has both syntactic and pragmatic processes. Some studies claim that the transition from a marked to an unmarked setting is more difficult than moving in the opposite direction, and therefore L2 learners may initially instantiate their L1 parameter setting in their L2, and L1

parameter settings become part of the interlanguage grammar, either briefly or for a long time until L2 input over time forces the learners to reconstruct their initial hypotheses about L2 constructions (e.g., White, 1988, 1989, 1994), although this result is not attested in the data of this study.

For L2 Turkish learners to make the transition from the marked to the unmarked setting or to reconstruct their initial hypothesis about English reflexive binding (that pragmatics plays some role in the target language), they may need some negative evidence in the form of overt instruction about the rules of English reflexive binding, although UG proponents claim that acquisition takes place based only on positive evidence. The results of this study indicate that L2 learners instantiate L1 pragmatic principles in their L2, and they have not received the right sort of L2 input which will prove that pragmatic constraints do not have any role in L2 reflexive binding.[7] In other words, positive evidence alone is not enough for Turkish learners to give up the pragmatic principles in the selection of antecedents for English reflexives, and hence their resetting of the parameter to the L2 value is still susceptible to pragmatic pressures.

An alternative and unexplored explanation of the results of this study relies on the notion, mentioned in chapter 4, that Turkish "kendisi" in not in fact a genuine reflexive, but rather has many properties of pronouns. It is conceivable that the Turkish learners of English in this study treated English reflexives much like Turkish "kendisi," since both are apparently morphologically complex, have agreement features, and contain a root meaning "self." Consequently, it could be argued that the Turkish learners simply allow long-distance binding of the English reflexive because the same is acceptable in their L1.

However, this explanation does not account for the fact that 69% of responses to biclausal neutral sentences correctly selected NP2, and that there was a statistically significant difference between the levels and the Turkish control group in terms of the number of correct responses. If the Turkish learners were simply behaving as they do in their L1, then we would expect no significant difference between the levels and the Turkish control. On the other hand, it could be argued that the fact that there was no difference between the proficiency levels and that so many learners readily selected NP1 in pragmatically biased sentences suggests that little or no learning has occurred for these learners. Furthermore, one might claim that the differences between the levels and the Turkish control group were not the result of differential knowledge, but rather an artifact of the experimental design of this study. In particular, the Turkish control group was responding to Turkish sentences which are very different in structure from the English sentences responded to by the proficiency levels, and thus one might question whether the results of the levels and the control group are comparable. In any case, future studies might attempt to test this possible alternative explanation.

6.5. Conclusion

6.5.1. Summary of results.

Given the result of this study that Turkish L2 learners of English select the nonlocal NP a large majority of the time in NP1 pragmatically biased sentences and prefer the local NP overwhelmingly in NP2 biased sentences, we may conclude that for these learners pragmatic constraints are stronger than L2 syntactic constraints in the interpretation of

reflexives in the L2. It has been argued that the interlanguage grammars of the L2 learners in this study include both syntactic and pragmatic constraints for English reflexives. This leads to the claim that subjects were not able to fully acquire the purely syntactic rules of English reflexive binding (there was a significant difference between the experimental group and the native control group), although many appear to have given up their L1 binding setting — there was a significant difference between the experimental group and the Turkish control.[8] The data indicate that L2 learners do not first adopt their L1 binding setting in their L2, because there was a significant difference between prep year and the Turkish control group. On the other hand, they do not also start the acquisition process with the L2 setting, because there was a significant difference between prep year and the native control.

The Turkish L2 learners of this study were not affected by the finite/nonfinite distinction, although their choice of antecedent was affected by the level of embedding — they were more successful with biclausal sentences than triclausal ones. Finally, there was not a significant difference between the proficiency levels in terms of their selection of antecedents for the different sentences types.

6.5.2. Implications.

Although the results of the study indicate that pragmatics plays a significant role in the interlanguages of L2 learners with regard to reflexive binding, the learners still obey principles of UG in constructing the binding system of the L2. For example, the learners of this study know that a reflexive must be bound by an antecedent (a UG constraint).

Therefore, they almost never selected the "neither" choice in any sentence type. Flynn and Marthohardjono (1994) state that adult L2 knowledge is structure dependent and certain errors are never made by adult L2 learners. Moreover, adult L2 learners are involved in an acquisition process which goes beyond any explicit teaching. The Turkish L2 learners in this study never received any explicit teaching about the binding of English reflexives. However, 69% of the responses to neutral sentences correctly selected the local NP. Thus, these points suggest that the apparent progress toward the acquisition of English reflexive binding on the part of the Turkish L2 learners might be attributed to UG.

Another implication of this study is that the acquisition of reflexives seems to be the result of many processes. Klein (1990, 1991) states that learning a language involves acquiring rules of the domains of phonology, syntax, lexicon, etc. The acquisition process simultaneously affects all of these domains, but not one domain after the other. According to Klein, the development in language acquisition is interactive as well as parallel. Therefore, the development of some syntactic rules, for example, is dependent on the development of a rule in some other domain. Klein states that "the fascination which the idea of a 'modular approach' has created — whether justified or not — must not make us forget that these modules, in whichever way the border lines between them are drawn, interact at some point" (p.182). In this study it was shown that the acquisition of reflexive binding for Turkish learners of English involves the interaction of the distinct modules of syntax and pragmatics.

The interaction of pragmatics and syntax is even evidenced in the results of this study for native speakers of English — for the NP1 biased sentences, the native control

group correctly selected NP2 only 76% of the time, an interesting result given the fact that the local NP is (theoretically) the only acceptable one for English. It is plausible that the I-principle is at play in the interpretation for reflexives for native English speakers, and therefore even in this supposedly purely syntactic binding system, syntax interacts with pragmatics in an integral way. Indeed, Huang (1994, p.144) suggests that the same pragmatic principles which apply in so-called "pragmatic" languages may in principle apply in languages like English. The results of this study lend credence to this view. Moreover, it is possible that the I-principle is a linguistic universal, because it appears to be employed even by native English speakers in reflexive binding. As Hawkins (1988) states; "There is also ample evidence that *some* universals are motivated by discourse-pragmatic functions..." (p.4).

In terms of implications for pedagogy, the overall lack of success and the lack of a developmental sequence found in the results of this study may indicate that L2 learners require explicit instruction in order to successfully acquire the rules of reflexive binding in the L2. It is plausible that the subjects of this study were not exposed to a sufficient amount of occurrences of reflexives in the input (the English used in the classroom), especially reflexives occurring in complex constructions. In addition, the input occurring in the classroom may not contain any evidence to indicate that pragmatics does not play a role in the interpretation of reflexives in English. Thus, explicit instruction about the properties of reflexives in the L2 "can be used to draw the learners' attention to properties of the input that might otherwise be missed" (White et al, 1991, p.417).

Finally, it seems to be the case that reflexives are more frequently used in Turkish

where reflexives and pronouns are in noncomplementary distribution, and are used more infrequently in English where reflexives and pronouns are in complementary distribution. Thus, given the greater frequency of reflexives in their L1, the L2 learners may more readily carry over the binding properties of their L1 into their L2.

In terms of attention constraints, the data may indicate that it was difficult for the L2 learners of this study to deal with both pragmatic and syntactic constraints at the same time, because there is a limitation to the amount of information that the attention can manage (McLaughlin, 1987, 1990; Tomlin and Villa, 1994; Carr and Curran, 1994). It seems plausible that for the pragmatically biased sentences, the learners' attention was divided between syntactic and pragmatic cues, and therefore only a limited amount of information (pragmatic only) was selected for further processing. Carr and Curran (1994) state the same thing metaphorically by saying that "much the same thing happens when too many energy demanding appliances are turned on at once and the kitchen's supply of electric power is insufficient to run them" (p.219).

The attention of the L2 learners in this study might have been specifically directed to pragmatic information to the exclusion of syntactic information. According to Tomlin and Villa (1994), stimuli not receiving attentional orientation (in this study, syntactic cues), when other stimuli are receiving it (in this study, pragmatic cues), are inhibited such that their detection requires more than normal effort. In other words, specific focus of attention on certain stimuli has facilitative or inhibitory consequences for further processing of the sentence.

The task that the L2 learners of this study performed required them to give attention

to both pragmatic and syntactic information simultaneously. However, it is possible that trying to process both types of information at the same time resulted in the violation of the syntactic rules of reflexive binding when the pragmatic cues received the greater amount of attention. Carr and Curran argue that attention is necessary for keeping track of the position of elements within a sentence, which in return leads the learner to be more dependent on other noticeable cues. In this study, the L2 learners needed to pay full attention to syntactic cues in order to determine the local domain where the reflexive could be bound to an antecedent. However, at the same time, attention to syntactic cues could have led the learners to recognize pragmatic cues within the same sentence. For example, in a sentence such as "the president ordered his bodyguard to protect himself during the speech," the L2 learners, in attempting to locate the local domain, could in doing so recognize the pragmatic bias of the sentence and thus lose track of the syntactic information.

In conclusion, this study has demonstrated that a purely syntactic account of reflexive binding for Turkish L2 learners appears to be inadequate. Further, future studies which attempt to explain L2 reflexive binding should consider the role of pragmatics in the acquisition process. In order to further substantiate these claims, future studies employing a variety of methodologies should aim to determine whether the pragmatic principles discussed by Huang are also applicable in other domains which have traditionally been studied only from syntactic perspectives.

FOOTNOTES

FOOTNOTES

Chapter 1

1 According to Chomsky (1986), the I-language is some element of the mind of the person who knows the language, acquired by the learner and used by the speaker. According to this definition, the grammar would then be a theory of I-language, which is the object under investigation. On the other hand, E-language is externalized language in the sense that it is understood independently of the properties of the mind or brain.

2 Chomsky (1986) later uses the term "local domain" instead of "governing category" and he defines the "local domain" as the governing category of X, where a governing category is a maximal projection containing both a subject and a lexical category governing X, where a governing category is a "complete functional complex" in the sense that all grammatical functions compatible with its head are realized in it.

3 Chomsky (1986) redefines the governing category as a "complete functional complex" in the sense that all grammatical functions compatible with the head are realized in it — the complements necessarily by the projection principle, and the subject, which is optional unless required to license a predicate, by definition. According to Dalrymple, 1993, this reformulation gives the same results as the original formulation by Chomsky (1981) for many examples except for one case for which results are different from the 1981 binding conditions, as in;

the children$_i$ like each other's$_i$ friends

4 According to Huang (1994) long-distance reflexivization displays several properties cross-linguistically:

a) long-distance reflexives allow an antecedent outside their local domain.
b) long-distance reflexives allow only subject antecedents
c) long-distance reflexives are morphologically simplex
d) there is no complementary distribution between pronouns and long-distance reflexives

5 MacLaughlin (1995) claims that the parameterized binding theory crucially relies on the lexical parameterization hypothesis in order to capture the binding facts, since many languages have both local and long-distance anaphors. According MacLaughlin (p.64), this hypothesis poses interesting issues for second language research, especially with regard to language transfer, such as whether the parameter values associated with individual lexical items are transferred onto lexical items in the L2, and if so, which L1 items are transferred. For example, Turkish has two anaphors, "kendi" ("self") and "kendi-kendi" ("self-self"), whereas English only has one — "himself." Therefore Turkish speakers learning English as a second language potentially have a choice as to which anaphor's parameter settings will be transferred onto English "himself."

6 Within this theory, marked values are derived from the interaction of UG and the learning component, instead of being stipulated.

7 The issue of only positive data is controversial and open to debate. However, this debate is beyond the scope of this dissertation. See Gass (1992), Schwartz (1992) and White (1991, 1992) for details.

8 Dalrymple (1989, p.91) notes that Manzini and Wexler's analysis treats antecedent constraints as well as domain constraints since an anaphor may be marked as to its proper antecedent.

Chapter 5

1 Since Turkish does not make any finite/nonfinite distinction, all of the results to be presented in all of the Tables and Figures pertaining to the Turkish control group represent the result for just one type. In the tables to follow, NP1BNF means "NP1 biased nonfinite," NP1BF means "NP1 biased finite," NP2BNF means "NP2 biased nonfinite," NP2BF means "NP2 biased finite," BCNF means "biclausal nonfinite," and BCF means "biclausal finite."

2 The Z-test is a statistical test of significance which compares two population proportions. It may be used in place of a t test when the population is large (over 30):

$$Z = \frac{p_1 - p_2}{p(1-p)\,(n_1+n_2)}$$

$$p_1 = \frac{x}{n_1}, \quad p_2 = \frac{y}{n_2}, \quad p = \frac{x+y}{n_1+n_2}$$

With significance (α) set at .05, and where H_o: $P_1 = P_2$ vs. H_1: $P_1 = P_2$, then if $Z > 1.96$, reject H_o. In cases where three populations are being compared, the Z test may not be used, and in such cases in this study, a chi-square test was used.

Chapter 6

1 Thomas did not check the effect of pragmatic bias on the NP2 — the local NP, and so could not compare this type with neutral sentences.

2 Huang defines anaphora as a relation between two linguistic elements wherein the interpretation of one called an anaphor is determined by the interpretation of the other called an antecedent. Reflexives are a subset of the larger set of anaphors.

3 Grice's theory of conversational implicature

(a) The co-operative principle
Make your conversational contribution such as is required, at the stage at which it occurs, by the accepted purpose or direction of the talk exchange in which you are engaged.

(b) The maxims of conversation
Quality: Try to make your contribution one that is true.
(i) Do not say what you believe to be false.
(ii) Do not say that for which you lack adequate evidence.
Quantity:
(i) Make your contribution as informative as is required (for the current purposes of the exchange.
(ii) Do not make your contribution more informative than is required.
Relation: Be relevant.
Manner: Be perspicuous.
(i) Avoid obscurity of expression.
(ii) Avoid ambiguity.
(iii) Be brief (avoid unnecessary prolixity)
(iv) Be orderly.

4 According to Levinson (1991), reflexives are prolix and marked expressions, as opposed to pronouns, because reflexives are referentially dependent and logophoric. Levinson's analysis of reflexive binding treats < reflexive, pronoun > as a "Horn scale" (defined as a set of contrastive semantic alternates such as scalar or gradable predictions (Huang, 1991, p.305)), which operates in terms of semantic informativeness. This Horn scale creates a contrast in semantic informativeness between a pronoun and a reflexive in the domain where a reflexive is bound. According to this scale, reflexives are informationally stronger than pronouns.

5 In languages which allow long-distance anaphors, any NP is a "structurally possible" antecedent in that there is no syntactic rule limiting the binding domain to the local NP.

6 The question of the role of syntax, with respect to reflexive binding, in pragmatic languages seems to me to be an open one. That is, one might argue, as Huang seems to do, that the positing of syntactic rules or constraints in such languages is unnecessary. On the other hand, one might suppose that pragmatic languages have a syntactic binding domain which is the entire sentence, and therefore includes all NPs as possible antecedents for reflexives.

7 Huang (personal communication) accepts the interpretation that the Turkish L2 learners in this study transfer the I principle from L1 to L2.

8 This assumes that parameter settings, which are part of UG, are distinct from pragmatic strategies or the sort that Huang discusses. Therefore it is possible for learners to transfer the pragmatic strategies, but not the L1 setting.

APPENDICES

APPENDIX A

APPENDIX A

Test sentences given to experimental group.

Consider each sentence, and circle either YES or NO for each statement below it. Think about each statement separately, and don't assume that a YES answer for one statement means that the answer for the next statement must be NO. If you do not know the meaning of any word, you can ask your instructor.

Lutfen, asagidaki cumleleri ve cumleleri izleyen secenekleri dikkatle okuyun. Her secenegi EVET veya HAYIR olarak cevaplayip, cevabinizi daire icine alin. Her secenegin cevabini ayri ayri dusunun. Bir secenege verdiginiz cevap digerini etkilemesin. Cumlenin size verdigi anlama gore, butun seceneklerin hepsini 'evet' veya seceneklerden birini 'evet' digerini 'hayir' olarak cevaplayabilirsiniz. Himself ve herself'i izleyen "esittir (=)" isareti "olabilir" veya "gosterir" anlamina gelmektedir. Bilmediginiz kelimeleri hocaniza sorun. Bilmediginiz kelimeleri ogrenmeden sorulari cevaplamayin. Zaman siniri yoktur.

1. Ahmet said that Cem always criticized himself.

a.'himself' = Ahmet	YES	NO
b.'himself' = Cem	YES	NO

2. Fatih believed that Adnan asked Salih to respect himself more.

a.'himself' = Fatih	YES	NO
b.'himself' = Adnan	YES	NO
c.'himself' = Salih	YES	NO

3. Didem says that Hande thinks that Sibel bought herself a new coat.

a.'herself' = Didem	YES	NO
b.'herself' = Hande	YES	NO
c.'herself' = Sibel	YES	NO

4. The doctor wanted the sick woman to care for herself.

a.'herself' = doctor	YES	NO
b.'herself' = sick woman	YES	NO

5. Feride knew that Nur told Berna not to blame herself for the argument.

a.'herself' = Feride	YES	NO
b.'herself' = Nur	YES	NO
c.'herself' = Berna	YES	NO

6. Nergis told Suzan to introduce herself to the class.

a.'herself' = Nergis	YES	NO
b.'herself' = Suzan	YES	NO

7. The student was angry that his teacher gave himself alot of homework.

a.'himself' = the student	YES	NO
b.'himself' = teacher	YES	NO

8. Cem said that Ali asked Hasan to talk about himself.

a.'himself' = Cem	YES	NO
b.'himself' = Ali	YES	NO
c.'himself' = Hasan	YES	NO

9. During the interview, the employer asked Faik to talk about himself.

a.'himself' = employer	YES	NO
b.'himself' = Faik	YES	NO

10. Salim knows that Ahmet blames himself for the accident.

a.'himself' = Salim	YES	NO
b.'himself' = Ahmet	YES	NO

11. The little boy was angry that his father always hit himself.

a.'himself' = little boy	YES	NO
b.'himself' = father	YES	NO

12. Mustafa asked Hakan to trust himself more.

a.'himself' = Mustafa	YES	NO
b.'himself' = Hakan	YES	NO

13. Leman told Fatma to ask Sevgi to buy herself new clothes.

a.'herself' = Leman	YES	NO
b.'herself' = Fatma	YES	NO
c.'herself' = Sevgi	YES	NO

14. Serkan asked his brother to help himself with his homework.

a.'himself' = Serkan	YES	NO
b.'himself' = brother	YES	NO

15. Feride remembers that Nermin introduced herself at the meeting.

a.'herself' = Feride	YES	NO
b.'herself' = Nermin	YES	NO

16. Ali remembers that Hakan said that Can hates himself.

a.'himself' = Ali	YES	NO
b.'himself' = Hakan	YES	NO
c.'himself' = Can	YES	NO

17. Cem heard that the famous actor talked about himself on T.V.

a.'himself' = Cem	YES	NO
b.'himself' = famous actor	YES	NO

18. Nergis asked Oya to write a letter to herself.

a.'herself' = Nergis	YES	NO
b.'herself' = Oya	YES	NO

19. Orhan asked Ahmet to tell Tarkan to look at himself in the mirror.

a.'himself' = Orhan	YES	NO
b.'himself' = Ahmet	YES	NO
c.'himself' = Tarkan	YES	NO

20. Mesut discovered that Ali killed himself.

a.'himself' = Mesut	YES	NO
b.'himself' = Ali	YES	NO

21. Cem asked Tuncay to write about himself in the letter.

a.'himself' = Cem	YES	NO
b.'himself' = Tuncay	YES	NO

22. The president ordered his best bodyguard to protect himself very carefully during the speech.

a.'himself' = the president	YES	NO
b.'himself' = bodyguard	YES	NO

23. Oya wanted Feride to ask Lale to talk about herself.

a.'herself' = Oya	YES	NO
b.'herself' = Feride	YES	NO
c.'herself' = Lale	YES	NO

24. Leman asked Belgin to describe herself on the phone.

a.'herself' = Leman	YES	NO
b.'herself' = Belgin	YES	NO

25. Kenan said that Sinan thought that Ahmet saw himself in the mirror.

a.'himself' = Kenan	YES	NO
b.'himself' = Sinan	YES	NO
c.'himself' = Ahmet	YES	NO

26. The little girl was happy that her mother bought herself a nice toy.

a.'herself' = little girl	YES	NO
b.'herself' = mother	YES	NO

27. Saliha learned that Emel hurt herself in a car accident.

a.'herself' = Saliha	YES	NO
b.'herself' = Emel	YES	NO

APPENDIX B

APPENDIX B

Test sentences given to Turkish control group

Lutfen, asagidaki cumleleri ve cumleleri izleyen secenekleri dikkatle okuyun. Her secenegi EVET veya HAYIR olarak cevaplayip, cevabinizi daire icine alin. Her secenegin cevabini ayri ayri dusunun. Bir secenege verdiginiz cevap digerini etkilemesin. Cumlenin size verdigi anlama gore, butun seceneklerin hepsini 'evet' veya seceneklerden birini 'evet' digerini 'hayir' olarak cevaplayabilirsiniz.

ORNEK: Ali Ahmet'in kendini bicakladigini duydu.

a. 'kendi' = Ali	EVET	HAYIR
b. 'kendi' = Ahmet	EVET	HAYIR

1. Ali Ahmet'in her zaman kendisini elestirdigini soyledi.

a. 'kendi' = Ali	EVET	HAYIR
b. 'kendi' = Ahmet	EVET	HAYIR

2. Ali Ahmet'in Cem'den kendisine daha saygili olmasini istedigine inaniyor.

a. 'kendi' = Ali	EVET	HAYIR
b. 'kendi' = Ahmet	EVET	HAYIR
c. 'kendi' = Cem	EVET	HAYIR

3. Elif Oya'nin Sema'nin kendisine yeni bir ceket aldigina inandigini soyledi.

a. 'kendi' = Elif	EVET	HAYIR
b. 'kendi' = Oya	EVET	HAYIR
c. 'kendi' = Sema	EVET	HAYIR

4. Doktor hasta kadinin kendisine iyi bakmasini istedi.

a. 'kendi' = doktor	EVET	HAYIR
b. 'kendi' = hasta kadin	EVET	HAYIR

5.Ayse Belgin'in Oya'ya tartismadan dolayi kendisini suclamamasini soyledigini biliyordu.

a. 'kendi' = Ayse	EVET	HAYIR
b. 'kendi' = Belgin	EVET	HAYIR
c. 'kendi' = Oya	EVET	HAYIR

6. Elif Hulya'nin kendisini sinifa tanistirmasini istedi.

a. 'kendi' = Elif EVET HAYIR
b. 'kendi' = Hulya EVET HAYIR

7. Ogrenci ogretmeninin kendisine cok odev vermesine kizmisti.

a. 'kendi' = ogrenci EVET HAYIR
b. 'kendi' = ogretmen EVET HAYIR

8.Ahmet Ali'nin kendisi hakkinda konusmasini istedi.

a. 'kendi' = Ahmet EVET HAYIR
b. 'kendi' = Ali EVET HAYIR

9. Gorusme boyunca, isveren Ahmet'ten kendi hakkinda konusmasini istedi.

a. 'kendi' = isveren EVET HAYIR
b. 'kendi' = Ahmet EVET HAYIR

10. Cem kazadan dolayi Ali'nin kendisini sucladigini biliyor.

a. 'kendi' = Cem EVET HAYIR
b. 'kendi' = Ali EVET HAYIR

11. Yaramaz cocuk babasinin herzaman kendisini dovdugunu soyledi.

a. 'kendi' = yaramaz cocuk EVET HAYIR
b. 'kendi' = baba EVET HAYIR

12.Cem Murat'in kendisine daha cok guvenmesini soyledi.

a. 'kendi' = Cem EVET HAYIR
b. 'kendi' = Murat EVET HAYIR

13. Oya Zeynep'in Elif'e kendine yeni kiyafet almasini soylemesini istedi.

a. 'kendi' = Oya EVET HAYIR
b. 'kendi' = Zeynep EVET HAYIR
c. 'kendi' = Elif EVET HAYIR

14. Hakan kardesinden kendisine odevlerinde yardimci olmasini istedi.

a. 'kendi' = Hakan EVET HAYIR
b. 'kendi' = kardes EVET HAYIR

15. Ayse Oya'nin kendisini toplantidakilere tanistirdigini hatirladi.

a. 'kendi' = Ayse EVET HAYIR
b. 'kendi' = Oya EVET HAYIR

16. Adnan Can'in Ali'nin kendisinden nefret ettigini soyledigini hatirliyor.

a. 'kendi' = Adnan	EVET	HAYIR
b. 'kendi' = Can	EVET	HAYIR
c. 'kendi' = Ali	EVET	HAYIR

17. Ahmet unlu bir aktorun televizyonda kendi hakkinda konusmasini izledi..

a. 'kendi' = Ahmet	EVET	HAYIR
b. 'kendi' = Unlu Aktor	EVET	HAYIR

18. Zeynep Oya'nin kendisine mektup yazmasini istedi.

a. 'kendi' = Zeynep	EVET	HAYIR
b. 'kendi' = Oya	EVET	HAYIR

19.Cem Ali'nin Hakan'a aynada kendisine bakmasini soylemesini istedi.

a. 'kendi' = Cem	EVET	HAYIR
b. 'kendi' = Ali	EVET	HAYIR
c. 'kendi' = Hakan	EVET	HAYIR

20. Ahmet Mustafa'nin kendisini oldurdugunu ogrendi.

a. 'kendi' = Ahmet	EVET	HAYIR
b. 'kendi' = Mustafa	EVET	HAYIR

21. Adnan Cem'in mektupta kendisinden bahsetmesini istedi.

a. 'kendi' = Adnan	EVET	HAYIR
b. 'kendi' =Cem	EVET	HAYIR

22. Devlet baskani koruma gorevlisine konusmasi boyunca kendisini korumasi icin emir verdi.

a. 'kendi' = Devlet Baskani	EVET	HAYIR
b. 'kendi' = Koruma Gorevlisi	EVET	HAYIR

23. Ayse Belgin'den Oya'ya kendi hakkinda konusmasini soylemesini istedi.

a. 'kendi' = Ayse	EVET	HAYIR
b. 'kendi' = Belgin	EVET	HAYIR
c. 'kendi' = Oya	EVET	HAYIR

24. Canan telefonda Ayse'den kendisini tanitmasini istedi.

a. 'kendi' = Canan	EVET	HAYIR
b. 'kendi' = Ayse	EVET	HAYIR

25. Adnan Cem'in Mustafa'nin aynada kendisini gordugunu sandigini soyledi.

a. 'kendi' = Adnan	EVET	HAYIR
b. 'kendi' = Cem	EVET	HAYIR
c. 'kendi' = Mustafa	EVET	HAYIR

26. Kucuk kiz annesinin kendisine oyuncak almasina cok sevindi.

a. 'kendi' = Kucuk Kiz	EVET	HAYIR
b. 'kendi' = Anne	EVET	HAYIR

27. Belgin Ayse'nin kendisini kazada yaraladigini ogrendi.

a. 'kendi' = Belgin	EVET	HAYIR
b. 'kendi' = Ayse	EVET	HAYIR

BIBLIOGRAPHY

BIBLIOGRAPHY

Battistella, E. (1989). Chinese reflexivization: A movement to INFL approach. Linguistics, 27, 987-1012.

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

Berent, G.P. (1994). The Subset Principle in second language acquisition. In E.E. Tarone, S.M. Gass, & A.D. Cohen (Eds.) Research methodology in second language acquisition (pp.17-40). Hillsdale, NJ: Erlbaum.

Carr, T. & Curran, T. (1994). Cognitive in learning about structural sequences. Studies in Second Language Acquisition, 16, 205-230.

Chomsky, N. (1981). Lectures on government and binding. Dordrecht, Holland: Foris.

Chomsky, N. (1982). Some concepts and consequences of the theory of government and binding. Cambridge, MA: The MIT Press.

Chomsky, N. (1986a). Knowledge of language. New York: Praeger.

Chomsky, N. (1986b). Barriers. Cambridge, MA: MIT Press.

Chomsky, N. (1988). Language and problems of language:The Managua lectures. Cambridge, MA: The MIT Press.

Chomsky, N. (1995). The minimalist program. Cambridge, M.A.: The MIT Press.

Christie, K. & Lantolf, J.P. (1993). Bind me up, bind me down: A UG study of reflexives in the second language. Paper presented at the MIT workshop on recent advances in SLA, Cambridge, MA.

Clahsen, H., & Muysken, P. (1986). The availability of universal grammar to adult and child learners. Second Language Research, 2, 93-119.

Clahsen, H., & Muysken, P. (1989). The UG paradox in L2 acquisition. Second Language Research, 5, 1-29.

Cole, P., Hermon, G. & Sung, L.M. (1990). Principles of parameters of long distance reflexives. Linguistic Inquiry, 21, 1-22.

Cook, V. (1996). Competence and multi-competence in performance and competence in SLA. In G. Brown, K. Malmkjar, & J. Williams (Eds.), Performance and competence in SLA (pp.54-69). Cambridge: Cambridge University Press.

Coppieters, R. (1986). Competence differences between native and near-native speakers. Language, 62, 544-573.

Cowan, R. & Hataca, Y.A. (1994). Investigating the validity and reliability of native speaker and second-language learner judgments about sentences. In E.E. Tarone, S.M. Gass, & A.D. Cohen (Eds.), Research methodology in second language acquisition (pp.287-302). Hillsdale, NJ: Erlbaum.

Dalrymple, M. (1993). The syntax of anaphoric binding. Stanford, CA: Center for the study of language and information.

Demirci, M., Glew, M. & Yamagata, A. (1993). The subset principle and semantic bias in the acquisition of anaphoric reflexives in SLA. Unpublished manuscript.

Eckman, F.R. (1994). Local and long-distance anaphora in second-language acquisition. In E.E. Tarone, S.M. Gass, & A.D. Cohen (Eds.), Research methodology in second language acquisition (pp.207-226). Hillsdale, NJ: Erlbaum.

Ellis, R. (1991). Grammaticality judgements and second language acquisition. Studies in Second Language Acquisition, 13, 161-186.

Ellis, R. (1996).The study of second language acquisition. Oxford: Oxford University Press.

Finer, D. (1991). Binding parameters in second language acquisition. In L. Eubank (Ed.), Point Counterpoint: Universal Grammar in the Second Language, (pp.350-374). Amsterdam/Philadelphia:John Benjamins.

Finer, D. & Broselow, E. (1986). Second language of reflexive binding. Proceedings of the North Eastern Linguistic Society, 16, 154-168.

Flynn, S. (1987). A parameter setting model of L2 acquisition: Experimental studies in anaphora. Boston: D. Reidel.

Flynn, S. (1988). Second language acquisition and grammatical theory. In F.J. Newmeyer (Ed.), Linguistics: the Cambridge survey: Vol. 2. Linguistic theory: extensions and implications. (pp.194-209). Cambridge: Cambridge University Press.

Flynn, S. (1991). Government-binding: parameter-setting in second language acquisition. In T. Huebner and C. Ferguson (Eds.), Crosscurrents in second language acquisition and linguistic theories. (pp.143-168). Amsterdam/Philadelphia:John Benjamins.

Flynn, S. (1993). Interactions between L2 acquisition and linguistic theory. In F.R. Eckman (Ed.), Confluence: linguistics, L2 acquisition and speech pathology (pp.15-36). Amsterdam/Philadelphia:John Benjamins.

Flynn, S. & Martohardjono, G. (1994). Mapping from the initial state to the final state: The separation of universal principles and language-specific principles. In B. Lust, M. Suner, & J. Whitman (Eds.), Heads, Projections, and Learnability (pp.319-336). Hillsdale, N.J.: Lawrence Erlbaum.

Gass, S. M. (1983). The development of L2 intuitions. TESOL Quarterly, 17, 273-291.

Gass, S.M. (1994a). English language teaching from a learning perspective. Paper presented at the Chulalongkorn 2nd International Conference on Language Teaching, Bangkok, Thailand.

Gass, S.M. (1994b). The reliability of second-language grammaticality judgments. In E.E. Tarone, S.M. Gass, & A.D. Cohen (Eds.), Research methodology in second language acquisition (pp.303-322). Hillsdale, NJ: Erlbaum.

Gass, S.M. & Selinker, L. (1992). Language transfer in language learning. Amsterdam/Philadelphia:John Benjamins.

Hawkins, J.A. (1988). Explaining language universals. In J. Hawkins (Ed.), Explaining language universals, (pp.3-23). Oxford: Basil Blackwell.

Hirakawa, M. (1990). A study of the L2 acquisition of English reflexives. Second Language Research, 6, 60-85.

Huang, C.-T.J.& Tang, C.-C.J. (1991). The local nature of the long-distance reflexive in Chinese. In J. Koster & E. Reuland (Eds.), Long-Distance Anaphora (pp.263-282). Cambridge: Cambridge University Press.

Huang, Y. (1991). A neo-Gricean pragmatic theory of anaphora. Linguistics, 27, 301-335.

Huang, Y. (1994). The syntax and pragmatics of anaphora. Cambridge: Cambridge University Press.

Klein, W. (1991). SLA theory: prolegomena to a theory of language acquisition and implications for theoretical linguistics. In T. Huebner and C. Ferguson (Eds.), Crosscurrents in second language acquisition and linguistic theories (pp.169-194). Amsterdam/Philadelphia:John Benjamins.

Kornfilt, J. (1997). Local and long distance reflexives in Turkish. Abstract for Workshop on reflexives at LSA summer institute at Cornell University.

Lakshmanan, U. & Teranishi, K. (1994). Preferences versus grammaticality judgements: some methodological issues concerning the governing category parameter in second language acquisition. In E.E. Tarone, S.M. Gass, & A.D. Cohen (Eds.), Research methodology in second language acquisition (pp.185-206). Hillsdale, NJ: Erlbaum.

Lebeaux, D. (1983). A distributional difference between reciprocals and reflexives. Linguistic Inquiry, 14, 723-730.

Levinson, S. (1987). Minimization and conversational inference. In J. Verschueren & M. Bertuccelli-Papi (Eds.), The pragmatic perspective (pp.61-129). Amsterdam/Philadelphia:John Benjamins.

Levinson, S. (1991). Pragmatic reduction of the Binding Conditions revisited. Journal of Linguistics, 27, 107-161.

MacLaughlin, D. (1995a). Universal Grammar and L2 acquisition of reflexive binding: Some learners acquire a non-L1/non-target system. Paper presented at the Second Language Research Forum, Ithaca, New York.

MacLaughlin, D. (1995b). Language acquisition and the Subset Principle. The Linguistic Review, 12, 143-191.

Manzini, M.R., & Wexler, K. (1987). Parameters, binding theory and learnability. Linguistic Inquiry, 18, 413-444.

Matsumura, M. (1994). Japanese learners' acquisition of the locality requirement of English reflexives: Evidence for retreat from overgeneralization. Studies in Second Language Acquisition, 16, 19-42.

Mazurkewich, I. (1984). The acquisition of the dative alternation by second language learners and linguistic theory. Language Learning, 34, 91-109.

McLaughlin, B. (1987). Theories of second language learning. London: Edward Arnold.

McLaughlin, B. (1990). Restructuring. Applied Linguistics, 11, 113-128.

Pica, P. (1987). On the nature of the reflexivization cycle. Proceedings of the North-eastern Linguistics Society, 17, 483-499.

Roeper, T. & Williams, E. (1987). Parameter setting. Dordrecht, Holland: Reidel.

Schachter, J. (1988). Second language acquisition and its relationship to universal grammar. Applied Linguistics, 9, 219-235.

Schachter, J. (1990). On the issue of completeness in second language acquisition. Second Language Research, 6, 91-124.

Schachter, J. (1996). Learning and triggering in adult L2 acquisition. In G. Brown, K. Malmkjar, & J. Williams (Eds.), Performance and competence in SLA (pp.70-88). Cambridge: Cambridge University Press.

Schwartz, D.B. (1992). On explicit and negative evidence effecting and affecting competence and linguistic behavior. Unpublished manuscript.

Selinker, L. & Lakshmanan, U. (1994). Language transfer and fossilization: The 'Multiple Effects Principle.' In S.M. Gass & L. Selinker (Eds.), Language transfer in language learning (pp.197-216). Amsterdam/Philadelphia:John Benjamins.

Sharwood Smith, M. (1988). L2 acquisition: logical problems and empirical solution. In J. Pankhurst, M. Sharwood Smith & P. Van Buren (Eds.), Learnability and Second Languages (pp.9-35). Dordrecht, Holland: Foris.

Sharwood Smith, M. (1990). Second language learnability. In I.M. Roca (Ed.), Logical issues in language acquisition (pp.259-273). Dordrecht, Holland: Foris.

Sorace, A. (1988). Linguistic intuitions in interlanguage development: the problem of indeterminacy. In J. Pankhurst, M. Sharwood Smith & P. Van Buren (Eds.), Learnability and Second Languages (pp.167-189). Dordrecht, Holland: Foris.

Tarone, E.E., Gass, S.M., & Cohen, A.D. (1994). Research methodology in second language acquisition. Hillsdale, NJ: Erlbaum.

Teranishi, K. & Lakshmanan, U. (1991). Preferences vs. grammaticality judgments: Some methodological issues concerning the governing category parameter in second language acquisition. Paper presented at the Conference on Theory Construction and Methodology, East Lansing, MI.

Thomas, M. (1989). The interpretation of English reflexive pronouns by non-native speakers. Studies in Second Language Acquisition, 11, 281-301.

Thomas, M. (1990). Acquisition of the Japanese reflexive 'zibun' by unilingual and multilingual learners. In H. Burmeister & P.L. Rounds (Eds.), Variability in Second Language Acquisition: Proceedings of the Tenth Meeting of the Second Language Research Forum. Eugene, Oregon, 701-718.

Thomas, M. (1991a). Universal grammar and the interpretation of reflexives on a second language. Language, 67, 211-239.

Thomas, M. (1991b). Do second language learners have 'rogue' grammars of anaphora?. In L. Eubank (Ed.), Point counterpoint: universal grammar in the second language (pp.375-388). Amsterdam/Philadelphia: John Benjamins.

Thomas, M. (1993). Knowledge of reflexives. Amsterdam/Philadelphia: John Benjamins.

Thomas, M. (1995). Acquistion of the Japanese reflexive 'zibun' and movement of anaphors in logical form. Second Language Research, 11, 206-234.

Tomlin, R.S. & Villa, V. (1994). Attention in cognitive science and SLA. Studies in Second Language Acquisition, 16, 183-205.

Trahey, M. (1996). Positive evidence in SLA: some long-term effects. Second Language Research, 12, 111-139.

Trahey, M. & White, L. (1993). Positive evidence and preemption in the second language classroom. Studies in Second Language Acquisition, 15, 181-204.

Underhill, R. (1976). Turkish grammar. Cambridge, MA: MIT Press.

Van Buren, P. (1988). Some remarks on the Subset Principle in second language acquisition. Second Language Research, 4, 33-40.

Wexler, K. & Manzini, M.R. (1987). Parameters and learnability in binding theory. In T. Roeper & E. Williams (Eds.), Parameter Setting (pp.41-76) Dordrecht, Holland: Reidel.

White, L. (1985). The acquisition of parameterized grammars. Second Language Research, 1, 1-17.

White, L. (1988). Universal grammar and language transfer. In J. Pankhurst, M. Sharwood Smith & P. Van Buren (Eds.), Learnability and Second Languages (pp.36-60). Dordrecht, Holland: Foris.

White, L. (1989). Universal grammar and second language acquisition. Amsterdam/Philadelphia: John Benjamins.

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. (1992). On triggering data in L2 acquisition: A reply to Schwartz and Gubala-Ryzak. Second Language Research, 8, 120-137.

White, L. (1994). Universal grammar: is it just a new name for old problems?. In S.M. Gass & L. Selinker (Eds.), Language transfer in language learning (pp.217-232). Amsterdam/Philadelphia:John Benjamins.

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

White, L., Hirakawa, M., & Kawasaki, T. (1995). Second language acquisition of long distance reflexives: effects and non-effects of input manipulation. Unpublished manuscript.

White, L., Bruhn-Garavito, J., Kawasaki, T., Pater, J., & Prevost, P. (1995). The researcher gave the subject a test about himself: Methodological issues in the investigation of reflexive binding. Paper presented at the Second Language Research Forum, Ithaca, New York.

Yuan, B. (1994). Second language acquisition of reflexives revisited. Language, 70, 531-545.