## *The maintenance and loss of reflexive anaphors in L2 English*

McCormack, Bede G.

# The Maintenance and Loss of Reflexive Anaphors in L2 English

by

Bede G. McCormack

a thesis submitted in partial
fulfillment of the requirements
for the degree of Doctor of
Philosophy

University of Durham

Department of Linguistics and English Language

2001

Abstract of Thesis

by

Bede G. McCormack

*The Maintenance and Loss of Reflexive Anaphors in L2 English*

For the Degree of Ph.D., 2001

This thesis is a generative-based investigation of second language (L2) attrition. L2 attrition research to now has measured the loss of lexical items, morphology, word order, and so on. However, none to my knowledge has examined attrition from within an established theoretical framework such as Chomsky's theory of Government and Binding (GB). In particular, this study considers the loss of reflexive binding in proficient L2 English speakers. Informants are six Japanese university students who spent their junior (third) years abroad in the United States. These six informants consisted of two groups: three who had childhood exposure to English, and three whose first exposure to English in the L2 environment was as adults during their university stay overseas. In order to observe attrition which might occur shortly after exposure to the L2 ceased, an important aspect of this research was to begin data collection as soon as possible after the informants' returns to Japan from studying abroad. Data collections occurred at various intervals for each informant and lasted up to 16 months. Data for this longitudinal study were collected via two tests: a truth value judgment test and a grammaticality judgement test.

This study is unique in that it uses generative-based SLA research tools and methods to investigate L2 attrition. Furthermore, the truth value judgment test and the grammaticality judgment test provide results which support the hypothesis that principles of reflexive binding attrite in a manner not inconsistent with UG constraints. The general pattern exhibited by all six test subjects initially shows varying but high levels of knowledge of reflexive binding. Over the course of their data collection periods, the informants' knowledge of reflexive binding in English becomes unstable in the face of zero exposure to the target language. In particular, reflexives in finite subordinate clauses tend to remain bound grammatically to local antecedents to a greater degree than in nonfinite clauses. Reflexive binding in tensed clauses thus appears more resistant to attrition. Age at first exposure to the L2 was also considered as a factor in determining ultimate level of attrition. Evidence was found of a sensitive period up to age eight for the successful acquisition and long-term maintenance of knowledge of the principles of reflexive binding, even upon loss of exposure to the L2.

**Statement of Copyright**

## Acknowledgements

After two languages, three cultures, and five years, this is both the easiest and the hardest page to write.

First and foremost my thanks goes to my supervisor Martha Young-Scholten for her unfailing support, critical comments, and reliability which made the successful conclusion of this project certain.

My thanks too go to the faculty and staff in the Department of Linguistics and English Language at Durham. During my long association with the department, they have consistently provided both a stimulating academic environment within which to work, and a refreshing social atmosphere within which to relax. Their accessibility and encouragement over the years have kept me coming back.

The six Obirin students who worked with me as test subjects, and the six native speakers who served as my control group also have my sincerest thanks. Without their overall commitment, this study would never have been possible. Arigatou gozaimasita.

I must also thank my family for their support in this project. My parents George and Martha McCormack could not have provided a more conducive environment within which to complete this thesis. My brothers More, Aidan, Colm, and Fursey have all lent their energies and enthusiasm in more ways than they might imagine. Their families and friends have provided a normalcy and grounding which has kept life in perspective.

My colleagues at Teachers College have been unfailingly supportive and generous in providing time for me to complete this thesis. And my colleagues in Japan, especially Boyd Davis, Patricia Galien, Henri Hatayama, Ed Riley, and Terry Royce, also have my sincerest gratitude for their unstinting loyalty and support over the years.

But finally, I must turn and thank my wife Yasuko who, throughout this project, has endured far more than anyone should be asked to. This thesis is dedicated to you.

## Table of Contents

**Appendices**

# Chapter 1: Introduction

## 1.0. Towards a principled account of second language attrition

This thesis reports on a longitudinal study of adult second language attrition. More specifically, it examines the effect of lack of exposure to second language (L2) English on the binding constraints of reflexive pronouns in six proficient L2 English speakers. The lack of exposure in this case is the result of individuals moving from the United States, the L2 environment, back to their first language (L1) environment, Japan. The analysis of the data specifically focuses on the binding principles which govern the relationship between reflexive pronouns such as *himself* and *herself*, and their potential antecedents. The results of the study demonstrate that, despite the apparent attrition of control over some aspects of reflexive binding, the attrited forms which emerge remain within the bounds of binding types allowed by the principles and parameters framework of Universal Grammar (UG).

Language attrition may generally refer to "the loss of any language or any portion of a language by an individual or speech community" (Freed, 1982: 1). Attrition can occur in the first or second language of individuals, or in entire linguistic communities. Pathological language loss can occur in individuals as the result of damage to the brain due to stroke, accident, etc. Although still a minor field when compared with other areas of linguistic research such as first and second language acquisition and sociolinguistics, an increasing number of publications addressing language attrition have appeared since the early 1980s. The majority of these studies, however, have focused on the loss of first language skills. They address issues concerning factors related to bilingualism and immigrant first language loss (Maher, 1991; Silva-Corvalan, 1991); aphasic language loss (e.g., Paradis, 1977), and sociolinguistic studies of language shift and language death (Dorian, 1981; Schmidt, 1985).

Second language attrition has received considerably less attention than has first language attrition. Early isolated studies (Kennedy, 1932; Geoghehan, 1950) tended to focus on foreign language learners' loss of lexical items over summer vacations, but these studies failed to develop any wide-ranging research interest. The first significant publication on L2 attrition was the result of the May 1980 conference on foreign language loss at the University of Pennsylvania, the proceedings of which appeared in

Lambert and Freed (1982). Since then, refereed journal articles, doctoral theses and several books have appeared which address the loss of second or foreign language skills.

Research into second language attrition is significant for several reasons. At one level, issues concerning language teaching pedagogy, bilingualism, and sociolinguistics can benefit from L2 attrition studies. Understanding the degree to which different linguistic areas such as word classes or syntactic constructions are vulnerable to attrition can help L2 curriculum designers stem second or foreign language loss. Explaining the sociolinguistic causes and impacts of L2 loss can help bilingual L2 attriters understand their place and value in the L1-dominant society.

At another level, investigations of a psycholinguistic nature try to explain phenomena related to the cognitive processes responsible for second language attrition once L2 input ceases and the language becomes prone to fossilization or attrition. Investigations of lexical and morphological loss have described the symptoms of such deterioration, both in L1 and L2 attrition (see Yoshitomi, 1994). But little research has tried to explain how the underlying linguistic competence of an attriter is reshaped during attrition. Sharwood Smith (1983c), in discussing L1 attrition as a consequence of L2 acquisition, was one of the first to call attention to the psycholinguistic nature of diverging competences in language attrition, whereby competence in the attriting language diverges from the native-speaker norm to a new, deviant form. By investigating whether the process of attrition is somehow constrained and results in a restructured, principled, yet non-target like grammar, we can ask if Universal Grammar (UG) constrains not only acquisition, but also attrition. And by considering the data presented here in light of theories of L2 acquisition as proposed by researchers such as White (1989a; 1989b), this thesis will investigate the possibility that, despite deviations from the native norm, an attrited (inter-) language nonetheless conforms to UG sanctioned rules. In the next section I will outline the main research goals of the investigation.

### 1.1. Research approach and goals

This present study investigates six Japanese test subjects' decline of knowledge of English reflexive binding over various periods of time. These periods range from 9

months to more than 16 months, depending on the test subject. The decision to conduct a longitudinal study was made for several reasons. The most important reason was the lack of longitudinal L2 attrition studies. As the literature review in Chapter 3 will show, most L2 attrition studies are cross-sectional in nature, similar to many L2 acquisition studies. Although useful for examining broad-ranging phenomena such as lexical or morphological attrition within a sample population, cross-sectional studies cannot capture changes at the individual level. Another reason to conduct a longitudinal study was that the number of variables which need to be controlled for in an attrition study limit the number of suitable informant candidates. The length and nature of the pre-attrition L2 acquisition period, the linguistic environment in which L2 acquisition occurs, and the age and L2 proficiency level at the cessation of exposure to the L2 are some of the variables which complicate the L2 attrition research process. Two further complications are that the initial level of the test subjects must be high, and they must be willing to be involved with the study for at least a year. As it turned out, finding test subjects was difficult, and even then, English levels as reflected in TOEFL scores was slightly compromised in two cases. Nonetheless, a high level of proficiency with respect to binding was observed immediately following cessation of exposure to English in all six test subjects, and they generously participated in the research project.

The phenomenon under investigation here is the binding of reflexive pronouns (anaphors) to antecedents. Reflexive anaphors have received considerable attention since Chomsky's *Lectures on Government and Binding* (1981). L1 binding research examines the referential qualities of nominals such as pronomials (e.g., *she*, *you*, *they*), reflexives (e.g., *ourselves, himself)* and R-expressions (referents such as *John,* or *the teacher*) (Otsu, 1981; Aoun, 1985; Lasnik and Crain, 1985; Wexler and Manzini, 1987; Grodzinsky and Reinhart, 1993; Aikawa, 1999). Second language acquisition research also includes a considerable number of studies into L2 learners' acquisition, knowledge of, and adherence to binding principles (Thomas, 1989; 1991; 1995; Hirakawa, 1990; Finer and Broselow, 1986; Broselow and Finer, 1991; Yusa, 1998). This attention to reflexive anaphora in both L1 and L2 research reflects the importance modern linguistics places on it as an observable and testable principle of UG.

In the current investigation into L2 reflexive binding, tests were designed and written which asked test subjects to choose whether a particular binding pattern was acceptable or not. There were two test types, truth value judgement tests and

grammaticality judgement tests. Both attempted to test the same two phenomena. These were adherence to the local binding restriction in English biclausal sentences such as (1), and adherence to the Proper Antecedent Principle outlined by Wexler and Manzini (1987) which allows either a subject or object NP to serve as the proper antecedent of a reflexive in monoclausal sentences such as (2).

(1) $Jack_i$ thought $Sam_j$ blamed $himself_j$

(2) $Alice_i$ showed $Rebecca_j$ some photographs of $herself_{i/j}$.

In (1) *Sam* is coindexed with *himself* and serves as the correct antecedent of *himself*. *Jack* is not a potential antecedent as English allows only local binding and not long distance binding (back to the matrix subject NP in (1)). Example (2) is a monoclausal sentence in which either *Alice*, the subject NP, or *Rebecca*, the object NP, can serve as the grammatical antecedent of *herself*. These facts will be further discussed in Chapters 4 and 5.

An important goal of this study is to address three questions regarding L2 attrition. When exposure to the L2 ceases,

Question 1

Do adult L2 English speakers exhibit evidence of loss of knowledge of English-specific principles of reflexive binding?

Question 2

Do patterns involved in this loss of reflexive binding conform to UG constraints?

Question 3

Does age at first exposure to the L2 seem to play a role in attrition patterns?

Question 1 asks if test subjects lose knowledge of English binding principles. Specifically, control of aspects of principle A of the binding theory is targeted in this study. This is significant in that previous L2 attrition literature has never, to my knowledge, addressed the effects of lack of L2 exposure on UG-established properties of language.

Question 2 addresses the nature of the process of the loss of adult L2 binding. Specifically we are looking to see if the attrited grammar will conform to UG constraints on binding.

Question 3 addresses the role of age in second language acquisition. Differences in the patterns and degrees of loss exhibited by the test subjects may offer insights into how exposure before and after the presumed critical period effects the robustness of L2 UG generated knowledge .

## 1.2. Outline of the thesis

The thesis is organized into five main parts. Chapter 2 outlines and discusses aspects and issues of the generative approach to the acquisition and knowledge of a second language. In this chapter we will review some of the central claims made in generative L2 acquisition research, especially arguments about the extent to which L2 learners have access to UG and the influence the L1 has on L2 acquisition. This research will be central to the arguments made concerning the attrition data. In particular, it will be used to support my proposal that attrition of second language binding principles results in a grammar which diverges from the native-like norm, but which is nonetheless UG-sanctioned.

Chapter 3 reviews the literature on language attrition. The purpose of this review is to familiarize the reader with both the types of investigations which have been made in the field and the explanations of attrition which have been put forth. This review will draw attention to the lack of a theoretical basis for most research carried out on attrition, and proposes a framework for investigating L2 binding attrition.

Chapter 4 reviews work which has been carried out on L2 reflexive binding. In this chapter we will look at various studies and consider the phenomena investigated, the

subjects involved, the tests used, and the results obtained. This review will provide the necessary background to the experiment reported on in Chapters 5 and 6.

Chapter 5 provides a detailed description of the two tests used and the subjects included in this experiment. This chapter gives the rationale behind each test and includes examples of the various stimulus sentence types used in each test.

In Chapter 6 I give a descriptive analysis of the data results, followed by a discussion of the findings. This analysis includes an exhaustive account of each of the six test subjects' results by test and by sentence type. The results of the tests are then combined and discussed in a way which hopefully provides a meaningful portrait of the effects of under-exposure on the test subjects' L2 English reflexive anaphors.

Chapter 7 concludes with a brief summary of the findings, some problems observed with the study, and suggestions for possible future investigation into the field of second language attrition.

This study is significant for two reasons. One reason is that the longitudinal nature of the study makes it unique in L2 attrition research, a field dominated by cross-sectional studies. Secondly, and more importantly, it is the first example, to my knowledge, of an L2 attrition study carried out within a generative framework.

# Chapter 2: Universal Grammar and Second Language Acquisition

## 2.0. Introduction

In order to answer the questions posed in Chapter 1, a clear account of the theory of language adopted in this thesis must be offered. This chapter begins by reviewing the historical context from which the field of generative linguistics emerged in the late 1950s as an explanation of first language acquisition. The chapter will then consider how aspects of this theory have been applied to second language acquisition, and how the theory can be used to account for second language attrition data.

## 2.1. The early years

Linguistic research throughout the nineteenth and early twentieth centuries was dominated by the comparison and classification of mostly Indo-European languages. Grammar was considered a branch of logic, and Latin and ancient Greek, as well as Sanskrit, were thought of as close to the ideal grammar. As new languages such as native American languages were 'discovered' in the twentieth-century, a more practical interest led linguists to identify common levels of phonetic, phonological, morphological and syntactic structure by which languages are organized. By the 1940s, structuralist linguists such as Preston (1948) and Hockett (1958) were convinced that languages should be analyzed according to these different levels of organization based on data collected from spoken language.

Despite the confidence of structuralist linguistics in its goals and methods at the time, relatively little attention was given to how language was learned. It was assumed that language acquisition was entirely an inductive process of building up the various structural elements from phonemes, to morphemes, etc., in a hierarchical fashion. Behaviorist psychology was one explanation offered for the processes by which acquisition occurs. A child's final linguistic ability was thought to be simply the result of the language he or she was exposed to. Spoken language was viewed as a response to a stimulus, either external or internal, which triggered the appropriate language from the corpus of language acquired during childhood. Appropriate language was then reinforced

through encouragement and rewards. An external stimulus could be the words spoken by another person, or it could be something which was read, seen, felt, smelled, etc. An internal stimulus was considered as coming from inside a person, such as a feeling of hunger which would elicit an oral response. But no explanation was offered as to how a "response" such as "I'm hungry" was formed.

Language is, however, more complicated than just phonemes and morphemes. And speaking is more than simply a stimulus/response/reinforcement mechanism. By the late 1950s, structural linguistics was faltering. Endless lists of phonemes and morphemes had no unifying theory and were unable to offer any new insights into language. In 1957, Chomsky published *Syntactic Structures* in which he outlined his intent to arrive at a "formalized general theory of linguistic structure and...explore the foundations of such a theory." (p. 5). In 1959, Chomsky's critical review of B.F. Skinner's (1957) *Verbal Behavior* caused the goals and methods of structuralist linguistics to be seriously questioned and eventually end the dominance structural linguistics had had for some forty years.

## 2.2. Generative linguistics

The knowledge a native speaker has of his or her language implies knowing a system of linguistic rules. For generative linguistics, knowledge of language is an adult's L1 linguistic competence, defined as "the speaker-hearer's knowledge of his language" (Chomsky, 1965: 4). The acquisition of this competence is guided by a "genetically determined language faculty..., an innate component of the human mind that yields a particular language through interaction with presented experience, a device that converts experience into a system of knowledge attained: knowledge of one or another language" (Chomsky, 1986b: 3).

### 2.2.1. *Knowledge of language*

Knowledge of language, or linguistic competence, is thought of as "a particular generative grammar, a theory concerned with the state of the mind/brain of the person

who knows a particular language" (Chomsky, 1986b: 3). "Generative" here is used simply to mean 'describe', 'define', 'give an explicit account of', or 'analyze' a particular grammar. Green and Morgan (1996) illustrate this meaning by referring to the formula for the graph of a circle as (1):

(1) $(x-a)^2 + (y-b)^2 = c^2$

This formula

> generates (i.e., describes) a set of circles in a plane defined by the *x* and *y* axes, a potentially infinite set, depending on the values assigned to *a* and *b*. ...[T]he formula is not a circle-producing machine; it doesn't produce circles, it only defines them. (Green & Morgan, 1996: 6)

In the same way then, the grammar of a particular language can be thought of as a formula which defines or 'generates' a potentially infinite set of grammatical sentences for a language. It is the knowledge of this unconscious formula, or grammar, which underlies language use and which generative linguists attempt to discover.

The principles and rules of this abstract grammar account for such formal properties of language as phonology, morphology, and syntax. White (1989b) considers several examples of this unconscious knowledge in English. These are briefly reviewed below.

Pronunciation of the plural form in English is governed by a rule which requires word-final voiced sounds to be followed by the voiced form of the regular English plural [z], as in *bags*, *cabs*, or *cads*. On the other hand, word-final voiceless sounds must be followed by the voiceless plural form [s], as in *backs*, *caps*, or *cats*.

An example of a morphological rule is the distribution of the suffix *-able*. This suffix forms an adjective when attached to a verb, but such attachment is generally limited to transitive verbs. Thus, while forms such as *breakable*, *usable*, or *drinkable*

are perfectly acceptable, native English speakers unconsciously know when *-able* attachment will not work: **dieable*, **goable*, **sleepable*, etc. Note that these words are not rejected because of meaninglessness, but rather on grounds of the transitive/intransitive rule governing *-able* attachment.

As for a syntactic rule, English assigns different functions to morphemes such as *up* or *down* as in:

(2) a. John looked up the telephone number.
b. John looked the telephone number up.

(3) a. John looked up the hill.
b. *John looked the hill up.

(4) a. Mary turned down my idea.
b Mary turned my idea down.

(5) a. Mary turned down the street.
b. *Mary turned the street down.

In these examples *up* and *down* serve as prepositions in (3) and (5) and as particles in (2) and (4). Only when they function as particles can they be separated from the verb.

These three simplified examples demonstrate the type of systematic and rule governed knowledge adult native English speakers know at an unconscious level.

A main goal of generative linguistics is to understand and explain the rules underlying steady-state adult native speaker grammar. But another goal of linguistics is to also explain how children acquire this knowledge. Chomsky argues that the acquisition of a grammar is guided by an innate, genetically determined mechanism or system, usually referred to as Universal Grammar (UG). Rather than containing all the rules of all languages, UG can be described as "the set of properties, conditions, or

whatever, that constitutes the 'initial' state of the language learner, hence the basis on which knowledge of language develops" (Chomsky, 1980b: 69).

UG, then, is 'universal' in the sense that it is broad enough to underlie all natural languages and is not specific to particular languages. Exposure to a particular language acts as a trigger on UG, providing it with the language specific information necessary for full L1 development.

### 2.2.2. *Acquisition of language*

The task of all children when learning their native language is, through primary linguistic data (PLD) (utterances the child is exposed to) in the target language, to acquire an adult grammar which allows them to understand and speak the language with native fluency. A logical problem with this is that the linguistic competence children eventually acquire goes beyond the PLD. This does not simply mean that children don't hear all the words they eventually come to know as adults, or that they are not exposed to, say for English, all the irregular verb or plural noun forms. Rather, children come to acquire the ability to distinguish between grammatical and ungrammatical sentences, to recognize ambiguity in sentences, or to understand that several sentences are paraphrases of each other. And they acquire this knowledge despite there being many "legitimate and acceptable sentence-types (which) may never occur in a child's linguistic experience" (Hornstein and Lightfoot, 1981: 10)

A key motivation in arguing for innate linguistic knowledge - UG - is that, despite receiving impoverished input, all child L1 acquirers of the same L1 come to know a grammar which goes beyond the input they hear. Specifically, three logical problems are associated with impoverished input:

- underdetermination
- degeneracy
- lack of negative evidence

Let us look briefly at each of these in turn.

2.2.2.1. *Underdetermination*

Underdetermination, also known as the poverty of the stimulus argument, is the term used to describe the paradox in which the nature and amount of linguistic data a child encounters during the acquisition years is nowhere near the range of utterances the child comes to produce and understand as an adult.

One example of underdetermination is the structure-dependence of rules, outlined by Chomsky (1965: 56). Structure dependence is knowledge of language based on rules which operate on structural units (phrases), rather than on linear sequential units. A linear rule system would operate by specifying whether the first, second, last word, etc. was to be moved to form a question, passive, etc. Presumably children are never taught this rule, and are at least certainly not taught that a linear rule system will not produce grammatical sentences.

Example (6) is an example of a "left-most verb preposing" rule for question formation, which produces (7).

(6) Jill's dress is green.

(7) Is Jill's dress green?

But this would lead children to wrongly transform sentences such as (8) into questions such as (9).

(8) The woman who is tall is Jill's mother.

(9) *Is the woman who tall is Jill's mother?

Children seem to grasp the concept of subject – auxiliary (AUX) inversion (Radford 1988), whereby the auxiliary verb is inverted with the subject NP to produce (10).

(10) Is the woman who is tall Jill's mother?

Inversion questions created from sentences such as (8) rely on knowledge of structure, not just on word order. Thus, despite the underdetermined nature of input that children receive, their knowledge of structure dependence lends support to the existence of an innate language faculty.

In a study designed to critically examine the structure dependence principle, Crain and Nakayama (1987) conducted a series of elicited production experiments which tested three- to five-year-old children's adherence to the subject/auxiliary inversion rule. Their initial test used sentences similar to (8) above and supported structure dependence. But Crain and Nakayama doubted the reliability of the tests, as the children produced errors such as (11).

(11) *Is the woman who is tall is Jill's mother?

Because the embedded verb and the AUX are both *is*, and each is left in the question by the child informants, the experimenters cannot be sure which verb has been moved. To control for this, they designed another test which included stimulus sentences in which the embedded verb and AUX were different in each clause, as in (12).

(12) The boy who was working hard is asleep.

Even if children fail to delete the moved verb during question formation, it is clear which verb was moved. Reliance on a linear analysis will result in (13), while a structure dependent analysis will result in (14).

(13) *Was the boy who working hard is asleep?

(14) *Is the boy who was working hard is asleep.

Despite the ungrammaticality of (13) and (14), it is clear which verb was moved. A linear analysis results in *was* being moved in (13) and a structure dependent analysis results in *is* being moved in (14). Crain and Nakayama's (1987) results clearly show inversion of AUX and no instances of fronting of the embedded verb. Their results using this type of sentence support "Chomsky's claim that only structure-dependent rules are formulated in language acquisition" (p. 542).

#### 2.2.2.2. *Degeneracy*

Another argument for the existence of an innate mechanism which aids the L1 acquisition task is that the learner receives inaccurate (degenerate) input as well as accurate input. Although research has shown that language directed at children is largely accurate (Newport, Gleitman and Gleitman 1977), the language of other speakers (children or adults) overheard by child learners may often be ill-formed grammatically or phonologically. Yet despite degenerate input, children's errors are such that each successive child grammar is thought to be UG constrained, and all normal children ultimately acquire their complete L1 grammar.

#### 2.2.2.3. *Lack of negative evidence*

A third input-based problem which supports the argument for UG is the question of how children come to distinguish between grammatical and ungrammatical forms in the L1. One possibility is that they are explicitly told something is ungrammatical when they make an error. This is known as negative evidence, which contrasts with positive evidence, evidence derived from the L1 environment about what is possible in the language. There are, however, three problems with negative evidence in that it assumes:

a) children consistently get negative evidence from adults

b) children make all relevant errors and are corrected

c) children remember and make use of the corrections

Research suggests, however, that negative evidence is haphazard and unreliable at best and contributes little to linguistic knowledge. Braine (1971) points out that most correction of child language is of a semantic, rather than grammatical/phonological nature, and even when children are corrected, they usually pay no attention to the correction.

Thus, an inductive theory of L1 acquisition which argues that all linguistic knowledge is derived from input seems unable to account for why children come to recognize what is ungrammatical in their language, and the 100% success rate of native speakers to acquire their L1. The input problem seems to point to some innate facility/propensity which facilitates L1 acquisition: UG.

### 2.3. The role of UG in second language acquisition

As discussed above, behaviorist psychologists such as Skinner (1957) believed that language was empirically-based habit formation which led to the production of patterns which made up a particular language. The language could be the L1 or an L2. To the behaviorist, L2 errors were attributed to the old habits of the L1 interfering with the forming of new L2 habits. From this view of behaviorist language acquisition came the Contrastive Analysis Hypothesis of second language learning which, in its strong form, stated that all L2 errors were caused by interference from the L1 (Fries, 1945; Lado, 1957).

This view of second language acquisition was questioned in the wake of Chomsky's revolutionary views on first language acquisition. In 1967 Corder argued that many L2 errors were unexplainable in terms of Contrastive Analysis. He suggested that the processes responsible for L1 acquisition may also be responsible for L2 acquisition. He proposed that, rather than stemming only from differences between the L1 and the L2, many errors which occurred during the SLA process originated with the L2 learners themselves as they tested ideas —hypotheses— about the target language.

Studies by researchers such as Gass (1979), Ioup and Kruse (1977), Kellerman (1979), and Schachter (1974) began investigating not only the role of the L1 in L2 acquisition, but also "ways in which principles independent of L1 knowledge determine patterns of acquisition" (Flynn, 1989: 89).

SLA theory-building during the last fifteen years has focused increasingly on examining whether UG might be at least partly responsible for second language acqu\isition, especially by adult learners. The 1980s saw researchers such as Cook (1985), Flynn (1986, 1987), Liceras (1986, 1989), Van Buren and Sharwood Smith (1985), and especially White (1985, 1986a, 1989a) investigate the possibility of UG playing a role in adult L2 acquisition. These studies identify a UG determined phenomenon and test the accuracy with which L2 learners have acquired that construction. L2 acquisition investigations have included UG structures such as the prodrop parameter (Phinney, 1987), headedness (duPlessis, Solin, Travis and White, 1987), Subjacency (Bley-Vroman, Felix and Ioup, 1988; Schachter, 1989), as well as reflexive binding (Finer and Broselow, 1986; Thomas, 1989).

More recent SLA studies have focused on the various states of an L2 learner's language-interlanguage (IL). Since Selinker (1972) introduced the term, IL research investigates the hypothesis that L2 acquisition consists of a series of different states of language development which are distinct from the L1, and that these states roughly correspond to each other among learners of a given L2 (Bailey, Madden and Krashen, 1974). Here again we see a correlation with L1 acquisition in that the L2 acquisition process also exhibits various stages of development, different from the final-state L1, and which are thought to proceed in a roughly similar sequence among acquirers with the same L1 background, acquiring a particular L2.

The following sections will present arguments for investigating what role UG might have in SLA. These arguments will be considered in light of significant UG-based SLA studies that have appeared in the literature.

### 2.3.1. *The logical problem of SLA*

Chomsky (1986b: 3) asks "*What constitutes knowledge of language?*" Although he is referring to L1 knowledge, we may also ask this question of L2 speakers. Just

like children acquiring their native language, L2 learners of all ages also face the problem of making sense of the linguistic input they are exposed to. If adult L2 acquisition were to proceed through similar stages as child L1 learners, and if the adults arrived at a final steady state comparable to native speakers, L2 acquisition research would have a strong argument for adult access to UG. However, few documented cases of adult L2 acquisition result in target-like levels of competence. Despite this general failure for adult L2 learners to attain native-like proficiency, they nonetheless can and do acquire *complex* knowledge of a second language, in many cases to high levels. We can refer to this as the logical problem of L2 acquisition (Bley-Vroman, 1989).

A language spoken as an L2 must share linguistic properties of that language spoken as an L1. And just as L1 knowledge of that language is represented in the mind as a system of principles and rules, so too must knowledge of that language as an L2 be represented in the mind as a system of principles and rules. However, the end steady-state of the L2 will invariably differ from the L1 end-state. Researchers have found that L2 acquisition is systematic and rule governed, and culminates in a steady state grammar which, although most likely different from the target language grammar, nonetheless represents knowledge of an abstract rule system which accounts for the L2 learner's competence.[1] The problem that L2 learners face, however, is qualitatively different from that which children face when learning their native tongue. L2 learners already have a fully formed language in place and adult learners also have highly developed skills for abstract problem solving.

### 2.3.2. *Access to UG in second language acquisition*

Generative research into adult L2 acquisition addresses three possibilities for the role of UG in L2 acquisition in older learners. The first is the no-access hypothesis which claims that no aspect of UG is available to the L2 learner beyond a certain age. The second is the indirect-access hypothesis. This claims that only UG principles common to all languages, such as structure dependence are available to adult learners,

[1] White (1989b) provides an account of the possible causes of differences between first and second language acquisition. These differences include degree of success, L1 influence, fossilization, input, and age. See White (pp. 41-45) for discussion.

but parametric variation is not. The third proposal, the full-access hypothesis, argues that UG in its entirety is available to learners during L2 acquisition no matter what their age. For the full access hypothesis, other cognitive processing variables such as memory are held responsible for general failure to achieve native-level proficiency.

Empirical research has provided evidence that L2 learner age is a critical factor in ultimate L2 attainment and that younger learners acquire a more complete L2 than do older learners. Since Lenneberg (1967), reports in the literature have supported the Critical Period Hypothesis. Studies of L2 phonology (Asher and Garcia, 1969; Oyama, 1976; Long, 1990; Flege, 1999) show that first exposure after as early an age as six years may lead to non-native-like pronunciation. Studies of L2 syntax (Coppieters, 1987; Johnson and Newport, 1989; Patkowski, 1980; Schachter 1990) indicate that the critical period for native-like attainment of morphology and syntax may be later, possibly around age 15. One counter example is Ioup, Boustagui, Tigi, and Moselle (1994) who found near-native proficiency in the L2 Arabic of two L1 English-speaking women whose L2 acquisition had not started until both were in their early 20s. Ioup et al attribute their success to high motivation and L2 environmental exposure, and an attention to grammatical form. Nevertheless, researchers such as Bley-Vroman (1989; 1990) and Clahsen and Muysken (1986) argue that, given the almost guaranteed failure of complete native-like acquisition of the target language, adult L2 learners have no access to UG and rely on general principles of information processing and problem solving (Meisel, 1991: 243). These arguments are partially based on Lenneberg's (1967) Critical Period Hypothesis which claims that after puberty the brain loses its "plasticity" and is unable to acquire language to a native-speaker level.

Long (1990), in a review of maturational constraints on L2 development, takes this distinction further. He suggests that rather than one or two specific ages after which native-like attainment is unlikely, there is a series of "sensitive" periods, each period corresponding to various aspects of the target language. Long suggests these aspects include phonology, morpho-syntax, lexis, collocation, etc.

For SLA researchers who claim UG is no longer available, L2 learners must, in part, make analogies from their L1 grammar to the L2 grammar and in part use general cognitive learning skills such as memorization and logical reasoning to acquire the language (Bley-Vroman, 1989).

In denying the L2 acquisition process access to UG, Bley-Vroman (1989) argues for a Fundamental Difference Hypothesis which states that L1 acquisition and L2 acquisition differ on internal, linguistic and qualitative grounds. He argues that the difference is internal in that it is "caused by differences in the internal cognitive state of adults versus children, not by some external factor or factors (insufficient input, for example)" (p. 50). This difference in cognitive states he continues, is linguistic as it is the language faculty which undergoes change, rather than general changes in general learning abilities. For Bley-Vroman the differences between L1 and L2 acquisition are qualitative rather than simply quantitative because "the domain-specific acquisition system is not just attenuated, it is unavailable" (p.50).

Clahsen and Muysken also argue that UG is unavailable to L2 learners. Rather than a conceptually-based argument like Bley-Vroman's, however, theirs is empirically based. Clahsen and Muysken (1986, 1989), and also Clahsen (1988) base their claims on a comparison of child L1 German acquisition data and adult L2 German acquisition data. They find that certain properties such as verb second and subject-verb agreement in L1 acquisition occur "in one single developmental phase" (Clahsen and Muysken, 1989: 24). Their argument against L2 access to UG is that L2 learners fail to acquire these constraints in clusters and that "the placement and inflection of verbs in German are separate acquisitional tasks" (p. 51).

Although Bley-Vroman and Clahsen and Muysken approach the issue quite differently, they both argue that UG is not available to constrain adult second language acquisition. But they fail to offer adequate, alternative explanations of L2A. Even if adult L2 learners no longer have access to UG, they nonetheless do come to know a system of morphosyntax and phonology which approximates the target language to varying degrees. And in some instances, the L2 is seemingly effortlessly acquired such as Obler's (1989) report of "CJ", who apparently "picked up" various languages as he traveled from country to country, or Asher and Garcia's (1969) experiment which claimed learners using the Total Physical Response teaching method outperformed control students in "regular" language classrooms over the same period of time.

The second proposal regarding the role of UG in SLA is that UG is only partially available. The access that is available will either be a principle of UG which exists in all languages, or parameter settings which coincide with the L1 of the learner. Two studies which claim to support this position are Bley-Vroman, Felix, and Ioup (1988) and

Schachter (1989). Bley-Vroman et al. use grammaticality judgements to test acceptance and rejection of Subjacency and ECP violations in proficient Korean speakers of L2 English. Subjacency restricts syntactic movement in such a way that phrases are constrained from moving over more than one bounding node in a single operation. (Sportiche, 1981; Rizzi, 1982b). The bounding nodes which limit movement differ from language to language. In English, bounding nodes are thought to be DP and AGRP. In a sentence such as (15)

(15) who$_i$ did [$_{AGRP}$Carol think [$_{CP}$[$_{AGRP}$[Joe loved $t_i$]]]]

the specifier in the lower CP is empty and thus prevents an illegal movement because it provides a stopping-off point for the wh-element along the way up the sentence. In this way, the sentence does not violate Subjacency.

A sentence such as (16), however,

(16) *who$_i$ did [$_{AGRP}$Carol think [$_{CP}$when$_j$ [$_{AGRP}$ $t_i$ loved Joe $t_j$]]]

has the lower CP specifier already filled by another wh-element (*when)* and cannot provide a stopping-off point for the other wh-element. (16) thus violates the Subjacency principle. In (17), the lower CP is empty to cross, but the movement of the *wh-* element is blocked from crossing two bounding nodes, DP and AGRP, causing (17) to be ungrammatical.

(17) *who$_i$ did [$_{AGRP}$Carol believe [$_{DP}$ the possibility [$_{CP}$ $t_i$ that [$_{AGRP}$ Joe loved $t_i$]]]]

When an item such as a *wh*-word is moved, it is assumed to leave a *trace* at the site from which is moved. These traces are said to be coindexed with the moved item and must be licensed by being properly governed by either lexical heads or the

antecedent of the trace. This is known as the Empty Category Principle (ECP) (Postal 1967).

Proper government here is defined according to Chomsky (1981): α properly governs ß if and only if α governs ß and α is a lexical category (N, V, A, P) (head government), or α and ß are co-indexed (antecedent government) (p. 250).

In English, depending on whether the object or the subject in a lower clause is questioned, there are differences in the behavior of complementizers. Consider the four sentences in (18):

(18) a. Who$_i$ do you think that Paul met $t_i$ last week?
b. Who$_i$ do you think Paul met $t_i$ last week?
c.*Who$_i$ do you think that $t_i$ met Paul last week?
d. Who$_i$ do you think $t_i$ met Paul last week?

In sentences such as (18a) and (18b), the object *wh*-phrase can move out of the lower clause and the complementizer (COMP) *that* may be optionally present or absent. In sentences (18c) and (18d), however, we can see that the subject *wh*-phrase can only move out of the lower clause if the complementizer *that* is absent. This is referred to the *that-trace effect*. The ungrammaticality of (18c) is caused by proper government being blocked by the presence of *that* in COMP.

Although Bley-Vroman et al.'s results show their test subjects perform better than chance on recognizing grammatical Subjacency and ECP restrictions and rejecting ungrammatical ones, but they still perform significantly below the native speaker control group. This leads the researchers to conclude that although UG seems available to adult L2 learners, it is not fully available, possibly being blocked or impaired by other cognitive systems such as general problem solving skills, or even the L1 itself. This study will be discussed in more detail below in section 2.4.1.2.

Schachter's (1989) study also investigates Subjacency, but with a variety of adult L1 speakers: Korean, Chinese and Indonesian. Her results do not show overwhelming evidence of obedience to Subjacency restrictions in English and she therefore argues

they cast doubt on claims that adult L2 learners have full access to UG. Schachter's results will also be further discussed below.

The results of both Bley-Vroman et al. and Schachter point to some role for UG in adult L2 acquisition and constitute a challenge to the position of researchers such as Clahsen and Muysken.

The third position on the role of UG in SLA is that the second language acquisition process is fully constrained by UG (White, 1992). Within this position, there are two distinct approaches based on differing views of first language acquisition. L1 acquisition researchers such as Hyams (1994) and de Villiers and Roeper (1995) argue for a Strong Continuity Approach, under which all functional projections, including CP, are in place at the start of L1 acquisition. The Strong Continuity Approach assumes that child language is fundamentally the same as adult language. Under the Weak Continuity Approach, researchers such as Clahsen, Eisenbeiss & Vainikka (1994) argue that, although certain aspects of UG such as functional categories are available to child learners, they need positive input to project these categories.

Following the Strong Continuity Approach, Schwartz & Sprouse (1996) and Schwartz (1998) argue for a Full Transfer/Full Access hypothesis in which "the initial state of L2 acquisition is the final state of L1 acquisition" (Schwartz and Sprouse, 1996: 40-41). According to Full Transfer/Full Access, an L2 learner initially assumes the L1 principle and parameter settings for the target language and in the face of positive evidence, gradually restructures this grammar to suit that of the L2. This restructuring results in an interlanguage continuum which is "determined in part by the initial state, in part by input, in part by the apparatus of UG and in part by learnability considerations" (Schwartz and Sprouse 1996: 41).

According to the L2 version of the Weak Continuity Approach proposed by Vainikka & Young-Scholten (1994, 1996), second language learners initially transfer only the lexical categories from their L1 to their L2 and not the fully projected trees Schwartz and Sprouse argue for. For Vainikka & Young-Scholten, 'Minimal Trees' explain developmental stages of L2 acquisition in that learners posit only the minimum amount of linguistic structure needed to analyze the L2 input (Vainikka & Young-Scholten, 1994).

Whether one subscribes to Full Transfer/Full Access or to 'Minimal Trees', both proposals assume a significant role for UG in second language acquisition. Both

positions not only assert that adult L2 acquirers have access to UG, but they also rely on L2 learners initially using their knowledge of L1 to establish the L2 though this is limited to lexical projections for Vainikka & Young-Scholten. And in both cases, when positive evidence clearly points to a different setting for a principle or parameter, or for positing a projection, acquisition will occur.

## 2.4. Investigating UG in SLA

An empirical investigation of any theory needs to focus on observable phenomena which can be described and measured. Especially when looking at something as abstract as UG, identifying such phenomena is crucial.

The principles and parameters theory identifies a wide range of such phenomena which researchers can use to investigate the role of UG in language acquisition. This theory recognizes that languages differ "in terms of which of the set of possible universals they actually realize" (White, 1989b: 29). An example of this is that, in principle, a language could incorporate all the possible sounds which occur in all the world's languages (a finite set), from the clicks of !Xhosa, to the tones of Chinese and the /l/ and /r/ distinction of English. But this does not happen. Rather, UG selects from the universal set of phonetic features and establishes for a particular language a subset of this larger set.

Parameters explain differences between languages in terms of a set of UG determined linguistic principles which can be set one way or another. The setting of a parameter is thought to trigger other linguistic phenomena, and a particular setting is established – triggered by – the positive input the child receives during L1 acquisition.

One example of a parameter is the *prodrop* or *null subject* parameter which accounts for the largely optional nature of subject pronouns in some languages such as Spanish or Italian. Other languages, such as English, always require subject pronouns. An example from [+prodrop] Italian and [-prodrop] English in (19a) and (b) illustrate this.

(19) a. Va al cinema sta sera
'goes to the movies tonight'

b. She is going to the movies tonight
*goes to the movies tonight

An interesting aspect of the prodrop parameter is that other properties are associated with the particular setting a language has. In particular, [+prodrop] languages allow subject-verb inversion as in (20a) and *that*-traces as in (20b. Languages which are [-prodrop] do not allow these phenomena

(20) a. E arrivata Laura
'is arrived Laura'
*has arrived Laura

b. Chi hai detto che e venuto/
'who ou have said that is come?'
*Whom did you say that came

In addition to these clustering effects, prodrop languages tend to be morphologically rich in order to compensate for the lack of subject pronouns, the person and number of the subject being carried by the verb morphology. As children are presumably not taught these rules when learning their L1, input from their linguistic environment is assumed to trigger the UG constrained mechanism (i.e., the prodrop parameters) which results in the observable phenomena.

The prodrop parameter will be examined in more detail below, but this brief illustration provides an example of the type of phenomena linguists use to explain the role of UG and input in L1 acquisition.

In the next sections, I will briefly discuss various principles and parameters which have been investigated in second language acquisition research. In the same way

that second language researchers have adopted first language acquisition research tools and phenomena, so too can second language attrition researchers use SLA research to investigate L2 loss.

### 2.4.1. *Principles*

As noted above, all languages are potentially constrained by principles defined by UG. These include structure-dependence, Subjacency, and the Empty Category Principle.

#### 2.4.1.1.*Structure dependence*

In section 2.4.1 above, we noted how linguistic rules operate on structural units (constituents), rather than on linear sequential units.

Turning to second language knowledge, we can ask if learners follow a similar UG driven structure-dependent rule system. One study which looks at this in connection with second language acquisition is Otsu and Naoi's (1986) investigation. They looked at whether the structure-dependent generalization was employed by 11 Japanese high school EFL students of approximately 15 years of age. Japanese has no syntactic movement in interrogative formation and structure-dependence plays no role in this construction. Test subjects were asked to take both a syntax test to check for presence of related structures such as simple declarative sentences, and a UG test which asked them to translate English declarative sentences into interrogatives. These tests support Otsu and Naoi's claim that adolescent L2 learners adhere to UG driven structure dependence restrictions in English. Two criticisms of this study made by White (1989b: 66) were that the learners are still quite young to consider as adult learners, and although Japanese does not have *wh-* movement, other aspects of Japanese structure dependence may guide the test subjects' English interrogative formation. Additionally, one must wonder what kind of primary linguistic data these children had prior to this experiment. Presumably their classroom English instruction was very grammar translation oriented. Perhaps this focus-on-form type of teaching methodology allowed learners to make use of

deductively learned aspects of English to correctly create interrogatives from declarative sentences, obviating any need for UG constraints.

2.4.1.2. *Subjacency and the Empty Category Principle*

In addition to structure dependence, two other principles of UG which have been investigated by SLA researchers are Subjacency (Schachter, 1988, 1989; White, 1992) and the ECP (Bley-Vroman et al, 1988). These have been described above in Section 2.3.2.

Schachter (1989) looked at Dutch speakers' adherence to Subjacency in English. Dutch and English obey Subjacency in the same way in that they observe the same bounding nodes. The Dutch speakers in Schachter's study consistently obeyed English Subjacency constraints. In Schachter's (1989) study, she again tested the Subjacency principle, but this time with speakers of Korean (n = 21), Chinese (n = 20) and Indonesian (n = 20), who were .all proficient ESL speakers. Schachter also included a control group of native English speakers (n = 19). Korean allows no *wh*-movement and therefore subjacency restrictions do not apply. Chinese allows some relative pronoun extraction, but no *wh*- movement. Indonesian allows *wh*- movement. The experiment consisted of both grammatical and ungrammatical versions of four different sentence types: sentential subject sentences (SS), relative clause sentences (RC), noun phrase complement sentences (NC), and embedded question sentences (EC). Examples of these are respectively given in (21a, b, c, and d) below.

(21) a. That oil prices will rise again this year is nearly certain.
*Which party did [for Sam to join *t*] shock his parents?

b. The theory we discussed yesterday will be on the exam next week.
*What did Susan visit the store [that had *t* in stock]?

c. There is a good possibility that we can obtain the information elsewhere.
*Who did the police have evidence [that the mayor murdered *t*]?

d. The dorm manager asked me who I wanted to have as a roomate.
*Who did the Senator ask the President [were he would send *t*]

Test subjects judged 24 of each sentence type to demonstrate their knowledge of Subjacency constraints in English. The results of the test complement the results of Schachter's test with the Dutch test subjects in that each L1 test subject group performed at a level which reflected the degree of Subjacency in their L1. Namely, the Dutch speakers showed the highest adherence to English subjacency rules, Chinese and Indonesian speakers showed some evidence, and the Korean speakers showed very little evidence of the principle. Schachter concludes that although the notion of UG driven second language acquisition is appealing, the results of her experiment do not support it (p. 85).

The study of the adherence to English Subjacency restrictions by Bley-Vroman et al. draws different conclusions to Schachter's. Their subjects are L1 Korean speakers (n = 92) with advanced L2 English abilities who had been living in an English speaking environment (Texas) for some time. Bley-Vroman et al. used a grammaticality judgement test which included sentences examining four contrasts related to Subjacency and the ECP. The contrasts included in Bley-Vroman et al.'s grammaticality judgement test are summarized in (22) and (23) below.

(22) Contrasts related to Subjacency

a *wh*-question Islands
What did Bill think that the teacher has said?
*What does Tim wonder where Nancy put?

b. Factives
What did Bill think that the teacher had said?
*What did Sam believe the claim that Carol had bought?

c. Relative Clauses
Where is the person that I want you to talk to?
*Where did Bill visit a friend who had just arrived from?

d. Coordination

What did John find the ball and?

*What does John like to eat tomatoes and?

(23) Contrasts related to the Empty Category Principle

a Superiority

I can't remember who did what.

*She forgot what who said.

b. *That*-trace Effect

What did Frank say that Judy would like to read?

*What did John say that would fall on the floor, if we're not careful?

c. Sentential Subject Islands

What kind of book is it necessary to read?

*What sort of food is to digest easy?

d. PP Islands

Which bed does John like to sleep in?

*What did Albert put money in the box during

e. Specified Subject Constraint

What did Mike see pictures of?

*What did Mary hear Bill's stories about?

(Bley-Vroman et al., 1988: 22-23)

Test subjects could choose from between three judgements - possible, impossible and unsure, although the unsure judgement was almost never used by either the test subjects (3.6% of answers) or the native speaker controls (2.2%). The results of this experiment combined for both the Subjacency sentence results and the ECP sentence results clearly show that both the test subjects (n=92) and the native speaker control group (n=34) perform at better than chance (50%) levels (test subject average = 75%, control group average = 92%). However, when the results for grammatical versions of a sentence type

are compared to results for the ungrammatical version of that sentence type, a different picture emerges. For example, if half the test subjects correctly judge grammatical *that*-trace sentences correctly, but misjudge ungrammatical *that*-trace sentences, and the other half misjudge all grammatical versions of this sentence type but correctly judge the ungrammatical version, then none of the subjects will have demonstrated the correct contrast in judgements (p. 21). Bley-Vroman et al. compare the results of their different stimulus sentences in this way and conclude that as a population, just more than 50% of the test subjects actually exhibit target-like judgements on any of the sentence types they included. Because of the low frequency with which the 'unsure' judgement was made, informants may have had a bias to reject sentences they were not sure of, somewhat skewing these results. However, on subjecting their results to a *chi*-square analysis where chance performance would have been 25%, they show that test subjects' results are in fact somehow principled and not just guesses.

The question Bley-Vroman et al. ask then is why the non-native speaker results are not better than they are if the principled nature of their judgements was UG driven? One possibility they suggest is that UG may operate in adult L2 acquisition, but in some "in some attenuated form" (1988: 27). Another possibility, proposed by Felix (1985), is that UG may be available, but is in competition with and somehow blocked by other cognitive systems such as general problem solving strategies.

A second study which investigated adult L2 acquisition of the *that*-trace ECP violations is reported in White (1990: 58). White looked at adult Dutch learners ($n$=62) of English to determine if they recognize the distinction between extractions of object *wh*-phrases and subject *wh*-phrases in English, as in (18) above, repeated here as (24a, b, c, and d).

(24) object extractions

(a) Who$_i$ do you think that Paul met $t_i$ ?

(b) Who$_i$ do you think Paul met $t_i$ ?

subject extractions

(c)*Who$_i$ do you think that $t_i$ met Paul?

(d) Who$_i$ do you think $t_i$ met Paul?

If UG were no longer available to the learners, White assumed that they would allow sentences like (24c). If UG were available, White reasoned that the ECP should be available to the learners and that proper government would not allow subject *wh*-phrases to move out of lower clauses with overt COMP. White's data indicated that her Dutch learners of L2 English do in fact disallow sentences like (24c) and led White to conclude that UG is available to these adult learners.

In a more recent, study of empty categories in adult L2 acquisition, Kanno (1996, 1998a, 1998b) investigates the possibility that to some degree, variation observed in adult L2 acquisition is "due to a lack of stability in learners' use of UG principles" (1998a: 1126). In particular, she looks at whether these learners display 'lateral' and 'longitudinal' consistency in their knowledge of the ECP (1996) and Overt Pronoun Constraint (OPC) (1998a).

Kanno's (1996) ECP test targeted case-particle deletion in Japanese. In Japanese, the nominative case particle *ga* is compulsorily suffixed to the subject, as in (25), while the object case particle *wo* can be optionally deleted, as in (26).

(25) Gakusei wa anno hon ga omosiroi to iimasita
student-TOP that book SUBJ interesting COMP said
'The student said the book is interesting'

(26) Gakusei wa anno hon (wo) yonda yo
student-TOP that book (OBJ) read EXP
'The student read that book'

Kanno's subjects were university students, approximately 20 years of age and studying Japanese as a second language (JSL) at the University of Hawai'i. A group of native Japanese speakers was also included as controls. Kanno finds that approximately half of her JSL test subjects could perform at a native like level on the ECP regarding *ga* and *wo*, claiming that there was no Japanese input from which they could have acquired the rule, nor had they been taught that *wo* could be deleted. Kanno's results lead her to conclude that UG continues to be available to L2 learners.

Kanno (1998a) considers another UG constraint, the OPC. In languages such as Japanese, which has both overt and null pronouns, overt pronouns cannot have quantified NPs such as *everyone* or *who* as an antecedent, and only null pronouns can be bound by such a quantified NP. The OPC disallows (27), while allowing (28).

(27) Dare$_i$ ga [$e_i$ sore o katta] to itta no?
who Nm that Ac bought that said Q

(28) Dare$_i$ ga [*kare*$_i$ ga sore o katta] to itta no?
who Nm he Nm that Ac bought that said Q
'Who$_i$ said that he$_i$ bought that?'

(Kanno, 1998a: 1126).

In languages such as English that do not allow null subjects, the OPC does not apply, as in (29a) and (b).

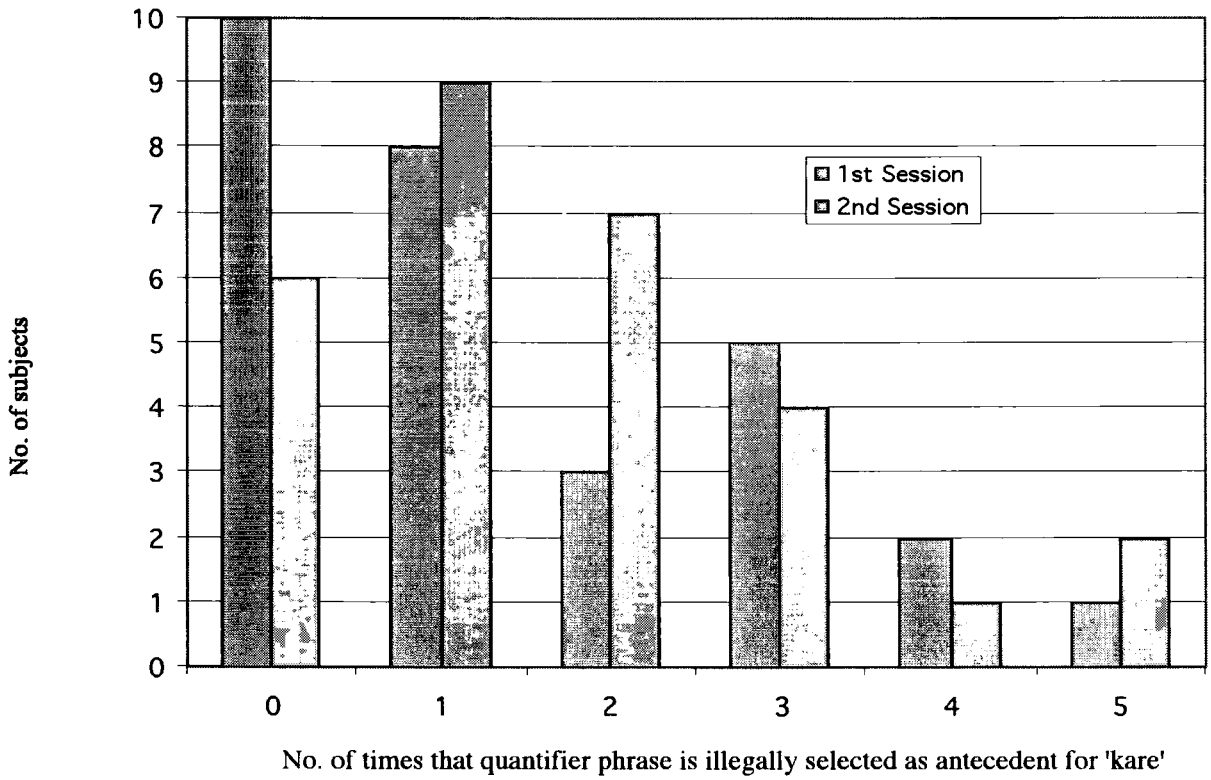
(29) a. Everyone$_i$ wishes that he$_i$ could be the winner
b. Who$_i$ thinks he$_i$ is the winner?

(Kanno, 1998a: 1127)

As the OPC is presumably not deductively learned in JSL classes, and as it does not apply to the learners' L1, English, it is underdetermined in adult L1 English speakers' acquisition of Japanese. Kanno thus identifies an interesting phenomenon to test for availability of UG in adult L2 acquisition. Her study tested 29 native English speaker intermediate-level JSL learners on the OPC twice- a pre-test at the start of their fourth semester of Japanese studies and a post-test at the end of the semester. All subjects were undergraduate college students. Kanno compared the results of the two test sessions to each other and also to the results of 12 adult native Japanese speakers. In analyzing the native Japanese speakers' results, it is clear that they "consistently and

categorically" reject the link between the overt pronoun and the quantified NP. (1998a: 1131). Turning to the JSL learners' two sets of test results, Kanno finds they are much more likely to select the quantified NP as antecedent when the pronominal is null (pre-test = 73%; post-test = 72%) than when it is overt (pre-test = 29%; post-test = 34%). Although the results don't indicate improvement over the time, they do exhibit similar preferences in both test sessions. Additionally, as an examination of the teaching materials indicated no overt teaching of the OPC, Kanno concludes that her results constitute evidence that adult JSL learners do have continued access to UG. Figure 2.3 illustrates Kanno's results for the overt pronoun *kare* ('him').

*Figure 2.1. Kanno's (1998a: 1135) results of JSL learners' first and second OPC tests, showing maintenance of restriction on selecting a quantifier phrase as the antecedent of the overt pronoun* kare.




Kanno (1998b) examines whether native-like knowledge of one principle (ECP) will predict native-like knowledge of the other (OPC) - 'lateral' consistency. The results of her OPC study again show native-like accuracy in her test subjects at the 50% level,

but only half of these - 25% - perform well on both tests. These findings do not lend strong support to her lateral consistency hypothesis.

Kanno's (1998a) study also tests whether the OPC remains constant over time - 'longitudinal' consistency. 12 weeks after her first test session, she again administers the same OPC test to a sub-group of the original test takers. Again Kanno's results are somewhat disappointing as she finds that only about one third of those who performed at native like levels on the first test maintain that level on the second test. These results are illustrated in Figure 2.1. Nevertheless, Kanno's tests are intriguing and rather than rejecting her results as support for a role access to UG in adult L2 learners, her results merely reflect lateral and longitudinal variation among her population. She might find more consistency both latitudinally and longitudinally were she to administer the test several time. Furthermore, Kanno does unarguably find native like levels of Japanese OPC and ECP in learners who were neither taught these principles, nor exposed to significant amounts of the target language outside the classroom. Although this last point regarding exposure has been questioned because of her subjects' possible exposure to Japanese in the local community (Hawai'i), (Kellerman and Yoshioka, 1999: 107) she still demonstrates that adult L2 Japanese learners can make native-like judgements about forms for which there is no evidence in the input.

### 2.4.2. *Parameters*

While principles are constant across languages, parameters can be set for different values, thereby accounting for major structural differences found among languages. As a child is exposed to a particular language, certain aspects of the input trigger the parameter to the setting for that language. By attributing various clusters of properties to a single parameter, the parameter setting model helps explain how typologically different languages can be acquired on the basis of a fixed initial state with equal ease despite limited and degenerate input. In the next sections, the prodrop parameter and the head-position parameter will be examined in some detail as these two parameters have proven productive to generative L2 acquisition research.

2.4.2.1. *The prodrop parameter*

The prodrop parameter has received considerable attention from both L1 and L2 researchers. It was originally proposed to distinguish some Romance languages such as Spanish and Italian, which allow phonetically null subjects in sentences, from languages such as English and German, which do not allow null subjects (Jaeggli, 1982; Rizzi, 1982c). As noted above, a prodrop language, [+prodrop] has rich verbal agreement morphology which allows sentential subjects to be identified, as in (30).

(30) a. Spanish: Estoy aqui.
(I) am here

Esta aqui.
(he/she) is here

b. Italian: Sono qui.
(I) am here

E qui
(he/she) is here

Non-prodrop languages [-prodrop], typically have less verbal agreement morphology and require overt subjects, as in (31).

(31) a. English: I am here
*am here

b. French: Je suis ici.
*suis ici
I am here

In addition to null subjects, the prodrop parameter is also thought to be associated with other properties which are triggered when the parameter is set to

[+prodrop] (Rizzi, 1982c). One such property is an obligatory null-subject for sentences which in English require pleonastic pronouns such as "It's fine". Compare to the Spanish equivalent "Este tanbien". Other properties include free inversion in simple sentences such as Italian "ha mangiato Giovanni", "has eaten Giovanni"; long *wh*-movement of subjects: "l'uomo che mi domando chi abbia visto", " the man who I wonder who [someone has] seen"; resumptive pronouns in embedded clauses: "ecco la ragazza che mi domando chi crede che possa..., "this is the girl who I wonder who thinks that she may..."; and apparent violations of the *that*-trace filter: "chi credi che partira", "who do you think that will leave".

L2 research into the prodrop parameter focuses on the presence or absence of null subjects, subject-verb agreement, use of pleonastic ("dummy") pronouns. Hyams (1986) has noted that L1 children's early stages in the acquisition of [-prodrop] L1 languages share certain similarities to adult [+prodrop] languages, such as the presence of null subjects and the absence of pleonastic pronouns. This comparison has led Hyams and others (Hilles, 1986; Phinney, 1987; Liceras, 1989; Al-Kasey and Perez-Leroux, 1998) to assume that the initial or "default" setting for UG in L1 acquisition is [+prodrop].

In contrast to this, White (1985), following Berwick (1985) and Rizzi (1986) claims that [- prodrop] is the default setting for child L1 acquisition. White argues that for L2 acquisition, the only evidence L1 speakers of a [- prodrop] language need to correctly reset their [- prodrop] parameter to the [+ prodrop] setting is to hear sentences with null subjects. L1 speakers of a [+ prodrop] language acquiring a [- prodrop] L2 such as English would presumably need negative evidence to recognize that a language like English does not allow null subjects. In support of this, White (1986a) tests whether [- prodrop] L1 speakers can easily acquire the [+ prodrop] setting through exposure to positive evidence alone. She investigated ([+prodrop]) Spanish and Italian speaking learners of ([- prodrop]) English. She also included French-speaking subjects as a [- prodrop] control group. Her data were collected via grammaticality judgement tests to determine whether second language learners would transfer their [+ prodrop] L1 settings to the [- prodrop] L2. White's results indicated significant differences between the Spanish and Italian L2 English learners' judgements on subjectless sentences and those of the French control group. The L1 Spanish and Italian learners were much more likely

to accept subjectless sentences in English than the L1 French speakers, leading White to conclude that learners transfer the L1 value of the prodrop parameter to their L2.

Phinney (1987) used open-ended writing samples to investigate the prodrop parameter. She looked at subject pronouns and agreement in both directions: L1 Spanish speakers learning L2 English, and L1 English speakers learning L2 Spanish. Phinney's results indicated that both groups accurately produced subject-verb agreement forms. This is not unexpected as both [+ prodrop] and [- prodrop] languages have subject-verb agreement. Of interest in Phinney's results is that the L1 Spanish speakers had a high rate of omission for pleonastic pronouns in their L2 English, while the native English speakers correctly omitted both referential and pleonastic pronoun subjects in Spanish. This suggests that the Spanish speakers were unable to completely reset the parameter and transferred their L1 parameter setting, while the native English speakers succeeded in resetting the parameter to the [+ prodrop] setting. This would seem to support White's (1986a) suggestion that [- prodrop] is the unmarked, more difficult to acquire, setting.

Hilles (1986) used longitudinal data from a study by Cancino, Rosansky, and Schumann (1978) to examine prodrop in a 12-year-old Spanish-speaking boy acquiring L2 English. Hilles' analysis proposes that [+ prodrop] languages characteristically lack pronominal subjects, lack lexical material in AUX, and lack expletives. Hilles examined the data from Cancino et al. (transcripts of spontaneous English conversations) for instances of the emergence of [- prodrop] effects proposed by Hyams. Hilles claims that the appearance of such [- prodrop] phenomena coincides with the appearance of expletives, supports her proposal that the [-] setting for the prodrop parameter is triggered by the appearance of expletives. Despite finding little evidence of spontaneous expletives, and only minimal evidence of expletives in elicited imitation tasks, Hilles nonetheless suggests that [+prodrop] effects seem to drop off with the acquisition of expletives.

A more recent study by Al-Kasey and Perez-Leroux (1998) looks for evidence of the resetting of the prodrop parameter in the L2 acquisition of Spanish by native English-speaking first and second year university students. Al-Kasey and Perez-Leroux collected data to examine two aspects of the prodrop parameter in L2 Spanish: null expletives and the omission of optional subject pronouns (p. 165). Their data collection methods consist of general classroom observation, a comprehension test, and a

controlled production test. Al-Kasey and Perez-Leroux find that the L2 Spanish data do indeed show that learners recognize that null expletives and null subjects stem from the prodrop parameter set to [-prodrop]. Their conclusion however, is not convincing. The lack of subjects and the null expletives could be the result of a learner tactic which says something like "delete all pronominal subjects" which would result in subjects never appearing anywhere, even in thematic subject positions where they are optionally allowed. As Thomas (1998) points out, "L2 learners might simply grasp that the subject position in Spanish may be phonetically empty, and then apply this insight across the board, to both thematic and non thematic subject positions" (p. 267).

2.4.2.2. *The head-position parameter*

Another binary parameter of UG, the head-position parameter has settings of head-initial where heads occur before complements, and head-final, where complements occur before heads. As with the prodrop parameter, research done on the head-direction parameter in L2 acquisition has led to different claims.

Arguing against access to UG, Clahsen and Muysken (1986, 1989) point to results which show that adult learners of L2 German do not easily acquire German word order. Clahsen and Muysken examine German L1 acquisition studies (Roeper, 1973; Clahsen, 1982) and show German L1 learners establish an SOV word order from the initial stages of L1 acquisition. Clahsen and Muysken then argue that findings from German L2 acquisition studies (Clahsen, Meisel, and Pienemann, 1983; Clahsen, 1984) show L2 learners tend to favor SVO word order. Based on the lack of clause-final verbs they find in their data, Clahsen and Muysken argue that German L2 learners have a canonical subject-verb-object word order strategy even if their L1 is verb final. Since L2 German learners do eventually produce nonfinite verbs in final position in main clauses, and finite verbs in final position in subordinate clauses, Clahsen and Muysken argue that learners elaborate "a series of complicated rules to patch up this (SVO) hypothesis when confronted with conflicting data" (1986: 116). As a result, they can only arrive at L1-like verb-final utterances by developing "alternative hypotheses, which lead to outputs more closely resembling the target language patterns" (1989: 24).

Arguing against Clahsen and Muysken's conclusion that UG is no longer available to adult L2 learners, duPlessis, Solin, Travis and White (1987) propose an alternative explanation for adult L2 German in terms of the resetting of several parameters of UG - the headedness parameter, the proper government parameter, and the adjunction parameter. In addition to reanalyzing Clahsen and Muysken's data, they examine university-age anglophone and francophone Canadians enrolled in an advanced German language course. One suggestion made by duPlessis et al. is that certain properties of German might initially mislead adults into analyzing German as an SVO (i.e., head initial) language, in that main clause word order is SVO, prepositions occur to the left of NPs, head nouns occur to the left of relative clauses, etc. (p. 62). This would account for early SVO patterns emerging. As learners are exposed to more input and parameters are gradually reset, they pass through different IL stages and move towards an SOV order. However, as duPlessis et al. stress, some parameters may reset to the German values more quickly than others. This variation in resetting may result in intermediate settings which are neither L1 nor the TL, but which are still within the range of settings found in natural languages. An SVO order such as observed by Clahsen and Muysken can be attributed to an intermediate stage of adult L2 acquisition when only some of the parameters may be set correctly.

Vainikka and Young-Scholten (1994) lend further support for the head-final VP in adult German L2 learners. They analyze four levels of cross-sectional data collected from 11 Turkish and 6 Korean adults acquiring German in Germany. Informants ranged in age from 28 to 60 years and had no formal instruction in German. Data was collected via several elicitation tasks including providing dialogue for comic strips, describing pictures, and describing a set of procedures. Each task attempted to elicit a variety of sentence types in order to examine verb placement, subject usage, and subject-verb agreement.. The three languages involved are interesting in that all have a head final VP, although in German, unlike in Turkish and Korean, the finite verb of the matrix clause does not occur at the end. Compare (32), (33), and (34).

(32) Turkish

Helmut [$_{vp}$ şimdi Istanbul-da Türkçe ögren-iyor]

Helmut now Istanbul+LOC Turkish learn+PROG

'Helmut is learning Turkish in Istanbul now.'

(33) Korean

Helmut-ka [$_{vp}$ Peter-eke chaek-uo chu-oss -ta]

Helmut+NOM Peter+DAT book +ACC give+PAST+DECL

'Helmut gave the book to Peter.'

(34) German

Ich habe [$_{vp}$ Klaus heute in Grafenberg gesehen]

I have Klaus today in Grafenberg seen

{I have seen/I saw} Klaus today in Grafenberg.'

(Vainikka and Young-Scholten, 1994: 269)

Vainikka and Young-Scholten identify four syntactic phenomena with which they create an implicational table they use to assess each L2 German learner's acquisition of German word order. These four phenomena are the German head-final VP; the fact that German does not allow empty (thematic) subjects – unlike Turkish and Korean; verb raising to the left of the VP; and the German subject-verb agreement paradigm. Each learner was given one point for each form acquired, for a maximum score of four. The data show that all learners have acquired the head-final VP. Additionally, over 60% of the sentences considered showed the learners had acquired a verb-final word order, regardless of overall level. This is in contrast to Clahsen and Muysken's (1986, 1989) analysis of data from Turkish speakers learning L2 German. Clahsen and Muysken concluded that their learners do not adopt a head-final VP for German.

As for Vainikka and Young-Scholten's null subject criterion, they found that the usage of overt subjects in sentences with raised VPs was very predictable, depending on overall level. The higher the level, the fewer the instances of null subjects being used.

Test subject's instances of verb raising to the left of the VP, as illustrated in (34) above, again increased with speaker level. Two informants raised verbs outside the VP more than 80% of the time, while for the three lowest-level learners, they seemed to analyze sentences as bare VPs 80% of the time, with no verb raising.

Acquisition of the agreement paradigm required a learner to correctly use German verb morphology over 60% of the time on raised main verbs. Vainikka and Young-Scholten conclude 6 out of the 17 informants exhibited productive agreement.

Vainikka and Young-Scholten's results have several implications. Their cross-sectional data shows the development of phrase structure in L2 German acquisition. They note that it follows a pattern seen in first language acquisition, namely that learners start off with bare lexical projections, then develop underspecified functional projections, finally filling in the specific features of the functional projections. Additionally, both L1 and L2 German learners' acquisition patterns coincide in their acquisition of agreement, avoidance of pro-drop, and obligatory verb raising. Considering the uninstructed nature of the Turkish and Korean informants' L2 German, the similarities between first and second language acquisition of phrase structure Vainikka and Young-Scholten find in their results constitute evidence against the position that UG is no longer available to adult learners.

### 2.4.3. *Markedness theory*

At a linguistically global level, the notion of markedness distinguishes between core grammar which is constrained by UG, and peripheral grammar which is not directly UG related (Mazurkewich, 1984; Liceras, 1985; 1986; Phinney, 1987). In this case, peripheral aspects of language include politeness level, register, use of discourse markers, etc.

Within the generative approach to language acquisition, one view of the natural state of UG prior to exposure to input is that all parameters are unset, or open. Exposure to input forces the parameters to set one way or the other. Under an alternative view, parameters are preset in one direction – the unmarked setting. A child learning a language which involved that particular setting would easily learn that aspect of the language. Learning a language which had the opposite setting would require resetting that particular parameter to the marked setting as the result of positive input from the linguistic environment (Atkinson, 1992: 102). One example of this second view we have seen is White (1986a) who argues that [-prodrop] is the unmarked setting of the prodrop parameter.

Markedness theory has been applied to second language acquisition research in several ways: to predict acquisition sequences (Mazurkewich, 1984), to account for language transfer (Liceras, 1985; 1986), and to predict which way [+/-] a parameter will be set (Phinney, 1987). Unmarked forms are assumed to be easier to acquire and to be acquired first in a learner's interlanguage grammar. Mazurkewich states that unmarked forms will be identified as an acquisition stage even if they are not in fact involved in the target language. Bardovi-Harlig (1987), however, counters this claim with her study of mixed L1-speaking L2 English learners. The L1s represented in her study include both languages which allow preposition stranding, and languages which do not. Bardovi-Harlig's results show that preposition stranding, a marked phenomenon, occurs before unmarked piedpiping, which she claims is a result of the prevalence of preposition stranding in the primary input.

Directionality of acquisition refers to which way parameter settings are more easily transferred from the L1 to the L2 during L2 acquisition: from marked to unmarked, or from unmarked to marked. Liceras (1985) claims the unmarked aspects of the L1 will be more likely to transfer than the marked aspects. Phinney (1987) argues along similar lines, claiming that it will be easier for a learner to move from a marked L1 form to an unmarked L2 form than the other way around. She bases this claim on a marked L2 value being more difficult to acquire and therefore requires a greater amount of positive evidence to set that marked version. Her data find that L1 Spanish speakers learning L2 English failed to produce pleonastic subjects such as *it* or *there*. These results complement her data for English speaking learners of Spanish which indicate they allow null subjects. White (1986b), however, argues that where positive evidence is available, changing from the unmarked version to the marked version is easier than from the marked to unmarked because there is no positive evidence to force the change going the opposite direction.

Despite White's argument on directionality, most researchers who consider there to be a relationship between markedness and transfer believe that unmarked aspects of the L1 will be more likely to be transferred to the L2 than marked ones. Mazurkewich's (1984) study of the acquisition of English dative structures by L1 French speakers and L1 Inuktitut speakers seems to show that L2 learners will learn the unmarked piedpiped setting such as (35) before the marked preposition stranding setting such as (36).

(35) To whom did John give a book? (entire PP is piedpiped to front)

(36) Whom did John give a book to? (preposition stranding)

In Mazurkewich's study, at least for the French speakers, this preference for the sentences like (35) could be caused by transfer from their L1, which shares the same unmarked form (piedpiping) as English.

Liceras (1985, 1986) also argues that there is a greater chance for unmarked aspects of the language to appear in a learner's interlanguage than marked forms. She suggests that because of learners' sensitivity to L1 marked structures, they can notice if they are missing in the L2 target language and transfer the unmarked version over.

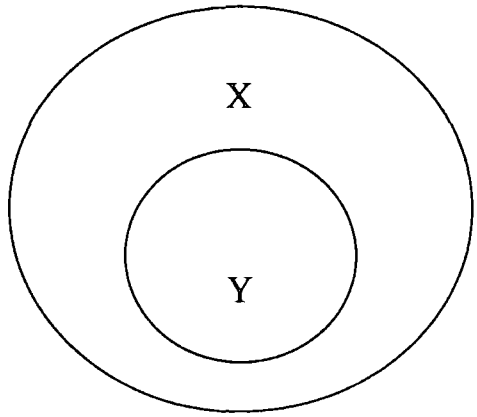
### 2.4.4. *The Subset Principle*

The Subset Principle attempts to explain how L1 acquisition can occur with positive evidence only (Berwick, 1985; Manzini and Wexler, 1987). Wexler and Manzini (1987: 61) describe the Subset Principle as (37).

(37) 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.

To see how this works, consider two grammars which generate the same subset of sentences, but one of the grammars produces additional sentences to those in the subset. Grammars in this kind of relationship are said to meet the Subset Condition, illustrated in figure 2.2.

Figure 2.2. Grammars X and Y in a subset condition




The grammar which generates X-type sentences also generates Y-type sentences, but grammar Y only produces Y-type sentences. Y is a proper subset of X. The Subset Principle addresses the following learnability problem: a child learning a language which contains only Y-type sentences cannot overgeneralize and produce X-type sentences, even though Y-type sentences are included in grammar X. To prevent this overgeneralization, the Subset Principle instructs the learner to adopt the most restrictive grammar which is consistent with the input. In the case of a child learning grammar X, input will provide positive evidence which will cause him or her to adopt the broader grammar.

As we will see in the discussion of the parameterization of binding constraints in Chapter 4, Manzini and Wexler develop an example of the Subset Principle central to their theory which attempts to account for the behavior of anaphors and pronouns cross-linguistically (Atkinson, 1992: 136). Their analysis proposes that the notion of governing category is not the same for all languages and that this can effect anaphor-antecedent relationships. Some languages, such as English, are quite restricted in what may serve as an antecedent for an anaphor. Other languages such as Japanese are much freer and allow a wider range of antecedent candidates to be associated with a particular anaphoric expression. For example, English anaphor-antecedent rules which restrict binding to the local antecedent as in (38), are a subset of the Japanese rule system which allows a reflexive to be bound locally as well as long distance, across a clause boundary or boundaries, to the subject of the matrix clause as in (39).

(38) $Mary_i$ thought that $Susan_j$ blamed $herself_{*i/j}$

(39) $Mary\text{-}ga_i$ $Susan\text{-}ga_j$ $zibun\text{-}o_{i/j}$ semeta to omot-ta
Mary-NOM Susan-NOM self-ACC blamed COMP think-past
'Mary thought that Susan blamed herself'

Although binding theory and the notion of a parameterized governing category will be discussed in detail in Chapter 4, they are briefly introduced here because Wexler and Manzini's ideas about binding will be used to examine and discuss the data in Chapter 5. Of interest in the current chapter is the possibility of the Subset Principle operating in second language acquisition. If the principle is fully operational, L2 learners should initially hypothesize that the most restricted grammar consistent with the data is correct, regardless of their L1, and only adjust their hypothesis in the face of positive evidence.

White (1989a) examines how the Subset Principle works in regard to the adjacency condition on case assignment. She looks at L1 French speakers learning English (ESL), and L1 English speakers learning French (FSL). In English, the adjacency condition requires "that a noun phrase receiving case must be next to its case assigner" (White, 1989b: 149) as in (40).

(40) a. Mary ate her dinner quickly.
b. *Mary ate quickly her dinner.
(White, 1989a: 136)

In French, certain items such as a "manner adjunct" like *quickly* may intervene as in (41).

(41) a. Marie a mange le diner rapidement.
b. Marie a mange rapidement le diner.
(p. 137)

But not other material, as in (42)

(42) a. *Jean a bu hier du café.

b. *Marie a place sur la table le livre.

(p. 137)

White's data results seem to suggest that the both types of learners, ESL and FSL, transfer their L1 settings to their respective L2s. In the case of the FSL learners, errors on adjacency were minimum, but at the same time learners were not allowing manner type adjuncts between the verb and the object NP. ESL learners, however, did make adjacency errors, inserting material between the verb and the NP. These results would seem to indicate that the Subset Principle is not available to learners. For adjacency, English is a subset of the French superset, and would predict that the ESL learners would have difficulty acquiring the English subset of their L1 superset and the FSL learners would more readily acquire the L2 French superset grammar through positive evidence. But as White points out, the results are actually inconclusive. Failure to obey the Subset Principle with regard to adjacency may not be due to the principle being unavailable to the learners, but rather to it being blocked by transfer from the L1. Alternatively, transfer from the L1 may occur precisely because the principle does not operate.

### 2.4.5. *Parameter resetting*

L2 acquisition involves not only transfer of UG parameters settings from the L1 to the L2, but necessitates that the learner eventually establish new settings for L2 instances where the L1 setting and the L2 setting differ, i.e., resets the parameter. Using the head parameter as an example, we can see that this parameter would need to be reset in an English speaker learning Japanese and vice versa, where the head initial setting initially triggered during the L1 acquisition of English would need to change to accommodate the head-final setting needed for Japanese (see 2.4 above)

White (1986a) highlighted the importance of parameter resetting by investigating the differences between Spanish/Italian (prodrop language) and French (non-prodrop

language) speakers' recognition of English grammaticality in terms of subject-verb inversion, *that*-trace, and optional or obligatory subjects. White found that Spanish and Italian (prodrop language) speakers made many more errors in null subject sentences and instances of *that*-trace effects than did the French speakers, but no difference was found for Subject-Verb inversion.

Arguments can also be made against parameter resetting. Eubank (1993/1994; 1996) has argued for access to UG in L2 acquisition on the one hand, but also that parametric values do not transfer in full from the L1 to the L2. Based on investigations into verb raising in L2 German acquisition, Eubank proposes that L2 learners arrive at the L2 acquisition task with only a partial transfer of L1 parameter settings. Some settings, such as those associated with directionality (head initial, final, etc.) do transfer. Others, such as those responsible for inflection, are inert, or "valueless" at the L2 initial state. Eubank claims that it is particularly this lack of parametric values at the initial stages of L2 acquisition which is the essential difference between L1 and L2 acquisition (Eubank 1996: 97).

## 2.5. Conclusion

This chapter has presented some of the central arguments generative linguistics uses to posit the existence of an innate language faculty, UG. We have discussed the major positions SLA researchers hold regarding the role of UG in second language acquisition, from no role (UG is dead) to full access to UG.

The studies reviewed in this chapter demonstrate the types of UG-driven phenomena which lend themselves to investigative research. One instrument often used to test these principles is some type of judgement task. Martohardjono (1998) questions the reliability of these tests to examine UG competence. She suggests that this testing methodology wrongly assumes "a direct relationship between access to UG (competence) and the ability to perform at native-like levels on a grammaticality judgement task (performance), so that we expect UG accessibility to be reflected in high scores" (p. 155). She goes on to argue for a more eclectic approach to research in this area, one which would include different types of tasks to test the same phenomena. And these tasks should not merely compare test subject results to a control group standard.

Using a variety of tests to find patterns of acceptability across different L1 populations, rather than absolute rates of levels of accuracy, can perhaps reveal trends showing adult L2 learners do in fact have access to UG constraints regardless of the other variables influencing their acquisition.

Whatever role UG may or may not have in L2 acquisition, the parameter theory developed by mainstream generative linguistics provides SLA researchers with a very productive means of investigating how second languages are learned. This has resulted in an increasingly large generative-based corpus of literature on L2 acquisition. In contrast to this, relatively little work has been done to examine the results of language attrition. What work has been done typically describes loss of lexical items, morphology, and word order, and fails to address underlying principles to explain why certain aspects of an L2 are lost. In the next chapter we will review various studies which have been done on L2 attrition, noting in particular the hypotheses which have been proposed to explain L2 attrition.

# Chapter 3: Language Attrition

## 3.0. Introduction

A study of second language attrition must consider a wide range of variables. Developmental factors such as the age at initial exposure to the L2, the type and amount of exposure involved, the linguistic environment in which acquisition took place, and affective variables such as attitude and perceived prestige of the language all contribute to the final state of the L2 before attrition sets in. In addition to these L2 acquisition variables, L2 attrition must also consider the amount of time elapsed since exposure to the attriting language ceased.

In this chapter, I will initially discuss these variables by looking at attrition from both a diachronic, intergenerational perspective, as well as from the perspective of individual, or intragenerational attrition. The discussion will then consider hypotheses - principally the regression hypothesis - which have been proposed to describe and explain attrition. This discussion will focus on second language attrition and will include a review of L2 attrition studies. Finally, this chapter proposes that despite the growing presence of L2 attrition as a sub-field of SLA, attrition studies lack a theoretical framework within which to conduct investigations. To address this, the framework I adopt to investigate phenomena related to L2 attrition is the generative approach to SLA studies as proposed and described by Flynn (1987) and White (1989b).

## 3.1. Typology of attrition

Language attrition is most commonly caused by a combination of under-exposure and competition from a dominant language in bilingual individuals. Language loss in communities is a diachronic phenomenon, the result of a minority language being replaced by a dominant language over several generations. This is known as intergenerational language attrition (Weltens, 1989: 3) and is associated with first language loss. In individual situations, attrition is more synchronic in nature and occurs when an individual's exposure to the language in question ceases, or the individual

suffers physical damage to the brain. This type of attrition can occur in the L1 or the L2 and is known as intragenerational attrition (op.cit.).

### 3.1.1. *Intergenerational language attrition*

Intergenerational language attrition can occur in immigrant communities as subsequent generations acquire a more and more imperfect variety of the original immigrant L1. While the present study is *intra*generational (see 3.1.2. below) in nature, it shares many of the same variables and research techniques with intergenerational studies. One study which looks at Finnish speaking communities in northern Minnesota, USA, is Larmouth (1974). He finds that the morphologically rich structure of Finnish gradually disappears from generation to generation because of its "syntactic congruence with English sentences" (p. 356). Sentence (1) below illustrates the word-for-word congruence of a simple Finnish sentence with English. Larmouth claims that the attrition of Finnish leads to an increasingly fixed word order in the language, that case endings become less obligatory, and adjective-noun concord disappears. Larmouth notes that Finnish interrogative pronouns are marked for both number and case in first-generation informants. His research shows that in later generations, case marking is prone to loss, but number is consistently maintained. For example, in (1) the case marker *-n* survives through the first (G1) and second (G2) generations, but has disappeared by the third generation (G3). Certain Finish verbs, such as *ost-* ('buy') obligatorily take nominative case on their objects. Evidence of loss of the determiner *se* is also apparent.

(1) G1: Kuka osti se huone(n)?
'Who (sg.) bought the house?' (Nominative case)
G2: Kuka osti se huone(n)?
G3: Kuka osti se juone?
G4: Kuka osti (se) hone?

(Larmouth, 1974: 360)

Intergenerational language attrition can also occur in indigenous communities which gradually become marginalized by a dominant linguistic group. Extensive research into this type of attrition has been done by Dorian (1973, 1978, 1981) who investigated language shift (one language replacing another in a community) and death (no more native speakers of a language) of a variety of Scots Gaelic (East Sutherland) in Scotland. One example of attrition Dorian (1981: 51) identified is the loss of the passive. East Sutherland Gaelic has two forms of the passive, each incorporating a different verb. Dorian finds older speakers maintain this distinction, but younger speakers tend to use just one form which includes elements of both. This type of change is not necessarily a result of transfer from the dominant language, English, but rather a simplification or convergence of two competing forms within the attriting language as a result of underuse.

Other studies of intergenerational attrition have investigated many minority languages including Native American languages (Hill and Hill, 1980; Miller, 1971), the loss of Hungarian in Austria (Gal, 1979), and of aboriginal languages in Australia (McConvell, 1991).

### 3.1.2. *Intragenerational language attrition*

Intragenerational language attrition is restricted to language loss at the individual level and can occur in three different ways depending on the language lost (L1 or L2) and the linguistic environment within which the language is lost (L1 or L2).

Intragenerational studies of the loss of the L1 in the L1 environment are relevant to intergenerational attrition as a means of identifying specific linguistic aspects undergoing attrition which contribute to language shift and possible death. An example of one such aspect is Dorian's work cited above, in which she identifies changes in passive constructions in younger speakers of the Scots Gaelic.

Another type of L1 attrition in the L1 environment can also happen in individuals as the result of pathological events such as stroke, or diseases which lead to dementia such as Alzheimer's or Parkinsons's (cf. Hyltenstam and Stroud, 1989). Studies of polyglot aphasics are perhaps the oldest type of attrition research, dating from the late

19th century (Pitres, 1895; Ribot, 1882). Obler and Mahecha (1991) analyze historical data on polyglot aphasic patients who lost their L1s as the result of sudden damage to the brain. Their analysis included a variety of variables such as age, gender, education level attained, manner in which languages were acquired, handedness, and location of brain damage. Their results indicated that permanent or temporary L1 loss appeared to be a consequence of "a unilateral right hemispheric lesion and left handedness" (p. 63). Paradis (1983, 1987) has also done extensive research on polyglot aphasics, investigating language loss and recovery patterns

The second type of intragenerational attrition, loss of the L1 in an L2 environment, is known to happen with immigrants who have lived in their host country for many years and gradually lose their mother tongue. Vago (1991) for example, looks at the loss of an L1 morphophonemic system in a Hungarian-Hebrew bilingual Israeli woman. Vago's informant had emigrated from Hungary to Israel at age 5;10 and her L2 (Hebrew) had gradually become her dominant language. The investigation looked at the loss of nominal and verbal morphology which constitute paradigmatic sets in Hungarian. Vago first prompted his informant with a noun or verb and asked her to produce a linguistic paradigm such as the singular possessive for a given noun. In subsequent re-testing sessions, the informant was presented with the paradigms she had created earlier and asked to judge their accuracy. Vago's results indicate evidence of phonological rule simplification and loss, as well as lexical restructuring.

Two other studies which have investigated L1 loss in the L2 environment are de Bot, Gommans and Rossing (1991), who looked at the L1 Dutch of immigrants in France, and Waas (1993), who looked at L1 German in Australia. Another study of this type is the 16-year longitudinal investigation conducted by de Bot and Clyne (1989, 1994). In 1971 they tested over 200 older (mean age = 65;4) Dutch immigrants who had moved to Australia after World War II. They recorded interviews with the informants in which they related their experiences about living in Australia. In addition, several formal language tests were also given. The nature of these tests is not described. In 1987, 40 of the original 200 immigrants were re-tested using the same methods and procedures. In considering their data, de Bot and Clyne (1994) selected a number of linguistic variables to examine. These included lexical richness, lexical borrowing from English, SVO word order in subordinate clauses (Dutch is SOV in subordinate clauses), adverbial placement, and overgeneralization of the two Dutch definite articles. They

also considered mean length of utterance (MLU). The results of this study suggest that the informants' L1 Dutch underwent minimal attrition over the 16 year period. Lexical borrowing from English, verb placement, article usage and MLU show no statistically significant differences between the two data collections. The only significant attrition of L1 Dutch discovered in this study appears to be limited to adverb placement, which increasingly took on the English place-time-manner order (Dutch = time-manner-place). Although de Bot and Clyne do not comment on why their subjects are able to maintain their L1 Dutch, two possible reasons are support from the community formed by the Dutch immigrants, especially the church, and the status they perceive in maintaining their L1.

The third type of intragenerational attrition is the loss of the L2 in the L1 environment. This type of attrition is of interest to the present thesis and will be discussed in detail in the next section. Briefly however, this is typically the loss of school learned languages which occurs once exposure through studying ceases. This may be the most common realization of language attrition. It has been the topic of numerous studies, several of which will be discussed below. Weltens (1989) in particular looked thoroughly at the loss of school learned French in Dutch L1 speakers. His cross-sectional study included 150 secondary school and university students who were divided into six groups based on number of years spent studying French, and number of years of non-use before Weltens' test (the "incubation" period). These time variables allowed Weltens to compare secondary school students who were just finishing their fourth year of French studies to University students with four years of secondary school French, but no university French, thus providing a longitudinal-like view of their attrition. The university students varied in the number of years since their secondary school studies, from two to four years. In addition to a self assessment test, all informants took the same battery of multiple choice tests which measured general proficiency, listening comprehension, reading comprehension and morpho-syntactic accuracy. Weltens concludes that his results are mixed. At the phonological and receptive skills (listening and reading) levels, he finds his test subjects' receptive proficiency in French is quite resistant to attrition, even after four years of non-use. However, attrition at the morpho-syntactic level "sets in rather quickly, and then levels off" (p. 92). In particular, he notes his university age informants lose between 9% and 14% of their accuracy on the morpho-syntactic test, especially pronoun comprehension.

In conclusion, Weltens suggests that his results do not exhibit a correlation between training level and period of non-use. That is, attrition is independent of training level - informants lose a fixed amount of their L2 knowledge rather than a proportion of that knowledge, independent of their original level (p. 92).

Of these three types of intragenerational language attrition, loss of the L2 in the L1 environment is of most interest to the present thesis. The next section will briefly describe the history of the regression hypothesis which has been used to explain L2 attrition, and will then examine studies which have used it to explain second language loss.

### 3.2. Second language attrition

Early research into language attrition tended to focus on students forgetting classical and foreign languages over summer vacation. Studies by Kennedy (1932), McMahon (1946), Geoghegan (1950) and Pratella (1969) fall into this category and will be discussed below. The 1982 appearance of Lambert and Freed's collection of papers on foreign language loss was one of the first major steps in establishing L2 attrition research as a sub-field of second language acquisition and initiated a more diverse and investigative interest in the field. As with earlier studies, Lambert and Freed's book focuses on classroom or foreign learned languages in an effort to stem loss. Considering the vast amount of energy, time and money invested in acquiring a second or foreign language at school, we can see that developing theories and methods aimed at curbing loss is one valuable reason to investigate L2 attrition (Weltens 1989:4).

As interest in L2 language attrition grew in the early 1980s, researchers recognized the need to develop relevant and testable hypotheses in order to research phenomena associated with attrition. Freed (1982: 6) categorizes L2 attrition hypotheses into the three areas included in (2).

(2) a. hypotheses based on regression theory: that is, a view of language loss as an unfolding or unraveling in reverse order of previously acquired forms.

b. hypotheses based on affective variables related to language learning and language maintenance.

c. hypotheses based on normative data of linguistic features controlled by proficient users of a language.

These three areas of hypotheses are useful distinctions with which to review and discuss attrition studies and will each be addressed in turn.

### 3.2.1. *Regression hypothesis studies*

Freed's first category - regression - refers to an ordered pattern of loss which is the reverse of the acquisition process. The regression hypothesis originates from Jakobson's (1941) research into the relationship between historical (diachronic) language change, language acquisition and sudden, pathological attrition (Welten's 1989: 5). Jakobson (1941) suggested that individual (synchronic) language acquisition may follow diachronic language change (i.e., from simple one word utterances to fully developed languages). This hypothesis is referred to as the recapitulation hypothesis, a term borrowed from 19th century biology which referred to the proposal that individual physical development from conception to full maturity (ontogeny) mimics the development of the species (phylogeny) (Haekel, 1910). Jakobson further proposed that the process of language attrition mirrors language acquisition in that loss follows the reverse path of acquisition. His concept of the regression hypothesis was based on evidence of phoneme loss found in child aphasic patients. As de Bot and Weltens (1991) correctly point out, however, aphasia is fundamentally different from child language acquisition or diachronic language change. Aphasia is typically the result of local damage to the dominant hemisphere of the brain which immediately leads to "specific deficits of parts of the language system" rather than gradual, global deterioration (p. 39).

Despite being inappropriate for investigating aphasia, the regression hypothesis has been applied to numerous L2 attrition studies. DeBot and Weltens (1991) identify three levels at which the regression hypothesis can be applied:

(3) a. between languages: with respect to the order of acquisition and loss of languages in multilinguals

b. within languages: in acquisition, perception precedes production, and spoken language precedes written language, in language loss, the sequence is reversed

c. within skills: as far as phonology, morphosyntax and lexicon are concerned. (p. 38)

Levels (3a) and (3b) apply to more sociolinguistic studies of language loss found in immigrant communities and instances of language shift and death (see above). These two levels will not be addressed. Level (3c) of de Bot and Weltens' distinctions is of particular interest to the present investigation, as it suggests linguistic aspects which can be tested for in attrition studies. The following review and discussion of the regression hypothesis will consider studies which address some of these phenomena.

L2 attrition studies identify two versions of regression hypothesis. One version proposes that the order in which attrition occurs is opposite to the order the language was acquired. That is, first learned, last lost. This is the historical concept of the regression hypothesis and we will follow Yoshitomi (1994: 12) in identifying this as the "reverse order hypothesis." The second version of the regression hypothesis claims that what has been learned best will be least forgotten. Yoshitomi refers to this as the "inverse relation hypothesis" (p. 12).

These two versions of the regression hypothesis - the reverse order hypothesis and the inverse relation hypothesis - clearly overlap. This overlap is expected as, to some degree, knowledge of the language acquired early in the acquisition process has had a long time to become deeply rooted in the mind and is unlikely to be readily lost. Conversely, knowledge acquired late in the acquisition process has not had as much time to become so firmly fixed and is more likely to succumb to attrition. Generally, this is to say that the vocabulary items and aspects of phonology, syntax, inflectional paradigms, etc. last learned will also be the least well learned and may be the items first forgotten. Conversely, the first items learned will be firmly established and be resistant to attrition.

One major problem with these two hypotheses is that the actual order of acquisition is difficult to establish. Unless a complete longitudinal corpus of acquisition data is available to complement an individual's attrition data, support for L2 regression is largely circumstantial. Furthermore, as de Bot and Weltens (1991: 46) point out, the regression hypothesis predicts a gradual process of loss which occurs in a fixed sequence. The hypothesis can therefore only apply to linguistic phenomena such as morpho-syntax and phonology which are gradually acquired over time and for which we know the acquisition sequences through previous research. Lexical items are much less suitable to a regression hypothesis as vocabulary acquisition sequences are hard to predict and can occur over short periods of time.

In the next section we will review studies which use the reverse order hypothesis to explain L2 attrition.

#### 3.2.1.1. *Reverse order hypothesis studies*

Cohen (1975) looked at the loss over summer recess of L2 Spanish in three L1 English-speaking kindergarten children in a Spanish immersion program in California. Cohen's study is rare in that the L2 acquisition process is not based on generalizations made from other L2 acquisition studies, but rather on data collected from the actual attrition subjects themselves while they were learning the language. From the beginning of the three children's kindergarten Spanish classes, data on their acquisition of Spanish verb morphology, articles and adjectives had been collected at regular intervals. For the actual study itself, data was collected via picture description tasks just before and just after the three month summer break. Cohen's results indicate a reverse order-type of regression, with the children exhibiting loss of later-learned verb agreement morphology such as the third person present tense inflection. Definite articles are also identified as being acquired later in the acquisition process and Cohen also finds evidence of them attriting early on. Additionally, evidence of the retention of earlier learned items such as first person inflection and indefinite articles is also found.

Two additional observations were made by Cohen. One is that "new incorrect patterns" (p. 137) (i.e. errors) emerged after the break which hadn't existed before it. One example of this was the use of the copula form *es* with the present participle to

incorrectly form the present progressive. This form had not been previously observed in the test subject's acquisition data and the subject had established stable control of the correct form of the present progressive prior to the break. Cohen's other observation is that subjects appeared to undergo "residual learning" in that there was "some sort of unlearning of incorrect patterns" during the break (p. 137). One example of this is the loss of one test subject's hypercorrection of the definite article over the recess. Cohen equates the loss of hypercorrection with the "unlearning" of incorrect patterns. But another possibility is that Cohen's test subjects underwent a "last learned first lost" pattern of attrition which initially manifested itself as a correction of overuse of a form, but given continued lack of exposure eventually led to deterioration of the correct use of the form.

Cohen's study is of interest to the investigation described in Chapter 5 of this thesis because of the careful investigative methods used. Specific forms were identified and tests which targeted those forms were given prior to and after the summer vacation.

Another study which targeted specific linguistic forms is a longitudinal investigation started by Jordens, de Bot, van Os and Schumans (1986), with results reported on by Jordens, de Bot and Trapman (1989). This study compares the loss of L1 German to the loss of L2 German in a Dutch-speaking environment. The L2 German speakers are L1 Dutch speakers with secondary school learned German who are divided into two main groups. One group consisted of final-year secondary school students who were completing their studies of German as a foreign language (GFL). The second group consisted of first-year and third-year Dutch university students who had not studied GFL since secondary school. The third year students were included in order to control for any residual learning in the first year students. The L1 German speakers were L2 Dutch speakers who immigrated to The Netherlands.

Jordens et al. (1989) find that the attrition of case marking differs between the L1 and L2 informants. The L1 German speakers appear to make an increasing number of mistakes, but do not default to one particular case. Rather, they use an increasingly simplified system which assigns case morphology according to perceived semantic function. The L2 attriters exhibit a more predictable reverse order pattern and use the nominative case as a default setting when they are unsure of the correct case. This is also the pattern observed by early GSL learners observed by Jordens' (1986) acquisition research.

Jordens et al. (1986: 173) propose two hypotheses for the regression of German case marking. One, the *Linguistic Hypothesis* states that "the process of attrition is seen as a reduction of the linguistic, i.e. both the morphological and the semantic system" (p. 173). This they propose will account for L2 attrition, and Jordens et al. (1989) indeed find this to be the case. The second, the *Cognitive Hypothesis* states that "attrition phenomena are due to the interaction between underlying conceptual structures and the way in which sentences are produced" (p. 173). This distinction however, does not seem supported by the evidence. First, Jordens et al. do not specifically define what is meant by "underlying conceptual structures." Presumably they mean the semantic relationships between agents and patients. But were these underlying conceptual structures to erode, then they would also erode in the L2 which is replacing the L1. And taken to the extreme, such erosion would ultimately lead to an individual's inability to perceive relationships within their surrounding physical environment. This is clearly unlikely to occur in non-pathological attrition. As Jordens et al. (1986: 161) themselves point out, "there is no simultaneous regression of cognitive abilities" in the normal processes of L1 and L2 attrition. That is, only the language attrites and other cognitive abilities are left untouched. Rather than cognitive versus linguistic differences between L1 and L2 attrition, first and second language loss can best be attributed to differences in the acquisition process. Yoshitomi (1994: 15) points out that L1 acquisition studies have observed that children "initially base their choice of form on a one-to-one correspondence to meaning" which is what Jordens et al. (1989) found in their German L1 attriters. The reverse order hypothesis therefore appears to account for both L1and L2 attrition.

In this study by Jordens et al. (1986, 1989) a specific grammatical system - case marking - of the attrition language is identified and tested. This is the approach I adopt in my investigation discussed in Chapters 5 and 6.

Another longitudinal study of child L2 attrition is reported on in Berman and Olshtain (1983), Olstain (1986), and Olstain (1989). These researchers used written and oral tests as well as story telling tasks to investigate the attrition of L2 English. Their informants were L1 Hebrew-speaking Israeli children who had recently moved from the United States back to Israel. Their interest was in the effects lack of spoken exposure to English and transfer from the L1 (Hebrew) would have on the children's L2 English. Three groups of children were used, aged 5 to 15. Results from the study show "a

limited reversal of the acquisition process, particularly with young children 5 to 8 years old" (Olstain,1989: 151). In particular, older children above 8 years seemed to maintain more stable levels of L2 English which were less prone to attrition.[1] Olstain (1989) suggests this stability is the result of the subjects having acquired a "useful level of literacy" (p. 157). By reading English in the Hebrew environment, the older children were able to maintain exposure to English and prevent it from suffering severe attrition

In proposing an explanation for attrition, Berman and Olshtain (1983) distinguish between language loss which occurs as a "general phenomenon", and loss which is the result of "L1 interference" (p. 233). But not knowing what is meant by "general phenomena" prevents such a distinction from being tested, except in the case of pathological attrition where the cause of attrition is identified as a brain lesion, or in a "Robinson Crusoe" situation, where the subject has no interlocutors for an extended period of time.

Two problems with the Olshtain et al. study are lack of detailed control for age and for level of English attained before re-immersion in the Hebrew environment. Olshtain (1989) identifies age as a variable which can effect ultimate L2 attainment and the articles associated with the study refer to "older" and "younger" informants. But for Olshtain, rather than reflecting cognitive development, age only signals whether a potential attriter is literate or not. Literacy, as noted above, provides exposure to the target language and aids maintenance. Another problem with this study is the lack of critical reference to the English acquisition process. Olshtain (1989) "assumes a normal sequence of acquisition to have had the following three stages: (a) the child uses some common irregular forms correctly; (b) regularization was applied to most verbs (including many irregulars); (c) a distinction was finally made between regular and irregular forms" (p. 163). However, a rigorous assessment of attrition needs to control more carefully for age, correlating particular ages with particular aspects of attrition.

Another study which addresses the reverse order hypothesis is Kuhberg's (1992) broad-ranging discussion of his longitudinal study of the attrition of L2 German in two L1 Turkish girls (Nayla and Nimet) after they moved from Germany back to Turkey at ages nine and seven, respectively. A feature of Kuhberg's study is that prior to collecting attrition data from the two Turkish informants in Turkey, Kuhberg also

[1] Pan and Berko-Gleason (1986) echo Olshtain's (1986) ideas on stability. They propose that once L2 learners acquire a "critical mass of language," the risk of attrition is greatly reduced (p. 204).

collected L2 acquisition data from a Turkish boy (Nadir) living in Germany (Kuhberg, 1988). Data collected from Nadir provided Kuhberg with evidence of acquisition sequences to support his claims of reverse order attrition. Kuhberg collected his attrition data by having his two attrition informants retell stories about pictures. Data collection sessions were tape-recorded and lasted about 40 minutes each. Data was collected regularly (every 6 to 8 weeks) for 15 months for Nayla and 20 months for Nimet. Both girls are reported to have been fluent in German upon their move back to Turkey. According to Kuhberg, the two girls' German was native-like in speed, communicative competence, and "included those modal particles which are so typical of German" (p. 140). Over time, however, Kuhberg's data show "(s)implification, overgeneralization, and over-regularization" of the girls' German (p. 138). He identifies several stages of attrition, including loss of categories such as verbs, codeswitching between German and Turkish bound morphemes, loss of object position pronouns and overall loss of vocabulary. Kuhberg's results lead him to suggest that very systematic language processes are at work which support a limited reverse order hypothesis. The loss of lexical items such as nouns, verbs and pronouns attrited in a manner which mirrored the L2 acquisition in the boy Nadir. But Kuhberg's data also show that some aspects of the two girls' language remain intact over time including verbal tense morphology and prepositions of direction (p. 152). This runs counter to the reverse order hypothesis as these aspects are not observed in the early acquisition stages of his GSL informant, and are presumed to be later-learned by the girls. Kuhberg concludes by advising caution in reading too much into the regression patterns he observes, noting that L1 influence also plays a role in attrition. He notes that German L2 attrition, in the face of an L1 other than Turkish, would most likely produce different attrition results.

Kuhberg's (1992) study raises two important issues. One is that attrition can occur selectively. That is, not all aspects of a language such as lexicon, morphology, tense, or phonology will attrite evenly. Secondly, he demonstrates the value of conducting a multi-informant longitudinal study as a means of providing a detailed description of individuals' changing language systems which can be compared.

A more recent study of the reverse order hypothesis is Hansen (1999a). Her study is different from the other reverse order studies presented here in that her informants were away from the L2 environment for decades, and whose L2s have presumably undergone attrition and fossilized. What language remained became part of

their long-term memory and will not suffer further attrition unless for pathological reasons. In this respect Hansen's is a study of long-term attrition whereas the other studies here look at attrition "in progress."

Hansen's cross-sectional study examines a specific system of a language, the attrition of negation in L2 Japanese among a population of L1 English-speaking adults. As Hansen points out (p. 143), Japanese is a highly agglutinative language and negation is expressed through a wide variety of morphology, the choice of suffix being determined by tense (non-past and past) and politeness level (formal, polite, and informal). Verbs, adjectives, as well as nouns and nominal adjectives (adjectives which behave morphologically like nouns) are all subject to this range of morphological marking. Compared to the relatively analytic nature of English negation, Hansen provides an interesting study of two typologically different languages. Her informants are three groups of American missionaries, each with a different length of time living in Japan and different incubation periods (the length of time since exposure ceased, i.e., since they returned from Japan to the United States). Table 3.1 summarizes information regarding her informants.

*Table 3.1. Hansen's (1999a: 146) three groups of informants*

| Group | n | Age | Sex | Months in Japan | Years since mission |
|---|---|---|---|---|---|
| 1 | 10 | 48-60 | Female | 24 | 25-34 |
| 2 | 10 | 46-53 | Male | 30 | 25-32 |
| 3 | 10 | 56-62 | Male | 36 | 33-37 |

Data was also collected from American missionaries in their early twenties who had recently (within two months) returned from two-year missions in Japan, and from Japanese university students as controls. A picture-stimulated role-play task was used to elicit negative answers in both the past and non-past, as well as in the various politeness levels.

In testing for the reverse order hypothesis, Hansen establishes the acquisition of L2 Japanese negation as proceeding from the negation of verbs (V-Neg), then the negation of nouns (N-Neg), followed by the negation of nominal adjectives (NA-Neg),

with the negation of adjectives (A-Neg) being acquired last. This order is based on L2 Japanese acquisition research by Kanagy (1994), as well as her own unpublished research. The regression hypothesis would then predict that attrition in L2 Japanese negation would initially occur with adjectives, then with nominal adjectives, then nouns, and finally with verbs. Hansen (1999) claims her results support this order of attrition. She finds that all her recently returned twenty-year-olds perform at native-like levels on all forms when compared to the controls' results. All her elder L1 English informants perform well on verbal negation, with the other three areas of negation tested for showing attrition to greater and lesser degrees, depending on amount of time spent in Japan. Table 3.2 summarizes, in a hierarchy of competence, Hansen's informants' levels of loss of negation in the various constructions she targeted, namely negation of adjectives (A-Neg), negation of nouns (N-Neg), and negation of nominal adjectives (NA-Neg). Also included at the bottom of the hierarchy are extreme cases of attrition in which informants overgeneralize a simplified system of negation for all three areas, and also those informants who overgeneralize the negation system and use L1 (English) word order. The table thus represents the most proficient informants at the top, and the least proficient at the bottom.

*Table 3.2. Hansen's (1999a: 149) degree of loss of L2 Japanese negation of adjectives (A-Neg), negation of nouns (N-Neg), negation of nominal adjectives (NA-Neg), and reversion to overgeneralization of simplified negation to in all three categories, including use of L1 word order*

| | Number of informants from each group performing at native-like levels for each targeted form. See Table 3.1 for summary of each group | | |
|---|---|---|---|
| | Group 1 (n=10) | Group 2 (n=10) | Group 3 (n=10) |
| Degree of attrition | | | |
| 1. Near-native forms | — | — | 1 |
| 2. A-Neg erosion | — | 3 | 4 |
| 3. A-Neg,loss N-Neg NA-Neg erosion | 6 | 7 | 5 |
| 4. Overgeneralization | 3 | — | — |
| 5. Overgeneralization and L1 word order | 1 | — | — |

Only one elder informant performed at a native-like level on all four forms. This was a Group 3 member (longest exposure), and the only one who continued his exposure to Japanese through reading. This supports Olshtain's (1989) suggestion that her older, literate child informants maintained their L2 English through reading. Table 3.2 shows that three Group 2 informants and four Group 3 informants underwent erosion (but not total loss) of only the A-Neg construction, but seemed to have maintained their NA-Neg, N-Neg, and V-Neg constructions. None of the Group 1 informants were able to maintain any approximation of the A-Neg construction (i.e., showed A-Neg loss), and exhibited erosion of the N-Neg and NA Neg. In four instances Group 1 informants' use of negation in Japanese was reduced to overgeneralization and transfer of L1 word order.

Despite the apparently strong support for the reverse order hypothesis Hansen finds in her results, she questions whether the acquisition process alone was responsible for the reverse order of attrition. Hansen notes that in addition to acquisition order, other factors need to be considered such as the "frequency of input, perceptual saliency, features of the attrition as well as the replacing language, and considerations of markedness and language universals" (Hansen, 1999: 150). Hansen's elder informants form in fact a very homogeneous group. All 30 missionaries were men who went to Japan with similar levels of education and motivation. Their pre-mission instruction in Japanese was also very similar (though Hansen doesn't elaborate on instructional methods or techniques). Hansen suggests these similarities contributed to their convergence on a similar learning sequence, and that the more similar this sequence was, "the greater the tendency will be for that sequence to be inversely mirrored in the course of attrition" (p. 150).

One problem with Hansen's study is that she can only assume her informants had in fact acquired the full system of Japanese negation. The other "in progress" studies reviewed above all found evidence of the linguistic aspects under investigation in their informants' production data immediately prior to testing for attrition. It is possible, at least for Hansen's Group 1 and possibly Group 2 informants, that they never acquired a complete grammar of negation and that she is testing for aspects of a system which they had never acquired.

Despite this problem, Hansen's study demonstrates the types of variables which need to be controlled for in a careful experiment of L2 attrition. The homogeneity of her informant groups, her use of native speaker controls, the use of a controlled experiment,

and her narrow focus on a specific system of grammar all reflect the type of experiment which should increasingly contribute to the L2 attrition field. The present thesis also attempts to control for these same variables.

The L2 attrition studies described above all claim to find evidence supporting the reverse order version of the regression hypothesis. The studies by Cohen (1975), Berman and Olshtain (1983), Olshtain (1986, 1989) and Kuhberg (1992) look at global attrition by describing the loss of a variety of lexical and grammatical aspects of the attriters' language. Cohen investigates paradigmatic verb morphology for both agreement and tense, and articles. Olshtain's studies looked at plural noun forms, past tense verb endings, loss of vocabulary, and L1 interference of word order. Kuhberg also looks at tense marking and vocabulary as well as numbers, prepositions, articles and pronouns. These studies provide a global picture of individual attrition, but do not closely examine a specific aspect of the underlying processes of attrition.

On the other hand, the studies by Jordens et al. (1986, 1989), and Hansen (1999) examine the attrition of specific aspects of their informants' grammar. Jordens looks at German case marking, and Hansen examines the attrition of negation in Japanese. The present thesis is more similar to these later types of studies in that it will investigate a specific grammar point, namely reflexive binding.

A common feature of studies testing for the reverse order hypothesis reflected in all these cases except Cohen's is their lack of actual acquisition data for the informants being tested. Cohen's study targets kindergarten children and collects data from them for a year prior to summer recess at which point attrition is expected to occur. The other studies, however, generalize from acquisition data collected in studies using different informants. Given the relative infrequency with which acquisition data is collected from the future attriters in question, reverse order hypothesis studies must rely on unrelated acquisition studies.

In the next section we will review studies which show support for the other version of the regression hypothesis, the inverse relation hypothesis.

3.2.1.2. *Inverse relation studies*

The second version of the regression hypothesis - the inverse relation hypothesis- states that, rather than *order* of acquisition, *quality* of acquisition is the crucial factor in determining what is lost. This is to say, attrition occurs in order of "best learned - last forgotten" (Weltens, 1989:6-7).

Two studies which have investigated the inverse relation hypothesis are Bahrick (1984a, 1984b) and Moorcroft and Gardner (1987). Bahrick's cross-sectional study included over 700 informants who either were currently studying, or who had studied Spanish as a foreign language. The current students were used as a control group. All informants were given reading, vocabulary and grammar tests. Depending on which group they were in - current or former students - informants were also asked to answer questions about their Spanish studies regarding variables such as length of time studied, type of instruction, length of time since the cessation of studies, highest level of proficiency attained, etc. Bahrick's results point to a large level of L2 skills being lost soon after cessation of studies. The type of language which is not lost, however, appears to be meaning (vocabulary) based and remains for long periods of times (more than 25 years), and possibly indefinitely, even with little or no exposure. Bahrick (1984a) refers to this long-term knowledge of the L2 as "permastore content" (p. 111) and concludes that learners with greater levels of language in the permastore also studied the longest and achieved the highest levels of Spanish. This suggests that, rather than simply quality of language, what influences quality of learning also effects degree of retention in permastore.

Bahrick does not discuss the possibility that a critical period exists after which access to the permastore is unavailable for the long-term maintenance of language. Presumably such a critical period would stretch well beyond puberty and into adulthood.

The correlation Bahrick finds between level of target language attained and degree of attrition experienced over long periods of time not only supports the inverse relation hypothesis, but points to the possibility that the level to which the L2 can be lost over time is limited. In addition, Bahrick's finding that vocabulary tends to exhibit a lower rate of loss than forms such as morphology lends support to the reverse order hypothesis as child L1 acquisition generally supports the theory that meaning based knowledge (vocabulary) is acquired before grammatical knowledge (Ingram, 1989).

A study which looks at the effects of summer vacation on elementary level L2 French is Moorcroft and Gardner's (1987) study of over 100 L1 English-speaking high school students. They based their results on both oral and written proficiency tests. Mean length of utterances, response times for oral answers, and number and length of pauses between utterances were measured. Grammatical accuracy and number of vocabulary items produced were also measured. The results from these tests indicate a reduction in all areas of proficiency after the three-month vacation. Only vocabulary and knowledge of idiomatic expressions seemed to be retained to a degree comparable to pre-vacation levels. Moorcraft and Gardner summarize their findings by pointing out that despite the maintenance of vocabulary and idioms, informants exhibited overall attrition because of the reduction in the grammatical patterns needed to use the vocabulary items. This overall loss is a reflection of the low level of French attained prior to the vacation period, a claim in support of the inverse relation hypothesis; higher level learners would not have lost grammatical constructions to the degree that it prevented them from using vocabulary items. Furthermore, in support of the reverse order hypothesis, specific grammatical items which were learned just before the vacation such as past tense negative forms are identified as undergoing substantial attrition during the course of the vacation period. Other forms which they identify as having been learned during earlier stages of the informants' L2 French acquisition period such as present tense negative forms seem to have withstood the effects of the three-month period of disuse.

Despite the findings of these studies which point to a correlation between order of L2 acquisition and order of attrition, and between knowledge of the L2 and degree of attrition, these two versions of the regression hypothesis fail to address the problem of adequately explaining the processes underlying L2 attrition. The reverse order hypothesis and the inverse relation hypothesis do not propose linguistically-based explanations of attrition patterns. Furthermore, while extent of L2 knowledge may be a reliable indicator of who will suffer greater or lesser overall loss of performance in the language, the inverse relation hypothesis cannot adequately predict what will be lost as L2 learner differences will cause different learners to acquire aspects of language other learners do not.

Several studies reflect the inconclusive nature of the two versions of the regression hypothesis. Bahrick's (1984a) study (described above) of L2 Spanish acquisition concludes that

> "the total *amount* of content to be forgotten during the five years following training is relatively constant for individuals at different levels of training, but this amount becomes a progressively smaller portion of total knowledge with higher levels of training" (p. 116).

This finding shows that lower level learners only *appear* to have lost more than higher level learners because loss at a low level may well result in communicative incompetence. Higher level learners who lose the same amount (but not necessarily the same type) of linguistic material will still be able to function in the L2. This distinction demonstrates one problem with relying on vocabulary-based production tests for measuring attrition. A test which probed Bahrick's lower level learners' knowledge of a particular grammatical system such as Hansens's (1999) test on negation (discussed above) might show that Bahrick's informants had actually maintained a higher level of language at the syntactic level, a linguistic aspect not addressed by his study.

Other studies which have found little or no differentiation in attrition rates among L2 learners with differing lengths of acquisition periods are Flaugher and Spencer (1967), Pratella (1969) and Kennedy (1932). Kennedy's results indicate that students with two years of exposure to Latin actually lost more than those students with only one year of exposure. On an individual level however, those students who initially learned the most seemed to retain the most. One of Kennedy's many conclusions states that "initial achievement is the significant factor in retention both from an absolute and relative viewpoint" (p. 146). Scherer (1957), in discussing his study of the loss of FL German in L1 English speakers, reports that learners with higher levels of acquisition seemed to forget more of their German over the summer than students with lower levels. He attributes this to the possibility that "good students had more to forget" (p. 276). And as we saw in Kuhberg (1992: 152), there are cases which exhibit "out-of-order" attrition. That is, supposedly simpler forms are forgotten while more difficult forms remain under control. Kuhberg cites the maintenance of tense marking in his

informants' later stages of attrition despite it not appearing in his early acquisition data. Or, items learned early in the acquisition period might be lost before items learned later. Another type of evidence which goes against the inverse relation hypothesis is instances of residual learning in which certain aspects (notably grammar-related items) of a subject's L2 appear to actually improve during the period of disuse. One example of this is observed by Cohen (1975: 136) who notes that one of his L2 Spanish informants tended to assign incorrect gender morphology to adjectives during a pre-summer-vacation test, but used correct forms on the post summer test. No mention was made about whether the informant may have had additional Spanish input during the summer vacation .

One study which reflects ambivalence about the regression hypothesis is Yoshitomi (1994, 1999.) This was a broad study of the attrition of L2 English in four L1 Japanese-speaking returnee girls aged between 9;6 and 11;3 at the start of data collection. The choice of subjects is interesting as two of the girls had short incubation periods (i.e., the time elapsed between returning to the L1 environment and the start of the data collection) of just several weeks, and two had longer, 13 and 15 month, incubation periods. This made her longitudinal study cross-sectional as well and provided her with data projecting out to over two years, as the data collection period lasted twelve months. The two recently returned informants provided Yoshitomi with attrition data for the first year after returning to Japan, and the two informants with longer incubation periods gave her data for the second year. The data collection was conducted via a variety of methods including free interaction, the telling of stories based on a looking at picture books , short, planned speeches, and a listening task requiring the selection of corresponding pictures. All these tasks were conducted by a native English speaker (not the investigator). Yoshitomi's tests targeted both receptive and productive skills in a variety of linguistic areas including phonology, verb morphology, article usage, lexicon, communicative strategies, use of complex syntactic structures such as biclausal sentences with embedded clauses, and mean length of utterance (MLU) (pp. 68-70). Despite the wide variety of tests Yoshitomi uses, she does not report including a native English speaking control group to norm her tests and seems to rely on her own (presumably native/near native) L2 English to judge accuracy.

The results of Yoshitomi's investigation lead her to believe that the discreet subsystems tested do show some signs of attrition. But she claims that knowledge of

language consists of connections and associations of information and that "a regression in one part of the connection or association would definitely affect other parts of it" (Yoshitomi, 1994: 222). For Yoshitomi, the global picture of attrition is important; to what level or why the subsystems erode is secondary. "Regression" is a step-wise process of overall erosion of whole-language rather than the ordered (reverse-order/regression) loss of sub-systems seen in the studies discussed above. She questions the results of previous L2 attrition studies which focus on discreet sub-systems, suggesting that initial stages of attrition can only be revealed "in the form of accumulative defects" (Yoshitomi, 1999: 93). Yet what Yoshitomi seems to ignore is that it is only through the detection and detailed analysis of "defects" in precisely those subsystems she cursorily examines that the accumulative effect of attrition can be not only observed, but meaningfully explained. In contrast to Yoshitomi, the present thesis limits itself to the detailed investigation of just one subsystem (anaphoric binding).

This discussion of the regression hypothesis suggests that in some instances it can predict general linguistic areas such as lexicon and verb morphology which may be subject to attrition. It is interesting to note the wide variety of L2 learners included as informants in the studies mentioned above, from Cohen's (1975) kindergarten-aged children, to Bahrick's (1984) and Hansen's (1999) older informants. Taken as a whole, however, attrition studies which use the regression hypothesis seem to find a wide range of results which often fail to concur with each other. Although there does seem to be a certain level of correlation between proficiency/order of acquisition and degree/order of attrition, the lack of consistency among findings points to the need for a more principled, linguistically-based means of explaining the attrition process on a subsystem by subsystem basis.

### 3.2.2. *Affective variable hypothesis studies*

The second attrition hypothesis category outlined by Freed (1980: 6) - hypotheses based on affective variables related to language learning and language maintenance - refers to research based on the role of attitudes and motivation in L2 attrition. These ideas have been borrowed from L2 acquisition research and applied to L2 attrition research. Gardner (1982) comments that "since attitudinal/motivational

characteristics are related to the level of second language proficiency, they will relate to second language retention" (p. 31). Variables associated with Krashen's (1981) affective filter hypothesis such as motivation, self-confidence, anxiety, attitude towards the target language, and attitude of the local linguistic community can contribute to the degree of retention or attrition which occurs.

Most evidence for a role of affective variables in language attrition is anecdotal in nature. Kennedy (1932) for example notes that "intention to continue with the study of Latin is a very important factor in terms of the amount of initial knowledge retained over the summer vacation" (p. 135).

Sharwood Smith (1983c: 223) describes a project by van Vlerken (1980) which looked at the attrition of L1 English in children living abroad in the Netherlands. van Vlerken reports that despite the high prestige value of English in The Netherlands, and its widespread use in the media, "clear symptoms of loss will appear at a relatively early stage" (p. 223). Loss in this case was lexical as well as syntactic. In a sense, this result runs counter to expectations as one would think an affective variable such as prestige value would go some way towards preventing attrition. Van Vlerken's study shows this is not always the case and hints at the powerful influence the dominant language, in this case Dutch, can have over the L2.

Kuhberg's (1992) study of L2 German loss notes that in addition to a lack of exposure to German upon returning to their native country (Turkey), German was completely marginalized both at home and at school. He notes that "[b]oth girls were under strong pressure at school to reach an advanced level of [Turkish] competence quickly [and] the family members gave absolute priority to Turkish" (p. 139). After 15 months for one and 20 for the other, each girl underwent almost complete attrition of her L2 German production and comprehension.

Reetz-Kurashige (1999) reports on L2 English attrition in 18 Japanese children (returnees) who have returned to Japan from lengthy overseas stays of from two to four years. She notes that, prior to attrition, the quality of their educational experience abroad varied widely. Those who enjoyed their lessons and had close rapport with their teachers generally excelled and exhibited lower attrition rates upon their return to Japan. Others who had less than ideal situations abroad tended to have higher levels of attrition (p. 47).

Two studies which address affective variables formally are Gardner, Lalonde and MacPherson (1985) and Gardner, Lalonde, Moorcroft and Evers (1987). These studies attempt to correlate learner attitude with motivation to study and maintain the language. Gardner et al. (1987) investigate the effects of summer vacation on L2 French in L1 English speaking Canadians. Motivation in this study is "shown to play a role in how much students attempt to use the language during the summer, and it is the 'Use' and the prior achievement that is responsible for individual differences in proficiency at the end of the summer" (p. 42).

The cumulative impression of L2 attrition studies is that they provide broad descriptions of linguistic items such as vocabulary and verb endings which are lost. The general claim seems to be that the better a learner's attitude towards the language is, the more he or she will attempt to seek out opportunities to use and maintain the target language. What they fail to attempt, however, is to provide explanations of *why* certain forms are lost. Are these losses caused by simple forgetting? Are changes in certain forms caused by transfer from the dominant language? Why does this transfer exhibit the patterns it does? Or perhaps such changes are caused by language-internal processes, such as an L2 speaker might make when recasting hypotheses about linguistic structures. One study we have seen which does attempt a close examination of a specific linguistic construction is Hansen's (1999) treatment of negation. The present study is a further attempt to focus narrowly on a specific linguistic phenomenon – reflexive binding – and explain the change and loss exhibited by a group of informants.

### 3.2.3. *Linguistic features hypothesis studies*

Freed's third category refers to language attrition hypotheses which compare attrition data to data from proficient speakers of the language. In this way, linguistic areas of the competing languages can be compared and potential areas of attrition in similar cases (same L1/L2, environments, etc.) can be identified. Weltens (1989: 8) refers to hypotheses developed from such research as "linguistic feature hypotheses."

Andersen (1982) developed an elaborate set of hypotheses to account for the overall phenomenon of L2 attrition. This set included hypotheses on what might be lost, the cognitive processes responsible for attrition, and the eventual fossilized state of

the attrited language. Preston (1982), in modifying Anderson's list, distinguishes three broad categories derived from Andersen's hypotheses. These are presented in (4).

(4) a. Attrition sites
b. Processes of attrition
c. Results of attrition

(Preston, 1982: 68)

"Attrition sites" refer to those areas or points at which attrition is likely to occur. These may include low frequency or uncommon items, low-content items, or irregularities within the language.

"Processes of attrition" include phenomena such as overgeneralization, transfer, mis-analysis and forgetting. Preston argues that only these four phenomena "conspire" to cause competence changes, regarding other processes such as paraphrase and avoidance as merely "on-line" compensatory performance strategies (Sharwood Smith, 1989: 192). But this is backwards from UG theory which holds that performance stems from competence and doesn't create competence as implied here. Perhaps we could say that Preston's processes of attrition are themselves caused by competence changes and that performance strategies are responses to that loss of competence.

A theory of second language acquisition applicable to processes of observed language attrition is Bialystok and Sharwood Smith's (1985) distinction between analyzed and unanalyzed knowledge. Analyzed knowledge refers to "the extent to which the learner has organized the linguistic information into structural categories" (p. 107). Such analyzed knowledge lets the speaker manipulate the language for use in a variety of syntactic, lexical and phonological ways. Unanalyzed knowledge is represented as limited "chunks" of language which are automatically processed. As the language is lost, the degree of analysis lessens, causing a slowing of access, which leads to hesitant speech patterns. Unanalyzed "chunks" of language may, however, linger on indefinitely with a relatively rapid degree of access.

The third distinction derived from Andersen's hypotheses identifies the 'results of attrition'. This refers to the degree of attrition obtained after the processes responsible

for attrition have eroded the particular aspects of language vulnerable to attrition and the L2 has essentially become fossilized in its attrited form.

Elaborating on these three distinctions - attrition sites, processes and results - Sharwood Smith repeats his (1983b) inventory of twelve linguistic variables which act as "a preliminary step towards a psycholinguistically motivated research programme" (1989: 193). These variables conspire to influence divergent change in one or the other of a bilingual's languages, a phenomenon Sharwood Smith (1983a:193) describes as "crosslinguistic influence" (CLI). Taken together, these twelve CLI variables provide attrition researchers with a useful tool for identifying sites, processes and results of attrition. These twelve variables are summarized below in (5a) to (5l).

(5) a. Typological Proximity: the bilingual's perception that constructions involving such aspects as morphology or syntax in the L1 are related to these aspects in the L2, resulting in a relatively high probability of transfer in those areas; idioms and bound forms are typically perceived as cross-linguistically unrelated and probably undergo minimal transfer.

b. Structural Similarity: transfer due to similarities between specific structures in the L1 which are perceived as similar to these structure in the L2 rather than to broad typological similarities.

c. Crosslinguistic Support: knowledge that a particular item is related in two languages can cause an individual to try to use a similarly based item in a third language. Although this may seem a performance tactic and not relevant to competence change, Sharwood Smith (1983b) found an informant who consciously borrowed an item despite being aware that it was deviant, but some time later denied that the borrowing was deviant. The deviant item became part of that speaker's competence.

d. Iconicity: refers to an individual who uses a semantically transparent word from one language as a model, or icon, for creating an alternative to a semantically equivalent but more obtuse word in another language. An example of this is the Japanese dominant Japanese/English bilingual who,

rather than producing the English "dentist," simply translates, on a word for word basis, the Japanese semantic equivalent "ha-isha" and comes up with "tooth-doctor."

e. Familiarity: how much exposure an individual has had with any given item. This is dependent on sociolinguistic factors such as linguistic environment and attitude towards the language.

f. Coding Efficiency according to Sharwood Smith, has to do not with the communicative effect on the interlocutor, but with the ease of processing experienced by the speaker (1989:196). This refers to one language having a supposedly "simpler" form than another; the simpler form is re-lexified and used in place of the more "complex" form.

g. Comprehensibility of the same languages shared by two or more people is a factor which causes forms from one language to be shifted to and used in a speech event primarily in the other language.

h. Solidarity among speakers may evidence itself as divergent forms being used to facilitate comprehensibility. On the other hand, solidarity may have a fossilizing effect on the language in that it will inhibit individuals from experimenting with new forms.

i. Input Sensitivity seems to refer to the degree of ease with which an individual takes on new linguistic information, either correct or divergent. Thus we can infer that a person who "picks up languages well" is sensitive to input, while a person who doesn't, isn't.

j. Associative Triggering is caused by what has been said or in anticipation of what will be said. That is, an item transferred into the language being used "can lead the speaker's mind into [i.e., can trigger the speaker to use further instances of] the language which is the source of the transfer and can therefore cause further instances of transference" either prior to or

after the actual occurrence in the utterance of the initially transferred item.

(Clyne, 1967:84)

k. Semantic Enrichment refers to the transfer of an item which has no semantic equivalent in the language it is transferred to. At the extreme end of this we may find that the "lacking" language actually takes over the word, e.g., the German *Gemutlichkeit* is said to have no one word equivalent in English and is now formally recognized as a loan word, particularly in literature (Merriam-Webster's, 1985: 510).

l. Ludic Potential is one of the enjoyable points of being bilingual; it is the ability to creatively play with '(p)erceived relations...' between the two languages (Sharwood Smith 1983b: 56). Such play may lead to some items or structures becoming processed as competence.

(Sharwood Smith, 1989: 194-198)

Of the twelve CLI variables summarized above, studies which focus on (5a), Typological Proximity, and (5b), Structural Similarity, may provide the most interesting results for the linguist interested in a theoretical account of L2 attrition. An example of (5a) above, Typological Proximity, can be found in Olshtain (1989). She finds her child L2 English attriters increasingly adopting the word order of their L1 Hebrew in their L2 English, producing sentences such as *Go upstairs, on your table the gum should be*, or *The students brought from home some materials* (p. 158). An example of (5b), Structural Similarity, can be found in Berman and Olshtain (1983). In their results, they find evidence of certain Hebrew lexical patterns transferring to their L2 speakers such as *Give me* (= let me) *to go to sleep*; *That pudding has the taste of bananas* (p. 229). In addition to descriptive studies such as these, identifying learners' perceptions and knowledge of their L1 and attriting L2 may provide support for a role for UG in L2 attrition. For example, understanding how a speaker of a [+ *pro*-drop] L1 comes to perceive *pro*-drop-related phenomena such as subject pronouns, *that*-traces and subject verb inversion when suffering from attrition in a [- *pro*-drop] L2 may provide support

for the continued operation of the subset principle. Evidence that the attriter increasingly exhibits principled [+ *pro*-drop] phenomena in the L2 as part of the attrition process could constitute evidence that the subset principle was responding to input from the [+ *pro*-drop] L1.

Research of a generative nature has until now, not been forthcoming. As we have seen, most work done in second language attrition in the 1970s and 1980s concentrated on collecting and quantifying performance level attrition data. They look at loss of features particular to the individuals undergoing attrition such as lexical items, inflectional features, and so on, without a rigorous effort being made to understand an attriter's underlying competence and the mechanisms contributing to L2 loss. In all cases, the regression hypothesis remains the favored descriptive tool of these studies. While this work can be invaluable to L2 teaching practices and government policies on the languages of minority groups, for the theoretical linguist, this approach lacks the kind of rigor and scientific approach necessary to understand and explain the processes of the attriter's mind as the language is being lost.

What seems necessary, then, is a new program of research into L2 attrition. L2 attrition researchers need to investigate the underlying L2 grammar by collecting and analyzing data in ways similar to how SLA research such as the acquisition of functional projections in L2 German by Vainikka and Young-Scholten (1998). Sharwood Smith and van Buren (1991: 17) recommend just such a framework noting that "however satisfactorily situated attrition data were with current **sociolinguistic** models, they could not be properly understood without considering their **psycholinguistic** status." Indeed, as far back as 1983, just a year after Lambert and Freed (1982), Sharwood Smith noted: "It would be interesting to see to what extent the tendency to restructure grammars in cases of loss was dictated by principles of markedness and core grammar as propounded in Chomsky (1980b)" (1983c: 229). In particular, then, it seems important to know whether an individual has lost or is even able to lose those kinds of underlying mental representations of his or her attriting language that may be identified as Chomskian style competence. Or perhaps the competence remains, but the ability to access it is lost. This could be thought of as performance attrition.

It seems that a new approach, similar to that carried out in current L2 acquisition research, would provide revealing results of L2 attriters' competence, independent of performance, if applied to language attrition data.

### 3.3. UG constrained attrition

If we accept Sharwood Smith's claim that we need to consider the attriting L1's psycholinguistic status and extend this consideration to second language attrition, what proposals can we offer? If adult learners do have access to UG, we might surmise that in the absence of exposure to the target language, any UG constrained aspects of the L2 such as parameterization of headedness or constraints on movement might behave differently from knowledge of non-UG constrained items such as lexical items or inflectional morphology. We might find that the UG driven aspects exhibit differing patterns of loss. Finding evidence of such a distinction would not only support a role for UG in adult second language acquisition, but would contribute to a more soundly based hypothesis of language attrition than the regression hypothesis seen in much of the attrition literature as discussed above.

In the event that lack of exposure did have some effect on parameter settings, several results seem possible. One possibility is that, just as L2 learners may initially transfer L1 parameter settings into the L2 and subsequently reset those parameters to the L2 settings in the face of positive evidence (White, 1989a), attrition may cause those L2 settings to switch back to the L1 setting in the face of "non-exposure." That is to say, if the L2 initial state is understood to be the L1 final state as proposed by Schwartz (1998), then attrition would result in L2 parameters reverting to those L1 settings. This could be referred to as parameter desetting.[2] This particular model of desetting might also be thought of as a form of UG constrained regression.

Another possibility is that through lack of exposure, parameters would deset, but not to an L1 setting. Rather, we might find that the grammar reconfigures itself to a third (i.e., neither L1 nor L2) setting. This third setting could either be UG sanctioned, or it could result in a non UG sanctioned "rogue" grammar. A UG sanctioned grammar would incorporate forms which, although not occurring in the L1 or the L2 are nonetheless UG permitted/constrained. Although not a UG-based study, Cohen (1975) did find "new forms" in his informants attrition data which had not been seen in the acquisition data (see Section 3.2.1.1 above). It is unclear whether these forms consisted of a so-called "rogue" grammar (Thomas, 1993) would exhibit non-UG constrained forms.

During language attrition, parameters may be affected in any of the four ways outlined in (6).

(6) a. No change
b. UG constrained reversion to L1 settings
c. UG constrained development of IL settings
d. Non-UG constrained change

In addition to the effects attrition may have on a specific parameter, we might also consider the repercussions a change in parameter setting would have on other aspects of an L2 grammar. If parameter setting is assumed to occur in conjunction with other linguistic attributes during acquisition, we might assume that parameters do not deset in isolation, but in conjunction with cluster attributes associated with a particular parameter. In a study of the attrition of pro-drop in a language such as Spanish by a non-pro-drop L1 speaker, for instance, we might expect to find that as pronouns cease to be omitted, other phenomena associated with pro-drop as investigated by White (1985) disappear as well. This might include the loss of subject-object inversion, and the insertion of pleonastic subjects such as *it* in English, or *es* in German. By considering specific parameters within the principles and parameters (PPT) framework, we may begin to develop a theory of attrition which predicts how such parameterized aspects of language might behave under the pressure of attrition. Desetting of certain parameters in a significant pattern in the face of disuse would lend support to the proposition that UG may be available to the L2 learner.

### 3.3.1. *Selecting a testable aspect of UG*

In attempting to explain language loss, attrition studies have largely ignored generative-based claims about language. One theoretical discussion article about UG-based L1 attrition is Sharwood Smith and van Buren (1991), mentioned above. In this article, the researchers recognize the importance of knowing whether "a given subject

[2] The term "desetting" was suggested to me by Bonnie Schwartz August, 1996.

has lost **or is even able to lose** those kinds of underlying mental representations of his or her first language that may be referred to as L1 competence" (p. 17). Another discussion of UG-based attrition is Ioup (1994) who asks "what is the logical problem of language attrition?" By this she turns Bley-Vroman's (1989) question on its head by asking how the learner, despite having been provided with "rich, complex data" (p 13) fails to maintain the L2. To my knowledge, these are the only UG-based attrition discussions to appear in the SLA literature. No L2 attrition investigations have been conducted from within a generative framework. To initiate such a program of investigations, a specific testable aspect of UG should be identified, such as a principle or a parameter. As with acquisition studies, a battery of tests would need to be drawn up to measure the specific principle or parameter in question. In the case of an attrition study, rather than increases of proficiency, the degree of retention or loss of that aspect would be measured. White (1985), Hilles (1986) and Schachter (1989) describe investigations into specific UG-determined aspects of second language acquisition. Both White and Hilles look at the prodrop parameter, and Schachter looks at subjacency. Schachter takes up a challenge laid down by Bley-Vroman (1989) to show that non-native speakers of a language show the same intuitions about a UG instantiated aspect of grammar as L1 speakers of that language (p. 62-63). In selecting her UG constrained aspect of language to investigate, Schachter notes that "not all principles are appropriate candidates for testing since not all principles are accessible to surface level analysis" (p. 76).

Another principle which lends itself to testing is binding principle A of the binding theory as described by Chomsky (1981). Principle A - the principle of anaphoric binding - has been the subject of numerous L2 acquisition studies which will be discussed in more detail in Chapter 4. These studies provide the attrition researcher with a wide range of examples of different learners and learner variables, and examples of how the principle has been tested for. By applying the research methodology of second language acquisition studies to a sample of attritiers, I hope to present an alternative type of attrition study to the usual regression-style approach, an alternative which focuses closely on one aspect of the attriters' underlying knowledge of language.

## 3.4. Conclusion

Just as a language learner's grammar goes through various stages, or interlanguages along his or her unique path towards native like competence, the grammar of a language attriter also goes through various stages along a path to complete loss. Both processes typically stop before the end of their respective end points. That is, few L2 learners acquire a native command of their target language, and language attriters rarely if ever lose 100% of their attriting language. Evidence that the two paths - language acquisition and language attrition - mirror each other would support a regression-type hypothesis. Differences between the acquisition and attrition paths such as suggested by Cohen (1975), however, could generate interesting findings. From within a generative framework, we could ask if L2 parameter settings change to the dominant L1 settings? Do they revert to a "default" or "unmarked" UG determined setting? Or do they stabilize and remain part of the L2 grammar? If, for example, we found an informant's L2 to initially adhere to UG-sanctioned constraints, but subsequently to violate those constraints in the face of disuse, we would be forced to account for this "rogue" grammar and explain why the acquisition path and the attrition path are not mirror images of each other. We might also find that certain aspects of the grammar appear more robust than others in the face of attrition. Or we might find that parameterized aspects of the grammar change in a UG sanctioned manner, but that the new grammar approximated neither the L1 nor the L2. This third possibility would be of most interest as such a change would raise the possibility that attrition could somehow be UG-guided as L2 acquisition research argues (i.e., Finer and Broselow, 1986).

In choosing to develop tests to investigate the loss of L2 English reflexive binding in Japanese speaking adults, the following three hypotheses were developed:

> Hypothesis A: Principles of reflexive binding already instantiated in the test subjects' English will undergo change due to lack of exposure.
>
> Hypothesis B: Changes observed in the test subjects' control over reflexive binding due to lack of exposure will be UG constrained.
>
> Hypothesis C: A correlation will be exhibited between age at first L2 exposure and degree of retention of L2 reflexive binding principles.

The next Chapter will review the main aspects of binding theory and discuss how it has been investigated in second language acquisition research. By discussing the investigative goals and methodology of studies on L2 binding, the appropriacy of applying this line of investigation to L2 attrition will hopefully be made clear.

# Chapter 4: Defining a Parameter to Test for Attrition: Binding

## 4.0. Introduction

This chapter discusses the major aspects of the binding theory, a theory of the relationship between anaphors, pronominals and their antecedents. One purpose of this chapter is to define the aspect of UG investigated in Chapters 5 and 6. Another purpose is to describe and discuss investigative techniques and goals included in L2 binding studies in order to create a framework for the tests and procedures used in this investigation.

The first section below describes binding theory as proposed by Chomsky (1981). In particular, the parameterized model of L1 binding as proposed by Manzini and Wexler (1987) and Wexler and Manzini (1987) will be included to illustrate the variation in binding constraints found across languages. Section 2 will discuss studies which investigate the L2 acquisition of binding. Section three will consider alternatives to the parameterized model of binding. Throughout this chapter, the emphasis will be on principle A, the UG aspect under investigation in Chapters 5 and 6.

## 4.1. A parameterized model of binding theory

Binding is a sub-theory of Chomsky's Government and Binding theory (1981, 1986a) which concerns anaphora. Anaphora refers to the relation of a category that lacks independent reference (an anaphor or a pronominal) with a category that has independent reference (an antecedent). Binding theory is therefore a theory which accounts for "the relations, if any, of anaphors and pronominals to their antecedents" (Chomsky 1982: 6). Anaphors here include reflexives, such as *herself*, and reciprocals, such as *each other*, which always refer back to a c-commanding antecedent. Pronominals refer to pronouns such as *him*, *her*, etc. which must be free in their governing category, but can be either free or bound within other clauses of a sentence. Binding theory also addresses NPs which always have independent reference. These are R-expressions ("referring expression") such as *Julia* or *her Majesty*. Chomsky (1986a: 166) summarizes the three binding principles as:

(1) Principle A: An anaphor is bound in a local domain

(2) Principle B: A pronominal is free in a local domain

(3) Principle C: An R-expression is free

"Bound" means the anaphoric expression and its antecedent are co-indexed, and that the antecedent c-commands the anaphor, as in (4).

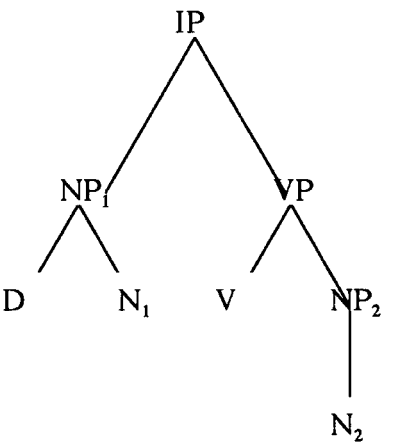
(4) Alex knew $\text{Jason}_i$ deceived $\text{himself}_i$

In (4), *himself* can refer back to *Jason*, but not to *Alex*. For this thesis, the definition of c-command proposed by Reinhart (1976) and adopted by Chomsky (1986a) will be used. This definition assumes that

(5) A node $\alpha$ c-commands $\beta$ iff (i) $\alpha$ does not dominate $\beta$ and (ii) the first branching node dominating $\alpha$ also dominates $\beta$ (Chomsky 1986a: 8)

The relationship defined in (5) is illustrated in the simple structure in (6).

(6)



In (6), the first branching node above V is VP, making any node dominated by VP C-commanded by V, including $NP_2$ and N demonstrating the general definition that an item can only C-command other items it is higher than or equal to in the tree.

"Local domain" in (1) and (2) above refers to the governing category, which is defined as "the minimal governing category of $\alpha$, where a governing category is a maximal projection containing both a subject and a lexical category governing $\alpha$" (Chomsky, 1982: 169).[1]

Let us briefly look at principles B and C before beginning a detailed discussion of principle A. Principle B states that a pronoun must be free, i.e., not co-indexed with a c-commanding NP, in its governing category. Examples (7), (8), and (9) illustrate principle B in English and Japanese, the two languages under consideration in this thesis.[2]

(7) a. $John_i$ hit $him_{*i}$

b. $John_i$-ga $kare_{*i}$-wo tatai-ta
John-NOM him-ACC hit-PAST

(8) a. $John_i$ said that $Bill_j$ hit $him_{i/*j}$

b. $John_i$-wa $Bill_j$-ga $kare_{i/*j}$-wo tataita-to itta
John-TOP Bill-NOM him-ACC hit-COMP say-PAST

(9) a. $John_i$ told $Bill_j$ to hit $him_{i,*j}$

b. $John_i$-ga $Bill_j$-ni $kare_{i,*j}$-wo tatake-to itta
John-NOM Bill-DAT him-ACC hit-COMP tell-PAST

In (7a-b), *John* c-commands the pronoun, which must be free in its governing category. Thus, neither *him* nor *kare* can take *John* as its antecedent. In (8a) and (b) and (9a) and (b), *him* and *kare* are free in their governing category, Such that *him* and *kare* cannot be

[1] A governing category is a maximal projection such as a VP, NP, PP, AP, or an IP, and contains a subject in the IP and NP, and a lexical category such as N, V, A, or P which governs the NP.

[2] Japanese case-markers are identified as: NOM = nominative, TOP = topic, ACC = accusative, DAT = dative.

co-indexed with *Bill* which is included in the governing category. English and Japanese share the same interpretation of principle B.

Principle C of binding theory states that R-expressions should always have independent reference. This is apparent in sentences such as (10) and (11).

(10) a. He hit the president$_j$

b. Kare-ga$_i$ dai-touryo$_j$-wo nagutta
he-SUBJ president-ACC hit-PAST

(11) a. Mari said she respects her

b. Mari-wa$_i$ kanojo$_j$-wo sonkei-shiteiru-to itta
Mari-TOP she-ACC respect-PRES-COMP said

In (10) and (11), the referring expressions *the president* and *Mary* must be free and hence refer outside the sentence and would be ungrammatical were they co-indexed to their respective matrix objects.

Binding theory as described above originates largely from investigations into English. But it has been recognized that certain principles of the theory vary across languages in how they describe certain binding constraints (see Chomsky, 1981a: 229). In particular, the definition of "governing category" varies from language to language. Yang (1983) provides an in depth account of how principles A and B vary among almost 20 different languages. He attempts to parameterize the language specific variations in binding he finds among languages such as Dutch, English, Hindi, Japanese and Russian.

Although of interest, principles B and C have received much less attention in L2 acquisition studies. This may be due to the greater amount of variation found in different languages in principle A-related phenomena. For example, the study described in Chapters 5 and 6 of this thesis investigates L2 English binding in Japanese speakers. English and Japanese represent languages which are very different regarding what is allowed as a governing category, but are the same for principles B and C as illustrated earlier. For these reasons we will now focus on principle A.

Let us begin by contrasting the English and Japanese governing categories. In English, a reflexive must be bound to an antecedent within the minimal clause

containing a subject, regardless of whether the clause is finite or non-finite. Japanese, on the other hand, does not require the reflexive to be bound to a local antecedent, but allows co-reference with either the "local" or "long-distance" (matrix) antecedent. Examples (12) through (15) compare English and Japanese and illustrate this point.

(12) a. Alice$_i$ blamed herself$_i$

b. Alice$_i$-ga zibun$_i$-o semeta

Alice-NOM self-ACC blame-PAST

(13) a. Alice$_i$ said that Sue$_j$ blamed herself$_{*i/j}$

b. Alice$_i$-wa Sue$_j$-ga zibun$_{i/j}$ o semeta-to itta.

Alice-TOP Sue-NOM self-ACC blamed-COMP say-PAST

(14) a. Alice$_i$ told Sue$_j$ to blame herself$_{*i/j}$

b. Alice$_i$-ga Sue$_j$-ni zibun$_{i/j}$ o semete-to itta.

Alice-NOM Sue-DAT self-ACC blame-COMP say-PAST

(15) a. Alice$_i$ said that Sue$_j$ said that Mary$_k$ blamed herself$_{*i/*j/k}$

b. Alice$_i$-wa Sue$_j$-ga Mary$_k$-ga zibun$_{i/j/k}$-wo semeta-to itta.

Alice-TOP Sue-NOM Mary-NOM self-ACC blamed-COMP say-PAST

In (12a) *herself* refers to the local antecedent *Alice*, which also happens to be the matrix subject. The Japanese equivalent of this monoclausal sentence, (12b), corresponds to the English interpretation. In (13a) however, *herself* refers to *Sue*, but not to *Alice*, as *Alice* is outside the local clause. In the Japanese equivalent of this biclausal sentence, (13b), however, the reflexive *zibun* is bound by an NP that can appear in any position in the sentence that c-commands *zibun*. In this way the antecedent of the reflexive is not limited to the nearest clause and may refer back to *Sue* or *Alice*. Since the antecedent here may be in the matrix sentence subject as well as in the local clause, the Japanese range of potential antecedents is greater than that of English.

(14a) and (b) illustrate the same phenomena as in (13a) and (b), but here the subordinate clause in (14) includes a non-finite verb form whereas (13) includes a finite form. Despite the tense difference, the grammatical antecedent of each reflexive in each language respectively remains the same. As we will see later, this finiteness distinction does affect reflexive binding in other languages such as Russian in that long-distance binding is allowed out of a clause with a non-finite verb, but not out of one with a finite verb.

(15a) and (b) illustrate more extreme forms of (13) and (14). Despite the multiple NPs, the reflexive *herself* in English only refers back to the nearest antecedent, *Mary*. In Japanese however, *zibun* in (14b) can refer back to any of the three possible antecedents - *Alice*, *Sue*, or *Mary*.

The above examples illustrate how English and Japanese differ with respect to what clause may function as a governing category. English is more restricted, Japanese freer. This distinction has been captured by Manzini and Wexler's (1987) Governing Category Parameter (GCP) which they describe as "a single definition of governing category with a parameter having five different values" (p. 419). This definition of governing category appears in (16).

(16) $\gamma$ is a governing category for $\alpha$ iff $\gamma$ is the minimal category that contains $\alpha$ and a governor for $\alpha$ and has

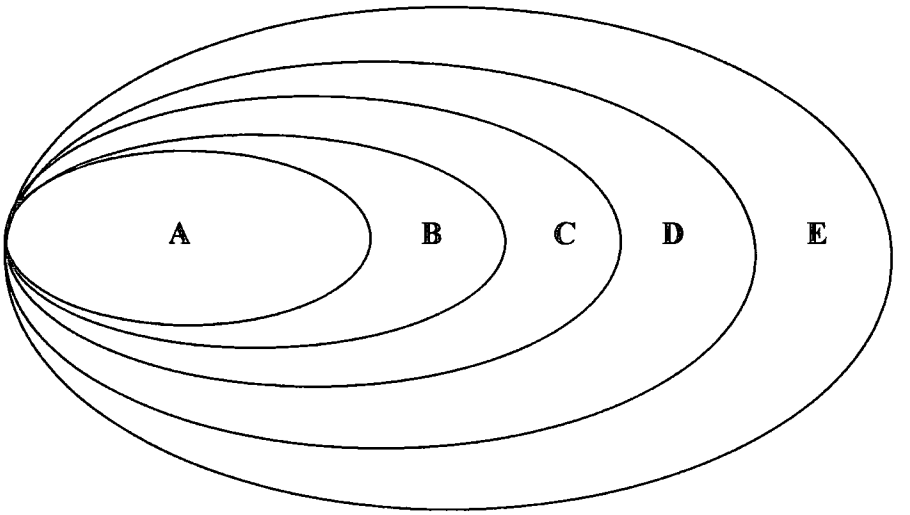
a. a subject; or
b. an Infl; or
c. a Tense; or
d a "referential" Tense; or
e. a "root" Tense.

(Manzini and Wexler, 1987: 419)

Each setting in a-e in (16) corresponds to the way the parameter may be realized in different languages. Type E languages like Japanese include every NP that is allowed and therefore is the most inclusive language type. Type D languages such as Icelandic include the possibilities of the C, B, A types. A type C language like Russian includes B

and A types. Type B languages like Italian include type A. Type A languages such as English are the most restrictive. The values for this increasingly strict value for governing category show an entailment relationship as diagramed below in Figure 1.

*Figure 4.1. Entailment relationship between possible governing categories*




In this way, the GCP distinguishes languages on the basis of how much linguistic material can separate the antecedent from the reflexive. Consider the following sentence (17) which illustrates the range of variation captured by the GCP, from the most restrictive to the least restrictive domain claimed by Manzini and Wexler.

(17) *Martin* mentioned that *Al* was asking that *Sam* convince *Ken* to consider *Mark* doubtful of **himself**. (after Finer, 1991: 355)

In (17), a type A language only allows *himself* to refer back to *Mark*. A type B language would allow **himself** to refer back to *Mark* or *Ken*. Type C languages would allow reference back to either *Mark*, *Ken*, or *Sam*. Type D languages would allow *himself* to refer back to either *Mark*, *Ken*, *Sam*, or *Al*. And a type D language would allow reference back to any of the five possible antecedents. This example demonstrates how the range of interpretations increases in an ever-widening relation of proper

inclusion, with every interpretation higher on the hierarchy included in the set of interpretations lower in the hierarchy.

This entailment relationship establishes English as an example of the most restrictive setting, and Japanese as the least restrictive in that any NP in a Japanese sentence can serve as the antecedent of the reflexive. The contrast between the two languages is illustrated in (13) above, repeated here as (18).

(18) a. Alice$_i$ said that Sue$_j$ blamed herself$_{*i/j}$

b. Alice$_i$-wa Sue$_j$-ga zibun$_{i/j}$ o semeta-to itta.

Alice-TOP Sue-NOM self-ACC blamed-COMP say-PAST

In addition to differences in GCP, the two languages differ in their choice of what can count as a proper antecedent for the reflexive. This parameter is called the Proper Antecedent Parameter (PAP) and was proposed by Manzini and Wexler (1987). It has two values with respect to what is allowed as the antecedent of the reflexive. These are represented in (19).

(19) A proper antecedent for $\alpha$ is

a. a subject $\beta$; or

b. any element $\beta$

(Manzini and Wexler, 1987: 431)

In setting A, only the subject is allowed to serve as the antecedent (subject orientation). Setting A can be found to accurately describe a language such as Japanese which requires the antecedent to be in subject position. In setting B, any NP is allowed as the antecedent. This setting describes a language such as English in which any NP within the restricted governing category may serve as the antecedent. Consider the following pair of sentences in (20).

(20) a. Mary$_i$ talked to Su$_j$ about herself$_{i,j}$

b. Mary$_i$-ga Su$_j$-ni zibun $_{i/*j}$ nitsuite hansita.

Mary-NOM Su-DAT self about talked

The examples in (20) illustrate how English *herself* can take either potential antecedent, *Mary* or *Su* and is a PAP type B language, while Japanese *zibun* can only take *Mary*, the subject antecedent, making it a PAP type A language. This forms a subset relationship between the two languages; Japanese, the type A language, is the unmarked setting of the PAP and is a subset of the less restrictive, more marked type B language, English. In this way, we can see that the subset relationship of the PAP in English and Japanese is reversed from that of the Governing Category Parameter. According to the GCP, a language such as English is the unmarked setting and is a subset of Japanese, the most marked. For the PAP, the opposite holds; Japanese is the unmarked subset of marked English.

The parameter setting model developed by Manzini and Wexler captures and predicts much of the variation found in the relationships between reflexives and their antecedents in native speakers. The next section will review studies which have used this parameterized model of binding to conduct research into the L2 acquisition of binding.

### 4.2. L2 acquisition of binding

The research reviewed here concerns the interpretation of reflexives by L2 English learners. The relationships which hold between anaphors and their antecedents is subject to principle A of the binding theory which is thought to be representative of knowledge of language attributed to UG. It has been proposed that acquisition of this knowledge, in particular what may be interpreted as the governing category, is subject to parametric variation (Wexler and Manzini, 1987).

This section reviews investigations into the L2 acquisition of English reflexives. These include studies by Finer and Broselow (1986), Cook (1990), Finer (1991), Thomas (1989, 1991, 1993, 1995), Hirakawa (1990), Maclaughlin, (1995), and Christie

and Lantolf (1998). Some studies indicate that, where the L1 and L2 settings differ, L2 learners can reset their L1 parameter to the L2 setting. Other studies find evidence against this and claim L2 learners are incapable of resetting parameters.

Because there is no body of literature on the L2 attrition of binding, I hope that a discussion of studies of L2 acquisition of binding will serve to establish the framework from within which I conduct my investigation. In particular, test design and collection procedures common to acquisition studies and used in the investigation presented in Chapters 5 and 6 will be highlighted.

### 4.2.1. *A parameterized explanation of L2 binding*

One of the earliest generative studies on the L2 acquisition of English reflexives is Finer and Broselow (1986). This study investigated how Korean adult L2 learners of English determine the binding domain of English reflexives. English and Korean are in a subset/superset relation with respect to reflexive binding domain. Korean, like Japanese, takes Manzini and Wexler's maximum superset value (Setting E) and English takes the lowest subset value (Setting A). Finer and Broselow attempt to show whether or not their test subjects apply their L1 parameter setting to the L2, or reset the parameter to the proper L2 setting ("A"). Despite the small scale of their study, its implications are of value and the study is often discussed by researchers. Finer and Broselow tested six Korean speakers learning English in an intensive English program at a university in the United States. In an oral elicitation task, subjects were shown pairs of pictures that depicted two easily identifiable characters (Mr. Fat and Mr. Thin) and asked to identify which picture contained the relationship between the two characters as indicated by oral prompts. They heard sentences such as (21).

(21) a. Mr. Fat thinks that Mr. Thin will paint himself
b. Mr. Thin asks Mr. Fat to paint himself

For Korean learners who apply their L1 binding setting to these sentences, the interpretation of these sentences should be ambiguous because their L1 allows either the

local or non-local antecedent to be bound with the reflexive *himself*. In both sentences (21a) and (b) the reflexive can be bound either to *Mr. Fat* or to *Mr. Thin*. In English, however, the reflexive can only be bound to the local antecedent, i.e. *Mr. Thin* in (21a) and *Mr. Fat* in (21b).

Excluding four control sentences, Finer and Broselow's test included 16 complex sentences, eight of which contained reflexives and eight of which contained pronominals. Four of the sentences containing reflexives had finite (tensed) embedded clauses like (21a) and the other four infinitive (non-tensed) clauses like (21b).

Finer and Broselow's results indicate that although 92% of the Korean learners' responses chose the local NP as the antecedent of the English reflexive in the finite embedded clauses (21a), only 58% of them chose the local antecedent in sentences containing an infinitive in the embedded clause as in (21b). Furthermore, as many as 38% chose the non-local antecedent in sentences containing infinitival constructions like (21b). Table 4.1. summarizes their results.

*Table 4.1. Finer and Broselow (1986: 165) responses to tests on L2 English reflexives by 6 Korean speakers*

| | Local | Nonlocal | Either |
|---|---|---|---|
| Tensed Clause | 92% | 8% | 0% |
| Infinitive Clause | 58% | 38% | 4% |

Finer and Broselow note that there is a sharp distinction between judgements on sentences which contain a tensed complement and those with an infinitival complement. Tensed sentences are mostly judged according to English binding principles with the anaphor bound to the local antecedent. However, judgements on sentences involving infinitival complements conflict with English binding principles. Finer and Broselow initially observe that it seems the learners follow English binding principles for sentences like (21a), and Korean principles for sentences like (21b). But they later conclude that the infinitival sentences actually do not follow Korean binding principles because the tensed/infinitival distinction does not apply to Korean., and neither does the distinction play a role in the distribution of reflexives in English. This leads Finer and

Broselow to conclude that "the distinction that emerges cannot be attributed to either the native language or the target language grammar" (p. 160). They suggest that their test subjects may have converged on a GCP setting for an intermediate category such as type C (Russian-like languages) in Manzini and Wexler's hierarchy. Though inconsistent for both Korean and English, such a result would be consistent with binding principles and would indicate "principles of UG are clearly playing a role here" (Finer and Broselow, 1986: 161).

An additional result of this finding concerns the Subset Principle (Berwick, 1985; Manzini and Wexler, 1987; Wexler and Manzini, 1987) and how it operates on the GCP. The Subset Principle is motivated by the possibility that a child L1 learner, presumably exposed only to positive evidence, might overgeneralize a particular parameter and adopt an incorrect setting. To prevent this so called *subset problem,* (MacLaughlin 1995: 145), the Subset Principle requires that a learner choose the parameter setting which results in the smallest grammar compatible with the input. Without sufficient positive evidence, an L1 learner could never choose a setting which overgeneralized the grammar.

Considering Finer and Broselow's results, had their Korean learners' L2 acquisition been fully constrained by UG, and had the Subset Principle been fully operational, they should have chosen the most restricted setting of the GCP, setting A, which happens to be the setting for English. Alternatively, had the learners had no access to UG or the Subset Principle and adopted their L1 Korean setting for English, Finer and Broselow's results would have shown an IL grammar with the widest GCP setting. In the end, however, their informants opted for a mid-range type C setting. This raises an interesting question. If L2 learners don't receive meaningful negative evidence, how is it that Finer and Broselow's learners came to partially reset the GCP to a more restricted setting? The Subset Principle operates under the assumption that learners move from a restricted setting to increasingly open settings; resetting parameters in the opposite direction is not an available option. One possible explanation is that learners do in fact receive negative evidence which allows them to move towards a more restricted setting. Another possibility is that the Subset Principle, at least as it applies to the GCP, was partially available to the learners at the time Finer and Broselow collected their data, but, perhaps because their learners' GCP was at the Korean (i.e., widest) setting, was unable to fully restrict the setting.

Finer (1991) again tested the acquisition of reflexives by L2 learners. In this study he includes not only a greater number of Korean test subjects (n=30), but also Japanese (n=20) and Hindi (n=29) speakers. Hindi exhibits GCP Type C binding constraints, similar to Russian (governing category has either a subject, Infl, or tense). All subjects were ESL students in the United States. He examined the same kinds of structures with the same type of task as in Finer and Broselow (1986) and finds similar results for the Korean and Japanese learners as the (1986) study found. However, for the Hindi speakers, he finds a very high degree of accuracy for local binding, regardless of embedded clause type (tensed or infinitive). This may be attributed to the test subjects' pre-puberty exposure to English, as well as the prestige status of English in India and its status as an official language there. As for the PAP, the Korean and Japanese learners show a greater tendency to bind to the matrix subject than to the object, but they still have some instances of binding the reflexive to the object, showing they have acquired the English setting for the PAP to a limited degree. The Hindi speakers almost exclusively bind to the subject, the choice of their L1. Finer remarks that this may indicate that the subjects are resorting to their L1 setting to interpret the test sentences, or that "they have retreated to the unmarked setting of the PAP and have not yet assimilated the positive evidence showing object antecedents" (p. 362).

Finer's investigations have been criticized for ignoring interpretations of the data which would indicate the presence of rogue grammars in test subjects' ILs (Thomas, 1993: 58). Rogue grammar here refers to a grammar not sanctioned by UG, specifically within the GCP/PAP framework (see Thomas, 1991, 1993). An example of this would be an illegal setting of the PAP in which a test subject allowed only subject antecedents while disallowing subject antecedents. Manzini and Wexler's parameterized binding principles require wider settings which allow matrix subject antecedents to also allow matrix object antecedents. Allowing only the subject is not an option (Thomas, 1991: 377). Despite this, Finer and Broselow and Finer's work is important because they raise the possibility that L2 learners may actually form a UG sanctioned grammar which is derived neither from L1 transfer, nor from complete acquisition of the L2 grammar. This would mean UG plays some role in the formation of an adult learner's interlanguage.

Finer and Broselow (1986) and Finer (1991) address the acquisition of binding in learners whose L1s do not share the same types of binding relationships which English

has. Thomas (1989) differs from the two previous studies in that she includes one group of learners whose grammar - Spanish - does include an English-like interpretation of binding characteristics. Spanish binds the reflexive with only antecedents in the same clause (setting A). The other group in her study was made up of Chinese speakers. Chinese is like Korean and Japanese in its interpretation of governing category, allowing non-local binding (setting E).

Thomas used a written multiple-choice elicitation task which included two sentence types, those involving GCP and those involving PAP. All sentences contained finite embedded clauses only. An example of each type is included in (22).

(22) a. Ernie was sorry Cookie Monster hurt *himself* (GCP)
b. Susan gave Mary three photographs of *herself* taken last summer (PAP)

The learners were asked to choose the antecedent for the reflexive in each sentence from a set of given choices, allowing more than one antecedent. This arrangement allows for ambiguity in learners' answers on the (22b), PAP-targeted-type sentences, an important consideration as both Chinese and Japanese allow this type of ambiguity. Table 4.2 below summarizes results from Thomas' study.

*Table 4.2. Thomas (1989:291-292) responses to GCP and PAP tests on L2 English reflexives containing finite subordinate clauses. (NES = Native English speaking control group)*

| GCP | Local | Nonlocal | Either local or nonlocal |
|---|---|---|---|
| Spanish | 55% | 28% | 17% |
| Chinese | 59% | 21% | 20 |
| NES | 95% | 8.7% | 1% |
| PAP | Subject only | Object only | Either subject or object |
| Spanish | 32% | 39% | 29% |
| Chinese | 29% | 44% | 27% |
| NES | 47% | 28.3% | 25% |

Thomas' results indicate that for the GCP test, both the Spanish and the Chinese learners fail to accurately choose the local antecedent for the reflexive. Both groups allow non-local antecedents, a binding domain wider than is acceptable for English. Although the difference between the Chinese speakers and the control group may be due to L1 Chinese transfer, the explanation does not hold for the low Spanish percentages. This is surprising in that, if L1 transfer really does influence antecedent choice, we would expect the Spanish speakers to approximate English binding principles since Spanish, like English, requires local antecedents.

On the PAP test, learners generally favored subject antecedents over object antecedents, a preference also observed in the native speakers.

Considering the failure of Thomas' Spanish speakers to perform convincingly better than the Chinese speakers on the GCP test, it seems the L1 actually has little influence on the binding domains they choose. The same can be said of the PAP test, on which the two groups of ESL learners perform very similarly. The data in this study therefore do not convincingly demonstrate evidence of L1 transfer or of operation of the Subset Principle.

Hirakawa (1990) looked at the GCP and the PAP with 65 Japanese junior and senior high school students (ages 15 to 19) learning English as a foreign language in Japan. The students ranged over four EFL levels. Also included was a native English speaking control group. Hirakawa's administered a two-part test. the first part identified learners awareness of the vocabulary and structures included on the actual test. The actual test was multiple-choice judgement task which was designed to examine the test subjects' interpretation of English reflexive with respect to the GCP and the PAP. It included five stimulus sentence types which ranged from multiple clause to monoclausal sentences, and varied by finite or infinitival verb in the embedded clauses of the multiple clause sentences. These five types are similar to the types used in the tests created for this thesis and are summarized in (23). Types A, B, C, and D test the GCP and Type E tests the PAP.

(23) Hirakawa (1990: 70) test sentence types for investigating the GCP and the PAP.

Type A: Two-clause sentence (finite)

John said that Bill hit *himself.*
[NP1 [NP2 refl]].

Type B: Three-clause sentence (finite)

Mary remembers that June said that Alice blamed *herself*
[NP1 [NP2 [NP3 refl.]]]

Type C: Two-clause sentence (infinite)

John told Bill to hit *himself.*
[NP1 [NP2 refl]].

Type D: Three-clause sentence (infinite)

Ann knows that Mary told June not to blame *herself*
[NP1 [NP2 [NP3 refl.]]]

Type E: One-clause sentence

Bob talked to Paul about *himself.*
[NP1 [NP2 refl]].

Potential antecedents were listed under each sentence and informants were asked to choose the appropriate antecedent for the reflexive. Hirakawa's results from this test found no improvement with increasing "grade level" (= Hirakawa's interpretation of proficiency level). The combined results of the finite and infinite sentences are shown in Table 4.3.

*Table 4.3. Hirakawa (1990: 73-74) percent of acceptance of each sentence type by Japanese test subjects' (acceptance of local antecedents is correct, acceptance of nonlocal antecedents is incorrct)*

| | Local | Nonlocal |
|---|---|---|
| Tensed Clause | 72% | 27% |
| Infinitive Clause | 55% | 45% |

In general there appears to be a greater propensity for the test subjects to choose the local antecedent in tensed clauses than in infinitival clauses. Of particular interest are the results for nonlocal binding where only 27% of the test subjects accepted the nonlocal antecedent in sentences with finite verbs in the subordinate clauses but 45% accepted the nonlocal antecedent in sentences with non-finite verbs. As Hirakawa points out, it seems her subjects are settling on an intermediary setting between Japanese and English in support of findings by Finer and Broselow. However, rather than a UG-based account of how her learners arrive at this setting, Hirakawa suggests it is the result of L1 transfer and claims that the learners have adopted the widest value of the GCP, the value required by their L1. In further support of this for the PAP, Hirakawa's subjects' responses to the monoclausal Type E sentences strongly favored the subject antecedent (74%) over the local antecedent (20%), again explicable as the result of transfer from subject-oriented L1 Japanese.

The studies discussed above generally investigate L2 learners who haven't attained a high level of proficiency in the target language. Perhaps the intermediate GCP and PAP settings found by Finer and Broselow and by Hirakawa reflect the intermediate proficiency levels of their test subjects.

One study that looks at advanced L2 learners of English is Cook (1990). Cook collected data from test subjects from a variety of different L1 backgrounds to see how they would interpret English reflexive anaphors. He carried out an experiment which tested the interpretation of *himself* and *him* in sentences such as (24).

(24) a. John said Peter helps himself
b. John said Peter helps him

47 advanced learners from three different language backgrounds - Romance languages, Japanese, and Norwegian - participated in the experiment. 14 native English-speaking control subjects were also included. Cook included four sentence types based on GCP and PAP. These are summarized with examples in (25).

(25) Cook (1990): Test sentence types for investigating the GCP and the PAP.

Types A/B: Simple sentence with no embedded clauses

*John shot him/himself*

Types C/D: Biclausal with tensed embedded clause

*Peter said that John voted for him/himself*

Types E/F: Biclausal with infinitival embedded clause

*Peter asked John to include him/himself*

Types G/H: Noun phrase sentence ("picture NP")

*John reported Peter's criticism of him/himself*

The task was a computer-controlled comprehension test consisting of four items for each type giving a total of 40 items. For each item, subjects had to decide whether *him* or *himself* referred to *John* or *Peter* by pressing the appropriate key. The results of the sentences containing reflexives are summarized in Table 4.4.

*Table 4.4. Cook (1990: 582) percentages of L2 learners' comprehension errors of English reflexives by type and language (in percent).*

| Language | Type B | Type D | Type F | Type H |
|---|---|---|---|---|
| NES (n=14) | 1.8% | 8.9% | 7.1% | 16.0% |
| Ro (n=14) | 3.6% | 7.1% | 30.4% | 35.7% |
| Ja (n=16) | 1.6% | 23.4% | 40.6% | 43.7% |
| No (n=17) | 0.0% | 2.9% | 19.1% | 33.8% |

Key: NES = Native English Speaker; Ro = Romance language speaker;
Ja = Japanese speaker; No = Norwegian speaker

Table 4.4. shows that L2 learners may set the governing category parameter to a GCP Type C setting which has an intermediate value requiring a reflexive to be bound within a tensed clause. Accuracy rates drop for Type F (non-finite) sentences and Type H (noun phrase) sentences for all learners. Cook's results are somewhat ambiguous. At first it seems the Japanese informants' Type F sentences, with 40.6%, suffer from L1 interference, as a contrastive analysis would claim. However, as Thomas' (1989) found with her Spanish speaking L2 English learners, Cook's Romance speakers perform poorly on Type F sentences as well, despite his expectation that their L1 would help them learn English. This would seem to disprove the transfer theory and result in

Cook's data failing to provide a clear explanation of how L2 learners interpret English reflexives. Despite this, the results of Cook's Japanese test subjects on biclausal sentences (Cook's Types D and F sentences) support Hirakawa's results with her intermediate proficiency learners with respect to the finite vs. infinitive subordinate clause. Where Cook's data reflect a 77% accuracy rate on local binding in finite subordinate clauses, Hirakawa's data show 72%. And for local binding in infinitival subordinate clauses, Cook shows 59% and Hirakawa shows 55%. In both studies informants performed better on sentences with finite clauses.

Thomas (1991, and in more detail 1993) describes an interesting study which looked at whether adult L2 learners conform to binding principles, or choose antecedents based merely on linear proximity. In this event, learners would choose the antecedent which is linearly closer to the anaphor. Such a process could lead to UG-like test performance even if UG played no role in L2 acquisition. Evidence of a proximity-based antecedent selection process would provide supporting evidence in arguments of no access to UG. In contrast to this, evidence of conformity to binding principles would indicate access to principles of UG in adult L2 learners. Thomas' study is interesting in that she tested and compared test results for two target languages - English and Japanese. Similar to her 1989 study, her test subjects included native speakers of Japanese and native speakers of Spanish who were studying English and were tested on the acquisition of English reflexive pronouns. Additionally, native English speakers and native Chinese speakers who were studying Japanese were tested on the acquisition of the Japanese reflexive *zibun*. For this cross-sectional study, Thomas tested 132 adult learners of English and 41 adult learners of Japanese. Three levels of proficiency were identified in the English (low, mid, high). Control groups of both native English speakers and native Japanese speakers were also given the tests. Thomas administered elicited imitation sentence tasks and multiple-choice comprehension tests of pronoun and anaphor interpretation and reports on the comprehension tests and their results. I will briefly review the tests and results for L2 English. The three types of grammaticality judgement sentences given to the Japanese and Spanish speakers are presented in (26).

(26) Thomas (1991: 227) test sentence types for investigating the GCP, the PAP, and linear proximity selection.

Type 1: Biclausal: sentential complement

Mary heard that Sue told the doctor about herself

*Sue* = Local NP (subordinate clause subject)

*Mary* = Long-distance NP (matrix clause subject)

Type 2: Biclausal: relative clause

The man who John met wrote a story about himself

*The man* = Local NP (matrix clause subject)

*John* = Long-distance NP (relative clause subject)

Type 3: Single clause; subject vs. nonsubject NPs

Mary told Karen one more story about herself

*Mary* = Subject NP

*Karen* = Nonsubject NP

Sentence types 1 and 2 test for the governing category principle and type 3 tests for the proper antecedent parameter. The possibility that linear order is responsible for target-like behavior is also tested for in sentence types 1 and 2. Table 4.5. (on the following page) repeats Thomas' (1991) results.

Thomas' results indicate that most L2 learners at all three levels chose a correct GCP value; that is, they locally bound reflexives in finite sentences. With respect to the PAP, learners observed subject-orientation. As for linear ordering, Thomas argues on the basis of the results of type 2 sentences that L2 learners do not determine coreference by linear proximity. Her study is unique and useful in that she investigated the reverse phenomenon of how English L1 speakers learn languages with Japanese-type settings. The test data suggest that overall, L2 learners do not interpret English reflexives linearly, but assign c-commanding antecedents to reflexives. Thus, Thomas concludes that adult L2 learners do have access to principles and parameters of UG when interpreting the referential properties of L2 English anaphors.

*Table 4.5. Thomas (1991: 228) responses to English reflexives arranged by L1, proficiency level, and sentence type. Percentages reflect (within-group) consistent (≥66%) interpretation of reflexives.*

| L1 | Japanese | | | Spanish | | | English (control group) |
|---|---|---|---|---|---|---|---|
| Proficiency level | Low | Mid | High | Low | Mid | High | n/a |
| *n* = | 20 | 25 | 25 | 21 | 20 | 21 | 21 |
| Type 1: Biclausal; sentential complement | | | | | | | |
| Local NP | 80.0 | 76.0 | 84.0 | 90.5 | 70.0 | 81.0 | 100.0 |
| Long-distance NP | 5.0 | .0 | .0 | 4.8 | 5.0 | .0 | .0 |
| Local OR long-distance NP | 5.0 | 16.0 | 16.0 | 4.8 | 20.0 | 9.5 | .0 |
| Type 2: biclausal; relative clause | | | | | | | |
| Local NP | 60.0 | 80.0 | 96.0 | 85.7 | 70.0 | 95.2 | 100.0 |
| Long-distance NP | 10.0 | 8.0 | .0 | .0 | .0 | .0 | .0 |
| Local OR long-distance NP | 10.0 | 8.0 | 4.0 | 4.8 | 15.0 | 4.8 | .0 |
| Type 3: Single clause; subject vs. nonsubject | | | | | | | |
| Subject NP | 50.0 | 52.0 | 76.0 | 61.9 | 75.0 | 81.0 | 52.4 |
| Nonsubject NP | 15.0 | 4.0 | .0 | .0 | .0 | .0 | .0 |
| Subject OR nonsubject NP | 5.0 | 24.0 | 24.0 | 23.8 | 20.0 | 19.5 | 47.6 |

4.2.2. *Summary and critique of studies on GCP and PAP*

The studies discussed in the previous section all support a role for access to UG for their respective L2 learners. Finer and Broselow, Finer, and Hirakawa claim that ESL learners reset their GCP to a position midway between the target language and the L1 when one represents the least restricted GCP setting (Korean and Japanese) and the other represents most restricted setting (English). However, as Thomas (1991) points out, merely preferring local binding without actively disallowing long-distance binding of reflexives does not necessarily mean the learners have not instantiated the L1 setting in the L2 (p. 231).

Thomas' (1989) results from Chinese and Spanish learners of L2 English are somewhat mixed. Despite her informants' L1s having different GCP settings, no significant differences are observed which show that the Spanish learners bind locally more than the Chinese learners in their interpretation of English reflexives. Both groups prefer long-distance over local antecedents in non-subject position within finite clauses (p. 282). However, none of her informants show evidence of making hypotheses about binding outside the range of settings in Wexler and Manzini's model.

Cook's (1990) results indicate informants performed better on sentences with finite clauses than on sentences with nonfinite clauses. This leads him to speculate that GCP settings play a role in language processing which influence L2 learners' grammatical competence and performance.

Thomas' (1991) study indicates that L1 learners are able to reset the GCP for the L2, leading her to see this as evidence that principles and parameters of UG are directly accessible to L2 learners.

Evidence of adherence to the Subset Principle is problematic due to the multiple-setting nature of Manzini and Wexler's GCP and the PAP. For example, the studies done by Finer (1991), Finer and Broselow (1986), and Hirakawa (1990) indicate that L2 learners tend to choose an intermediate setting and bind English reflexives to matrix subjects in which the subordinate clause contains a finite verb. However, this seems to violate the Subset Principle which requires more restricted (less marked) permutations to be included in less restricted (more marked) settings. That is, an intermediate setting such as GCP "C" should also allow settings "B" and "A", "A" being the setting corresponding to English.

A research methodology-related problem common to the studies discussed above concerns L2 learners' judgements on tests used to collect data. Merely choosing one answer over another doesn't necessarily mean the test subject rejects the answer not chosen. He or she may merely *prefer* one answer over another. Native speakers of English tend to prefer subject antecedents over object antecedents in PAP sentences like (27)

(27) Alfred gave Victor a photograph of himself

where *Alfred* is the subject NP and *Victor* the object NP. By prompting NS informants, however, they should come to see that Victor, the object NP, is also acceptable, given the right context (Read and Chou Hare,1979).

Korean and Japanese speakers tend to prefer the non-local subject antecedent in GCP sentences like (28), though they can bind locally as well (Finer, 1991; Hirakawa 1990).

(28) Japanese:
Howard-wa$_i$ Hong-sul-ga$_j$ zibun-wo$_{i,j}$ nutta-to itta
Howard-TOP Hong-sul-NOM self-ACC painted-COMP said
Howard said that Hong-sul painted self

More rigorous test methods which force test subjects to recognize more than one potential antecedent would address this problem of learners preferring one antecedent while not necessarily rejecting the other. One possible solution used by Crain and McKee (1986) and White, Bruhn-Garavito, Kawasaki, Pater & Prévost (1997) is for test items to include a short description of some action or relationship between two or more characters followed by a statement which purports to summarize the description. Test subjects decide whether or not the statement is true or false. Truth value judgements like this are used to develop one section of the test questions for the experimental section of this thesis (see Chapter 5).

A third issue, suggested by Yusa (1998: 218, 224), is that these studies fail to thoroughly address L1 transfer. In addition to the standard reflexive *zibun* ('self'), Japanese also has special morphologically complex reflexives - *kare-zisin* (himself) and *kanajo-zisin* ('herself') - which can only be bound locally. Although rarely used in Japanese, these compound reflexives behave similarly to English reflexives in that they must be bound locally (Thomas, 1995: 212-213; Yusa, 1998: 218, 224). The difference in interpretations between the simple *zibun* and compound reflexives can be seen in (29).

(29) a. Howard$_i$ said that Hong-sul$_j$ painted himself$_{*i/j}$
b. Howard-wa$_i$ Hong-sul-ga$_j$ zibun-wo$_{i/j}$ nutta-to itta
Howard-TOP Hong-sul-NOM self-ACC painted-COMP said
c. Howard-wa$_i$ Hong-sul-ga$_j$ kare-zisin-wo$_{*i/j}$ nutta-to itta
Howard-TOP Hong-sul-NOM himself-ACC painted-COMP said

English *himself* in (29a) only refers back to *Hong-sul*. The morphologically simple Japanese reflexive *zibun* can be co-indexed locally with *Hong-sul*, or long-distance with *Howard*. However, the morphologically complex Japanese *kare-zisin* can only be co-indexed with the local antecedent *Hong-sul*. Other examples of such compound reflexives which behave this way can be found in Chinese and Korean (*ta-aji* and *cu-casin*, respectively).

The problem this point raises for the studies discussed here is that, if L1 transfer does in fact influence L2 performance, we can't be sure which L1 anaphoric paradigm the test subjects are transferring from - *zibun* or *kare-zisin*. If learners adopt an intermediate GCP setting as Finer and Broselow (1986) and Hirakawa (1990) argue, the learners should also accept lower (more restricted) settings, including the most restrictive (English) setting. This would assume the Japanese learners had transferred their morphologically simple L1 *zibun* paradigm which allows the minimally constrained setting. However, if the test subjects happened to transfer the restricted, English-like, *kare-zisin*, paradigm and *still* adopted a more unrestricted mid-way GCP setting, then it cannot be claimed that they had chosen an intermediate setting as a result of splitting the difference between Japanese and English. If Chinese speaking learners of English transfer their short distance setting for English reflexives, they would bind the reflexive

locally. By the same token, if they transfer a long-distance setting, they would be more likely to allow illegal nonlocal binding. Since Japanese-type languages have both short distance and long-distance reflexives, it is very hard to know which values they would transfer, or even if transfer is involved at all.

Finally, the issue of language proficiency is not adequately addressed in the studies. Finer and Broselow's (1986) study included both intermediate and advanced learners, but the effects of proficiency on results is not thoroughly discussed. Hirakawa's (1990) study includes four groups of learners ranked by grade in school, the underlying assumption apparently being that higher grades will reflect higher proficiency levels. Although she expected an effect for grade level in her results, she found none. This surprises Hirakawa, but she fails to discuss why her subjects' proficiency fails to increase from the lower to the higher grades. Additionally, and most curiously, Hirakawa actually creates a hierarchy of conformity to the GCP based on test subjects grade levels. She claims low-level learners fail to set the subset value which results in them choosing long-distance antecedents in English. Intermediate level learners choose an intermediate GCP value, and advanced learners set the correct values for English. She creates this hierarchy by assuming that Finer and Broselow's test subjects have a higher level of English because "they were exposed to English in the United States" (p. 79). Yet this is not clear from Finer and Broselow's study and Hirakawa never discusses what she means by high and low level, nor what kind of exposure she assumes Finer and Broselow's test subjects had in the States. A clearer picture of the test subjects' proficiency levels through a standardized test, or some integrated skill assessment would ensure that cross study comparisons of results are meaningful. Differences observed in the studies may reflect test subjects' overall English proficiency and not necessarily just their knowledge of the GCP.

### 4.3. An alternative approach to parameterized binding

The relationship between anaphors and their antecedents as described by Manzini and Wexler (1987) has served as the basis for a considerable number investigations into L2 binding acquisition. Most of these studies have found that L2 learners' behavior

regarding binding is sanctioned by Manzini and Wexler's model. This has led SLA researchers to claim that L2 acquisition of anaphoric binding is UG-constrained.

Despite the apparent adequacy with which the GCP and PAP account for binding relationships, a number of problems with this theory have been noted. One problem is that it is overly complex and puts a high processing load on learners. Rather than catching an overarching, elegant generalization about the distance over which an anaphor can be bound and what type of NP (i.e., subject or object) can serve as the proper antecedent, Manzini and Wexler's solution presents "a potential undergeneralization problem as a result" (Safir, 1987: 80). An example of this is the exceptions which need to be made for specific lexical items. Japanese *zibun* for example, conforms to the GCP and PAP as outlined above. That is, it can be bound locally or long-distance to a subject NP. But if the morphologically complex *kare-zisin* discussed above is used, it must be locally bound. This reliance on specific lexical items as well as on parameter settings is "in direct contradiction to the spirit of the principles and parameters model" (Hermon, 1992: 78).

Another criticism of Manzini and Wexler's model is made by Reuland and Koster (1991). They argue that the definition of a governing category and a proper antecedent have no principled restrictions and are not derived from linguistic principles. They claim that the arbitrary nature of constraints on what can be a local domain allows "virtually unlimited possibilities for anaphors to differ" (p. 2).

A third difficulty with Manzini and Wexler's proposals concerns their reliance on the operation of the Subset Principle to correctly set the GCP or PAP. The role of the Subset Principle in language acquisition has been criticized as lacking convincing evidence and not being theoretically viable. (see for example Saleemi (1992) and MacLaughlin (1995)).

In light of these criticisms of the GCP and the PAP, one alternative explanation of binding which has been proposed is the movement at Logical Form (LF) analysis.

### 4.3.1. *LF-movement analysis*

Logical Form is held to be an abstract syntactic level of representation where such relationships as quantifier-variable and antecedent-anaphor are represented

(Chomsky 1986b: 68). The LF-movement approach to binding explains the binding domain in terms of how far an anaphor can move up a syntactic tree. Yang (1983) and Pica (1987) observe that the degree of morphological complexity of a reflexive seems to determine binding domain. Under this analysis, simple reflexives (e.g. *zibun* in Japanese), analyzed as $[_{NP}[_{N'}self]]$ ($X^0$) by Pica, are thought to undergo successive raising cyclically from INFL to INFL at LF. This allows them to be raised to INFL which is c-commanded by a subject, resulting in the subject-orientation effect seen in Manzini and Wexler's PAP. Compound reflexives (e.g., *kare-zisin* in Japanese), analyzed as $[_{NP}[_{SPEC}him][_{N'}self]]$ ($X^{max}$), are blocked from adjoining to a higher VP and can only take local antecedents. The movement analysis therefore assumes binding at LF and accounts for the subject-orientation of long distance-bound reflexives. By making this assumption, it attempts to correlate the three binding issues seen above, namely long-distance binding, morphological simplicity, and subject-orientation, into one united theory.[3]

Three studies which have used this analysis to look at L2 binding are Christie (1992), Thomas (1995), and Christie and Lantolf (1998). Christie's cross-sectional study examines L2 learners' knowledge of long-distance binding and subject orientation. In the same way that Thomas (1989, 1991) includes test subjects with different L1s, Christie includes Spanish-speaking learners of English, Chinese-speaking learners of English, English-speaking learners of Spanish, and English speaking learners of Chinese. These languages include both $X^{max}$ reflexives (English *himself*, Spanish *si mismo* and Chinese *taziji*) and one $X^0$ reflexive (Chinese *ziji*). Of these three languages, only Chinese allows long-distance binding. In order to try to control for the problem of preferences in the test subjects' grammars, test subjects were asked to complete picture identification truth value tasks in which they heard a statement and decided whether what was depicted in a picture presented to them corresponded to what they heard. Presenting only one choice showed the learners a potentially acceptable interpretation

[3] Another alternative to the parameter model of binding is Progovac's (1992) "relativized SUBJECT" analysis. This shares the observation by Yang (1983) and Pica (1987) that reflexives can differ in morphological complexity (i.e., Japanese $X^0$ *zibun* vs $X^{max}$ *kare-zibun*, etc.). Progovac proposes that for long distance reflexives the $X^0$ head AGR is the only SUBJECT in determining the binding domain. $X^0$ reflexives are bound to an $X^0$ head with pronominal features such as AGR and can only be long-distance bound across clauses which lack AGR. $X^{max}$ reflexives are thought to bind with an $X^{max}$ specifier such as the subject of a clause, and are therefore locally bound. This account of binding is explored in an L2 acquisition context by Bennett and Progovac (1998) in their study of the acquisition of L2 English by L1 speakers of Serbo-Croatian, a language which includes $X^0$ reflexives.

which they might not have thought of on their own. Biclausal items on the judgement task such as (30) were used to test for acceptance of long-distance binding domain.

(30) *Lisa said that Jim had hit herself on the head.

Monoclausal sentences which contained subject and object NPs, such as (31) examined whether test subjects would bind anaphors to nonsubject NPs.

(31) George gave Barbara a book about herself.

By looking at the results of these two sentence types together, Christie hoped to show a correlation between long-distance binding and subject orientation which would support the raising of anaphors at LF. Although Christie's results provide little evidence of such a correlation, Thomas (1995) points out that anaphor raising at LF doesn't necessarily require locally bound anaphors to accept non-subject antecedents – something Christie looks for and doesn't find strong evidence of. According to Thomas, there is nothing in the assumption that $X^{max}$ anaphors must be bound locally which also states they should accept non-subject antecedents. Similarly, subject orientation does not necessarily entail long-distance binding.[4] By expecting that the movement at LF approach will entail a strong correlation between governing categories and proper antecedents, Christie concludes that her data only show a weak correlation between domain and orientation and fail to give strong support to L2 raising of anaphors and thus access to UG.

In a similar study, Christie and Lantolf (1998) developed a cluster analysis to interpret data results. They claim using this analysis provides a graphic visualization of how close or disparate individual responses are. As in Christie (1992), they test for a correlation between domain (Governing Category) and orientation (Proper Antecedent). This time though, the correlation between the two takes the form of a clustering effect of reflexive binding properties. Christie and Lantolf interpret their results of the clustering

[4] For example, both Japanese *kare-zisin* and English *himself* are $X^{max}$ anaphors. But whereas English *himself* allows both subject and non-subject antecedents, Japanese *kare-zisin* allows only subject antecedents. And although English and Japanese both allow subject orientation, English prohibits long-distance binding, while Japanese allows it, except with *kare-zisin*.

as again failing to convincingly support UG constraint of L2 learners' knowledge of reflexives. As with Christie (1992), they do not address the problem of the LF movement analysis allowing binding to local subject and object NP antecedents despite some languages like Chinese and Japanese disallowing object NPs. Their clustering analysis only compares the learner's grammars to an idealized native speaker norm. A more meaningful analysis might identify those test subjects who allow long-distance binding and then further checking this group's judgements on subject orientation. (Thomas, 1998). Any test subject who allows long-distance binding with subject orientation would be following UG sanctioned rules, even though a language like English doesn't permit long-distance binding. Long-distance binding to an object NP would be out. Likewise, local binding to a subject or object would be UG-constrained, even for languages like Chinese or Japanese which happen not to permit binding to object NPs. It would be ungrammatical in that particular language, but not disallowed according to UG.

Thomas' (1995) study is motivated by Christie's, and Christie and Lantolf's questionable ability to explain the relationship between long-distance binding and subject-only antecedents using the movement at LF analysis. In trying to accommodate these two seemingly disparate aspects, Thomas asks if L2 acquisition data can provide evidence that non-native speakers raise anaphors at LF, i.e., "do long-distance bound reflexives in learners' interlanguages require subject antecedents, and must reflexives with nonsubject antecedents be locally bound?" (p. 215). In addressing these questions, Thomas finds previous studies lacking for three reasons. First, she claims data reported as group means score fail to reveal individual patterns of response. Secondly, most binding studies investigate the acquisition of the same L2, English. The subject orientation preference of ESL learners, and its acceptability in English might obscure "the issue of whether nonsubject antecedents entail local-only binding in L2 grammars" (p 215). Thomas' third problem with previous studies is that most of them use comprehension tasks, which she claims may not reliably control for the problem of test subjects' preference for subject antecedents. Identification of acceptable nonsubject antecedents as well will reveal their true, underlying grammars.

Thomas (1995) reexamines the data from her (1991) study of the acquisition of L2 Japanese binding and finds that the test stimuli are inadequate by the third fault outlined above. Her test items are single clause sentences such as (34) and (35) below.

Thomas felt that failure to include biclausal sentences prevented the researcher from knowing if a test subject "who allows nonsubject antecedents for locally bound *zibun* would also admit nonsubject antecedents where the reflexive is bound long-distance" (p. 220). Additionally, Thomas notes that her (1991) stimuli are comprehension tasks which may fail to see through test subjects' preferences, as noted above.

To conduct a more reliable study, Thomas develops a new set of stimuli. As Thomas' study served as an early basis for the test design used in the present thesis, it will be discussed here in some detail.

Thomas' test instrument consists of four types of truth-value judgement sentences with Types A and B being biclausal, and Types C and D being monoclausal. Examples of these are given in (32) to (35).

(32) Type A
Biclausal sentence, *zibun* in subordinate clause. Expected response: accept
A wa B ga zibun no kuruma o tukau koto o sitte imasu
A TOP B NOM self GEN car ACC use COMP ACC know be
'A knows that B will use self's car'

(33) Type B
Biclausal sentence, *zibun* in subordinate clause. Expected response: reject
C wa B ni A ga zibun no hon ga suki da to iimasita
C TOP B DAT A NOM self GEN book ACC like COP COMP said
'C told B that A likes self's book'

(34) Type C
Single-clause sentence. Expected response: accept
A wa B ni zibun no mondai ni tuite hanasimasita
A TOP B DAT self GEN problem about spoke
'A spoke with B about self's problems'

(35) Type D

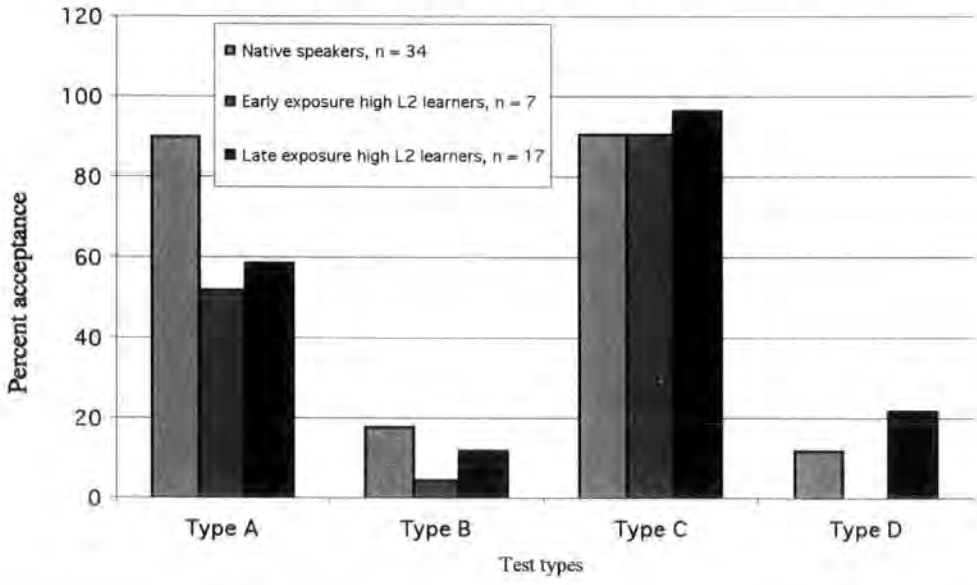
Single-clause sentence. Expected response: reject

Okaasan wa A o zibun no sensei ni syookai-simasita

'The mother introduced A to self's teacher'

Type A sentences test for acceptance of long-distance corereference; Type B for rejection of coreference with a long-distance non-subject antecedent. Type C sentences check whether test subjects allow coreference between *zibun* and a local subject, and Type D whether they reject coreference with a local nonsubject. Thomas' new study includes three types of informants: native Japanese speakers as a control group, and two levels (high and low) of L2 Japanese learners. Additionally, the high level learners are divided into two groups: those with early (pre 17 years of age) exposure to Japanese, and those with late exposure (post 17)[5] Of those groups, Thomas discusses the results of the adult high level learners. Within this group, she compares those learners with early exposure to those with late exposure. The NS results are also included. These results are presented in figure 4.2.

*Figure 4.2. Thomas (1995: 231) percent acceptance of test stimuli, by sentence type, for high-proficiency L2 learners with early vs. late exposure to languages with long-distance reflexives. (Early exposure for Type D = 0%)*




[5] Thomas bases 17 as the age cutoff on Johnson's (1988) study on critical period effects in L2 acquisition. Johnson's criteria require learners to be immersed in the L2 environment.

These results offer a mixed picture of L2 learners' ability to move *zibun* at LF, where movement is held to be in evidence in the absence of long-distance, nonsubject antecedents for *zibun*, represented in Type B sentences. Although both early and late exposure groups accept some long-distance nonsubject antecedents, the native Japanese speakers actually accept more. Thomas does not offer an explanation for this high level of acceptance, though she notes that it nullifies the difference on type D sentences between the early and late exposure learners and disallows any claim that the late exposure learners "are less able to capture UG-constrained aspects of the grammar of a second language" (p. 231).

Thomas' sophisticated study controls for variables earlier studies did not control for. Test subject target language level, age at first exposure, and the use of truth value judgements as a means of circumventing learner preference all mark her study as a step forward in binding research. These same variables will be controlled for in the experiment on L2 attrition presented in Chapters 5 and 6 below.

## 4.4. Conclusion

This chapter has reviewed several approaches to describing and explaining the properties of reflexive binding. Manzini and Wexler's governing category parameter model incorporates a range of settings from A to E, with setting A the most restrictive domain (associated with English) and setting E the least restrictive (associated with Chinese and Korean-like languages). Assuming L2 learners have access to UG, SLA research examines how learners react to evidence of a new governing category parameter setting and subsequently interpret the relationship between anaphors and their antecedents in the target language. The resetting of parameters is explained by recourse to the Subset Principle, which, when applied to L2 acquisition, is assumed to facilitate learning of settings less restricted than the setting a learner begins with, namely their L1 setting. Although the studies discussed in this chapter find evidence that adult L2 learners fail to achieve native-like levels of competence, their above-chance results suggest UG may be operational in their IL grammars.

In addition to Manzini and Wexler's model, Pica's movement at LF analysis draws attention to the morphological complexity of reflexives and accounts for subject orientation in long-distance binding.

The next two chapters will report on a study which investigates the degree to which advanced ESL speakers maintain their knowledge of English reflexive binding after leaving an English-speaking environment. The discussion of the data from this longitudinal attrition study will draw on the approaches and analyses outlined in this chapter.

# Chapter 5: Research Design and Method

## 5.0. Introduction

In Chapter 3 we saw the main themes that have been investigated in L2 attrition research. This literature review showed that much work in the field focuses on the sociolinguistic and lexical loss of second language, and that very little attention has been paid to investigating L2 attrition from within a formal theoretical linguistics framework. Chapter 3 concluded by outlining similarities and differences between anaphoric reflexives in English and Japanese, generally proposing that an examination of L2 attrition data from within Chomsky's Government and Binding theory would provide a more thorough understanding of an L2 speaker's attriting competence.

In Chapter 4 we looked more closely at the nature of binding in Japanese and English and saw how the two languages differ from each other. In particular, we saw that in English, reflexives are obligatorily bound to local antecedents, as in (1).

(1) $Mary_i$ thought that $Susan_j$ blamed $herself_{*i/j}$

Japanese, however, allows a reflexive to be bound locally as well as long distance, across a clause boundary or boundaries, to the matrix subject as in (2).

(2) Mary-ga $_i$ Susan-ga $_j$ zibun-$o_{i/j}$ semeta to omot-ta
Mary-NOM Susan-NOM self-ACC blamed COMP think-past
Mary thought that Susan blamed herself

We also saw that in mono-clausal sentences with two NPs and a reflexive, English allows either a subject or an object NP antecedent as in (3).

(3) $Mary_i$ showed $Susan_j$ a photograph of $herself_{i/j}$

Japanese, however, restricts the proper antecedent to a subject NP only, as in (4).

(4) Mary-ga$_i$ Susan-ni$_j$ zibun-no$_{i/*j}$ shasin-o mise-ta
Mary-nom Susan-dat self-gen picture-acc show-past
'Mary showed Susan self's picture'

The current chapter describes the design of a longitudinal investigation conducted to assess the degree to which adult L2 English speakers maintain control of these types of binding relationships between reflexives and their antecedents once exposure to English ceases. Section 5.1 outlines the goals and hypotheses of the investigation. Section 5.2 examines different types of data collection tools and techniques used in second language acquisition research. Section 5.3 describes the current investigation, including test types used, test subjects, and test collection procedures. The results are reported on in Chapter 6.

## 5.1. Explanatory Goals and Working Hypotheses

The primary goal of the current research project was to investigate how lack of exposure to English effected the test subjects' knowledge and use of English reflexive pronouns. Another related goal was to investigate the role of age at first exposure to the target language in determining ultimate level of maintenance or loss of the binding principle targeted.

The choice to conduct a longitudinal study rather than a cross-sectional study was made for two reasons. First, it was felt that as the attrition process itself occurs over time, a longitudinal study which closely examines individuals' performance data would best capture various aspects of L2 loss. Secondly, the length of time since exposure to the L2 ceases is a critical variable in the investigation, which requires the testing of subjects to begin as soon after their return to their (Japanese) L1 environment as possible. This requirement severely restricted the pool of potential test subjects, making a cross-sectional study with large groups of informants at different stages of attrition unfeasible.

By giving a syntactic explanation of an aspect of L2 attrition, I hope this thesis contributes to the understanding of L2 knowledge of language. In particular, three questions can be asked which relate to the issue of adult access to UG in adult second language speakers as discussed in Chapter 2. These questions are:

Question 1

Do adult L2 English speakers exhibit evidence of loss of knowledge of reflexive binding?

Question 2

If evidence of such loss is found, are there any patterns to this loss?

Question 3

Does age at first exposure to the L2 play a role in attrition patterns?

Question 1 asks if test subjects lose knolwedge of their ability to apply binding principles. Specifically, control of aspects of Principle A of the binding theory is targeted in this study. However, before loss of these forms can be investigated, it must be established that the test subjects included in the study exhibit native or near native competence of English reflexive binding at the beginning of the investigation. This was one criterion for a test subject to be included in the study and will be described in more detail below. Let us assume that the test subjects' knowledge of reflexive binding was established to some level through access to UG during their L2 acquisition periods. If evidence of loss of reflexive binding can be found, a hypothesis can be made that a UG based aspect of grammar can suffer attrition in the same way that other, non-UG-based aspects of language such as the lexicon have been shown in previous research to undergo attrition. This is significant in that previous L2 attrition literature has never, to my knowledge, addressed the effects of language disuse on UG established properties.

Question 2 addresses the nature of the process of the loss of adult L2 binding. If at any point the data indicate the development of random or "impossible" (Thomas, 1995: 42) grammars not sanctioned by UG, then we can assume that the UG constraints which

were initially established at the height of the speaker's knowledge of the L2 have been lost.

If, however, it can be shown that there is a UG-sanctioned pattern to the loss of binding which at all times falls within the bounds of UG constraints, then we can argue that these constraints continue to operate. And if they continue to operate on an attriting grammar, it might further be argued that UG plays a role in controlling the emerging attrited form of the language. This emerging form might be a closer and closer approximation of the speaker's L1, or it might be a third state, outside both the L1 and the L2 norm. The emergence of such a third state, such as a "Russian-like" pattern which only allows long distance binding outside of embedded clauses with nonfinite verbs, would provide evidence of some degree of adult L2 learner access to UG. The finite/nonfinite distinction plays no role in English or Japanese and other than access to UG, there would be no way for learners to know this.

Question 3 addresses the role of age in second language acquisition. Three of the test subjects for this study had childhood exposure to English and three did not have exposure until after puberty. Differences in the loss patterns exhibited by the test subjects by age at first exposure, and the degrees to which attrition of L2 binding occurs among the test subjects according to age at first exposure may offer insight into how age at first exposure effects the robustness of UG instantiated knowledge.

From these research questions, three working hypotheses, A, B, and C, can be stated.

Hypothesis A: Principles of reflexive binding already instantiated in the test subjects' English will undergo change due to lack of exposure.

Hypothesis B: Changes observed in the test subjects' control over reflexive binding due to lack of exposure will be UG constrained.

Hypothesis C: A correlation will be exhibited between age at first L2 exposure and degree of retention of L2 reflexive binding principles.

## 5.2. Test type selection

This section will review some of the significant data collection tools used in second language acquisition research. This review will help support the validity of the test types used in the present study.

Language acquisition research, whether L1 or L2, investigates the mental processes by which language is acquired and the knowledge a language speaker has of a particular language. There is, however, no way to directly know how an individual's knowledge of language is instantiated in the mind. The underlying rules which make up this knowledge or competence can only be studied indirectly by using a variety of methods to observe, collect, and analyze performance data.

Many researchers conducting second language acquisition research use controlled experiments (e.g., Finer and Broselow, 1986; Thomas, 1989, 1991; White, 1985, White et al., 1997). Other studies have used spontaneous spoken data collected during free or guided conversations. These data are recorded and then transcribed. Spontaneous data collection requires the researcher to comb through the data in search of target forms. In studies which investigate broad categories of lexical items or grammatical properties in areas such as morphology, spontaneous data collection can be an appropriate method to use, although it has limitations. Such a procedure has been used for attrition studies, for example Kuhberg, in his 1992 investigation of the attrition of L2 German mentioned in Chapter 2, collected spontaneous data monthly over a period of 20 months. This study looked at the loss of a broad spectrum of lexical items such as nouns, prepositions, pronouns and articles, as well as morphological and grammatical loss. Spontaneous data such as that collected by Kuhberg is not, however, appropriate for studies which target properties such as the type of antecedents a test subject's grammar allows. Thomas (1993: 62-63) points out several reasons why spontaneous data may not be appropriate for studies of binding. One problem is that, despite the relative frequency with which reflexive forms can come up in general conversation, finding an adequate number of targeted forms in such spontaneous data can make it difficult to create an adequate corpus. Knowing which antecedent or free R-expression an anaphor might refer to is rendered less certain with spontaneous data; experiments allow for this to be controlled for. Failing to produce a particular structure during a spontaneous data collection does not necessarily rule out that structure from the test subject's linguistic ability. And in the

same way, the inclusion of errors does not prove these "bad forms" are part of the test subject's grammar.

Experiments, if carefully developed and implemented, allow researchers to efficiently target and elicit specific linguistic forms. Two types of experiments are commonly used in linguistic research to elicit data: comprehension tasks and grammaticality judgments.

### 5.2.1. *Comprehension tasks*

Experimenters eliciting anaphoric data from test subjects have used a variety of comprehension tasks according to purpose and test subject type. Act-out tasks are one such tool, often used with children who are too young to reliably take pencil and paper tests. Direct questioning tests are another type of tool, and have several types. Direct questioning tests can be multiple choice tests such as traditional paper and pencil tests for adults, or picture or puppet identification tasks with children. Another method involving the direct questioning method is to use truth value judgments (Thomas, 1993: 64-65). When working with adults, more 'pencil and paper' forms of tests are typical, although pictures can play a role in these.

#### 5.2.1.1. *Act out tasks*

Act-out tasks require children to manipulate dolls and other realia, or to play games. Otsu (1981: 177), testing L1 acquisition of binding, used a set of toy zoo animals and had children manipulate them in response to statements such as (5).

(5) *The lions patted each other.*

Wexler and Chien (1985: 145) developed a multi-participant act-out game of "Simon says" in which children had to point at themselves or each other in response to

commands. This task was further developed by having children give out toys to each other (Chien and Wexler, 1987).

5.2.1.2. *Direct questioning tests*

Direct questioning tests can be multiple-choice tests, or truth-value judgment tests. There are two types of multiple choice tests: identification tests, usually used with children, or a traditional paper and pencil test, used with older test subjects. The identification task asks test subjects to choose which of several pictures or dolls (puppets) best represents a statement they hear from the investigator. One example of an identification task used to collect data on child L1 acquisition of binding comes from Read and Chou Hare (1979). In their study, children hear statements such as "*Big Bird told Oscar a story about himself*". The children then have to answer the question *Who was the story about?* by pointing to a puppet of either Big Bird or Oscar.

Picture identification tasks ask test subjects to choose from two or more pictures, each having a different interpretation of an anaphor-antecedent relationship stimulus. The stimulus can be provided in written form or stated orally by the investigator (e.g. Bennett, 1994; Eckman, 1994). Finer and Broselow's (1986) investigation into the acquisition of the Governing Category Principle (GCP) and Proper Antecedent Principle (PAP) for English in native Korean speakers and Finer's (1991) follow-up investigation, which added native Japanese and Hindi speakers, both used picture identification tasks. These tasks asked test subjects anaphoric questions about two men, Mr. Fat and Mr. Thin, who paint each other or themselves. Test subjects are then asked to choose the corresponding picture.

The second type of multiple choice test, the traditional paper and pencil test, is usually used with older test subjects, such as used by Hirakawa's (1990) study of L2 acquisition of the GCP and the PAP. Hirakawa conducted a cross-sectional study using native Japanese speaking ESL students at four different levels. Each subject read a set of twenty-five anaphoric sentences and was asked to indicate who the reflexive referred to in each sentence. Five choices were given for each sentence, as in (6). Choice (6d) provides space for the test subject to write in who *himself* might refer to.

(6) John said that Bill hit *himself.*

a. John

b. Bill

c. either John or Bill

d. someone else _____

e. don't know

The ability of this type of test to reflect a learner's linguistic competence seems questionable. The overtness of the forms being studied in a test like this may easily become apparent to the test subjects, who could then rely on deductive, non-intuitive strategies to make their choices. A more covert method of having test subjects produce direct responses would seem to increase the likelihood of test subjects relying on their intuitions when choosing an answer. This could be done by including non-target-like distractor sentences.

White, Bruhn-Garavito, Kawasaki, Pater & Prevost (1997) also question the validity of multiple choice tasks, arguing that in situations where more than one antecedent is acceptable, as in (7)

(7) Jack$_i$ showed George$_j$ a photograph of himself$_{i/j}$

test subjects may choose one antecedent over another, though this may be a result of the multiple choice test type and does not prove the alternate acceptable antecedent isn't part of a test subject's grammar.

White et al. (p. 148) point out that these types of multiple choice tests may elicit test takers' preference rather than reveal their competence (for examples see Hirakawa 1990; Lakshmanan and Teranishi 1994; Thomas 1989, 1991; Wakabayashi, 1996). Even training test subjects to take the tests has not proven successful in eliminating the preference-over-competence problem (Bennett, 1994; Thomas, 1991). As an example, a sentence like (7) above typically prompts English native speakers to choose *Jack* as the proper antecedent, and only when it is pointed out to them, do the test subjects recognize that the object *George* is also acceptable.

This leads us to the second type of direct questioning test, the truth value judgment test. Generally for this task, a test subject is asked to listen to or read a passage followed by a statement about the passage. The test subject then must decide if the statement accurately represents the information contained in the initial passage. In trying to overcome the antecedent preference limitation described above, Crain and McKee (1986) used a truth-value judgment task to show that L1 English children do permit backwards anaphora as in (8).

(8) When *she* was outside playing, *Strawberry Shortcake* ate an ice cream cone.

Previous research had led investigators (Tavakolian, 1978; Solan, 1983) to assume children only recognized the extrasentential interpretations of anaphors. Crain and McKee however, argue that although children do prefer a sentence-external reading, they will accept the backwards anaphora version when provided with an appropriate context. This context draws attention to the acceptability of a particular co-reference, such as that of *she* and *Strawberry Shortcake* in (8), which might otherwise be overlooked. Their study used a toy/character/puppet manipulation protocol as a truth-value judgment test. Following a puppet skit performed by one of the experimenters, the puppet characters made comments about the skit, as in (8), and the child judged the accuracy of these comments. By using a truth-value judgment test in this context, Crain and McKee found that children did in fact judge the backwards anaphoric reading of such sentences as (8) grammatical, accepting the co-reference of *she* and *Strawberry Shortcake*. Crain and McKee argue that earlier findings of non-co-reference were most likely the result of the children's preference for the sentence-external reading combined with a failure to establish an appropriate context for the co-reference to be recognized.

It would seem likely, then, that tests of reflexive binding which use multiple choice antecedents may not provide as accurate a picture of competence — L2 or adult native speaker — as would a truth-value test. Truth value judgment tests draw testees' attention to acceptable patterns of co-reference which might not be considered in multiple choice tests. Lakshmanan and Teranishi (1994) and Thomas (1993) make similar observations.

Crain and McKee (1987) again used a truth value judgment test, using an even more interactive format with the children. The investigators acted out a sketch using props and then asked a puppet to describe the action. If the children felt what the puppet said was accurate, they rewarded it by feeding it a cookie. If they felt the puppet was wrong, they fed it a rag.

Thomas (1993) recognizes the usefulness of this type of test when working with children as it "heightens the child's investment in the task experiment and thereby increases its reliability" (64). In other words, it keeps the children from becoming bored. She also sees the value of the truth value judgment test in adult data collection in her 1995 investigation of the acquisition of the Japanese reflexive pronoun *zibun* by adult native English speakers. In this experiment Thomas used 16 illustrated stories, each consisting of a group of pictures with captions under each picture. After the story was an anaphoric statement about the story and test subjects had to decide if the statement accurately reflected the content of each story. Thomas used the captioned pictures in a effort to make all possible antecedents obvious to the test subjects. A test subject who consistently chose subject antecedents for the Japanese reflexive *zibun* may not necessarily reject non-subject antecedents. The inclusion of the captions with the pictures was intended to encourage test subjects to consider all possible antecedents – both subject and object. One potential problem with using a picture-caption combination is that test subjects may base their choice of antecedent on a one-to-one lexical association with specific chunks of text in the captions. Though such an association might result in the correct judgment (accept or reject), it would not be based on the learner's knowledge of grammar. Another problem with the captioned illustrations is that test subjects might completely disregard the captions and base their judgment on the progression of the four images. A misinterpretation of the images might result in an accurate or inaccurate answer, but again, it would not be based on a learner's knowledge of grammar.

White (1995) and White et al. (1997) also use a truth value judgment test with Japanese native-speaking adults in an investigation of binding in L2 English. In this, they developed a battery of short two to three sentence stories, with a following statement which included an anaphoric expression. As with Crain and McKee, and Thomas, the test subjects had to indicate if the statement was an accurate summary of the main point of the story. White et al. propose that providing the context of a story from

which to judge the accuracy of the anaphoric sentence can avoid test subject preference for one antecedent over another in ambiguous monoclausal sentences where binding can either be to the subject or object antecedent. Consider (9), (10), and (11).

(9) Annie is a student. There was a new teacher in her class today. During class, the teacher asked Annie some questions about Annie's hometown. Annie told the teacher that she was born in Montreal.

*Annie gave the teacher some information about herself.*
(Subject antecedent, true)

(10) Susan is very good at her job. The supervisor thought Susan deserved a promotion but she wanted to know as much as possible about her. The supervisor asked Susan about her habits, her friends and her life in general. She questioned Susan for a long time.

*The supervisor questioned Susan about herself.* (Object antecedent, true)

(11) A young boy was looking at one of Mr. Robins' antique guns. The young boy accidentally pulled the trigger and the gun fired. Unfortunately, the bullet hit Mr. Robins in the arm.

*Mr. Robins realized that the boy shot himself accidentally.*
(Matrix subject (long distance) antecedent, false)
(White et al., 1997: 167)

These examples of truth value judgment (TV) test items demonstrate how the test subject is led to see the connection between the anaphor and the subject antecedent *Annie* in (9), the object antecedent *Susan* in (10), and the illegal long distance subject matrix subject antecedent *Mr. Robins* in (11). By lowering the possibility of test subject preference obscuring test subject knowledge of binding, more reliable test results may be obtained.

For the reasons listed here, the present study makes use of truth-value judgment tests as one means of collecting data to assess speakers' maintenance of L2 English reflexive binding over time. Another data collection method - a grammaticality judgment test - was also used to confirm findings from the truth value judgment test. Let us now look briefly at grammaticality judgment tasks.

### 5.2.2. *Grammaticality Judgments*

Grammaticality judgment tasks elicit a speaker's judgments about the well-formedness of utterances. Early examples of grammaticality judgments are found in Chomsky (1957), where he uses his own intuitions to make his seminal claims about the nature of language. Despite the reliance on grammaticality judgments, Sorace (1988: 168-169) notes that Chomsky (1965) points out that "although intuitions were not an "objective operational measure", the search for more reliable procedures was a matter of minor importance at that stage of research (Chomsky, 1965: 20)." Despite limitations, grammaticality judgments, with their associated acceptable or starred ("out") sentences, are an important tool which both L1 and L2 researchers use to understand the boundaries of linguistic competence and the rules which determine it.

Prior to the use of grammaticality judgments, early second language research during the 1960s and 1970s reflected the contrastive analysis approach to L2 teaching and acquisition. Errors in learner language were used as evidence that differences between the L1 and the L2 result in problems learning the target language. Beyond these structuralist-based analyses of language, no research was done to establish a model of what L2 knowledge of language might be until the late 1970s and 1980s. Bialystok (1979) was one of the first L2 researchers to use grammaticality judgments to elicit learners' implicit knowledge of the L2. Gass (1983) also used grammaticality judgments to distinguish between learners' ability to make implicit judgments about the well-formedness of a given sentence, and the explicit knowledge required to correct or comment on a deviant sentence. With the increasing use of GJs in second language research over the last several decades, Mandell (1999) recognizes the debate over whether grammaticality judgments can meaningfully reflect linguistic competence of a learner, and tries to test this. Mandell points out that many grammaticality judgment

tests focus on aspects of the language which are most likely part of L2 instruction. He argues that a test of grammaticality judgment reliability would investigate learner judgments of a UG constrained aspect of language, in particular a construction not overtly taught in the classroom. Additionally, the results of the grammaticality judgment test should be compared to another measurement of the same phenomenon. In his study, he uses native English speakers studying L2 Spanish in a cross-sectional test of knowledge of Spanish verb movement. He compares the results of the grammaticality judgments with a "dehydrated sentence test" which provides test subjects with a string of words which either need a word added or deleted to make a grammatical sentence. A statistical comparison of the results from the two tests indicates a high degree of correlation between them, and "that a definite relationship exists between a standard grammaticality judgment test and a dehydrated sentence test" (p. 93). Mandell's study concludes that grammaticality judgments have a better than chance degree of reliability and can serve as a measure of L2 syntactic competence when carefully designed and implemented.

In their study of knowledge of L2 binding in Japanese and French speakers described above, White et al. (1997) use both a truth value judgment test which they refer to as "story and picture tasks", and a grammaticality judgment test to collect their data. In their study, however, the grammaticality judgment test serves only to confirm the learners' sensitivity to grammaticality in monoclausal sentences controlled for gender as in (12).

(12) The man showed his daughter a photograph of herself as a baby.

(White et al., 1997: 155)

L2 learners who accept object NPs as antecedents in monoclausal sentences such as this will judge sentences like (12) to be correct. Other instantiations of reflexive binding which include embedded clauses with finite and non-finite verbs are not addressed by White et al. They do, however, note this and observe that "it would be desirable to test L2 learners with a grammaticality judgment test that covers the same sentence types used in the story and picture tasks" (p. 164). The inclusion of the various permutations of reflexive binding under consideration in the present investigation

address White et al.'s desire to test a wide range of reflexive binding using a grammaticality judgment test. To this end then, a grammaticality judgment test covering the same structures as the truth value test was given to the six test subjects in my study. This grammaticality judgment (GJ) test is described below (section 5.3.2.3.).

## 5.3. The present study

The present study is limited to an investigation of L2 reflexives for reasons discussed in Chapter 1. One reason is that there exist models of investigation into reflexive binding provided by the corpus of L1 investigations into reflexive binding. Another, more compelling, reason to choose reflexive binding is that the UG-based nature of the syntactic constraints which determine the interpretation of reflexives provides structures for investigation which are highly unlikely to be overtly taught to language learners. My investigation was conceived as a method of exploring the stability with which proficient young adult ESL speakers maintain their knowledge of the constraints on English reflexive binding in the absence of L2 exposure upon return to their L1 environment.

At the beginning of this study, when the test subjects had just arrived back in Japan, spontaneous data and test data were collected from informants. Depending on how long a test subject continued with the study, data continued to be collected over the subsequent 9 to 16 months. The spontaneous data collection sessions consisted of recorded free and guided conversations. Pictures and stories with multiple characters were used to guide conversations in an attempt to tease reflexive type forms from test subjects. The test data collection instruments consisted of truth-value judgment tests similar to those of Chien and Wexler (1990), and White et al. (1997). These tests were given during the same session as the spontaneous data was collected. It became clear after the first session that eliciting reflexive forms during spontaneous data collection was both time consuming and not productive. Few target forms were observed, and the decision was made to abandon spontaneous data collection.

Following the failure of the spontaneous data collection, a decision was made to add a grammaticality judgment test as a means of confirming the results from the truth value and grammaticality judgment tests, and in so doing, establish a correlation between

the results of the two test types. The remainder of the test subjects' data collection was conducted using truth-value judgment tests and grammaticality judgments, each targeting the test subjects' competence in L2 English reflexive binding. Section 5.2 explained the motivation for developing the two test types. The rest of section 5.3 will describe the informant selection process, explain the linguistic forms tested, and describe the testing procedure.

### 5.3.1. *The Informants*

Twelve informants participated in this study. Six Japanese university students at Obirin University in Machida City, Tokyo were the actual test subjects. Four native English-speaking adults in Tokyo and two in the United States served as the native speaker control group. The number of test subjects was arrived at as a result of the informant selection process described below in section 5.3.1.2. The number of native speakers was chosen to match the number of test subjects.

#### 5.3.1.1. *The native English speaker control group*

The six native English speakers who participated in this study were selected at random from among colleagues and professional acquaintances. The native speakers were all speakers of North American English as typically spoken by the university-educated middle class. Two of the native speakers were living in the United States at the time of the data collection, and spoke no foreign language. The other four native English speakers were educators who were living in Japan and spoke a limited amount of Japanese. However, as a result of regular contact with them, and by their own admission, none of the native speakers in Japan could be considered much more than incipient bilinguals (Diebold, 1964), and transfer from Japanese to their English (i.e., L1 attrition) was not considered a threat to their performance on judgment tests. Their reluctance to mix with the local Japanese community and resultant lack of exposure to Japanese prevented any possible attrition of their English in the Japanese environment.

A comparison of the results of the US based native speakers and the Japan based native speakers supports this claim.

5.3.1.2. *The Japanese L2 English test subjects*

The Japanese ESL speakers were all Japanese nationals and were students from Obirin University, where I was an English language instructor from 1991 to 1998. All test subjects were native Japanese speakers. These students were considered appropriate test subjects as they were in a position to commit to long-term volunteering for the study, and their proximity to the university enabled data collection sessions to be readily arranged. Most importantly, however, the university has an overseas study program which provided me with students returning to Japan after completing their overseas period of study. As in the United States, undergraduate degrees in Japan continue for four years. All test subjects selected for this study were either finishing their third year of university, or were beginning their fourth year, after six years in high school.

At the start of the data collection stage in 1997, potential test subjects were considered for inclusion in the study based on three initial criteria:

- a score on the Test of English as a Foreign Language (TOEFL) of at least 500
- a short interview
- the amount of time which had elapsed between their return to Japan and the beginning of the selection process

All test subjects initially selected as candidates for the study had high TOEFL scores, ranging from mid 500s to mid 600s. The maximum score possible on a TOEFL is 677 points. American universities typically require foreign, non-native English speaking students to score at least 500 points on the TOEFL to be considered for undergraduate enrollment. Scores in the mid 500s indicate accurate language use with a strong vocabulary. Scores in the high 500s and 600s reflect high levels of fluency and accuracy, in some cases approaching near-native levels of proficiency.

The interview confirmed that the candidates had a fluency in English which matched their ability assessed by their TOEFL scores. Having already taught EFL in Japan for over ten years, I felt my own opinion of the potential test subjects' English was accurate.

The elapsed time factor between return and data collection was important to ensure that any attrition which might occur soon after their return to Japan was observed. Thus, a cutoff of two weeks following their return to Japan was established. Within those two weeks a potential test subject had to be contacted, interviewed and given the first of the TV tests and have some spontaneous data collected from them. With test subjects who were recruited after spontaneous data collection had been abandoned, they had to begin the TV and GJ tests within two weeks. This two week cutoff point was met by all six test subjects except one (test subject R3).

The final selection of test subjects was determined by three additional criteria:

- evidence of accurate knowledge of binding in English
- little or no anticipated exposure to English subsequent to returning to Japan
- willingness to participate in the study for approximately one year

Knowledge of binding was determined by giving potential test subjects a test of the relevant aspects of binding under consideration, namely, local vs. long distance binding of anaphors to their antecedents in biclausal sentences, and treatment of subject vs. object NPs in monoclausal sentences. For subjects included in the study, this test became the first truth value judgment test (TV1) and the first grammaticality judgment test (GJ1) as described in section 5.2.1.2. above.

Ensuring restricted exposure to English once back in Japan was essential in order to reduce the chances that the subjects' English would be maintained outside of the L2 environment. All test subjects had studied at American universities during their third year (out of four) and had returned to Japan to complete their last year of university. Fourth year students were chosen as test subjects as they are typically involved with writing their graduation theses and looking for jobs. They therefore have little opportunity to come in contact with English after they return to Japan. This ensured that

no test subject had the advantage of maintaining their L2 through frequent contact with English, although one test subject initially included in the study later had to be excluded as she was found to have a monolingual English-speaking boyfriend.

Willingness to participate in the study was of course critical. As it turned out, because of language level and tentative commitment to the study, out of eleven students who initially met all the criteria,, only the six individuals whose data are presented here actually continued with the data collection sessions.

Three test subjects had lived overseas in an English-speaking environment as children. In all cases, they had lived in the United States. These three are known as "returnees", a term used in Japan to refer to children who have accompanied their parents on an overseas company transfer and then returned to Japan after several years. The time my returnee test subjects spent abroad as children varied from slightly less than 3 years to more than 8 years. All three of these test subjects returned to Japan at very similar ages, shortly before beginning junior high school in Japan (age 13). All three subsequently returned to the States approximately 7 years later for their junior ($3^{rd}$) year abroad and studied at universities there, enrolling in normal university courses. In this study these three test subjects are referred to as the 'R' (for Returnee) test subjects: R1, R2, and R3.

The other three test subjects did not have any childhood experience living overseas. Their pre-university exposure to English was limited to six years of Japanese junior and senior high school grammar translation English classes, and in one case included classes at an English conversation school. Despite the lack of childhood exposure to the L2, these three test subjects obtained a level of English sufficient for them to also attend universities in the United States. Two of these students attended pre-sessional English language courses in the USA prior to taking their university courses. In this study these three test subjects are referred to as the 'S' (for Student) test subjects: S1, S2, and S3.

The inclusion of three returnees and three non-returnees in this study allows us to investigate the effects of age at first exposure to the L2 on attrition. Including the three returnees in the group has provided data which shows mere childhood exposure in the L2 environment doesn't guarantee a robust knowledge of reflexives, nor does lack of such exposure predict a less robust knowledge. Even test subject R1 demonstrates that some

loss can occur within a relatively short period of time in an L2 speaker with a high level of English.

Table 5.1 outlines important periods in the acquisition of L2 English for each of the test subjects who participated in this investigation.

*Table 5.1. Ages of all six test subjects at significant periods of L2 acquisition, including type of exposure and TOEFL details.*

| | *Ages* | *Type of L2 Exposure* | *TOEFL Score* | *Age at TOEFL* |
|---|---|---|---|---|
| R1 | 0;0-8;10<br>13;1-19;11<br>20;4-21;4 | US local kindergarten and primary school<br>Compulsory EFL instruction in Japan, including 2 years at university<br>US university | 630 | 18 |
| R2 | 6;3-10;2<br>13;1-20;2<br>20;8-21;7 | US local primary school<br>Compulsory EFL instruction in Japan including 2 years at university<br>US university | 580 | 19 |
| R3 | 8;3- 11;1<br>13;1-20;1<br>20;8-21;8 | US local primary school<br>Compulsory EFL instruction in Japan including 2 years at university<br>US university | 533 | 20 |
| S1 | 13;1-20;6<br>20;11-21;8 | Compulsory EFL instruction in Japan including 2 years at university<br>US Pre-sessional language school and university | 610 | 20 |
| S2 | 13;1-19;5<br>19;6-20;7 | Compulsory EFL instruction in Japan including 1.5 years at university<br>US university | 593 | 19 |
| S3 | 13;1-20;8<br>21;1-22;3 | Compulsory EFL instruction in Japan including 2 years at university<br>US Pre-sessional language school and university | 520 | 19 |

### 5.3.2. *Materials*

#### 5.3.2.1. *The placement test*

As mentioned above, in addition to the TOEFL score and interview, potential informants were also given a placement test of reflexive binding on which informants were asked to consider and judge the accuracy of a variety of permutations of possible relationships between reflexive pronouns and their potential antecedents in English. These permutations are realized in the form of truth value judgments and grammaticality judgment tasks and include the following three basic sentence types:

(13) Type 1: biclausal sentences with a finite verb in the embedded clause:
*John knew Ben asked about himself*

(14) Type 2: biclausal sentences with a non-finite verb in the embedded clause:
*George wanted Arnold to buy himself a watch*

(15) Type 3: single clause sentences with subject and non-subject NPs (Pragmatically favors object antecedent.)
*The students asked the teachers about themselves*

The placement test included 30 TV items and 11 GJ items. For informants subsequently included in the study, the results of these tests were used as data. The tests, TV1 and GJ1, are included in full as Appendices C1 and E1.

Sentence types are described in detail below. It is by reference to these sentence types that the data analysis and discussion will proceed.

5.3.2.2. *The truth value judgment tests*

As discussed above, various L2 acquisition studies have made use of truth value judgment tests. These include White et al. (1997), Bennett and Progovac (1998), Christie and Lantolf (1998), and Thomas (1989, 1993, 1995).

The present study includes the basic truth value sentence types examples shown in (13), (14), and (15) above. Each token of the three basic sentence types is prefaced by a short scenario. Each scenario provides a pragmatic context which favors the selection of one candidate antecedent over the other in the stimulus sentences. The pragmatically determined antecedent can be either the local or long distance antecedent in the biclausal target sentences, and either the object or subject antecedent in the monoclausal target sentences. The biclausal sentences are further diversified by specifying either a finite or a nonfinite verb in the complement. This distinction was noted by Finer and Broselow (1986) who found that Korean learners of L2 English follow English binding rules in sentences involving tensed complements, but allowed long distance antecedents in sentences involving infinitival complements (159).

The combination of the two basic sentence types - biclausal with a finite verb in the embedded clause, and biclausal with a nonfinite verb in the embedded clause, offers

four types of stimulus sentence if both subject and object NPs are considered potential antecedents. These will be referred to as Types 1a, 1b, 2a, and 2b. Monoclausal sentences which distinguish subject and object NPs offer two additional stimulus sentences which will be referred to as Types 3a, and 3b. The TV tests thus have six types of stimulus sentence which are discussed below, with examples, in sections 5.3.2.2.1 through 5.3.2.2.3.

Informants were asked to judge if the stimulus, or target, sentence was true or false. "True" here thus refers to the accuracy - truthfulness - with which the target sentence represents the content of its scenario, rather than to the grammatical accuracy of the sentence itself. All target sentences in the truth value test items are, on their own, grammatically correct.

During the course of the 16-month data collection period, tests were created as needed. In the end, 13 TV test versions were written as the most test sessions an informant (S2) sat for was 13. Each TV test version included three tokens of each of the six TV types for a total of 18 items per test. Beginning with the fifth session, one earlier token of each sentence Type was included, starting again with master test item 1.

These repeat items were presented in later sessions along with two previously unseen items to maintain the total of three truth value test items of each type per test session. As some earlier tokens were repeated in later tests then, 30 tokens of each of the six TV types were created, giving a total of 180 distinct truth value test items. Appendix B cross references the TV test items by sentence type with the particular test session they appeared in. Appendix C lists all TV test items by type.

The six types of truth value (TV) judgment sentences will now be discussed in detail.

5.3.2.2.1. *Truth value Type 1 sentences.*

Truth value Type 1a and 1b sentences are biclausal structures which have a finite verb in the embedded clause. In both types, the reflexive in the embedded complement clause has two candidate antecedents. In Type 1a sentences, the associated story makes the local (c-commanded) antecedent the pragmatic (and in English the grammatical) choice of the two potential antecedents, and for Type 1b sentences, the story makes the

long distance antecedent the pragmatically correct (but for English ungrammatical) choice. Let us look at an example of each type.

For all Type 1a sentences, the correct answer is "True". This is illustrated in (16).

(16) Type 1a
The CIA officer captured a foreign spy who knew many secrets. The spy drank some poison and died.

*The officer knew the spy killed himself.* **True**

In (16) the two candidate antecedents of *himself* are *the spy* and *the officer*. Depending on which antecedent a test subject chooses, an aspect of her knowledge of how she defines the governing category in English will be revealed. If a subject maintains her knowledge of UG constraints on binding with respect to English, she should judge the sentence true and that it accurately reflects the content of the story, namely that *the spy* killed himself. In this case we can know she holds the governing category to be restricted to the lower IP. If, however, the test subject does not maintain her knowledge of UG-sanctioned constraints for English binding, then she may rely on her wider L1 (Japanese) setting which allows both local and long distance subject antecedents to be grammatical antecedents. Because *the spy* is an object NP, the test subject would reject it and require *the officer*, the (acceptable in Japanese) long distance subject NP to be the proper antecedent. But because *the officer* doesn't fit the story, she would judge the stimulus sentence *The officer knew the spy killed himself* to be false.

However, whether the test subject chooses true or false, the situation is more complex than this. Because of the nature of the subset principle, a choice of "True" does not necessarily rule out the test subject's excluding the (Japanese-like) marked long distance antecedent, but merely that she happened to choose the local antecedent because the content of the story pointed to it as correct. Only if we assume that a choice of "False" on TV Type 1a sentences is based on the test subject recognizing the long distance antecedent as grammatical as it would be in Japanese, do we have positive evidence of a UG determined constraint violation.

Stimulus sentences in the TV1a type sentence corpus which included bare infinitives in the embedded clause rather than finite forms of the verb, served as distractor sentences. These are discussed below in section 5.3.2.4. Informants correctly identified these distractors as being bound to the local antecedent at a level similar to sentences containing finite forms. See Appendix J.

Truth value Type 1b sentences are structurally the same as Type 1a sentences. They are biclausal structures with a finite verb in the embedded clause, and the reflexive in the embedded complement clause has two candidate antecedents. In Type 1b sentences, however, the associated story makes the long distance (matrix subject) antecedent the pragmatic (and in English the ungrammatical) choice of the two potential antecedents. The correct answer for all truth value Type 1b sentences is "False". This sentence type is illustrated in (17).

(17) Type 1b
One day, John was delivering letters when a big dog ran at him. John tried to stop the dog, but it was too late. The dog grabbed John's leg with his teeth.

*John knew the dog bit himself on the leg.* **False**

In (17) the two candidate antecedents of *himself* are *the dog* and *John.* Again, depending on which antecedent a test subject chooses, an aspect of her knowledge of how she defines the governing category in English will be revealed. If a subject maintains her knowledge of UG constraints on binding for English, she should believe the sentence is false and that it inaccurately reflects the content of the story. *Himself* is restricted to the embedded clause and cannot refer back across the clause boundary to *John*. If, however, the test subject does not maintain her knowledge of constraints for English binding, then she may believe the sentence to be true and to accurately reflect the content of the story. In this case then, the test subject tends to revert to her L1 setting, which for Japanese would allow *himself* in (17) to refer back long distance across the clause boundary to *John*. A choice of False in this instance would indicate a preference for the matrix subject as the correct antecedent, but which doesn't match the content of the story, resulting in her choice of False. The other possible outcome of failing to maintain binding constraints for English would be for a non-UG sanctioned grammar to emerge.

In the case of Truth value Type 1a sentences, this would mean only allowing long distance binding.

As with TV Type 1a sentences, sentences with bare infinitive verbs in the embedded clause were included in the TV Type 1b corpus and served as distractor sentences. Informants correctly identified these distractors as being impossible because of the local binding constraint at a level similar to sentences containing finite forms. See section 5.3.2.4 below and Appendix J.

5.3.2.2.2. *Truth value Type 2 sentences.*

Truth value Type 2a and 2b sentences are structurally similar to TV 1a and 1b sentences, differing in that TV Type 2 sentences have a nonfinite verb in the embedded clause. As with TV Type 1a, the stories for TV type 2a test items require the local antecedent as the pragmatic antecedent in the stimulus sentences, as in (18).

(18) Type 2a
Arnold always asked George what the time was. George got tired of this and told Arnold to buy a watch.
*George wanted Arnold to buy himself a watch.* **True**

Maintenance of reflexive binding should result in all TV Type 2a sentences being judged True, as the reflexive cannot be bound to the matrix subject in English. This is pragmatically established by the two sentence story in which *George* specifically tells *Arnold* to *buy a watch*. If the test subject allows the matrix subject to serve as the antecedent of the reflexive *himself*, and strongly prefers this option as is allowed in her L1, she should incorrectly judge sentences like (18) "False".

TV Type 2b sentences are similar to TV Type 2a sentences. They are biclausal sentences, with a nonfinite verb in the embedded clause, and the reflexive in the embedded complement clause has two candidate antecedents for the reflexive. The difference between TV Type 2a and 2b is the same as that between TV1a and 1b sentences — the associated story causes the matrix subject antecedent to be

pragmatically favored, thus rendering the correct judgment of the stimulus sentence false, as in (19).

(19) Type 2b
Ed was a customer at a fancy restaurant. He waited a long time for the waiter to take his order. Finally he got up and asked the waiter for service.
*Ed asked the waiter to serve himself.* False

In (19), candidate antecedents of *himself* are *the waiter* or *Ed*. English only allows *the waiter* to serve as the antecedent, but pragmatically the antecedent should be *Ed*. By marking sentences such as (19) False, a test subject demonstrates knowledge of English binding principles. Test subjects marking such sentences true indicate that they mistakenly allow Japanese-like long distance bound antecedents in their English grammar.

5.3.2.2.3. *Truth value Type 3 sentences.*

TV Type 3 sentences are monoclausal structures which have a subject NP and an object NP, either of which may serve as the antecedent of the reflexive in English. Type 3 sentences investigate test subjects' Proper Antecedent Parameter settings. Test subjects may use the minimal distance principle in which case they always bind the reflexive to the closest antecedent, namely the object antecedent. Or they may allow either the object NP *the teachers*, as in (20), or the subject NP, as in (21).

(20) Type 3a
On the first day of school the teachers wanted the students to relax, so they had an open meeting. The students could ask the teachers any question they wanted.

*The students asked the teachers about themselves.* True

(21) Type 3b
Mo wanted a job at the drugstore, so he told the owner about his background and skills. He told the drugstore owner many things.

*Mo told the drugstore owner about himself.* True

Acceptance of either the subject or object antecedent in TV Type 3 sentences reflects the operation of the English setting for the proper antecedent parameter. Were the test subject to undergo attrition or reset her proper antecedent parameter to the Japanese setting, she may come to allow only the subject NP to serve as the antecedent, in which case, (20) would be judged false.

The Truth value judgment test stimulus sentence types are summarized in Table 5.2.

*Table 5.2. Truth value judgment stimulus sentence features*

| *Sentence Type* | *Clause Type* | *Embedded Clause Verb Type* | *Anaphoric Orientation* | *Expected Correct Answer* |
|---|---|---|---|---|
| TV1a | Biclausal | Finite | Local | *True* |
| TV1b | Biclausal | Finite | Long distance | *False* |
| TV2a | Biclausal | Non-finite | Local | *True* |
| TV2b | Biclausal | Non-finite | Long distance | *False* |
| TV3a | Monoclausal | na | Object NP | *Pragmatically True* |
| TV3b | Monoclausal | na | Subject NP | *Pragmatically True* |

5.3.2.3. *The grammaticality judgment test*

In addition to the truth-value judgment test, data were also collected by conducting a grammaticality judgment test which asked informants to judge the grammaticality of sentences which reflect the three basic structure types in (13), (14), and (15) above. Test subjects read stimulus sentences and decided whether they were accurate or not, indicating their judgment by circling "OK" or "Wrong" following the sentence. By testing the same sentence types tested as in the truth value judgment tests, the grammaticality judgment test results act as a confirmation of the truth value judgment test results. Both biclausal and monoclausal sentence types were used in the GJ tests. Biclausal sentences of both Type 1 (finite verb in the embedded clause) and Type 2 (nonfinite verb in the embedded clause) were further developed by having the reflexive in half the test items agree with the local antecedent according to number and

gender, and half disagree. Reflexives which disagree with the local antecedent are made to agree with the long distance antecedent. In this way it could be determined if a test subject's knowledge of binding principles changed such that she came to accept long distance binding, a violation of English binding principles. Brief descriptions of the biclausal test types, with examples, are giving below in (22) - (29). Appendix D cross references the GJ test items with the particular test they appeared in by sentence type. Appendix E lists all GJ test items by type.

5.3.2.3.1. *Grammaticality judgment Type 1 sentences*

As with the TV Type 1 sentences, the grammaticality judgment Type 1 sentences are biclausal structures with a finite verb in the embedded clause. The reflexive in the embedded clause has two candidate antecedents.

In GJ Type 1a sentences, the reflexive agrees in number with the grammatical antecedent, as in (22).

(22) GJ Type 1a: Biclausal sentences, finite verbs in embedded clauses. Number agrees.
Example: *His parents knew Ben had asked about himself.*
Expected judgment: *OK* (i.e., grammatical)

An answer of *Wrong* on GJ Type 1a sentences would indicate the test subject is trying to bind back to the long distance antecedent, but an anaphoric relationship is blocked by the wrong number (plural in (22) above).

GJ Type 1b sentences are ungrammatical due to the antecedent and the local antecedent disagreeing in number, as in (23). An answer of correct would indicate the test subject is allowing ling distance binding back to the matrix subject which *does* agree in number, but is illegal in English.

(23) GJ Type 1b: Biclausal sentences, finite verbs in embedded clauses. Number disagrees.

Example: *The coach knew the players trusted himself.*
Expected judgment: *Wrong* (i.e., ungrammatical)

GJ Type 1c sentences are grammatical, the gender of the reflexive and the local antecedent agreeing, as in (24). Again, an answer of *Wrong* would suggest the test subject is attempting to bind long distance back to the matrix subject.

(24) GJ Type 1c: Biclausal sentences, finite verbs in embedded clauses. Gender agrees.

Example: *Adam thought his daughter had served herself.*
Expected judgment: *OK* (i.e., grammatical)

GJ Type 1d sentences are ungrammatical, the gender of the reflexive and the local antecedent disagreeing, as in (25). Again, an answer of *OK* would suggest the test subject is allowing the antecedent to bind long distance back to the matrix subject.

(25) GJ Type 1d: Biclausal sentences, finite verbs in embedded clauses. Gender disagrees.

Example: *John's mother knew Ben asked about herself.*
Expected judgment: *Wrong* (i.e., ungrammatical)

5.3.2.3.2. *Grammaticality judgment Type 2 sentences*

The GJ Type 2 sentences are biclausal and follow the same number and gender agreement patterns as the GJ Type 1 sentences. However, the Type 2 sentences all have nonfinite verbs in the embedded clauses, just as the TV Type sentences do. Brief definitions of the GJ Type 2 sentences along with an example appear in (26) to (29) below.

(26) GJ Type 2a: Biclausal sentences, non-finite verbs in embedded clauses. Number agrees.

Example: John's parents told him to wash himself.
Expected judgment: *OK* (i.e., grammatical)

(27) GJ Type 2b: Biclausal sentences, non-finite verbs in embedded clauses. Number disagrees.

*Example: The doctors expected the patient to see themselves.*
Expected judgment: *Wrong* (i.e., ungrammatical)

(28) GJ Type 2c: Biclausal sentences, non-finite verbs in embedded clauses. Gender agrees.

*Example: The policeman warned the woman to behave herself.*
Expected judgment: *OK* (i.e., grammatical)

(29) GJ Type 2d: Biclausal sentences, non-finite verbs in embedded clauses. Gender disagrees.

*Example: Mary's mother asked Fred to help herself fix her car.*
Expected judgment: *Wrong* (i.e., ungrammatical)

5.3.2.3.3. *Grammaticality judgment Type 3 sentences*

Monoclausal grammaticality judgment test items were also created to favor one antecedent over another. In this case, Type 3a grammaticality judgment sentences favored the object antecedent, while Type 3b sentences favored the subject antecedent. Test subjects who obey English binding Principle A should accept both types as correct. Rejection of the Type 3a, object antecedent sentences by a test subject would be evidence that she was reverting to her L1 – Japanese – binding principles which are subject oriented in monoclausal sentences. Rejection of Type 3b, subject antecedent sentences would reflect a "rogue" grammar not sanctioned by the PAP.

A third type of monoclausal sentence Type, 3c, was included in the grammaticality test. In GJ Type 3c sentences the antecedent of the reflexive was of an ambiguous nature. Type 3c sentences were retained throughout the testing period to monitor test subject performance. Were a test subject to begin consistently rejecting

Type 3c sentences, this would be an indication something was seriously wrong with their knowledge of L2 English as either NP can pragmatically and linguistically serve as antecedent. This situation never developed with any of the test subjects. Examples of the monoclausal grammaticality judgment test items by type are given in (30), (31), and (32) below.

(30) GJ Type 3a: Monoclausal sentences with 2 NPs. Pragmatically favors object antecedent.
Example: *Bill and June gave Sam a picture of himself.*
Expected judgment : *OK* (i.e., grammatical)

(31) GJ Type 3b: Monoclausal sentences with 2 NPs. Linguistically favors subject antecedent.
Example: *Sally faxed her teachers a report about herself.*
Expected judgment : *OK* (i.e., grammatical)

(32) GJ Type 3c: Monoclausal sentences with 2 NPs. Pragmatically ambiguous antecedent.
Example: *The bad boy told his father a lie about himself.*
Expected judgment : *OK* (i.e., grammatical)

The Grammaticality Judgment test stimulus sentence types are summarized below in Table 5.3.

*Table 5.3. Grammaticality judgment stimulus sentence features*

| | *Sentence Type* | *Embedded Clause Verb Type* | *Anaphoric Orientation & Agreement type* | *Expected Correct Answer* |
|---|---|---|---|---|
| GJ1a | Biclausal | Finite | Local; number agreement | *OK* |
| GJ1b | Biclausal | Finite | Long distance; number agreement | *Wrong* |
| GJ1c | Biclausal | Finite | Local; gender agreement | *OK* |
| GJ1d | Biclausal | Finite | Long distance; gender agreement | *Wrong* |
| GJ2a | Biclausal | Non-finite | Local; number agreement | *OK* |
| GJ2b | Biclausal | Non-finite | Long distance; number agreement | *Wrong* |
| GJ2c | Biclausal | Non-finite | Local; gender agreement | *OK* |
| GJ2d | Biclausal | Non-finite | Long distance; gender agreement | *Wrong* |
| GJ3a | Monoclausal | na | Object NP | *OK* |
| GJ3b | Monoclausal | na | Subjects NP | *OK* |
| GJ3c | Monoclausal | na | Ambiguous | *OK* |

5.3.2.4. *Bare infinitive distractor sentences*

In addition to the finite and non-finite embedded verbs used in the two tests described above, a third verb type – bare infinitives – was included in the embedded clauses of several biclausal sentences on both the TV and the GJ tests. Bare infinitives lack the *to* infinitive marker normally associated with infinitival forms, and occur as complements of verbs of perception such as *see*, *watch*, and *hear* as in (33).

(33) *The monkey saw the hunter kill himself.*

It has been observed that infinitival complements such as *the hunter kill himself* in (33) indicate a perfective aspect – a bounded or completed event (Comrie 1976).

Stimulus sentences such as (33) were included as distractors in the corpus of finite sentences (TV and GJ type 1 sentences) because, lacking the *to* infinitive marker, it was thought that informants might treat these sentence types as having finite

complements. As the informants were unlikely to have been taught to recognize bare infinitives, differences in acceptance of these bare infinitive sentence types from complements with finite clauses would constitute evidence of an unconscious, underlying distinction between the two complement types in their grammars. In particular, acceptance of the long distance antecedents in sentences containing bare infinitives would support findings by Thomas (1993) that non-native speakers tend to bind reflexive anaphors out of complements with non-finite verbs more readily than out of complements with finite verbs.

Bare infinitive distractor test items are included in the TV Type 1 section of Appendix C, and in the GJ Type 1 section of Appendix E and are each marked with an asterisk.

### 5.3.3. *Test administration*

It was assumed that following a period of overseas study in an English speaking country, each test subject's highest observable level of English would occur immediately upon return to Japan. It was also assumed that attrition would begin to set in upon cessation of exposure to the L2, in other words, from their return to Japan onwards. As stated above, testing of suitable test subjects therefore began as soon after their return to Japan as possible. Let us consider the specifics of the administration of the tests.

#### 5.3.3.1. *Frequency of collection sessions*

In Japan, the academic year begins in April and ends in February. This conflict with the American system, which begins in September and ends in May, means that Japanese foreign exchange students studying in the United States typically return to Japan in April or September to resume their university education. Although not a problem for the present study, this time frame did need to be considered to determine when data collection periods would begin. Once data collection had begun, collection sessions continued for various lengths of time, the longest being over 16 months, the shortest 9 months. Following their graduation from university, most test subjects

continued meeting for data collection sessions, though as they became busier with their new lives, meetings tapered off. The number of collection sessions thus varies from test subject to test subject. Despite the differences in the number of collection sessions per test subject, all but one attended sessions for over a year. Overall, the test sessions provided enough data from all test subjects to allow the study to proceed.

Collection sessions occurred approximately once a month for the first three or so sessions, subsequently reducing to every two or three months. This pattern reflects my interest in seeing if any attrition occurred soon after the test subjects' return to Japan, and also the time both the subjects and I had to arrange meetings. The most data collections were carried out with test subject S2 (13 sessions), and the least from subject S3 (7 sessions). Test subject S2 also provided data the most regularly, attending collection sessions once a month on average. Subject S3 was perhaps the most irregular, meeting only once every two or three months throughout the collection period. Data collection sessions for the Japanese test subjects are summarized in Table 5.4.

*Table 5.4. Test subjects' data collection sessions*

| *Collection sessions* | *R1* | *R2* | *R3* | *S1* | *S2* | *S3* |
|---|---|---|---|---|---|---|
| *Return Date* | 7/1/97 | 6/20/97 | 8/28/97 | 3/19/97 | 9/15/96 | 8/29/97 |
| *Session 1* | 7/11/97 | 7/8/97 | 9/29/97 | 4/4/97 | 9/27/96 | 9/29/97 |
| *Session 2* | 7/25/97 | 8/1/97 | 10/20/97 | 4/18/97 | 10/11/96 | 10/20/97 |
| *Session 3* | 9/26/97 | 9/23/97 | 11/17/97 | 5/16/97 | 11/1/96 | 12/15/97 |
| *Session 4* | 10/17/97 | 10/14/97 | 12/15/97 | 6/6/97 | 11/22/96 | 2/27/98 |
| *Session 5* | 11/21/97 | 11/18/97 | 1/20/98 | 6/27/97 | 12/13/96 | 4/13/98 |
| *Session 6* | 12/15/97 | 12/16/97 | 2/23/98 | 7/15/97 | 1/17/97 | 6/30/98 |
| *Session 7* | 1/23/98 | 1/20/98 | 4/13/98 | 9/26/97 | 2/21/97 | 9/21/98 |
| *Session 8* | 3/15/98 | 2/24/98 | 6/1/98 | 10/17/97 | 4/4/97 | NA |
| *Session 9* | 5/29/98 | 4/14/98 | 7/20/98 | 12/13/97 | 5/16/97 | NA |
| *Session 10* | 7/10/98 | 6/16/98 | 9/22/98 | NA | 6/27/97 | NA |
| *Session 11* | 9/25/98 | NA | 10/26/98 | NA | 8/1/97 | NA |
| *Session 12* | 10/23/98 | NA | NA | NA | 9/26/97 | NA |
| *Session 13* | NA | NA | NA | NA | 10/17/97 | NA |
| ***Total No. of Sessions*** | *12* | *10* | *11* | *9* | *13* | *7* |
| ***Length of Data Collection Period (Months from return)*** | *16.3* | *12.25* | *14.1* | *9* | *13.2* | *12.2* |

5.3.3.2 *Test scoring*

In both the TV and GJ tests, informants were asked to make a binary choice for each test item. A sliding scale reflecting degree of acceptability was avoided because of the vague and ambiguous results such a scale can produce. In the case of the truth value judgment test items, informants were expected to choose "True" if they felt the summary statement accurately reflected the content of the scenario and the corresponding antecedent was locally bound, as in (34) below. Informants were expected to choose "False" if they felt the summary was somehow reflected in the statement, but that the corresponding antecedent was long distance bound as in (35) below. Informants circled their choice on the test paper and correct answers were awarded 1 point each. (See figure 5.2. above for expected correct answers on TV test items.)

(34) TV Type 1a
Whales are very good swimmers, but during the summer, several whales swam onto the beach and died. It was a mystery why they did it.

*No one could understand why the whales killed themselves.*

**True/False**

(35) TV Type 1b
One day, John was delivering letters when a big dog ran at him. John tried to stop the dog, but it was too late. The dog grabbed John's leg with his teeth.

*John knew the dog bit himself on the leg*

**True/False**

In the case of the grammaticality judgment test items, informants were expected to choose "OK" if they felt the reflexive indicated the correct antecedent in the stimulus sentence. As noted above, the grammaticality of the anaphoric relationship in these sentences was established by number agreement as in (36) below, and by gender agreement as in (37).

(36) GJ Type 1a
*John's parents told him to wash himself.*

**OK / Wrong**

(37) GJ Type 2a
*Jim's uncle helped the woman set herself up in the new country.*

***OK / Wrong***

Informants circled their choice on the test paper and correct answers were also awarded 1 point each. (See figure 5.3. above for expected correct answers on TV test items.).

This concludes the description of the tests and data collection methods used in this study. In the next chapter we will examine the results produced by these tests and methods.

# Chapter 6: The Loss of Anaphoric Reflexive Binding in L2 English

## 6.0. Introduction

Chapter 5 reviewed the differences between anaphoric binding in English and Japanese and established the research questions for the study. The test materials used in the present study were reviewed and the rationale behind using them explained. We saw examples of truth value judgement questions and grammaticality judgement questions. The present chapter reports on the results of these tests. Section 6.1 reports on the results of the native English speaker control group. Section 6.2 analyses the test results for the six Japanese test subjects and Section 6.3 discusses those results. Throughout this chapter and Chapter 7, "test" will refer to the entire battery of "test sessions" given for the grammaticality judgement test and the truth value judgement test. Each test had a maximum of 13 test sessions.

## 6.1. Native English Speaker Control Group Results

In order to validate the TV and GJ tests as reliable tools with which to investigate the various permutations of reflexive binding under consideration in this study, this section will begin by examining the six native English speakers' test results. These results will be described in terms of degree of adherence to the principles underlying the test items described in Chapter 5. This descriptive analysis will also be used when examining the Japanese test subjects' results. The native speaker controls were given the entire battery of test sessions at once and responded to all 234 TV test items (217 target items and 17 distractors), and 143 GJ test items (136 target items and 7 distractors) over the course of one to two weeks. These same test sessions were administered to the Japanese test subjects over the course of 9 to 16 months. Section 6.1.1 will report the results of the truth value (TV) test, followed by an examination of the grammaticality judgement (GJ) test results in Section 6.1.2. Native speaker TV test scores are included in Appendix F, and their GJ scores in Appendix G.

### 6.1.1. *Native English speaker control TV test results*

Each TV subtest has 18 items and each GJ subtest has 11 items, all worth one point each. The test items on TV subtests 1 through 4 are unique. Starting from the fifth TV subtest, one earlier test item for each test type is included as a way of observing test subjects' responses to the same items later in the data collection period. All test items on all of the GJ subtests are unique; no items are repeated. The raw totals of the tests are listed by informant in Table 6.1. The sum of all scores, the percentage correct, the mean score, and the standard deviation are listed at the bottom of the table.

*Table 6.1. Overall native speaker TV test results*

| NS Informant | NS1 | NS2 | NS3 | NS4 | NS5 | NS6 |
|---|---|---|---|---|---|---|
| Score (out of a possible 234) | 200 | 207 | 217 | 206 | 222 | 208 |
| Percent | 84% | 88% | 93% | 88% | 95% | 89% |
| Mean | 15.15 | 15.92 | 16.69 | 15.84 | 17.07 | 16 |
| SD | 1.21 | 1.18 | 0.75 | 0.98 | 0.75 | 0.70 |

Compared with the maximum total possible score of 18 on each subtest (for a total of 234 for all tests combined), the differences between the NS mean scores and 18 are insignificant at the $p < .05$ level (one-tailed $t$ test: $t = -6.81$, $df = 10$). Furthermore, with a range of just 25, and a standard deviation of 8.1, the native speaker control group informants performed very similarly to each other. The NS scores are therefore considered reliable as a measure of fitness for the test subject results.

Table 6.2 outlines the NS's performance on the TV tests by sentence Type.

*Table 6.2. Degree of native speaker judgement accuracy of truth value test items by sentence Type*

| NS Informant/ TV sentence Type | NS1 | NS2 | NS3 | NS4 | NS5 | NS6 | Total |
|---|---|---|---|---|---|---|---|
| TV Type 1a (object) | 100% | 95% | 100% | 95% | 100% | 95% | 97% |
| TV Type 1b (subject) | 100% | 100% | 100% | 97% | 100% | 100% | 99% |
| TV Type 2a (object) | 95% | 95% | 97% | 95% | 97% | 92% | 95% |
| TV Type 2b (subject) | 97% | 100% | 100% | 95% | 100% | 97% | 98% |
| TV Type 3a (object) | 57% | 74% | 80% | 70% | 84% | 74% | 73% |
| TV Type 3b (subject) | 57% | 66% | 82% | 77% | 89% | 74% | 75% |

By TV sentence type, the NS informants' overall poorest performance was on the Type 3a sentences, monoclausal sentences with two NPs with the context favoring reflexive binding to the object NP antecedent. These sentences test the Proper Antecedent Parameter (PAP) as described in Chapter 4. Of a total of 234 Type 3a sentences among the six NS informants, they incorrectly judge 63 to be false. That is, the NS group rejects the object NP antecedent candidate such as *the man* in (1) below as the correct antecedent of the reflexive on 27% of the TV Type 3a test items.

(1) Jorge went to a party and met an interesting man. Jorge asked him many questions.

*Jorge asked the man about himself.*

Only slightly better than the Type 3a sentences are the NS control group's Type 3b sentences, monoclausal sentences with two NPs with the context favoring reflexive binding to the subject NP antecedent, as in (2).

(2) Jim wanted to know what his father was like as a child, so he asked him. He found out that his father had been a bad boy and made life very hard for his parents.

*Jim's father told Jim about himself.*

Of a total of 234 TV Type 3b sentences, 60 sentences, or 25% were judged false by the NSs. This slightly lower rejection rate of subject NPs in monoclausal sentences corresponds to other studies on L1 English binding such as Read and Chou Hare (1979), mentioned in Chapter 4. Their study of children's knowledge of reflexives in English reports that 81% of their native English speaking adult control group accepts the subject NP in Type 3a sentences as the correct antecedent.

That NS1 exhibits the lowest acceptance rate (57%) on both TV Types 3a and 3b is of interest in that of the six NSs, only his educational background was not in arts and humanities (BS in architecture). This has led me to speculate that because of his lack of extensive reading in the humanities, he might not be so used to considering different interpretations of a text. Such an openness to alternatives could be responsible for recognizing different antecedents as grammatical.

The similar rejection rates for subject NP and object NP antecedents in monoclausal sentences by adult native speakers of English point to the indeterminacy they experience in judging these sentences. On the one hand, NSs perform above chance on both sentence types (object antecedent = 73% acceptance, subject antecedent = 75% acceptance). On the other hand, the fact that the NSs' highest rejection rate of grammatical structures for all TV sentence types is with monoclausal sentences suggests that an analysis of NS judgements of Type 3 sentences compared to the Japanese test subjects may prove significant. Should the Japanese test subjects also reject grammatical monoclausal sentences the most, as in fact turns out to be the case, we have evidence that the test subjects' knowledge of reflexive binding parallels that of the native speakers. This can be seen in the item analysis presented in Appendices J and M.

The NSs' results on the other four TV sentence Types – 1a, 1b, 2a, and 2b – are very much in line with results predicted by the governing category principle. Overall accuracy on Type 1a sentences, biclausal sentences with a finite verb in the embedded clause, pragmatically favoring the local C-commanded antecedent, is 98%. For Type 1b sentences, which are the same structure as Type 1a, but pragmatically favor the long distance matrix subject for the antecedent, the accuracy is also high with a rejection rate of nearly 100%. Type 2 sentences also had high accuracy rates among NSs, and are only slightly lower than for the Type 1 sentences. For Type 2a, biclausal sentences with a non-finite verb in the embedded clause and pragmatically favoring the local antecedent, the overall accuracy is high with an acceptance rate among NSs of 95%. The rejection

rate of Type 2b sentences – biclausal, non-finite verb in the embedded clause, pragmatically favoring the long distance antecedent – is highly accurate at 98% rejection. The high accuracy rates with which NSs judged local vs. long distance antecedents not only confirms their adherence to binding principle A as realized for English, but also supports the validity of the stimulus sentences as accurate representations of the forms under investigation.

### 6.1.2. *Native English speaker control GJ test results*

The results of the Native English speaker control group on the grammaticality judgement test will now be considered in terms of degree of adherence to binding principle A. The structures tested on the GJ test correspond to the structures tested on the TV test. GJ Types 1a and 1c correspond to TV Type 1a (finite verb, local antecedent orientation). GJ Types 1b and 1d correspond to TV Type 1b (finite verb, long distance orientation). GJ Types 2a and 2c correspond to TV Type 2a (nonfinite verb, local antecedent orientation). GJ Types 2b and 2d correspond to TV Type 2b (nonfinite verb, long distance antecedent orientation). GJ Type 3a corresponds to TV Type 3a (monoclausal, object NP orientation), and GJ Type 3b corresponds to TV Type 3b (monoclausal, subject NP orientation). GJ Type 3c sentences are controls as explained above and have no corresponding TV type.

The NS results on the GJ test conform to the TV test results in terms of overall accuracy, and accuracy averages by type. For this study, 13 GJ test sessions were developed, each test session having eleven items. On each of the 13 test sessions, each accurately judged item is awarded 1 point, for a maximum score of 11. The raw totals for the test are listed by informant in Table 6.3 which also includes the sum of all scores, the percentage correct, the mean score, and the standard deviation

*Table 6.3. Native English speaker scores on the grammaticality judgement (GJ) test. Maximum score on each test session is 11.*

| NS Informant | NS1 | NS2 | NS3 | NS4 | NS5 | NS6 |
|---|---|---|---|---|---|---|
| Sum | 134 | 139 | 140 | 139 | 140 | 141 |
| Percent | 94% | 97% | 98% | 97% | 98% | 98% |
| Mean | 10.30 | 10.69 | 10.76 | 10.69 | 10.76 | 10.84 |
| SD | 0.94 | 0.48 | 0.43 | 0.48 | 0.43 | 0.37 |

As can be seen in Table 6.3, the NSs performed very well on the GJ test, and were accurate almost 98% of the time, a higher percentage than for the TV test. The differences between the NS mean scores and 11 (highest score possible) are insignificant at the $p < 05$ level (one-tailed $t$ test: $t = -4.18$, $df = 10$). Furthermore, with a range of just 7, and a standard deviation of less than 5.4, the native speaker control group informants performed very similarly to each other on these GJ test.

As on the TV test, we again see the NSs performing most poorly by sentence Type on the monoclausal structures as illustrated in Table 6.4.

*Table 6.4. Percentage of native speaker judgement accuracy of grammaticality judgement test items by sentence Type*

| NS Informant/ GJ Sentence Type | NS1 | NS2 | NS3 | NS4 | NS5 | NS6 | Total |
|---|---|---|---|---|---|---|---|
| Type 1a | 100% | 100% | 100% | 100% | 100% | 100% | 100% |
| Type 1b | 92% | 92% | 100% | 92% | 100% | 92% | 94% |
| Type 1c | 100% | 100% | 92% | 100% | 100% | 92% | 97% |
| Type 1d | 100% | 100% | 100% | 100% | 100% | 100% | 100% |
| Type 2a | 92% | 100% | 100% | 100% | 100% | 100% | 99% |
| Type 2b | 92% | 100% | 100% | 100% | 92% | 100% | 97% |
| Type 2c | 100% | 100% | 100% | 100% | 100% | 100% | 100% |
| Type 2d | 100% | 100% | 100% | 100% | 100% | 100% | 100% |
| Type 3a | 84% | 84% | 92% | 84% | 92% | 100% | 89% |
| Type 3b | 92% | 92% | 100% | 100% | 100% | 100% | 97% |
| Type 3c | 84% | 100% | 92% | 92% | 92% | 100% | 93% |

By sentence Type, the NS informants again perform slightly lower on the Type 3a sentences, monoclausal sentences with two NPs, grammatically favoring the object

NP antecedent. As mentioned above, these sentences test Wexler and Manzini's Proper Antecedent Parameter (PAP). Of a total of 78 GJ Type 3a sentences among the six NS informants, they judge 8 to be false. In other words, 11% of the time the NS group as a whole rejects the antecedent object NP candidate as the correct antecedent of the reflexive. This gives an accurate judgement rate of 89% on the GJ Type 3a sentences. This acceptance rate on GJ Type 3a sentences parallels the acceptance rate of 73% on TV Type 3a sentences in that they are both the most misjudged type on their respective tests. Again, this reflects the preference native English speakers have for subject NP antecedents over object NP antecedents in monoclausal sentences.

The next most rejected GJ sentence Type by the NS control group is Type 3c, monoclausal sentences with 2 NPs which had an accuracy rate of 93%. In this sentence type, the antecedent is ambiguous, the anaphor pragmatically being able to refer back to either the subject NP or the object NP. In a sense this makes these the easiest sentence type to judge accurately- they can select either the one NP or the other and mark their answer "OK" and not be wrong. So at first it is somewhat surprising four out of the six NSs would wrongly judge GJ Type 3c sentences. On closer inspection however, it becomes apparent that the 7% rate of misjudged sentences stems from just three tokens- two tokens each twice judged wrongly, and one token judged wrong once. The two commonly mistaken Type 3c tokens are presented in (3) and (4).

(3) *The bad boy told his father a lie about himself.*

(4) *The mother wrote to her daughter about herself.*

Both (3) and (4) were judged wrong by two NSs, who perhaps expected a pronoun such as *him* or *her* outside the matrix sentence, or they perhaps wanted to bind the reflexive back to the matrix subject. The demands of the processing load here may have caused them to wrongly interpret these antecedents as "long distance", and therefore not allowed. This may also explain the single GJ Type 3c wrong judgement in (5)

(5) *Bill's sick mother questioned her nurse about herself.*

NSs also exhibit some degree of error on GJ Type 1b sentences, with an accuracy rate of 96%. These errors are however, consistently stem from the same token, reproduced in (6).

(6) *The government officials hoped the army would defend themselves.*

GJ Type 1b items should be judged as wrong because the plural anaphor *themselves* doesn't agree in number with the local singular antecedent *army*. In judging (6) correct, informants appear to consider *army* a semantically plural form.

The NSs' results on the other seven GJ sentence Types – 1a, 1c, 1d, 2a, 2b, 2c, and 2d – are very much in line with results predicted by the Governing Category Principle, and further confirm the validity of the data collection instruments. Overall accuracy in Type 1a sentence judgements, biclausal sentences with a finite verb in the embedded clause favoring the local C-commanded antecedent through verb agreement, is 100%. For Type 1c sentences, which are the same structure as Type 1a, but favor the long distance matrix subject for the antecedent, the accuracy is also high with a rejection rate of 97%. The four GJ Type 2 sentences also have high accuracy rates among the NSs. For GJ Type 2a, biclausal sentences with a non-finite verb in the embedded clause favoring the local antecedent through verb agreement, the overall accuracy is high with an acceptance rate among NSs of 99%. The rejection rate of Type 2b sentences – biclausal, non-finite verb in the embedded clause, pragmatically favoring the long distance antecedent – is highly accurate at 97%. GJ Type 2c and 2d sentences, which respectively favor local and long distance binding domains through gender agreement, exhibit an accuracy rate of 100% each among the NSs. Performance on the GJ Type 3b sentences – monoclausal sentences favoring the subject NP as antecedent – is also high, at 97%. The high accuracy rates with which NSs judge local vs. long distance antecedents on the grammaticality judgement test not only confirms their adherence to binding principle A as realized for English, but also supports the validity of the stimulus sentences as accurate representations of the forms under investigation.

The NS scores on the GJ test are therefore considered to reliably correlate with the results of the TV test. The two tests are therefore considered reliable as tools with which to measure the test subjects' results. We will now turn to these results.

## 6.2. Japanese Test Subject Results

The results of the two tests will be considered here in light of the hypotheses formulated in Chapter 5 (section 5.1), and restated here:

Hypothesis A: Principles of reflexive binding already instantiated in the test subjects' English will undergo change due to lack of exposure to English.

Hypothesis B: Changes observed in the test subjects' control over reflexive binding due to lack of exposure will be UG constrained.

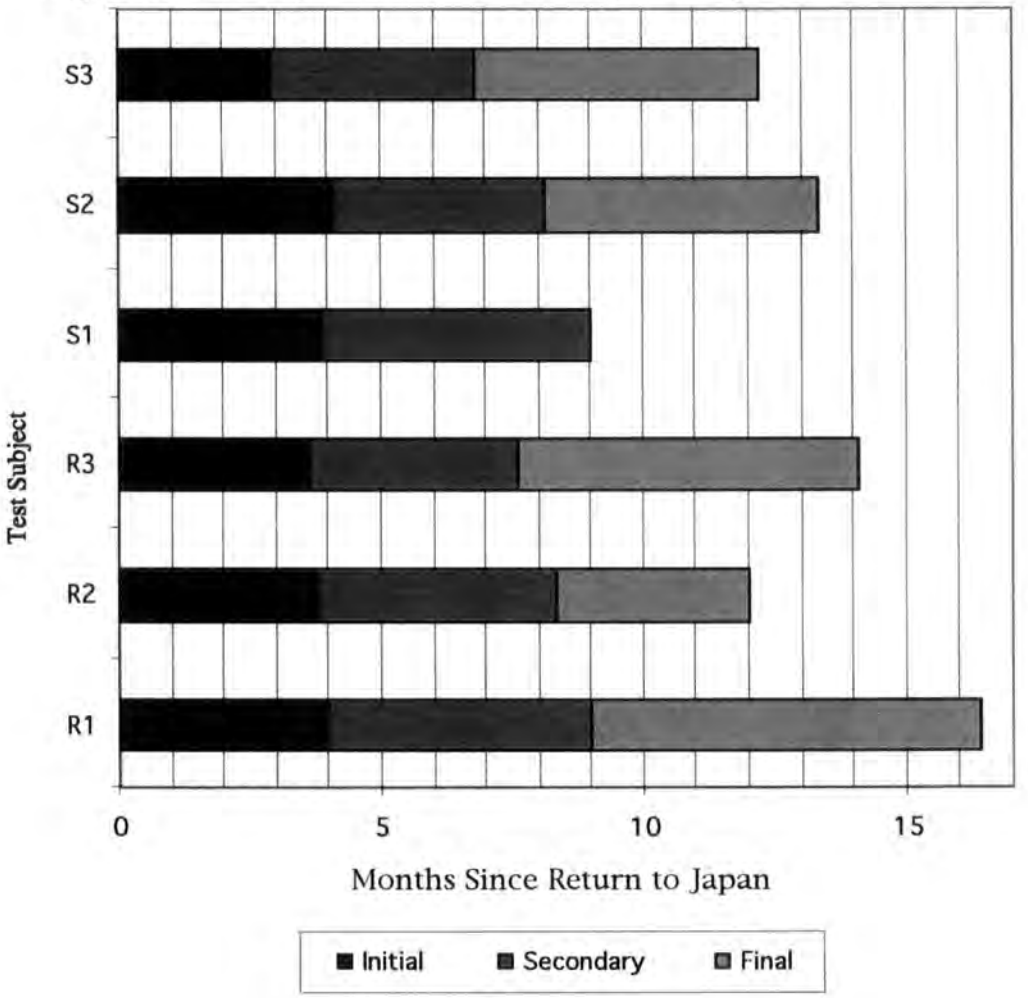
Hypothesis C: A correlation will be exhibited between age at first L2 exposure and degree of retention of L2 reflexive binding principles.

The results of each test will be considered in turn, first the truth value judgement test, followed by the grammaticality judgement test.

For each test subject, the period of time from her return to Japan to her last test session will be referred to as her 'data collection period'. Although test subjects had varying numbers of test sessions, all six had a data collection period lasting at least 1 year, except test subject S1 whose data collection lasted 9 months. In order to compare test subjects' results within a meaningful time-frame, each test subject's data collection period has been divided into approximately three equal sub-periods which will be referred to as the initial, the secondary, and the final data collection sub-periods. The exact length of each sub-period is determined by when test sessions took place with respect to time since last exposure (i.e., return to Japan), thus causing the same sub-period for each test subject to vary. The initial sub-period includes the point of re-entry into Japan to whichever test session is closest to 4 months since her re-entry. For example, test subject S2 had her 6th test session on the 4$^{th}$ month and 3$^{rd}$ day after her return to Japan. This test session will be referred to as Time Since Re-entry (TSR) 4.1, the number of months in months and tenths of months since re-entering the L1 Japanese environment. The secondary sub-period will include the approximately 4-month period

subsequent to the initial sub-period. For example in S2's case, her secondary sub-period lasts from TSR 4.1 to TSR 8.1, 8 months and 3 days after TSR 0, her re-entry date into Japan. These initial, secondary and final data collection sub-periods for each Japanese test subject are represented in Figure 6.1. Note that test subject S1's entire data collection period lasted only 9 months and thus has only an initial and secondary sub-period.

*Figure 6.1. Test subject data collection sub-periods by TSR*




As the results are analyzed and discussed, instability which becomes evident in the subjects' grammar be noted. Some test subjects show a weakness on certain sentence types from the start of the data collection period. This is reflected in fluctuating results from the first few test sessions and may point to a lack of knowledge of the principles underlying those sentence types from the outset. Also of interest are the

results of certain test subjects' other sentence types which initially show a consistently high level of accuracy in judging test items, but later become more erratic, with high and low scores mixed together. This change from stable to unstable results within individuals is captured by the longitudinal nature of this study and may serve as one indicator of attrition.

### 6.2.1. *The test subjects' truth value judgement test results*

In analyzing the six test subjects' test performance, their scores on the truth value judgement test over the data collection period are considered in terms of degree of adherence to the principles of reflexive binding in English. This degree of adherence was the standard used when examining the native English speakers' data.

Table 6.5 includes each test subject's score on each TV test session she took as percentages, along with the test session number and corresponding TSR number (i.e., the months elapsed at each test session since the test subject's return to Japan). The initial, secondary and final data collection sub-periods presented in Figure 6.1 are indicated by the color coding in Table 6.5.

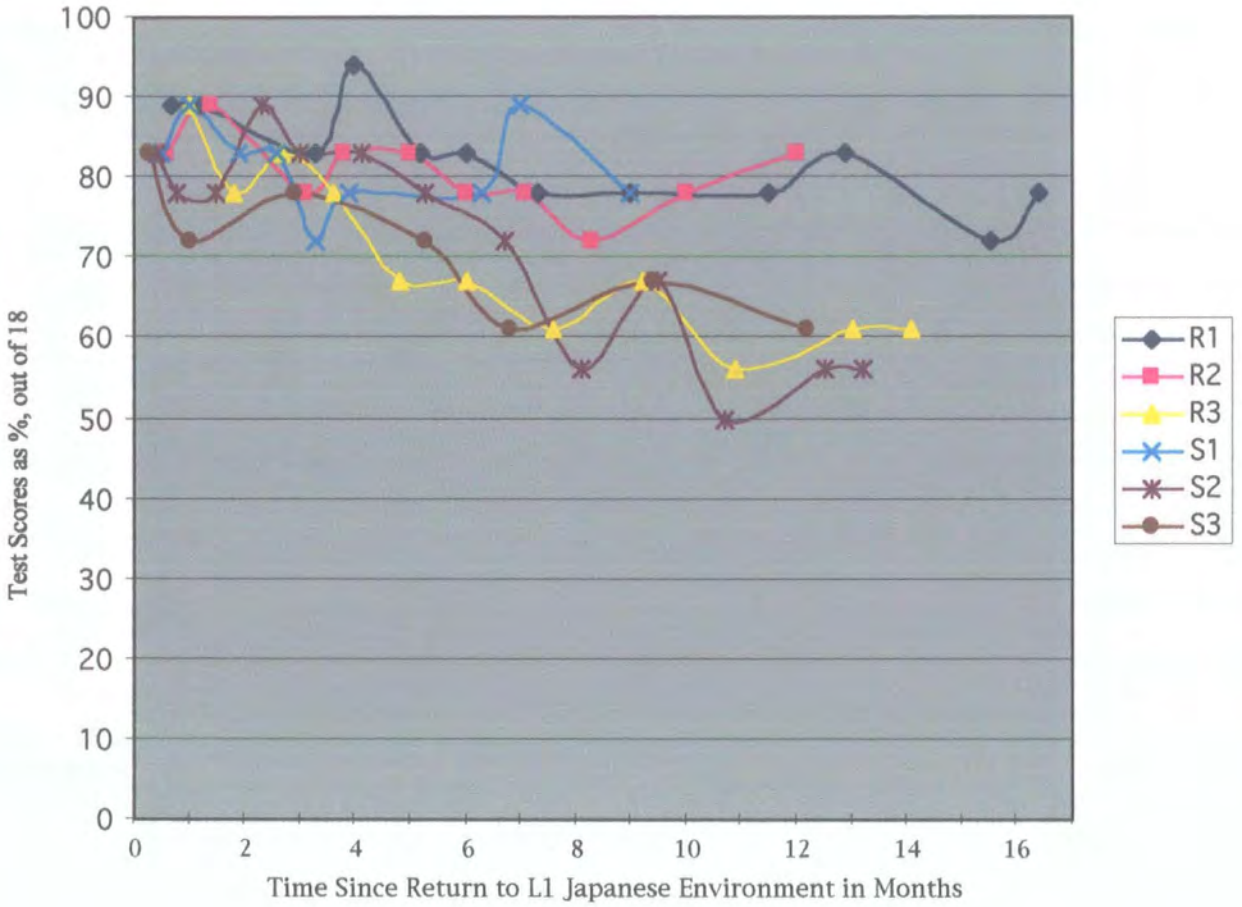
*Table 6.5. TV scores at time since L1 re-entry (TSR)*

| Test Subject/Test Session # | R1 | | | R2 | | | R3 | | | S1 | | | S2 | | | S3 | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Score | TSR | Date | Score | TSR | Date | Score | TSR | Date | Score | TSR | Date | Score | TSR | Date | Score | TSR | Date |
| (Return date) | na | 0 | 7/1/97 | na | 0 | 6/20/97 | na | 0 | 8/28/97 | na | 0 | 3/19/97 | na | 0 | 9/15/96 | na | 0 | 9/15/97 |
| #1 | 89% | 0.7 | 7/11/97 | 83% | 0.6 | 7/8/97 | 89% | 1 | 9/29/97 | 83% | 0.5 | 4/4/97 | 83% | 0.4 | 9/27/96 | 83% | 0.3 | 9/29/97 |
| #2 | 89% | 1.2 | 7/25/97 | 89% | 1.4 | 8/1/97 | 78% | 1.8 | 10/20/97 | 89% | 1 | 4/18/97 | 78% | 0.8 | 10/11/96 | 72% | 1 | 10/20/97 |
| #3 | 83% | 3.3 | 9/26/97 | 78% | 3.1 | 9/23/97 | 83% | 2.7 | 11/17/97 | 83% | 1.9 | 5/16/97 | 78% | 1.5 | 11/1/96 | 78% | 2.9 | 12/15/97 |
| #4 | 94% | 4 | 10/17/97 | 83% | 3.8 | 10/14/97 | 78% | 3.6 | 12/15/97 | 83% | 2.6 | 6/6/97 | 89% | 2.3 | 11/22/96 | 72% | 5.3 | 2/27/98 |
| #5 | 83% | 5.2 | 11/21/97 | 83% | 5 | 11/18/97 | 67% | 4.8 | 1/20/98 | 72% | 3.3 | 6/27/97 | 83% | 3 | 12/13/96 | 61% | 6.8 | 4/13/98 |
| #6 | 83% | 6 | 12/15/97 | 78% | 6 | 12/16/97 | 67% | 6 | 2/23/98 | 78% | 3.9 | 7/15/97 | 83% | 4.1 | 1/17/97 | 67% | 9.4 | 6/30/98 |
| #7 | 78% | 7.3 | 1/23/98 | 78% | 7.1 | 1/20/98 | 61% | 7.6 | 4/13/98 | 78% | 6.3 | 9/26/97 | 78% | 5.3 | 2/21/97 | 61% | 12.2 | 9/21/98 |
| #8 | 78% | 9 | 3/15/98 | 72% | 8.3 | 2/24/98 | 67% | 9.2 | 6/1/98 | 89% | 7 | 10/17/97 | 72% | 6.7 | 4/4/97 | na | na | na |
| #9 | 78% | 11.5 | 5/29/98 | 78% | 10 | 4/14/98 | 56% | 10.9 | 7/20/98 | 78% | 9 | 12/13/97 | 56% | 8.1 | 5/16/97 | na | na | na |
| #10 | 83% | 12.9 | 7/10/98 | 83% | 12 | 6/16/98 | 61% | 13 | 9/22/98 | na | na | na | 67% | 9.5 | 6/27/97 | na | na | na |
| #11 | 72% | 15.5 | 9/25/98 | na | na | na | 61% | 14.1 | 10/26/98 | na | na | na | 50% | 10.7 | 8/1/97 | na | na | na |
| #12 | 78% | 16.4 | 10/23/98 | na | na | na | na | na | na | na | na | na | 56% | 12.5 | 9/26/97 | na | na | na |
| #13 | na | na | na | na | na | na | na | na | na | na | na | na | 56% | 13.2 | 10/17/97 | na | na | na |

Initial sub-period  Secondary sub-period  Final sub-period

Figure 6.2 provides a general illustration of the range and decline of individual test subject scores on the TV tests over time, as outlined above in Table 6.5

*Figure 6.2. Test Subjects' individual TV scores over time*




In Figure 6.2, initial scores cluster between 83% and 89%, but widen over time. In particular, R1, R2, and S1 tend to maintain the levels initially obtained, whereas the levels of R3, S2, and S3 decline over time.

In the next sections we will consider each test subject's TV and then GJ test results by sentence type over time.

6.2.1.1. *Individual TV test results by Type over time*

In this section each of the six test subjects' scores on the six TV sentence types over time will be examined in order to identify the degree to which particular aspects of reflexive binding are susceptible to attrition in each individual.

As described in Chapter 5 (section 5.3.2.2.), each TV test session includes 3 tokens of each sentence type. Because test subjects varied in the number of test sessions they took, the total number of test type tokens for each test subject varies. For example, R1 took 12 test sessions and thus has 36 judgements for each sentence type, whereas S3 took 7 test sessions and has 21 judgements for each type. The total number of sessions conducted in this study is 62 (the sum of the test subjects' various number of sessions), and the total number of tokens of each TV sentence type included in the study is 186 (62 x 3 tokens of each type).

Throughout this section and Section 6.2.2, bar graphs such as in Figure 6.3 below are associated with each test subject and illustrate her performance throughout her data collection period. Within each graph, each group of six vertical bars represents one test session and each bar within a group represents performance on a sentence type. Below the groups of bars is a grid. The top row of the grid indicates the TSR- the amount of time elapsed since a test subject's return to Japan from the States. Next to each TSR number the TV test session number which corresponds to that TSR is indicated in parentheses. The numbers in the columns below each TSR represent the percentage of correct responses for each sentence type at that TSR for that test subject. These percentages are color keyed at the left and correspond with the colored bars in the graphs. The order of the sentence types in each group of bars is, from left to right, 1a (object antecedent), 1b (subject antecedent), 2a (object antecedent), 2b (subject antecedent), 3a (object antecedent), and 3b (subject antecedent). Each test session includes three tokens of each type. As each test session has six different sentence types this gives a maximum of 18 points per test session. Each test subject's TV test scores can be found in Appendix H, and results by sentence type can be seen in Appendix I. Appendix I lists the test subjects' results for each TV sentence type over time. TV test session numbers TV1 through TV13 are listed across. The sentence types are grouped into six main sections, each section corresponding to one of the six TV sentence types, biclausal finite local, biclausal finite long distance, etc.

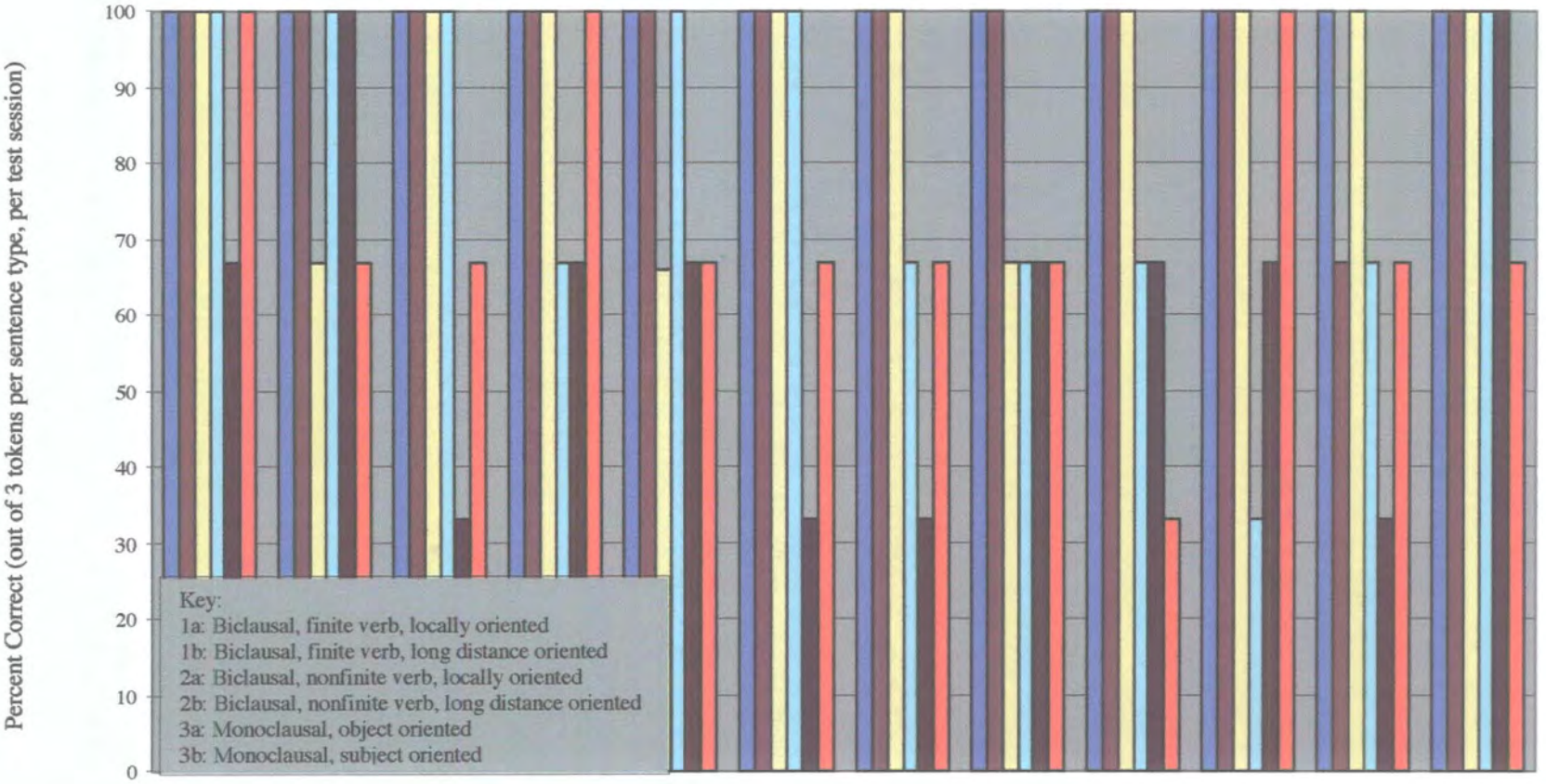
For each test subject, results for each data collection sub-period will initially be combined, averaged and compared to the secondary and final sub-period averages. Following this will be a more thorough examination of results by TV sentence type (i.e., Types 1a to 3b) for each test subject. Averages for each sub-period of a data collection period will be calculated by first determining the number of correctly judged tokens (out of 3) of each sentence type on each test session for that sub-period, and then averaging those scores.

Let us begin by looking at test subject R1's TV test data to see how this works.

6.2.1.1.1. *Test subject R1's TV test results by sentence type over time*

R1's collection period lasted just over 16 months from her time of arrival back in Japan. The overall average of her TV results for her initial sub-period (the first 4 months up to TSR 4.0) is 89%, the highest of all the test subjects for this test over this period of time. Over her secondary, 5 month long sub-period (up to TSR 9), R1's results remain high, but exhibit some attrition as the average declines to 81%. Over her final data collection sub-period, up to TSR 16.4, R1's results show further overall attrition, declining to an average of 78%. R1's results by sentence type are depicted in Figure 6.3.

Figure 6.3. R1 truth value test scores over time by sentence type




| | 1 (0.7) | 2 (1.2) | 3 (3.3) | 4 (4) | 5 (5.2) | 6 (6) | 7 (7.3) | 8 (9) | 9 (11.5) | 10 (12.9) | 11 (15.5) | 12 (16.4) |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1a Object | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 |
| 1b Subject | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 67 | 100 |
| 2a Object | 100 | 67 | 100 | 100 | 66 | 100 | 100 | 67 | 100 | 100 | 100 | 100 |
| 2b Subject | 100 | 100 | 100 | 67 | 100 | 100 | 67 | 67 | 67 | 33 | 67 | 100 |
| 3a Object | 67 | 100 | 33 | 67 | 67 | 33 | 33 | 67 | 67 | 67 | 33 | 100 |
| 3b Subject | 100 | 67 | 67 | 100 | 67 | 67 | 67 | 67 | 33 | 100 | 67 | 67 |

TV scores by sentence type. Top row of grid indicates test number, TSR in pararentheses.

R1's results exhibit high levels of consistent accuracy in sentence Types 1a, and 1b, respectively averaging 100% and 97% over the entire collection period. These mean scores are comparable to the NS control group average scores for these two types. R1's scores on Type 1a and 1b sentences reflect high degrees of continued acceptance of local antecedents and rejection of long distance antecedents in biclausal sentences with a finite verb in the embedded clause.

R1's Type 2a results average a high 91% over the entire data collection period and show only some slight fluctuation within the secondary sub-period, from TSR 5.2 to TSR 9. Her high averages on this sentence type reflect continued acceptance over time of binding to the local antecedent. R1's Type 2b sentence results, however, reflect attrition in biclausal sentences with a nonfinite verb in the embedded clauses. On Type 2b sentences, test subject R1's test results average 92% in her initial data collection sub-period (up to TSR 4.0) and decline during her secondary sub-period (TSR 4.0 to TSR 9) to 83%. From TSR 9 to the end of her data collection period at TSR 16.4, R1's 2b results show considerable attrition and average 67% over this final 7 month period. This change from relative consistency up to TSR 9 to a more erratic pattern points to an erosion of the knowledge of the principles underlying the restriction to object NP antecedents in biclausal sentences in English.

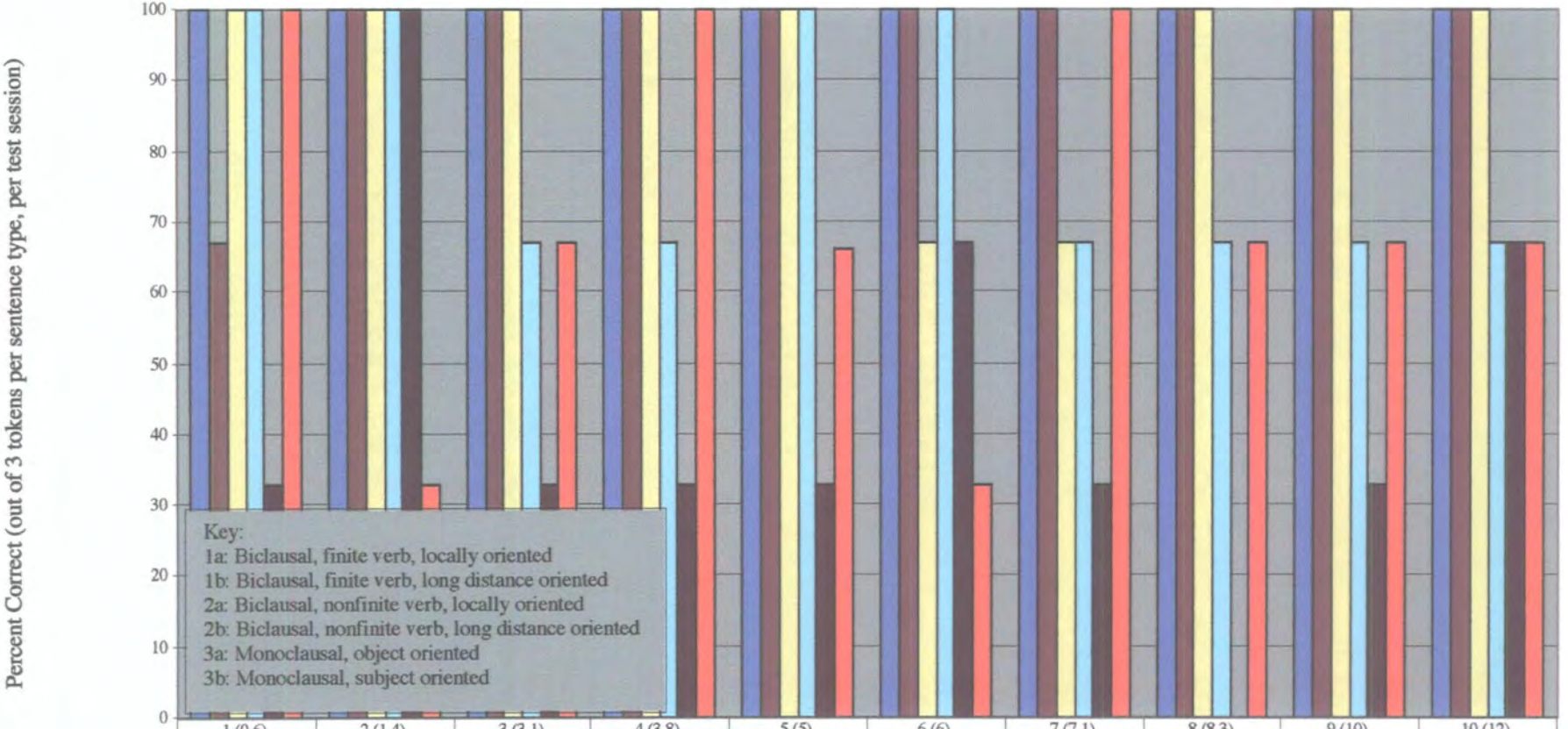
On Type 3a (object) sentences, R1's results fluctuate over time and average 61% for the entire test period. Despite being the highest Type 3a mean score of all the test subjects, R1's performance on Type 3a sentences is her lowest level of performance of all six TV Types. R1's low level of accuracy for Type 3a sentences is the result of consistently lower scores throughout her test period, rather than a cluster of low or fluctuating scores during the collection period. Up to TSR 4.0, her results average 67%. Over the next five months of her secondary sub-period, this declines to 50%, and over the remainder of her collection period (to TSR 16.4), R1's Type 3a results average 67% again. This broad pattern of fluctuation over the entire length of her data collection period indicates that test subject R1 had trouble recognizing object NPs as acceptable in English monoclausal sentence of the type represented by 3a even from the start of the data collection period. This is actually not so surprising as binding in Japanese is subject oriented and the changes seen here may be a result of L1 transfer. R1's results on her TV Type 3b (subject) sentences again show fluctuation throughout her data collection period. Her average over the initial four months (to TSR 4.0) is 83%. Following this we

can observe some attrition in her knowledge of the principles underlying sentence Type TV 3b as the average of her results declines to 67% for both the secondary TSR 5.2 to 9.0 period and the final TSR 11.5 to 16.4 period. These results apparently contradict the suggestion that R1 is transferring her L1 Japanese binding rules to English, because, as noted, Japanese is subject oriented. Perhaps the context of the story somehow blocked recognition of the subject NP as an acceptable antecedent.

R1's TV test results by type are included in Appendix I.

6.2.1.1.2. *Test subject R2's TV test results by sentence type over time*

Test subject R2's data collection period lasted 11 months from her time of arrival back in Japan and included 10 test sessions. The overall average of her TV results for the initial 3.8 months is 83%. Over the next four and a half months, to TSR 8.3, R2's results decline slightly to an average of 78%, but increase again to an average of 80% in the final three and a half months (to TSR 12). Although exhibiting some fluctuation, R2's results reflect an overall strong knowledge of English binding principles. Her TV test results by sentence type are depicted in Figure 6.4.




| | 1 (0.6) | 2 (1.4) | 3 (3.1) | 4 (3.8) | 5 (5) | 6 (6) | 7 (7.1) | 8 (8.3) | 9 (10) | 10 (12) |
|---|---|---|---|---|---|---|---|---|---|---|
| 1a Object | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 |
| 1b Subject | 67 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 |
| 2a Object | 100 | 100 | 100 | 100 | 100 | 67 | 67 | 100 | 100 | 100 |
| 2b Subject | 100 | 100 | 67 | 67 | 100 | 100 | 67 | 67 | 67 | 67 |
| 3a Object | 33 | 100 | 33 | 33 | 33 | 67 | 33 | 0 | 33 | 67 |
| 3b Subject | 100 | 33 | 67 | 100 | 66 | 33 | 100 | 67 | 67 | 67 |

TV scores by sentence type. Top row of grid indicates test number, TSR in parentheses.

Figure 6.4. R2 truth value test scores over time by sentence type

Though not quite as proficient as R1, R2's sentence type scores also exhibit a consistently high level of accuracy throughout her collection period. For Type 1a, and 1b sentences, test subject R2 scores nearly 100% throughout her data collection period. These match the averages of the control group scores of 97% and 99% for Types 1a and 1b respectively. These high percentages reflect R2's ability to accept local antecedents and reject long distance antecedents in biclausal sentences which include a finite verb in the embedded clause.

On TV Type 2a sentences, R2's results again show little fluctuation, averaging 100% to TSR 3.8, 83% to TSR 8.3, and increasing to 100% again during the last three and a half months of her collection period. Despite the dip in her results on Type 2a sentences in the middle of her collection period, R2 compares favorably with the NS results on this sentence type which average 99%. In contrast to this, the fluctuation and decrease in R2's TV Type 2b results show distinct evidence of attrition over time. On this sentence type, her results average 83% to TSR 3.8, 89% to TSR 7.1, and 67% over the last five months. The fluctuation and decline seen here clearly show a loss of the principles underlying the Type 2b sentences over the course of her data collection period.

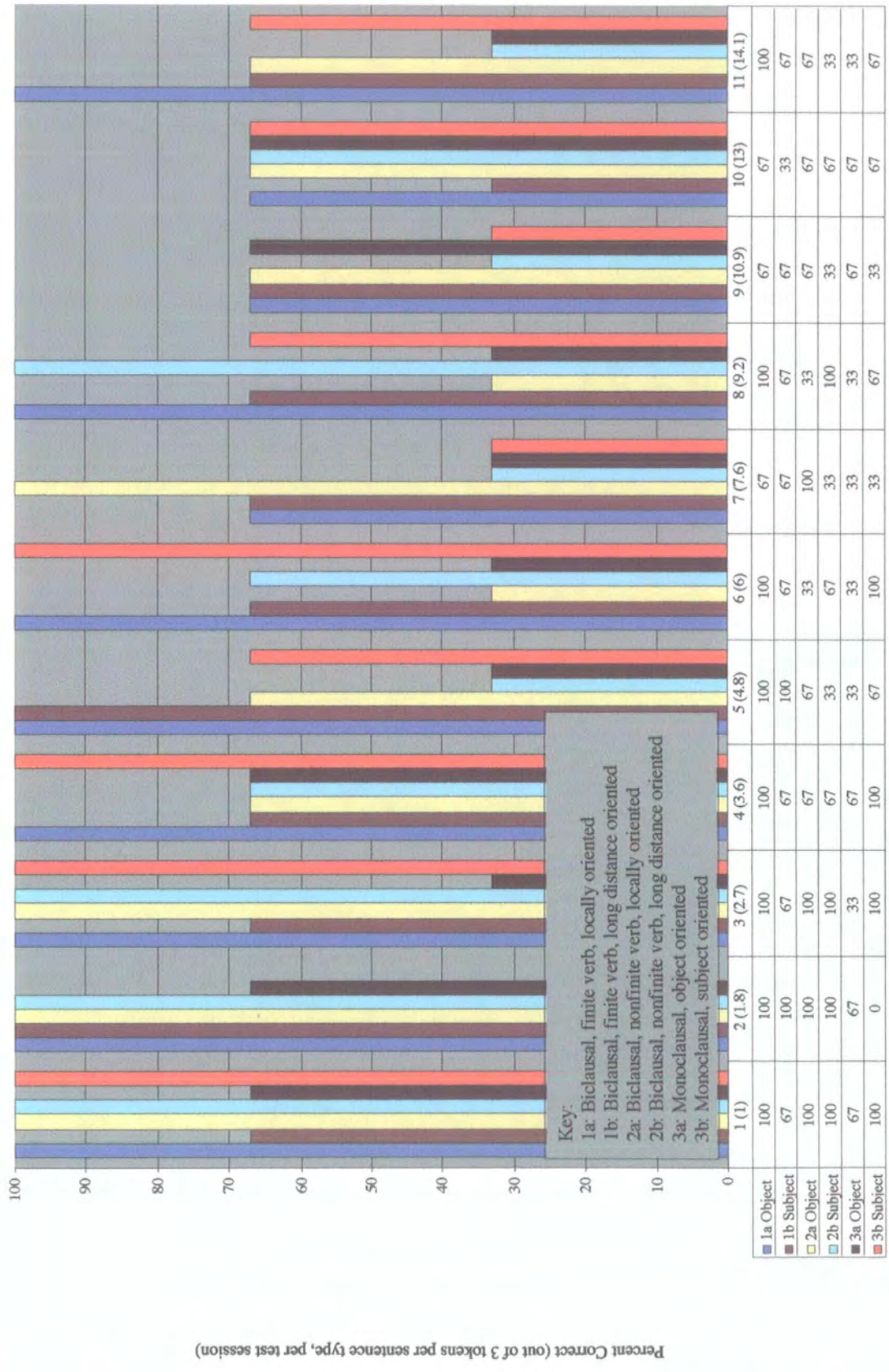
As with R1 and the NS controls, R2 performs most poorly on the monoclausal Type 3a and 3b sentences. On the Type 3a sentences R2's results clearly fluctuate and decline over time. Up to TSR 3.8 R2's results average only 50%. This is slightly lower than R1's 67% over a similar period of time (4.0 months). Over the next three and a half months, to TSR 7.1, R2's Type 3a results average just 33% and then fluctuate up to 50% over the last five months. These low averages and a single 100% average in the second test session at TSR 1.4 are evidence of the uncertain and fluctuating nature of her knowledge of English binding principles in monoclausal sentences of this type. On Type 3b sentences, R2 performs somewhat better than on her Type 3a sentences, but shows some signs of attrition. To TSR 3.8, R2's results average 75%, the same as the NS average for this sentence type. However, for the remainder of her data collection period, R2's results decline to 67% during both the TSR 3.8 to 7.1 period, and during the final stage of her collection period, TSR 7.1 to 12.

R2's TV scores by type are included in Appendix I.

6.2.1.1.3. *Test subject R3's TV test results by sentence type over time*

Test subject R3's data collection period lasted just over 14 months from her arrival back in Japan and included 11 test sessions. The overall average of her TV results for the initial 3.6 months is 82%. This approximates R1's and R2's overall performances during similar periods of time. R3's overall scores are only about 10% below the NS control group mean, and reflect an initial overall good knowledge of English reflexive binding. After TSR 3.6, however, R3's overall scores decline and the average of her results over her secondary period, from TSR 4.8 to TSR 7.6, is 66%, whereas R1 and R2 maintain high scores of 89% and 83%, respectively, at their corresponding secondary periods. For the remainder of their data collection periods, R1 and R2 maintain overall average percentages of accuracy of 80% and 75%, respectively, while R3 averages 62% on the remainder of her TV test sessions. This percentage is well below the NS control mean of 89% for these four test sessions (TV8 through TV11) and provides general evidence of attrition in R3's knowledge of English binding. R3's TV test results by sentence type are depicted in Figure 6.5.

Figure 6.5. R3 truth value test scores over time by sentence type




TV scores by sentence type. Top row of grid indicates test number, TSR in parentheses

At the sentence type level, test subject R3's strongest TV test type results are on Type 1a sentence, with an average of 90% throughout her data collection period. Over the initial four months, up to TSR 3.6, R3 averages 100% on her Type 1a sentences. During the next four months, up to TSR 7.6, R3's Type 1a scores remain high, averaging 92%. From TSR 7.6 to the end of her data collection period however, test subject R3' results begin to show signs of attrition on TV Type 1a sentences as her average for this period averages 78%. R3's TV Type 1b sentence results show a pattern similar to her 1a results, though at a lower level. The results of her Type 1b sentence up to TSR 3.6 averages 75%. From TSR 3.6 to 7.6, her average is also 75%. For the remainder of her collection period, to TSR 14.1, however, the average of R3's TV Type 1b results declines to 56%, reflecting an increasing inability to recognize the illegality of long distance binding in English biclausal sentences with a finite verb in the embedded clause.

R3's Type 2a and 2b sentence results also establish patterns of attrition. From the start of her data collection period to TSR 3.6, test subject R3's Type 2a results average 92%, comparable to the NS average of 95%. During the next 4 months however, R3's 2a results attrite dramatically to 58%. They subsequently recover slightly over the last six and a half months, but still only average 67% at TSR 14.1. This attrition reflects an increased propensity to reject locally bound antecedents in biclausal sentences with a nonfinite verb in the embedded clause. R3's results for sentence Type 2b parallel those of test subjects R1 and R2 who initially show evidence of adherence to the principles underlying Type 2b sentences, but decline after TSR 7.1 and 7.6 respectively. This attrition shows a marked decline in test subject R3's ability to reject long distance binding in English biclausal sentences with nonfinite verbs in the embedded clause.

R3's Type 3a sentence results fluctuate greatly across time and reflect an initial lack of understanding of the acceptability of object NPs as antecedents in English rather than attrition over time. Up to TSR 3.6, R3's results on this sentence type average 58%. This average drops over the next four months to 33% by TSR 7.6, and rises over the final half year to 56%. These results indicate overall uncertainty on the acceptability of object NP antecedents in monoclausal sentences, a fluctuating pattern of low scores similar to those observed in R1's and R2's results for these sentence types. On TV Type 3b sentences, R3's average of 75% up to TSR 3.6 equals the NSs' overall average of 75%. But this high average decreases steadily over time to 67% from TSR 3.6 to TSR

9.2 and then to 56% over the last five months. This decline indicates an increased propensity for R3 to choose the subject NP as the antecedent over the object antecedent. The attrition observed in R3's TV Type 3b sentences contrasts with the results obtained for R1 and R2 whose scores fluctuate throughout the collection period, but show little or no sign of attrition and obtain averages throughout their data collection sessions comparable to the NS averages.

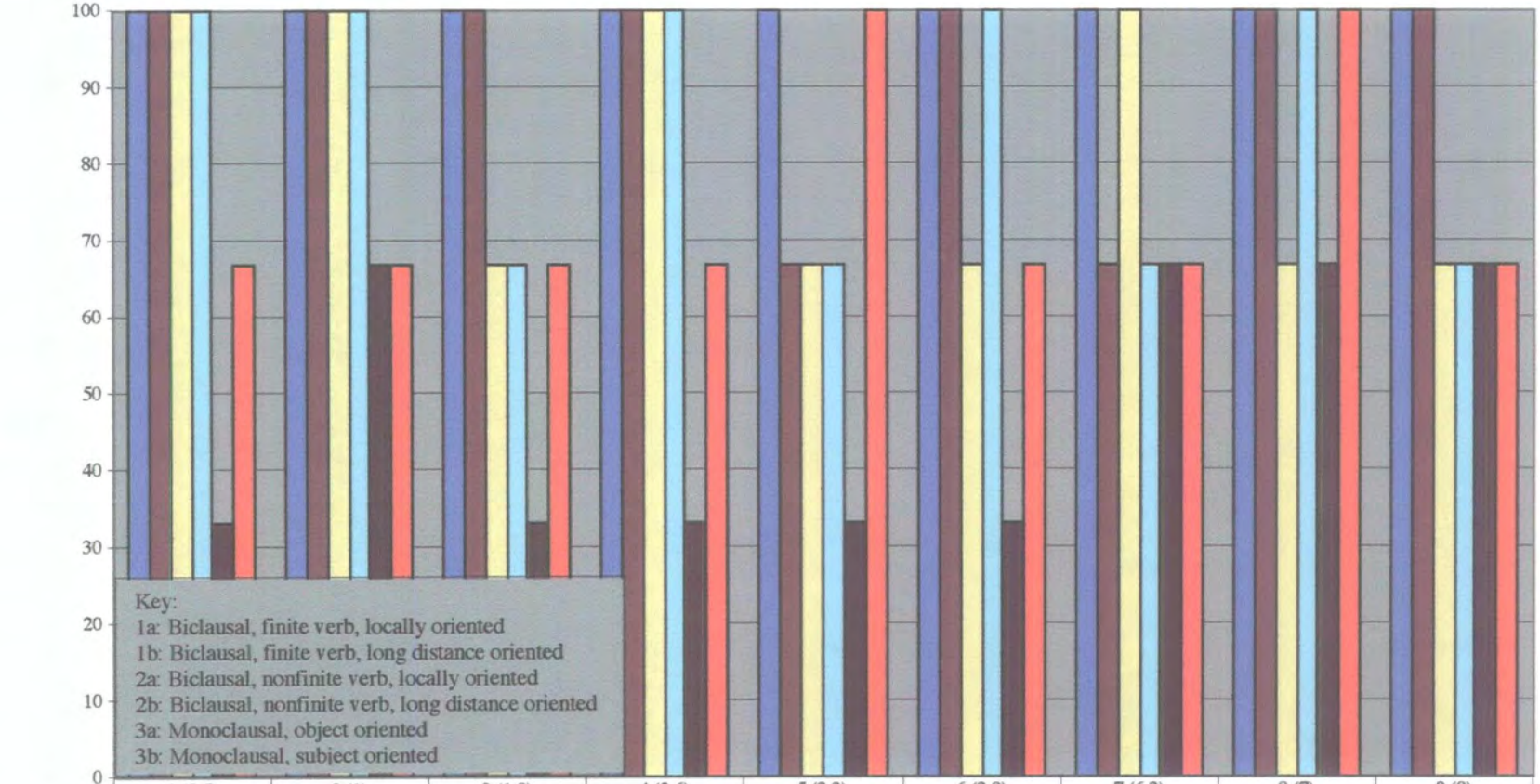
Taken on their own, R3's test results point to sentence-type-based performance differences. Additionally, when compared to the other 'R' subjects, her results indicate test subject based performance differences. These issues will be addressed in the discussion section.

R3's TV scores by type are included in Appendix I.

6.2.1.1.4. *Test subject S1's TV test results by sentence type over time*

Truth value test results by sentence type for the 'S' test subjects (i.e., those with no childhood exposure to English) largely parallel the patterns established in the 'R' subjects' results, though the 'S' scores are generally lower than the 'R' subjects' scores.

Test subject S1's data collection period lasted nine months from her return to Japan and included 9 data collection sessions. This nine-month period of time will be considered in two sections, the 4 months from TSR 0 to TSR 3.9, and the five months from TSR 3.9 to TSR 9. Despite her lack of childhood exposure to English, S1's performance on the TV test is comparable to that of R1 and R2, and she performs noticeably better than R3 and her own 'S' group mates, S2 and S3. For the initial 4 months (up to TSR 3.9), S1 has an across-type accuracy rate of 81%. These results closely resemble those of R1, R2, and R3 over their initial periods of testing and reflect the accuracy of S1's overall knowledge of English reflexive binding at the start of her data collection period. S1's overall results remain stable during her secondary sub-period, averaging 82% from TSR 3.9 to TSR 9. S1's TV test results by sentence type are depicted in Figure 6.6.




| | 1 (0.5) | 2 (1) | 3 (1.9) | 4 (2.6) | 5 (3.3) | 6 (3.9) | 7 (6.3) | 8 (7) | 9 (9) |
|---|---|---|---|---|---|---|---|---|---|
| 1a Object | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 |
| 1b Subject | 100 | 100 | 100 | 100 | 67 | 100 | 67 | 100 | 100 |
| 2a Object | 100 | 100 | 67 | 100 | 67 | 67 | 100 | 67 | 67 |
| 2b Subject | 100 | 100 | 67 | 100 | 67 | 100 | 67 | 100 | 67 |
| 3a Object | 33 | 67 | 33 | 33 | 33 | 33 | 67 | 67 | 67 |
| 3b Subject | 67 | 67 | 67 | 67 | 100 | 67 | 67 | 100 | 67 |

TV scores by sentence type. Top row of grid indicates test number, TSR in parentheses

At the sentence type level, S1 scores 100% on all the TV Type 1a sentences over the entire course of her data collection period, as do R1 and R2 and the NS control group. R3's results on this type, though quite strong and consistent up to TSR 3.9, fluctuate during the remainder of her collection period. On Type 1b sentences, S1's results show a slight decrease from an average of 100% up to TSR 3.9, to 91% by her last test session at TSR 9. In contrast, R1 and R2 maintain the 100% NS-like level right up to the end of their data collection periods. R3's results for this type again fluctuate considerably throughout her collection period. Despite the slight drop on Type 1b sentences, S1 exhibits a robust knowledge of reflexive binding in biclausal sentences with a finite verb in the embedded clause.

S1's Type 2a sentence results show some evidence of attrition on this type, with a steady average of 83% over the initial 4 month period, to fluctuating results over the next 5 months which result in an average of 78%. S1's results on her Type 2b sentences fluctuate throughout her data collection period, but surprisingly show a slightly higher degree of robustness than the results obtained by R1 and R2 for this sentence type. S1's results average almost 89% during the initial 4 months, and 78% over the remainder of her collection period. In contrast to this, R1 and R2 have mean scores of 89% over their initial test periods on Type 2b sentences, while over the remainder of their data collection periods, R1's results average 74% and R2's average 67%. R3's results show severe attrition on this sentence type when compared to S1. Up to TSR 3.6, R3's averages average 92%. During the remainder of her collection period, R3's widely fluctuating results only average 52%. Compared to R1, R2, and R3 then, S1 does seem to maintain the highest level of recognizing the illegality of long distance binding in English biclausal sentences with a non-finite verb in the embedded clause.

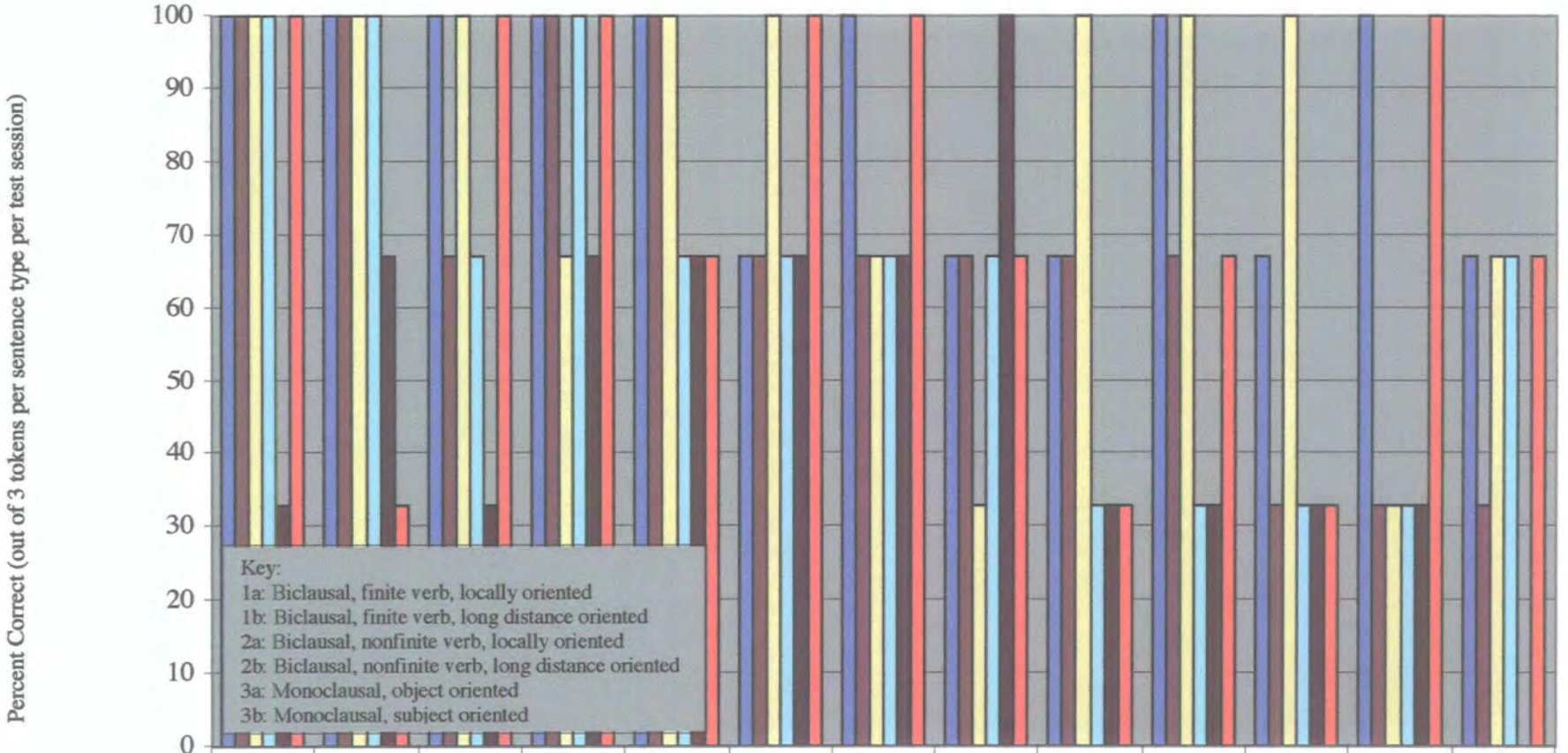
Test Subject S1's results for sentence Types 3a and 3b actually improve over the course of her data collection period, her 3a results in particular. Up to TSR 3.9 S1's TV Type 3a sentence results average just 39%. During her secondary (and final) test period, to TSR 9, S1's 3a results average increases dramatically to 67%. It is unlikely that this 28% increase was due to the test subject becoming familiar with the test as there was a two and a half month gap in test sessions between test session #6 (TSR 3.9) and test session #7 (TSR 6.3). As seen above, the TV Type 3a results for test subjects R1, R2, and R3 are quite mixed and all reflect various degrees of fluctuation. S1's Type 3b results show a trend similar to, though less dramatic than, her Type 3a results. S1'a TV

Type 3b results average 72% during the first four months after her return to Japan, and 77% over the remainder of her collection period. These averages are comparable to the NS average of 75% for this type. Overall, S1's Type 3a and 3b results reveal she strongly rejects object NPs as antecedents, but has a relatively stable knowledge of the principles underlying the choice of subject antecedent in monoclausal sentences in English.

S1's TV scores by type are included in Appendix I.

6.2.1.1.5. *Test subject S2's TV test results by sentence type over time*

Test subject S2's TV test results exhibit the most striking attrition of the 6 test subjects. Her data collection period lasted slightly more than 13 months from her return to Japan (TSR 0), during which time she had 13 data collection sessions. This 13-month period has been divided into 3 parts, from TSR 0 to 4.1, then to TSR 8.1, and then to TSR 13.2. Over the initial 4 months, S2's overall average for her TV results is 82% and is comparable to the results for R1, R2, R3, and S1 for the same approximate time period. However, throughout the remainder of her collection period, S2's TV results decline steadily. Her results for the secondary sub-period, from TSR 4.1 to 8.1, average 69%. And over the final 5 months, from TSR 8.1 to 13 .2, the average drops to 57%. This decline and the fluctuation evidenced in her scores throughout her collection period reflect an initial knowledge of English binding principles which subsequently attrites over time. S2's TV test results by sentence type are depicted in Figure 6.7.




| | 1 (0.4) | 2 (0.8) | 3 (1.5) | 4 (2.3) | 5 (3) | 5 (4.1) | 6 (5.3) | 7 (6.7) | 8 (8.1) | 9 (9.5) | 10 (10.7) | 11 (12.5) | 12 (13.2) |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1a Object | 100 | 100 | 100 | 100 | 100 | 67 | 100 | 67 | 67 | 100 | 67 | 100 | 67 |
| 1b Subject | 100 | 100 | 67 | 100 | 100 | 67 | 67 | 67 | 67 | 67 | 33 | 33 | 33 |
| 2a Object | 100 | 100 | 100 | 67 | 100 | 100 | 67 | 33 | 100 | 100 | 100 | 33 | 67 |
| 2b Subject | 100 | 100 | 67 | 100 | 67 | 67 | 67 | 67 | 33 | 33 | 33 | 33 | 67 |
| 3a Object | 33 | 67 | 33 | 67 | 67 | 67 | 67 | 100 | 33 | 33 | 33 | 33 | 0 |
| 3b Subject | 100 | 33 | 100 | 100 | 67 | 100 | 100 | 67 | 33 | 67 | 33 | 100 | 67 |

TV scores by sentence type. Top row of grid indicates test number, TSR in parenthes

Figure 6.7. S2 truth value test scores over time by sentence type

By sentence type, test subject S2's mean score on TV Type 1a sentences falls from 94% in the first 4 months to 78% by TSR 8.1. The average for her results over the last 5 months for this sentence type subsequently increases to 83%. This initial high average followed by fluctuating results reflects a loss of use of the principles underlying Type 1a sentences. The results of S2's Type 1b sentences also show attrition, declining from an average of 88% in her initial sub-period, to 67% by the end of her secondary sub-period at TSR 8.1, to 42% by the end of her collection period at TSR 13.2. This is the largest drop of all the test subjects for this TV sentence type. It clearly demonstrates that test subject S2 increasingly rejects local antecedents, and comes to perceive subject antecedents as acceptable in English biclausal sentences with finite embedded clauses.

S2's Type 2a sentence results show a drop of 27% over the first eight months of her data collection period, going from 94% in her initial sub-period to 67% in her secondary sub-period. The average over the last 5 months somewhat surprisingly increases 16% to 83%. Despite this increase, S2's results for Type 2a sentences show a clear decline which, together with the fluctuation exhibited in the later stages of her collection period, point to attrition of the restriction in English that antecedents in biclausal sentences be locally bound. The results from test subject S2's Type 2b sentences also show distinct signs of attrition. From her return to Japan to TSR 4.1, her average for this sentence type is 83%. During her secondary sub-period, to TSR 8.1, her average declines to 56%, and drops even further to 42% during her final data collection sub-period. This increased level of acceptance of illegal long distance binding, coupled with the increased level of rejection of locally bound antecedents observed in Type 2a sentences, indicates an erosion of knowledge of English reflexive binding rules in biclausal sentences with nonfinite embedded clauses.

S2's results on sentence Types 3a and 3b both exhibit percentages of loss similar to those found in her test-mates' results. S2's average in her initial sub-period is 56% which is comparable to the other test subjects' averages (except S1's) on this sentence type over similar time periods. S2's results on Type 3a sentences improve somewhat over the next several months, up to TSR 8.1, during which time her results average 67%. The final 5 months clearly show attrition of the principles responsible for accurately judging Type 3a sentences, with S2's results averaging a mere 25%. This last average is a clear indication of test subject S2's increased failure to recognize the acceptability of the local object NP as an antecedent in monoclausal sentences in English. S2's Type 3b

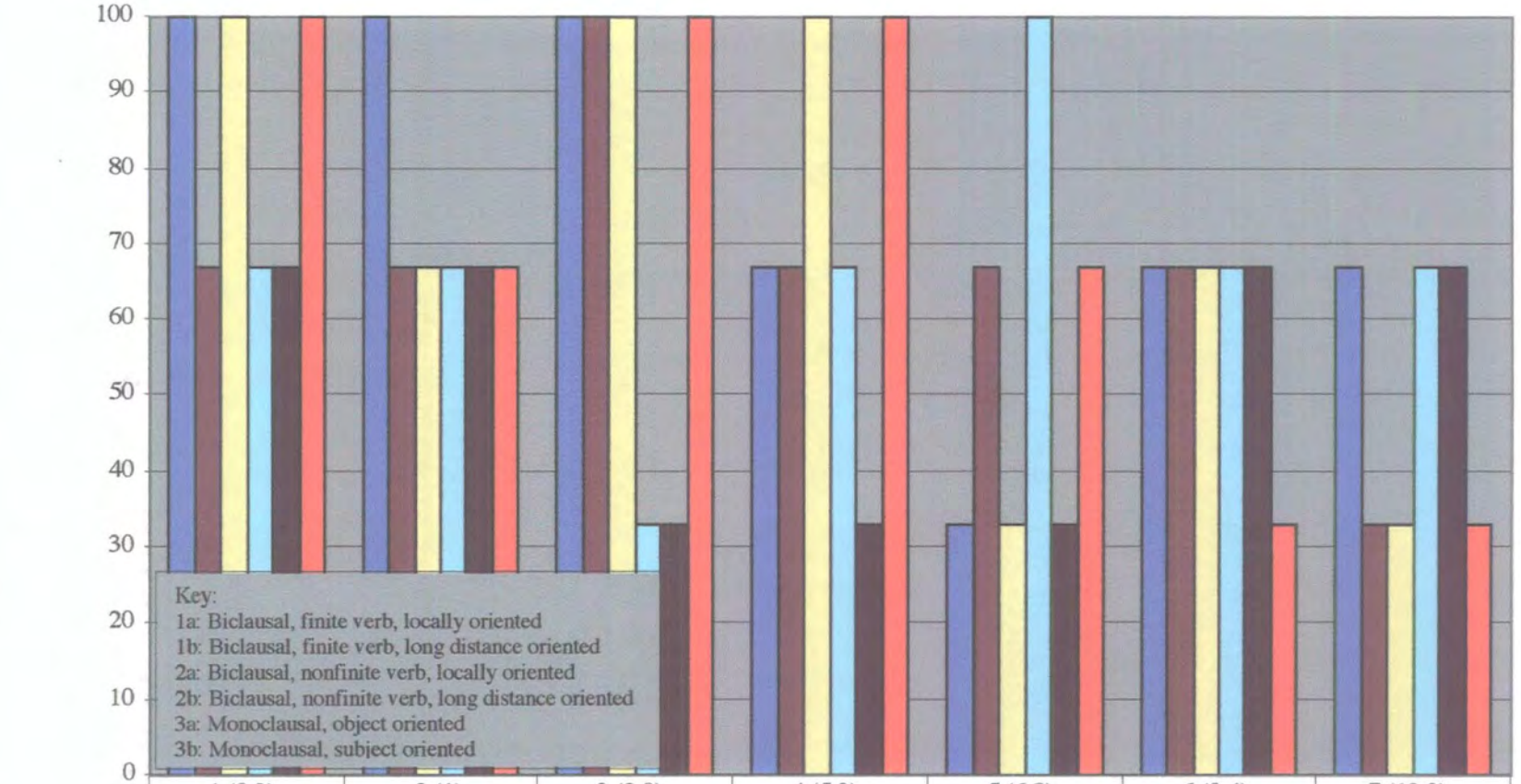
sentence results also show attrition, though not as severe as her Type 3a sentences. Over the first 4 months, S2 averages 83% on this sentence type. Over her secondary sub-period, to TSR 8.1, her average drops to 67%. She then maintains this average of 67% throughout the remainder of her data collection period. S2's TV Type 3b results reflect a reluctance for her to accept subject NPs in monoclausal sentences as the antecedent. This is surprising as the attrition observed in her 3a sentences would predict she has a strong propensity to choose subject NPs as antecedents.

S2's TV scores by type are included in Appendix I.

6.2.1.1.6. *Test subject S3's TV test results by sentence type over time*

The final set of truth value test results are test subject S3's. Although S3 had only seven data collection sessions, her data collection period lasted slightly more than 12 months. Because of the two and a half month gap TSR 2.9 and TSR 5.3, her initial data collection sub-period will only be about three months long, to TSR 2.9. Her secondary sub-period is 3.9 months long, to TSR 6.8, and her final sub-period lasts five and a half months, to TSR 12.2. In her sub-period, S3's combined results average 78%, somewhat below the other test subjects' results averages for this sub-period. S3's results subsequently decline and average 66% for the secondary sub-period, to TSR 6.8, and continue at this approximate level during the final sub-period, averaging 64%. These overall results indicate S3's knowledge of English reflexive binding is initially lower than the other test subjects', but that it attrites less than does R3's and S2's whose results have greater differences between their initial and final sub-periods (S3 with a 14% difference, but R3 with a 21% difference and S2 with a 25% difference)

S3's TV test results by sentence type are depicted in Figure 6.8.




| | 1 (0.3) | 2 (1) | 3 (2.9) | 4 (5.3) | 5 (6.8) | 6 (9.4) | 7 (12.2) |
|---|---|---|---|---|---|---|---|
| 1a Object | 100 | 100 | 100 | 67 | 33 | 67 | 67 |
| 1b Subject | 67 | 67 | 100 | 67 | 67 | 67 | 33 |
| 2a Object | 100 | 67 | 100 | 100 | 33 | 67 | 33 |
| 2b Subject | 67 | 67 | 33 | 67 | 100 | 67 | 67 |
| 3a Object | 67 | 67 | 33 | 33 | 33 | 67 | 67 |
| 3b Subject | 100 | 67 | 100 | 100 | 67 | 33 | 33 |

TV scores by sentence type. Top row of grid indicates test number, TSR in parentheses

Figure 6.8. S3 truth value test scores over time by sentence type

By sentence type, S3's TV 1a sentence scores decrease dramatically over time, averaging 100% between TSR 0 and TSR 2.9, but only 50% over the next almost 4 months to TSR 6.8. S3's 1a scores then increase slightly to 67% during her final sub-period. The drop and fluctuation in TV Type 1a scores reflect S3's increasing unwillingness to accept the local antecedent as correct. Her 1b scores decline steadily over the collection period, averaging 78% up to TSR 2.9, then 67% up to TSR 6.8, and finally averaging 50% by the final TSR, 12.2. This drop in Type 1b sentences reflects S3's increasing tendency to incorrectly accept the long distance bound antecedent in biclausal sentences with a finite verb in the embedded clause. Taken together, S3's attrition on Type 1a and 1b sentences indicates an increasing difficulty on her part to identify the correct antecedent of a reflexive pronoun in the finite embedded clauses of these two sentence types.

S3's TV Type 2a and 2b results show both decline and fluctuation throughout her data collection period. From her return to Japan to TSR 2.9, S3's TV Type 2a sentences average 89%. The averages for this sentence type drop dramatically to 67% by TSR 6.8, and then to 50% during her final sub-period. The steady decline of S3's results on the TV Type 2a sentences show an increasing reluctance to allow local object NPs as antecedents in biclausal sentences with non-finite verbs. On her TV Type 2b sentences, test subject S3 shows some increase in disallowing subject NPs as antecedents in biclausal sentences of this type. Although initially averaging only 55% from her return to TSR 2.9, S3's Type 2b results show a strong increase to an average of 83% during the following period TSR 2.9 to TSR 6.8. The average for this sentence type over the final 5.5 months of the collection period declines to 67% but corresponds closely to R1's and R2's final averages on this type which are 66% and 67% respectively. These results seem to reflect a growing certainty of the unacceptability of long distance antecedents in English up to TSR 6.8, but this certainty seems to lessen during the final sub-period.

S3's Type 3a sentence results follow a pattern similar to the other test subjects' except S2's. This pattern starts with an average (56%) somewhat lower than the NSs' 73% average for the sentence type, then declines over the next four or so months, finally increasing again during the final sub-period of the collection period. In S3's case, her results for the 3a sentences start with an average of 55% up to TSR 2.9, decline to 33% over the next period up to TSR 9.4, and increase to 67% from TSR 6.8 to TSR 12.2, the final part of her data collection period. S3's scores on her Type 3b sentences exhibit

signs of attrition. Her initial results average 89%, an average for this sentence type higher than any informant's, including the NSs'. This average drops to 78% over the next 5.5 months, an average higher than any other test subject for this sentence type over this time period, and again higher than the average of the NSs' results. S3's results for her final sub-period, from TSR 6.8 to TSR 12.2 average 67%, comparable with all the other test subjects'. These TV Type 3 results indicate test subject S3 has a preference for subject NPs as antecedents which can be seen in her rejection of object NPs in her 3a results, and her strong acceptance of subject NPs on Type 3b sentences.

S3's TV scores by type are included in Appendix I.

This concludes the analysis of the Japanese test subjects' truth value judgment test results. This analysis represents the consideration of 62 truth value tests which include 1,116 test items. Although the results are somewhat mixed, I do believe they provide evidence of the attrition of reflexive binding in the L2 English of the test subjects included in this study, especially in biclausal sentences having nonfinite verbs in the embedded clauses. This will be discussed at the end of this chapter in section 6.3. In the next section I will present the results from the grammaticality judgement test.

### 6.2.2. *The test subjects' grammaticality judgement test results*

The role of the grammaticality judgement test in this study is primarily to confirm the findings of the truth value test, which they do. The trends revealed in the truth value test such as overall and sentence-type-specific degree of adherence to principles of reflexive binding over time are also found in the results of the GJ test. Further, no wildly divergent results between the TV test results and the GJ test results are observed. Both tests confirm near native-like knowledge of most aspects of Principle A binding early on in the data collection period, and then show some propensity to digress from this knowledge later in the collection period. The GJ test also provides additional though limited data about the test subjects' knowledge of reflexive binding in English. Each test subject's GJ test scores can be found in Appendix K, and results by sentence type can be seen in Appendix L. An item analysis of all the GJ test items, including bare infinitive distractors, can be found in Appendix M.

Table 6.6 includes each test subject's score on each GJ test session she took, along with her TSR (the time elapsed at each test session since the test subject returned home to Japan). The analysis will proceed in the same way as the TV test results, that is, the results of the six test subjects on their GJ test over their data collection period are considered in terms of degree of adherence to the principles of reflexive binding in English. Each item on each GJ test session was awarded 1 point, with a total of 11 points per test session session. In order to compare the overall GJ test session results to the TV results, the TV results seen in Table 6.5. have been included in Table 6.6.

A comparison of the GJ scores to the TV scores in Table 6.6 indicates that the GJ scores are generally higher than the TV scores, but that both groups of scores decline over time.

*Table 6.6. TV and GJ scores at months since L1 re-immersion*

| Test Subject/Test Session # | R1 | | | | R2 | | | | R3 | | | | S1 | | | | S2 | | | | S3 | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | TV Score | GJ Scores | TSR | Date | TV Score | GJ Scores | TSR | Date | TV Score | GJ Scores | TSR | Date | TV Score | GJ Scores | TSR | Date | TV Score | GJ Scores | TSR | Date | TV Score | GJ Scores | TSR | Date |
| (Return date) | na | na | 0 | 7/1/97 | na | na | 0 | 6/20/97 | na | na | 0 | 8/28/97 | na | na | 0 | 3/19/97 | na | na | 0 | 9/15/96 | na | na | 0 | 9/15/97 |
| #1 | 89% | 100% | 0.7 | 7/11/97 | 83% | 100% | 0.6 | 7/8/97 | 89% | 91% | 1 | 9/29/97 | 83% | 91% | 0.5 | 4/4/97 | 83% | 82% | 0.4 | 9/27/96 | 83% | 91% | 0.3 | 9/29/97 |
| #2 | 89% | 100% | 1.2 | 7/25/97 | 89% | 100% | 1.4 | 8/1/97 | 78% | 100% | 1.8 | 10/20/97 | 89% | 100% | 1 | 4/18/97 | 78% | 91% | 0.8 | 10/11/96 | 72% | 82% | 1 | 10/20/97 |
| #3 | 83% | 100% | 3.3 | 9/26/97 | 78% | 100% | 3.1 | 9/23/97 | 83% | 82% | 2.7 | 11/17/97 | 83% | 100% | 1.9 | 5/16/97 | 78% | 91% | 1.5 | 11/1/96 | 78% | 82% | 2.9 | 12/15/97 |
| #4 | 94% | 91% | 4 | 10/17/97 | 83% | 100% | 3.8 | 10/14/97 | 78% | 91% | 3.6 | 12/15/97 | 83% | 91% | 2.6 | 6/6/97 | 89% | 91% | 2.3 | 11/22/96 | 72% | 73% | 5.3 | 2/27/98 |
| #5 | 83% | 100% | 5.2 | 11/21/97 | 83% | 100% | 5 | 11/18/97 | 67% | 82% | 4.8 | 1/20/98 | 72% | 91% | 3.3 | 6/27/97 | 83% | 73% | 3 | 12/13/96 | 61% | 73% | 6.8 | 4/13/98 |
| #6 | 83% | 91% | 6 | 12/15/97 | 78% | 91% | 6 | 12/16/97 | 67% | 73% | 6 | 2/23/98 | 78% | 73% | 3.9 | 7/15/97 | 83% | 73% | 4.1 | 1/17/97 | 67% | 64% | 9.4 | 6/30/98 |
| #7 | 78% | 91% | 7.3 | 1/23/98 | 78% | 91% | 7.1 | 1/20/98 | 61% | 82% | 7.6 | 4/13/98 | 78% | 91% | 6.3 | 9/26/97 | 78% | 82% | 5.3 | 2/21/97 | 61% | 73% | 12.2 | 9/21/98 |
| #8 | 78% | 100% | 9 | 3/15/98 | 72% | 82% | 8.3 | 2/24/98 | 67% | 64% | 9.2 | 6/1/98 | 89% | 91% | 7 | 10/17/97 | 72% | 73% | 6.7 | 4/4/97 | na | na | na | na |
| #9 | 78% | 91% | 11.5 | 5/29/98 | 78% | 82% | 10 | 4/14/98 | 56% | 73% | 10.9 | 7/20/98 | 78% | 91% | 9 | 12/13/97 | 56% | 73% | 8.1 | 5/16/97 | na | na | na | na |
| #10 | 83% | 82% | 12.9 | 7/10/98 | 83% | 82% | 12 | 6/16/98 | 61% | 55% | 13 | 9/22/98 | na | na | na | na | 67% | 64% | 9.5 | 6/27/97 | na | na | na | na |
| #11 | 72% | 82% | 15.5 | 9/25/98 | na | na | na | na | 61% | 64% | 14.1 | 10/26/98 | na | na | na | na | 50% | 55% | 10.7 | 8/1/97 | na | na | na | na |
| #12 | 78% | 91% | 16.4 | 10/23/98 | na | na | na | na | na | na | na | na | na | na | na | na | 56% | 55% | 12.5 | 9/26/97 | na | na | na | na |
| #13 | na | na | na | na | na | na | na | na | na | na | na | na | na | na | na | na | 56% | 55% | 13.2 | 10/17/97 | na | na | na | na |

Initial sub-period  Secondary sub-period  Final sub-period

Following Table 6.6, Figure 6.9 provides a general illustration of the range of individual test session subject scores on the GJ test session over time as outlined in Table 6.6. In Figure 6.9, initial scores can be seen to cluster between 80 and 100 percent. As with the TV test results illustrated in Figure 6.1 above and reproduced below as Figure 6.10, the GJ test results in Figure 6.9 can be seen to widen over time. The TV test results in Figure 6.1 are reproduced below Figure 6.9 for comparison as Figure 6.10.

*Figure 6.9. Test Subjects' individual GJ scores (%) over time*

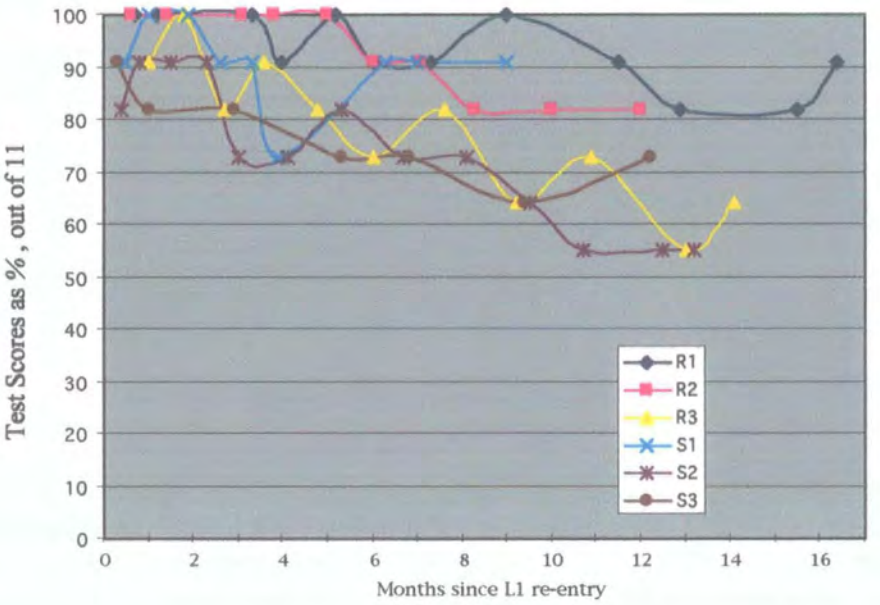
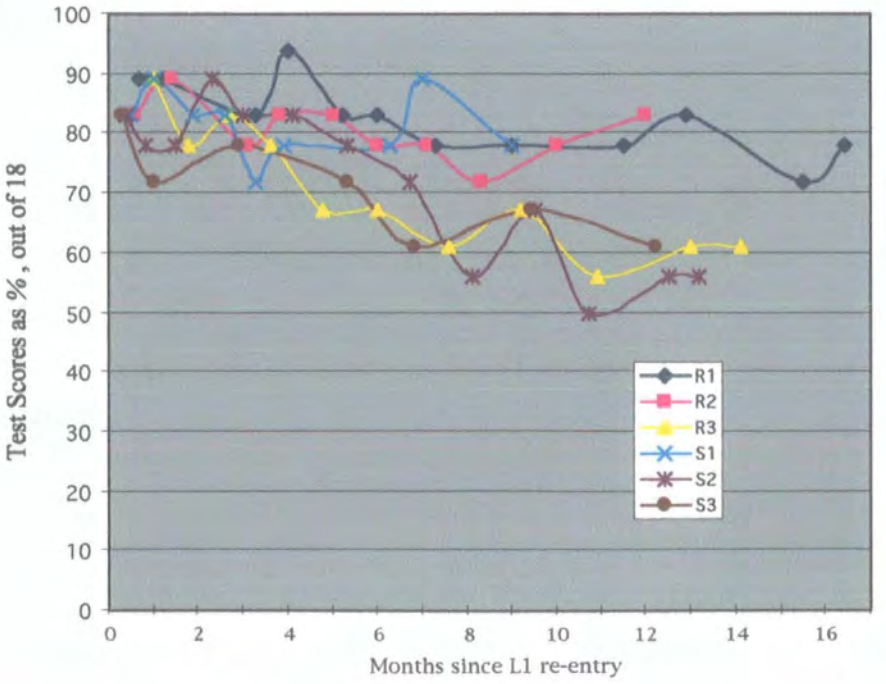



*Figure 6.10. Test Subjects' individual TV scores (%) over time*



By comparing the lines in Figures 6.9 and 6.10, we can see that R1, R2 and S1 again tend to maintain the levels initially obtained in both tests, and that the levels of R3, S2 and S3 decline over time. In the next section I will consider the individual results of the test subjects' GJ tests.

### 6.2.2.1. *Individual GJ test results by Type over time*

In this section, each of the six test subjects' scores on the GJ sentence types over time will be described. These descriptions will identify the degree to which particular sentence types are susceptible to attrition in each individual test subject.

Appendix L lists the test subjects' scores for each GJ sentence type over time. GJ test session numbers GJ1 through GJ13 are listed across. The sentence types are listed from top to bottom and are grouped into three main sections, each section corresponding to the broad sentence types biclausal finite, biclausal nonfinite, and monoclausal. Because of the limited number of type tokens per test session (one), the eleven GJ sentence types will generally be treated in these three groups. Only where the results of a particular sub-type of a broad sentence type show a divergent pattern will a sub-type be addressed. Also, as with the TV test results, GJ test results will be measured using the TSR numbers. The GJ sentence types are discussed in Chapter 5 (section 5.3.2.3.) and are summarized below.

*Grammaticality judgement sentence Types*

**Type 1 sentences**: biclausal, with a finite verb in the embedded clause:

*GJ Type 1a:* Biclausal sentences, finite verbs in embedded clauses. Anaphor and local antecedent number agreement. Correct answer is ***OK***.

*GJ Type 1b:* Biclausal sentences, finite verbs in embedded clauses. Anaphor and local antecedent number disagreement. Correct answer is ***Wrong***.

*GJ Type 1c sentences:* Biclausal sentences, finite verbs in embedded clauses. Anaphor and local antecedent gender agreement. Correct answer is ***OK***

*GJ Type 1d sentences:* Biclausal sentences, finite verbs in embedded clauses. Anaphor and local antecedent gender disagreement. Correct answer is ***Wrong***

**Type 2 sentences**: biclausal, with a nonfinite verb in the embedded clause:

*GJ Type 2a sentences:* Biclausal sentences, non-finite verbs in embedded clauses. Anaphor and local antecedent number agreement. Correct answer is ***OK.***

*GJ Type 2b sentences:* Biclausal sentences, non-finite verbs in embedded clauses. Anaphor and local antecedent number disagreement. Correct answer is ***Wrong***.

*GJ Type 2c sentences:* Biclausal sentences, non-finite verbs in embedded clauses. Anaphor and local antecedent gender agreement. Correct answer is ***OK***

*GJ Type 2d sentences:* Biclausal sentences, non-finite verbs in embedded clauses. Anaphor and local antecedent gender disagreement. Correct answer is ***Wrong***

**Type 3 sentences**: monoclausal sentences.

*GJ Type 3a sentences:* Monoclausal sentences with 2 NPs. Pragmatically favors object antecedent. Correct answer is ***OK***

*GJ Type 3b sentences:* Monoclausal sentences with 2 NPs. Linguistically favors subject antecedent. Correct answer is ***OK***

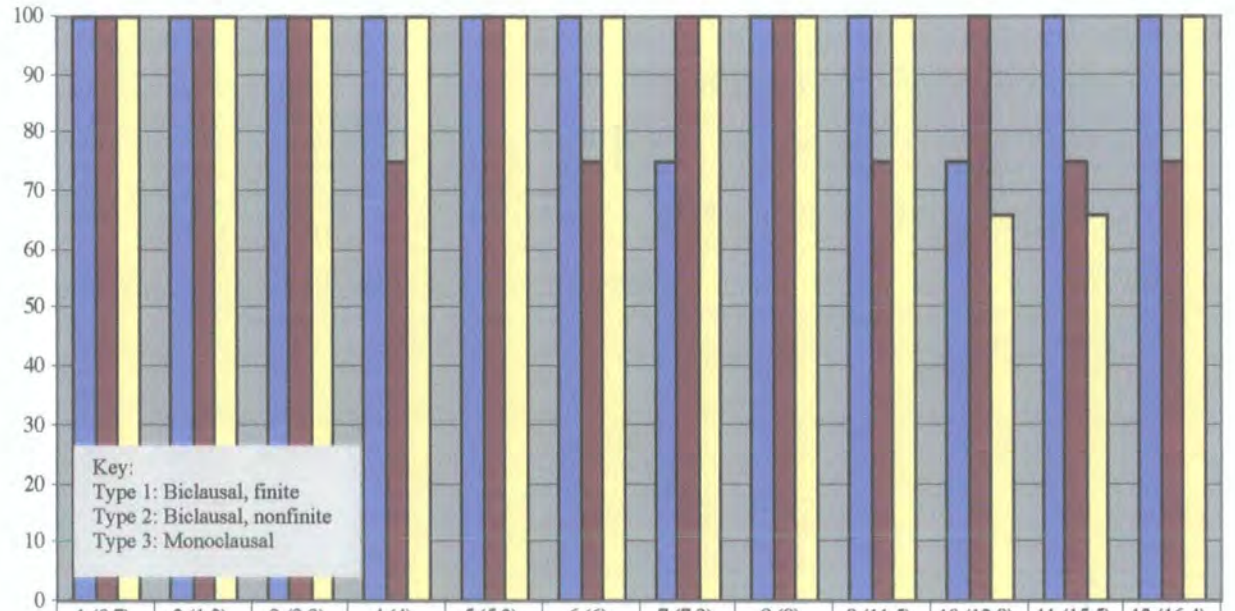
*GJ Type 3c sentences:* Monoclausal sentences with 2 NPs. Pragmatically ambiguous antecedent. Correct answer is ***OK***

As described in section 5.3.2.3, each GJ test session includes one token of each GJ sentence type. The total number of test type tokens for each test subject varies because test subjects varied in the number of test sessions they took. Except at some initial test session early in the data collection, all test subjects took both the TV test session and the GJ test session at each test session they attended.

6.2.2.1.1. *Test subject R1's GJ test results by sentence type over time*

Test subject R1's data collection period lasted just over sixteen and a half months. Up to TSR 4 her overall results average 98% and exhibit a high level of consistent accuracy across all GJ sentence types. These strong initial results indicate a good knowledge of the principles underlying reflexive binding in English. Despite signs of instability from TSR 4.4 onwards R1's results remain high up through TSR 9. From TSR 10 though, we can see increasing signs of instability and attrition of her use of reflexives. R1's GJ results are illustrated in Figure 6.11 and are included in Appendix L.

*Figure 6.11. R1 grammaticality judgement test scores over time by sentence type*




| | 1 (0.7) | 2 (1.2) | 3 (3.3) | 4 (4) | 5 (5.2) | 6 (6) | 7 (7.3) | 8 (9) | 9 (11.5) | 10 (12.9) | 11 (15.5) | 12 (16.4) |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Type 1 | 100 | 100 | 100 | 100 | 100 | 100 | 75 | 100 | 100 | 75 | 100 | 100 |
| Type 2 | 100 | 100 | 100 | 75 | 100 | 75 | 100 | 100 | 75 | 100 | 75 | 75 |
| Type 3 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 66 | 66 | 100 |

GJ scores by sentence type. Top row of grid indicates test number, TSR in parentheses

By sentence type, test subject R1 only misjudges two Type 1 sentences out of 48 possible, both misjudged sentences being Type 1b sentences. In GJ Type 1b sentences the number disagrees between the reflexive and the local antecedent. This corresponds with the native speaker results in which the lowest GJ Type 1 mean score was on the Type 1b sentences. R1's GJ Type 1 results also closely resemble the strong results of her TV Type 1 sentence results. Taken together, her GJ Type 1 results, and her TV Type 1a and 1b results reflect R1's high degree of long-term continued acceptance of local antecedents, and rejection of long distance antecedents in biclausal sentences with a finite verb in the embedded clause.

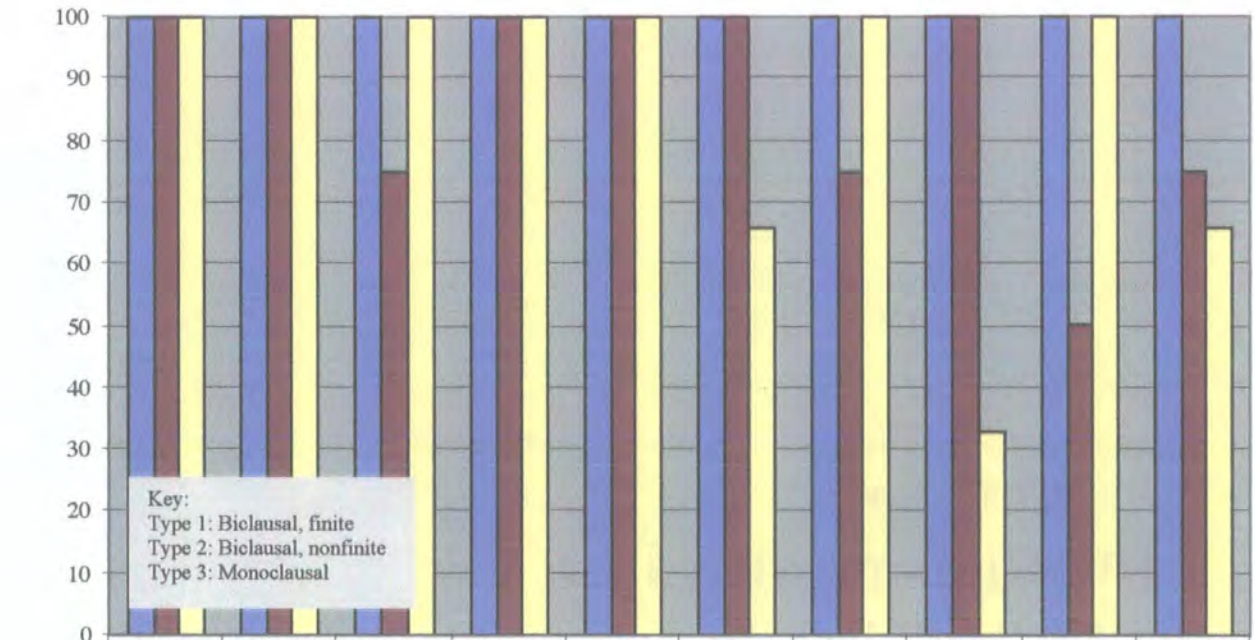
R1's results on the GJ Type 2 sentences are also strong, but show some signs of attrition. Up to TSR 3.3 R1's Type 2 results average 100%. From TSR 4.0. however, the results become less stable as she appears to become less sure of the principles underlying reflexive binding in sentences with non-finite embedded verbs. From TSR 4.0 to TSR 8, her results average 90% on the GJ Type 2 sentences, and from TSR 11.5 to the end of the collection period, R1 averages just 81% on the GJ Type 2 sentences. This trend of greater instability in sentences with nonfinite embedded verbs than in sentences with finite verbs is also evident in the TV test results and will be addressed in the discussion section below.

R1's Type 3 sentence results reflect a high level of accuracy in judging the acceptability of antecedents in these monoclausal sentences for almost a year after returning to Japan. R1's results on these sentences average 100% up through TSR 11.5. During the last three and a half months however, her results become unstable and her average on this sentence type declines to 77%. This trend is quite different from the lower averages and greater instability R1 exhibits throughout her TV Type 3a and 3b sentences results. This difference may be indicative of a preference for the GJ test over the TV test, the short scenarios presented in the TV test items interfering with the test subject's judgement of the antecedent.

6.2.2.1.2. *Test subject R2's GJ test results by sentence type over time*

Test subject R2's data collection period lasted 12 months. Up to TSR 3.8 she maintains an overall average of 100% on her GJ test across all sentence types. This initial high average corresponds with the high average seen on her TV test up for this time period. And as with her TV results, her GJ results decline slightly, but remain strong over the secondary period from TSR 5 to TSR 8.3. The overall GJ average over these three and a half months is 91%. Though still strong, during the final 4 months of the collection period R2's overall GJ results decline somewhat further, averaging 82%. Although her TV results increase slightly over this final period, her results on both tests become less stable as the collection period proceeds. R2's results are illustrated in Figure 6.12.

*Figure 6.12. R2 grammaticality judgement test scores over time by sentence type*




| | 1 (0.6) | 2 (1.4) | 3 (3.1) | 4 (3.8) | 5 (5) | 6 (6) | 7 (7.1) | 8 (8.3) | 9 (10) | 10 (12) |
|---|---|---|---|---|---|---|---|---|---|---|
| Type 1 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 |
| Type 2 | 100 | 100 | 75 | 100 | 100 | 100 | 75 | 100 | 50 | 75 |
| Type 3 | 100 | 100 | 100 | 100 | 100 | 66 | 100 | 33 | 100 | 66 |

TV scores by sentence type. Top row of grid indicates test number, TSR in parenthes

By sentence type, test subject R2 averages 100% on her GJ Type 1 sentences throughout her data collection period. This corresponds to the strong results on her TV test for this sentence type and further indicates her high degree of long-term continued acceptance of local antecedents, and rejection of long distance antecedents in biclausal sentences with a finite verb in the embedded clause.

Though not as consistent as her Type 1 results, R2's results on her GJ Type 2 sentences also reflect a strong knowledge of the principles underlying the choice of antecedent for this sentence type. From her return to Japan up to TSR 3.8, R2 averages 94% on GJ Type 2 sentences. Her 100% average on the TV Type 2a sentences reflect a very strong initial knowledge of the principles constraining binding in these sentence types. Her 83% average on the TV Type 2b sentences reflect an imperfect, but still strong initial knowledge of the principles constraining binding in these sentence types. From TSR 5, to TSR 8.3, R2's Type 2 results continue to average a stable 94% and compare favorably with her 83% average for the same period on both her TV Types 2a and 2b sentences. During her final sub-period from TSR 8.3 to TSR 12, test subject R2's GJ Type 2 results become unstable and decline sharply to 62%. This decline is due to low scores on her GJ Type 2b sentences and especially her 0% average on her GJ Type 2d sentences during this final sub-period. These two sentence types require test subjects to reject long distance (subject) NPs as the antecedent in sentences with non-finite verbs in the embedded clause. However, R2 tends to illegally accept these long distance antecedents here and also on her TV test. On the TV Type 2b sentences, test subject R2 declines from a consistent 83% over the first 8 months to an average of 67% over the final 4 months. R2's trend to accept long distance antecedents in biclausal sentences with non-finite verbs in the embedded clauses can thus be seen in the results of both her GJ and TV tests.

R2's GJ Type 3 sentence results are consistently strong up through TSR 3.8, averaging 100%. These results reflect a high level of accuracy in judging the acceptability of antecedents in these monoclausal sentences on the GJ test. Compared with these results however, R2 exhibits lower levels of accuracy on her TV tests for these monoclausal sentence types. R2's results for the same time period average only 63% for TV Types 3a and 3b combined. This trend of TV results being lower than GJ results was seen in R1 above and in fact can be seen in all test informants, both the Japanese test subjects and the native speaker control group. This tendency may indicate

that the GJ test items are easier to process and correctly judge than the TV test items. From her initial 100% average on the Type 3 results, R2's GJ Type 3 results show signs of instability by dropping 25% to an average of 75% during this secondary period. This instability continues in the final 4 months of R2's collection period during which she averages 83%. R2's TV Type 3 results also show signs of attrition in the later stages of the data collection period, averaging only 50% over the final four months. On both the GJ and TV tests, the declining averages on the Type 3 sentences results from a trend to increasingly reject object antecedents in these monoclausal sentences (TV Type 3a and GJ Type 3a). This trend may stem from the test informant increasingly approximating her L1 Japanese binding constraints, which do not allow binding to object NPs.

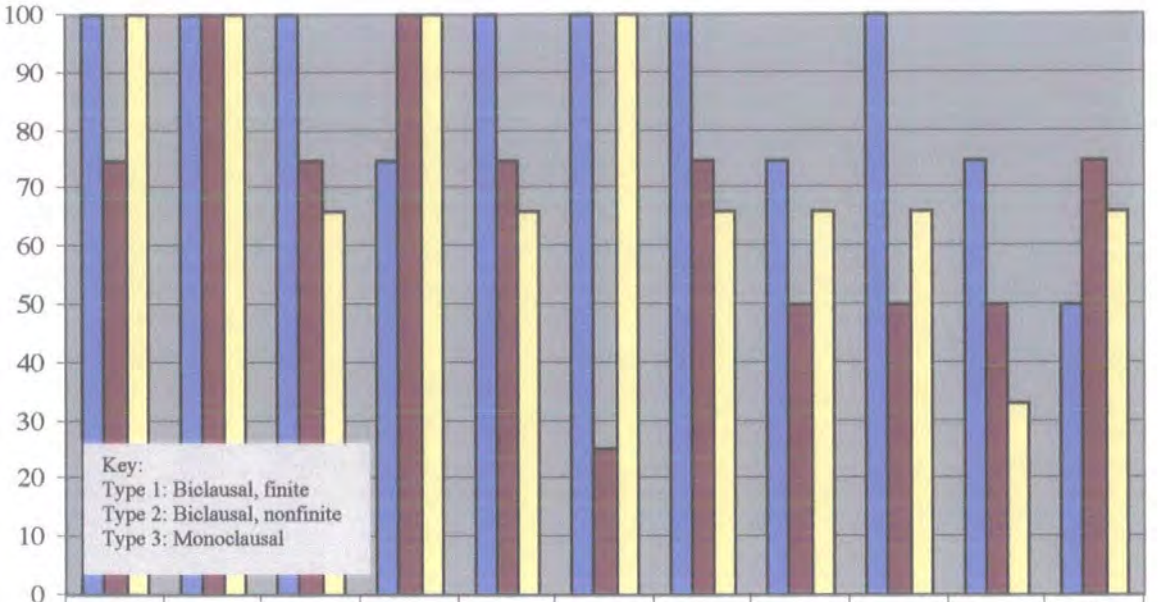
Test subject R2's GJ results by type are included in Appendix L

6.2.2.1.3. *Test subject R3's GJ test results by sentence type over time*

Test subject R3's data collection period included 10 test sessions and lasted just over 14 months, the second longest collection period of the test subjects after R1. R3's overall GJ test results indicate her knowledge of general binding constraints in English is initially strong, but undergoes attrition during the course of her data collection period. From her return to Japan to TSR 3.6 her combined average on the GJ test is 91%. This compares favorably with the overall average of her TV test results for the same period, 82%. After TSR 3.6 however, R3's overall GJ average declines sharply to 79%. This decline is again mirrored in her overall TV test results which drop 17% to 65% for the same period. This correspondence continues for the final phase of the collection period, TSR 7.6 to TSR 14.1, during which her overall GJ results average just 61% and her TV results drop a further 15% to 64%.

Test subject R3's GJ scores are depicted graphically in Figure 6.13 and are included in Appendix L.

*Figure 6.13. R3 grammaticality judgement test scores over time by sentence type*




| | 1 (1) | 2 (1.8) | 3 (2.7) | 4 (3.6) | 5 (4.8) | 6 (6) | 7 (7.6) | 8 (9.2) | 9 (10.9) | 10 (13) | 11 (14.1) |
|---|---|---|---|---|---|---|---|---|---|---|---|
| Type 1 | 100 | 100 | 100 | 75 | 100 | 100 | 100 | 75 | 100 | 75 | 50 |
| Type 2 | 75 | 100 | 75 | 100 | 75 | 25 | 75 | 50 | 50 | 50 | 75 |
| Type 3 | 100 | 100 | 66 | 100 | 66 | 100 | 66 | 66 | 66 | 33 | 66 |

TV scores by sentence type. Top row of grid indicates test number, TSR in parentheses

By sentence type, test subject R3's GJ Type 1 results show signs of attrition during her collection period. From her return to Japan to TSR 3.6, R3's results on the Type 1 sentences average 94%. This average increases to 100% during the secondary sub-period of her collection period, from TSR 3.6 to TSR 7.6, but declines to 75% during the final sub-period, from TSR 7.6 to TSR 14.1. R3's TV results for this sentence type over the same periods also show signs of instability and attrition. The combined averages for the TV Type 1a and 1b sentences are 91%, 82% and 71% for R3's the initial, secondary and final sub-periods respectively. A closer look at both her TV and GJ results indicates she has increasing trouble rejecting long distance subject antecedents in the biclausal sentences included in this type. On both the TV and GJ tests, R3 increasingly misjudges Type 1b sentences and allows long distance binding to subject NPs. This tendency seems to approximate her L1 Japanese binding constraints which allow long distance binding to subject NPs.

Test subject R3's results on her GJ Type 2 sentences show instability from the beginning of her data collection period, and attrition as the test continues over 14 months. During the initial sub-period of data collection, the results of R3's GJ Type 2 sentences average 88%. Only Type 2d sentences are misjudged at this initial stage which again indicates a tendency on R3's part to prefer her Japanese L1-like long distance binding to subject NPs. Throughout the remainder of her data acquisition period, test subject R3's GJ Type 2 results decline sharply to 58% during the secondary sub-period, TSR 3.6 to TSR 7.6 and 56% during the final sub-period, TSR 7.6 to TSR 14.1. Although she misjudges tokens from all 4 types of GJ Type 2 sentences, R3 increasingly accepts long distance antecedents. This supports the evidence of attrition found in R3's TV Type 2 sentence results which also indicate she increasingly accepts illegal long distance binding and rejects locally bound antecedents, preferring what appears to be her L1 Japanese acceptability of long distance binding.

R3's GJ Type 3 sentences also resemble her TV Type 3 scores in that she maintains accuracy in judging subject NPs as the correct antecedent, but increasingly fails to accept object NPs. Up to TSR 3.6 the results of R3's GJ Type 3 sentences average 92%. This declines to 77% during the next 4 months (up to TSR 7.6), dropping even further to 58% over the final six and a half months (to TSR 14.1). This decline is primarily due to misjudged GJ Type 3a sentences in which she increasingly rejects the object antecedent in these monoclausal sentences. And as with the TV sentences, some attrition is also observed in the Type 3b sentences, presumably due to her preference for subject NPs as antecedents.

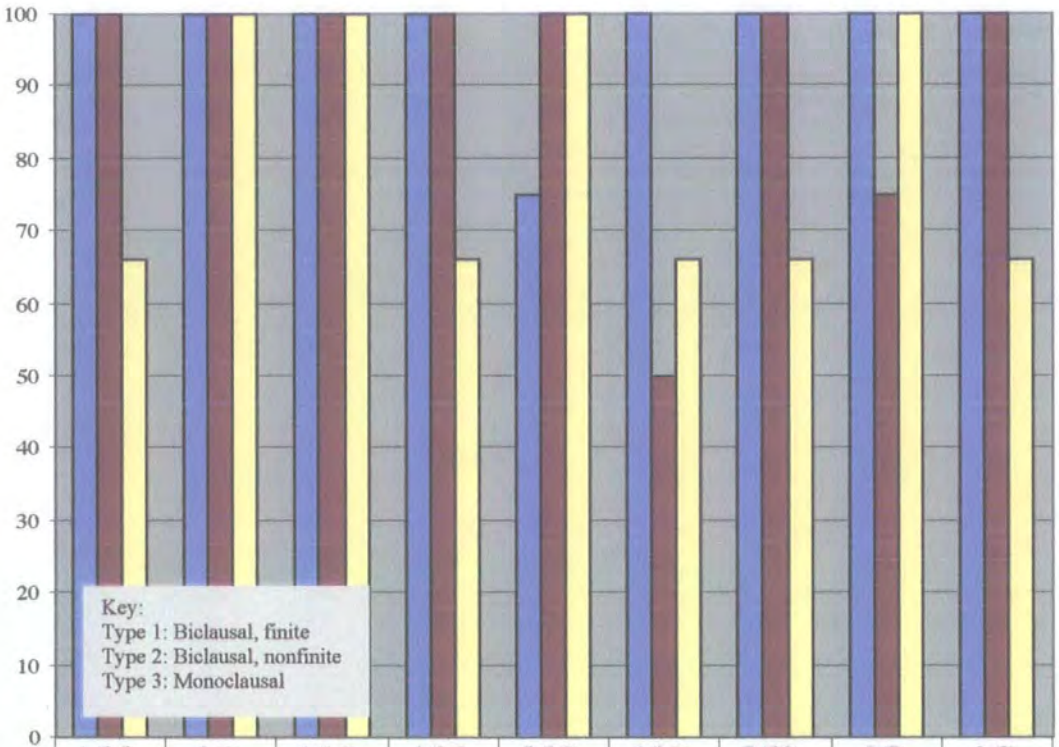
6.2.2.1.4. *Test subject S1's GJ test results by sentence type over time*

Test subject S1's data collection period lasted 9 months and included 9 test sessions. Her data collection period is considered in two sections, the 4 months from TSR 0 to TSR 3.9, and the 5 months from TSR 3.9 to TSR 9. S1's overall GJ results show a high level of maintenance of English reflexive binding principles. During both the initial 4 months of her data collection period and the subsequent 5 months, S1's GJ results average 91%, just 6% below the NS average of 97% on this test. This degree of maintenance is also reflected in S1's TV results which average 81% and 82% over the

same time periods. Test subjects R1 and R2's results also show similar levels of maintenance for both tests over similar periods of time.

Test subject S1's GJ scores by type are depicted graphically in Figure 6.14 and are included in Appendix L.

*Figure 6.14. S1 grammaticality judgement test scores over time by sentence type*




| | 1 (0.5) | 2 (1) | 3 (1.9) | 4 (2.6) | 5 (3.3) | 6 (3.9) | 7 (6.3) | 8 (7) | 9 (9) |
|---|---|---|---|---|---|---|---|---|---|
| Type 1 | 100 | 100 | 100 | 100 | 75 | 100 | 100 | 100 | 100 |
| Type 2 | 100 | 100 | 100 | 100 | 100 | 50 | 100 | 75 | 100 |
| Type 3 | 66 | 100 | 100 | 66 | 100 | 66 | 66 | 100 | 66 |

GJ scores by sentence type. Top row of grid indicates test number, TSR in parentheses

By sentence type, S1's GJ Type 1 sentences results are very stable throughout her data collection period, averaging 96% up to TSR 3.9, and 100% for the rest of her collection period, to TSR 9. These high averages are also seen in her TV test results and together these two test results point to a robust knowledge of the principles limiting the anaphoric binding of antecedents to local NPs in these biclausal sentences.

S1's results on Type 2 sentences average 92% for her entire data collection period (TSR 0 to 3.9 and TSR 3.9 to TSR 9). Despite these high averages, some signs of instability can be seen developing as S1 misjudges a GJ Type 2a and 2d at TSR 3.9 and another 2d and TSR 7. This slight instability is seen to a greater degree in her TV test results in which she initially averages 100% up to TSR 1.0, but subsequently declines, averaging 79% during her initial sub-period, and 77% during her final sub-period. The results presented here point to a greater degree of attrition in both the TV and GJ tests of the principles underlying the choice of antecedent in biclausal sentences with a non-finite verb in the embedded clause (Type 2 sentences), than in the Type 1 sentences which have finite verbs in the embedded clauses.

Test subject S1's GJ Type 3 results become less stable later in her data collection period, initially averaging 83%, and then declining to 77% during her secondary sub-period. Up to TSR 3.3, S1 only misjudges (i.e., judges as *false*) the ambiguous GJ Type 3c sentences, but from TSR 3.9 onwards consistently judges them correctly. This is of interest in that we would consider 3c the easiest of all the GJ sentence types as either the subject NP or the object NP can correctly serve as the antecedent of the reflexive. Of further interest is that 'R' test subjects never exhibit this trend, only misjudging the GJ Type 3c sentences later in their collection periods. This trend may indicate some level of 'S' subject reliance on the pragmatics of a sentence to choose the antecedent, rather than solely on the principles of binding which allow or prevent a particular NP from serving as the antecedent. The overall pattern this portrays of an initially unprincipled judgement process gradually becoming more principled may also be reflected in test subject S1's TV test results which increase from TSR 3.9 to the final months of her data collection period.

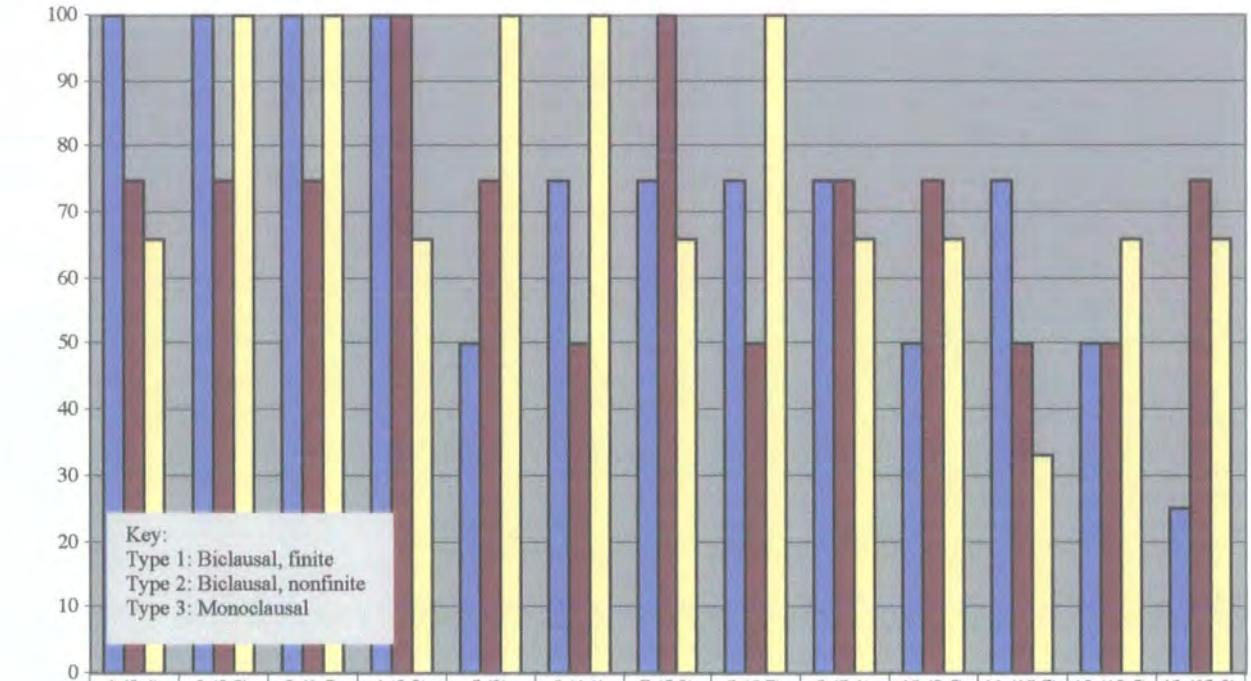
6.2.2.1.5. *Test subject S2's GJ test results by sentence type over time*

As with her TV test results, test subject S2's GJ test results reflect the most pronounced attrition of the six test subjects. Her data collection period lasted just over 13 months and has been divided into three parts, from TSR 0 to 4.1, then to TSR 8.1, and finally to TSR 13.2, the end of her collection period. Her combined average on the GJ test up to TSR 4.1 is 84%, the lowest average of the test informants for this initial period

of data collection. This approximates her average of 82% on her TV test over this period. Despite having a lower initial average than the 'R' informants and S1, these results still indicate S2 had a strong knowledge of English binding principles during the early stages of her collection period. However, this initial average drops to 76% during the secondary period up to TSR 8.1, and declines even further to 57% during the final five months. These overall averages alone indicate S2 undergoes severe attrition of English binding principles over the 13 month course of her data collection period.

Test subject S2's GJ results by type are depicted graphically in Figure 6.15 and are included in Appendix L.

*Figure 6.15. S2 grammaticality judgement test scores over time by sentence type*




| | 1 (0.4) | 2 (0.8) | 3 (1.5) | 4 (2.3) | 5 (3) | 6 (4.1) | 7 (5.3) | 8 (6.7) | 9 (8.1) | 10 (9.5) | 11 (10.7) | 12 (12.5) | 13 (13.2) |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Type 1 | 100 | 100 | 100 | 100 | 50 | 75 | 75 | 75 | 75 | 50 | 75 | 50 | 25 |
| Type 2 | 75 | 75 | 75 | 100 | 75 | 50 | 100 | 50 | 75 | 75 | 50 | 50 | 75 |
| Type 3 | 66 | 100 | 100 | 66 | 100 | 100 | 66 | 100 | 66 | 66 | 33 | 66 | 66 |

GJ scores by sentence type. Top row of grid indicates test number, TSR in parentheses

By sentence type, S2's GJ Type 1 sentence results are the lowest of the six test subjects, with an overall mean score of 73%. Up to TSR 4.1, her GJ Type 1 sentence average 88%, decline to 75% in the secondary sub-period up to TSR 8.1, and finally

drop to just over 50% in the final five months, a loss of 27% from her initial period. These results parallel her TV Type 1 results for these time periods. The corresponding TV results are are 92%, 72%, and 62%, a decline of 30% from the initial to the final collection period. On both tests, S1 shows an overall increasing preference to accept illegal subject antecedents in these biclausal sentences. On the GJ sentences in particular, her results on the Type 1b sentences drop from an average of 83% over the first four months, to 0% over the final five months. Her TV Type 1b sentences for the same periods also decline from 89% to just 41%, respectively, again demonstrating her preference for accepting subject NPs.

S2's GJ Type 2 sentence results are initially lower than the other test subject's results for this sentence type, and exhibit increased instability over time. Up to TSR 4.1 and then to TSR 8.1, S2's results average 75%. Over the final 5 months, to TSR 13.2, the results decline to an average of 63%. This increased instability and decline is also reflected in her TV scores which average 83%, 61% and 58% for the same 3 TSR periods. These declining averages for S2's GJ Type 2 sentences are further evidence of the attrition already seen in her TV Type 2 sentences. And as seen with the other test subjects, Type 2, nonfinite verb sentences, appear to attrite to a greater degree than the Type 1 sentences on both the GJ and TV tests. One explanation for this is that the non-finite verb in the embedded clauses of the Type 2 sentences may play a role in the test subjects' allowing subject NPs as antecedents.

Test subject S2's Type 3 sentence results also show clear signs of attrition, averaging 89% up to TSR 4.1, declining to 77% over the next 4 months and further declining to 50% by TSR 13.2. In particular, her GJ Type 3a sentences are most responsible for this attrition, suggesting S1's grammar gradually comes to reject object NPs in these monoclausal sentences. Further evidence of this can be seen in her subject NP oriented GJ Type 3b sentences which show less attrition and even a slight increase from 67% during the secondary sub-period, to 75% over the final 5 months. S2's TV Type 3 results also form a similar pattern and show a decline over the entire data collection period for the object oriented Type 3a sentences, but more stable results (though still attrition prone) subject oriented Type 3b sentences. As with her Type 1 and 2 sentences, S1's Type 3 results show a gradual rejection of object NPs as antecedents, again pointing to possible transfer of reflexive binding rules from her L1 Japanese into

her L2 English. As with the other two 'S' test subjects, S2's misjudged GJ Type 3c sentences occur within the first four months of data collection.

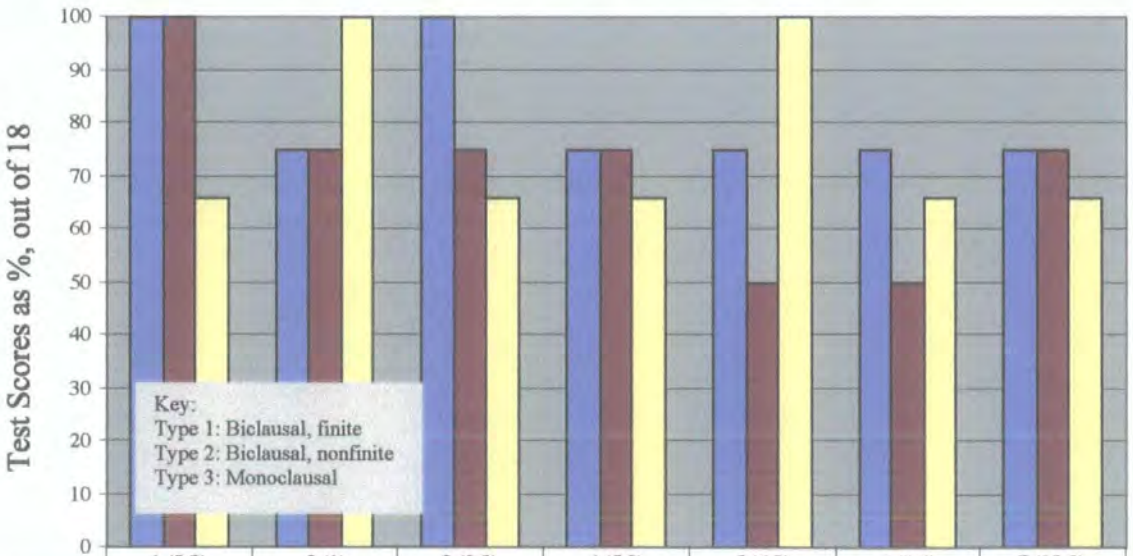
6.2.2.1.6. *Test subject S3's GJ test results by sentence type over time*

Test subject S3's data collection lasted slightly more than 12 months and included 7 collection sessions, the fewest number of collections of all the test subjects. Her data collection period has been divided into 3 parts for discussion, from TSR 0 to 2.9, then to TSR 6.8, and finally to TSR 12.2. S3's combined average on the GJ test up to TSR 2.9 is 84%. Her combined TV test average for this period is 78%. Both these averages drop about 10% during the secondary collection period, up to TSR 6.8 - the GJ average to 74%, and the TV average 67%. During the final 5 months of data collection, S3's Combined GJ results decline to an average of 68%. This is paralleled by the drop in her TV average to 64% for the same period.

Test subject S3's GJ results are depicted graphically by type in Figure 6.16 and are included in Appendix L.

*Figure 6.16. S3 grammaticality judgement test scores over time by sentence type*

By sentence type, S3's results for GJ Types 1a, 1c, and 1d sentences remain




| | 1 (0.3) | 2 (1) | 3 (2.9) | 4 (5.3) | 5 (6.8) | 6 (9.4) | 7 (12.2) |
|---|---|---|---|---|---|---|---|
| Type 1 | 100 | 75 | 100 | 75 | 75 | 75 | 75 |
| Type 2 | 100 | 75 | 75 | 75 | 50 | 50 | 75 |
| Type 3 | 66 | 100 | 66 | 66 | 100 | 66 | 66 |

GJ scores by sentence type. Top row of grid indicates test number, TSR in parentheses

strong throughout her data collection period. GJ Type 1b however, exhibits instability and attrition throughout her data collection period. Similar results on this sentence type were also seen for test subjects R3 and S2. S3's results average just 67% for her initial sub-period, 0% over the next four months, and 50% over the final five and a half months. This pattern is also exhibited in S3's TV Type 1b sentence results in which she averages 77%, 66%, and 50% over her three sub-periods. As with R3 and S2, S3's performance appears to indicate that her knowledge of the principles restricting antecedents to local NPs is initially uncertain and attrites in the final stages of her collection period. This attrition seems caused by her binding rule system moving towards an approximation of her L1 Japanese rule system.

S3's results on her GJ Type 2 sentences also show an increasing tendency to reject local NPs and accept long distance bound antecedents. Her GJ Type 2a (locally bound) results initially average 100% up to TSR 2.9. Over the next four month period this average declines to 50% as it does over the final five and a half month period. This attrition is also seen in her TV Type 2a sentence results which drop from an average of 89% over the initial sub-period, to just 50% in the final sub-period. Low averages on this sentence type can be interpreted as a rejection of the locally bound NP serving as the antecedent of the reflexive. Complementing this, S3's GJ Type 2b (long distance bound) results decline and show that she increasingly accepts the long distance antecedent. Her 2b results attrite very neatly, initially average 100%, then dropping to 50% during her secondary sub-period, and finally finishing out at 0% by her final sub-period. On the TV test, S3's results also indicate a strong rejection of the local antecedent, averaging 89%, 67%, and 50% on her TV Type 2a sentences for her initial, secondary and final collection sub-periods. Her TV Type 2b sentences show overall instability, averaging 55%, 89%, and 66% for the 3 sub-periods. Taken together, test subject S3's GJ and TV Type 2 sentences suggest her knowledge of the principles underlying the correct choice of antecedent in these sentence types is not well established from the beginning of her data collection period, and attrites to an approximation of her L1 binding system. That is, she increasingly tends to reject the local object NP antecedent and accept the long distance subject NP antecedent.

S3's GJ Type 3 sentences also decline over the course of her data collection period, though not as a severely as the declines found in test subjects R3 or S2. S3's GJ Type 3 results average 77% up to TSR 2.9. then actually increase to 83% from TSR 2.9

to TSR 6.8 before dropping to 66% over the final five and a half months. This decline in the final period is due exclusively to S3's average of 0% on the GJ Type 3b sentences, a further sign of her tendency to choose subject NP antecedents over object NPs as noted above for her Type 1b and 2b sentences. This trend is also evident in her TV Type 3b sentences which attrite from an initial average of 88% from TSR 0 to 2.9, to an average of just 33% during the final five and a half months of her data collection period.

This concludes the analysis of all six Japanese test subjects' longitudinal TV and GJ test results. Considering the results initially as indicators of how well the two tests correspond to each other, the averages of the three data collection sub-periods do appear to follow similar trends in two ways. Firstly, all averages except S1's decrease over time, and secondly, all GJ averages except S2's final data collection sub-period are higher than the TV averages. This can be seen in Table 6.7 which lists the averages for each test subject's initial, secondary and final data collection sub-periods. The first two sub-periods are each roughly four months long, with the third periods varying more from test subject to test subject. Table 6.7 summarizes the results of each of the three data collection sub-periods for each test subject.

Table 6.7. Comparison of test subjects' TV and GJ test result averages over time by data collection sub-periods

| | | First | Secondary | Final |
|---|---|---|---|---|
| R1 | TV | 91% | 81% | 78% |
| | GJ | 98% | 96% | 85% |
| R2 | TV | 83% | 78% | 81% |
| | GJ | 100% | 91% | 82% |
| R3 | TV | 82% | 65% | 61% |
| | GJ | 91% | 79% | 64% |
| S1 | TV | 81% | 82% | na |
| | GJ | 91% | 91% | na |
| S2 | TV | 82% | 67% | 57% |
| | GJ | 84% | 76% | 57% |
| S3 | TV | 78% | 67% | 64% |
| | GJ | 84% | 74% | 68% |

Through these comparisons, both the TV and GJ tests elicit comparable sets of results from the test subjects which indicate that the two tests successfully measure the phenomena being investigated in this study.

Section 6.2. has presented the breakdown of the data by test subject, by test type, and by sentence type. In the next section I will discuss these results in more detail and attempt to show that despite the attrition of certain aspects of their knowledge of reflexive binding, the test subjects in this study do not exhibit non-UG sanctioned binding constraints.

## 6.3. Discussion

In this study, truth value judgement tests and grammaticality judgement tests were used to collect data on L2 English reflexives. These tests were designed to examine the maintenance or loss of parameter settings in L2 English within the UG framework as it applies to principle A of binding theory. The data collected in this study and presented above do indicate attrition occurs over time. That is, the test subjects exhibited losses in their ability to identify correct antecedents in the tests conducted. The phenomena observed and reported here are reflected in the hypotheses stated in Chapter 5 and repeated here:

> Hypothesis A: Principles of reflexive binding already instantiated in the test subjects' English will undergo change due to lack of exposure to English.
>
> Hypothesis B: Changes observed in the test subjects' control over reflexive binding due to lack of exposure will be UG constrained.
>
> Hypothesis C: A correlation will be exhibited between age at first L2 exposure and degree of retention of L2 reflexive binding principles.

Hypothesis A has been supported by the analysis of the test results presented in section 6.2 above. The test subjects' principles of English reflexive binding do, in fact,

appear to undergo change in the months after their return to Japan when their exposure to English ceases.

To find support for Hypothesis B, let us consider the results in terms of a hierarchy of possible grammars. If UG is available to L2 learners, and positive evidence allows a learner to move from an unmarked subset setting to a more marked superset, as claimed by White (1986b), then in the case of L2 attrition, where there is no evidence positive or negative, we would expect no changes to occur in a test subject's grammar. But as we have seen, changes do indeed occur, changes which result in the test subjects' IL grammar becoming less target like. The question we must then ask is whether these changes are random or principled. Random changes might result from memory loss, test subjects forgetting aspects of the grammar. Principled changes, on the other hand, would reflect the effects of a rule-bound process. This process could be either L1-based, or UG-constrained. If we attribute attrition solely to L1 transfer we would expect only instances of L1-like structures to appear in the attriting L2. If, however, we were to find evidence of non-L1-like structures occurring in a principled manner, and were those structures UG sanctioned, then this could be taken as evidence of UG constraining the attriting IL.

In the following sections I will combine the test subjects' TV and GJ test results to obtain a broad impression of the degree to which test subjects' overall adherence to or rejection of English reflexive binding rules undergo attrition, and the degree to which that attrition is principled. In this way I will look for support for Hypothesis B.

### 6.3.1. *The attrition of adherence to reflexive binding in L2 English*

In this section I will first examine adherence to local binding in biclausal sentences, and then examine antecedent preference in the monoclausal sentences.

#### 6.3.1.1. *The biclausal sentence types*

In order to consider how well or poorly test subjects maintained their overall acceptance of local antecedents and rejection of long distance antecedents in biclausal

sentences over time, the data from all TV and GJ biclausal test types have been combined in two ways. Firstly, to examine the accuracy with which test subjects accept locally bound antecedents, the results from TV sentence Types 1a, and 2a, and GJ Types 1a, 1c, 2a, and 2c have been grouped together and discussed as averages for each data collection sub-period by test subject. These results will be referred to as the Overall Acceptance (OA) results. The data analyses in Appendices J and M illustrate the particular TV Type 1a and 2a sentences, and the GJ type 1a, 1c, and Types 2a and 2c sentences which contribute to any attrition which occurs. Secondly, to examine the accuracy with which test subjects reject long distance bound antecedents, the results from TV sentence Types 1b and 2b, and GJ sentence Types 1b, 1d, 2b, and 2d have also been grouped together and discussed as one average for each test subject's data collection sub-period. These results will be referred to as the Overall Rejection (OR) results. The data analyses in Appendices J and M illustrate the particular TV Type 1b and 2b sentences, and the GJ type 1b, 1d, and Types 2b and 2d sentences which contribute to any attrition which occurs. High OA and OR would be indicative of a robust knowledge of the GCP. These combined scores which form the OA and OR averages are presented in Appendix N and are summarized below in Table 6.8.

Table 6.8. Test subjects' Overall Acceptance (OA) results of local binding and Overall Rejection (OR) results of long distance binding over time

| Collection Sub-period/ Test Subject | | *Initial* | *Secondary* | *Final* |
|---|---|---|---|---|
| R1 | **OA** | 95% | 93% | 98% |
| | **OR** | 95% | 93% | 80% |
| R2 | **OA** | 100% | 95% | 100% |
| | **OR** | 90% | 93% | 75% |
| R3 | **OA** | 98% | 80% | 68% |
| | **OR** | 83% | 67% | 63% |
| S1 | **OA** | 92% | 93% | na |
| | **OR** | 93% | 87% | na |
| S2 | **OA** | 92% | 77% | 80% |
| | **OR** | 82% | 63% | 40% |
| S3 | **OA** | 97% | 70% | 65% |
| | **OR** | 70% | 65% | 60% |

One form of evidence of the attrition of a test subject's knowledge of reflexive binding would be a decrease in both OA percentages and OR percentages. As a subject's grammar attrites, she may increasingly stop accepting the local antecedent, thus lowering her OA, while her increasing failure to reject the long distance antecedent will in turn lower the OR. Results of this nature can be seen to varying degrees in the data. All the test subjects' OA scores in this study decrease from the initial sub-period to the final sub-period, with the exception of R1, R2, and S1, who have slight increases over time. This indicates that over time, three out of the six test subjects increasingly reject correctly formed and pragmatically accurate local NP antecedents.

In the case of R1, she had the most overseas living experience of all the test subjects, and that from the earliest age. Her consistent acceptance of the locally bound antecedents, never scoring lower than 90% on each OA average throughout her sixteen and a half month data collection period, reflects a robust, native-like acceptance of local binding. Despite this, R1's OR of long distance bound antecedents decreases over time. That is, she increasingly comes to accept illegal long distance antecedents. Her OR for her initial sub-period is 95%. For the final sub-period this declines to 80%. This 15% drop points to an increasing number of instances of anaphors being bound with a c-commanding NP outside its English, governing category. These violations increasingly occur with all six test subjects in sentences such as (7).

(7) A boy got lost in the big city. He didn't know how to get back home. Finally he saw a police officer and asked him for directions.

*The child asked the police officer to help himself.*

(7) establishes a scenario in which *the child* should be the antecedent of *himself.* But instead of *the child,* the test subject finds *police officer* as the local C-commanded antecedent, and clearly from the story the police officer doesn't need to help himself, making the correct judgement of (7) false. Judging a sentence like (7) true would reflect a belief on the test subject's part that long distance antecedents such as *the child* are allowed in English. Results of this nature would indicate that the tests subject has widened her governing category setting to a more marked, Japanese-like setting.

Increases in wrong judgements of this type are also evident in the other test subject's data and reflect change in their understanding of what constitutes acceptable antecedents in English. Let us look at each test subject in turn.

Test subject R2's data collection period was 12 months. Throughout her entire data collection period, she maintains very high OA percentages of 100%, 95% and 100% for her three sub-periods. The slight decline in her secondary sub-period to 95% is caused by lower scores on her TV Type 2a sentences. This is of interest, as Type 2a sentences have a non-finite verb in their embedded clause, a binding constraint typically acquired later than similar sentences with finite verbs. This is evidence of the last learned item attriting first and supports the regression hypothesis as outlined by de Bot and Weltens (1991).

R2's overall rejection of long distance antecedents shows signs of attrition, her initial, secondary and final data collection sub-periods averaging 90%, 93%, and 75% respectively. Misjudgments in R2's OR results are most evident in her Type 2 sentences on both the TV and GJ tests. In these cases, R2 increasingly accepts long distance bound antecedents in sentences which have a nonfinite verb in the embedded clause.

Test subject R3 shows strong evidence of attrition during her 14-month data collection period with both her OA and OR results declining over time. R3's OA results average 98% during her initial sub-period, 80% in her secondary sub-period, and then drop to 68% in her final sub-period. Although R3's Type 1 and Type 2 sentence results on both tests decline over the course of her entire data collection period, this sharp degree of attrition seen in her final sub-period is primarily caused by drops in her Type 2 sentences.

At 83% R3's OR averages start off lower than her OA results and subsequently show an earlier and sharper decline; her secondary sub-period OR scores average 67%, and decline further to 63% in her final sub-period. As with R1 and R2, R3's decline of her OR averages can be attributed to misjudged Type 2 sentences on both the TV and GJ tests. Considered together, test subject R3's OA and OR results thus show evidence of a stronger maintenance of knowledge of English binding in sentences with finite verbs than in sentences with non-finite verbs. Again, this provides evidence that the status of the verb — finite or nonfinite — plays a role in the degree of attrition which obtains.

Although test subject S1's data collection period lasts only nine months and thus has only two data collection sub-periods, her OA and OR results show the most robust

continued knowledge of binding principle A of the three 'S' subjects. Her initial and secondary OA sub-period averages of 92% and 93%, respectively, are comparable with those of R1 and R2. Although her OA averages increase by 1%, a closer look at her scores reveals that at the individual test item level, her TV Type 2a scores are only slightly unstable, declining from 83% in the initial sub-period to 78% in the secondary sub-period. This provides some further, if slight, evidence that Type 2 sentences — non-finite verbs in the embedded clauses — have a higher propensity to attrite than do Type 1 sentences with their finite verbs. This pattern can also be seen in S1's OR percentages which drop from 93% to 87%. This drop can again be attributed primarily to her Type 2 sentences, in particular TV2b and GJ2d, reflecting her increasing willingness to accept long distance antecedents in sentences containing non-finite embedded clauses.

As noted in the analysis of the test subjects' test results, test subject S2 exhibits the highest levels of attrition of all six test subjects during her 13 month data collection period. This can be seen in her OA results which average 92% for her initial sub-period, 77% for her secondary sub-period, and 80% for her final sub-period. Although this drop is actually attributable more to sentences with Type 1 (finite) sentences with Type 2 (non-finite), there is still some attrition of Type 2 sentences. In particular, The data analyses in Appendices J (TV items) and M (GJ items) show the particular test items which contribute to S2's OA drop.

Perhaps the strongest evidence of this study for attrition of anaphoric binding rules in English can be seen in test subject S2's OR percentages which are the result of S2's misjudged sentences indicated in the data analyses in Appendices J, TV Types 1a and 2a, and M, Types 1a and 1c. As shown in Table 6.8, S2's initial sub-period OR results average 82%. In the secondary sub-period, the average declines to 63%, while in the final sub-period it drops to just 40%. This dramatic level of attrition is due to a mix of misjudged sentence types. Both her TV Type 1b sentences TV Type 2b sentences undergo severe attrition, dropping from 89% to 67% to 42% and 83% to 56% to 42%, respectively, over the three sub-periods. On the GJ test, S2's 1b sentences exhibit massive attrition, dropping from 83% to 67% to 0% over the sub-periods. This attrition is far more extreme than her GJ Type 1d sentence which average 100% for the initial and secondary sub-periods, and only decline to 75% during the final sub-period. In GJ Type 1b sentences, number between the reflexive and the local antecedent disagrees, but agrees with the long distance antecedent, as in (8)

(8) *The coach knew the players trusted himself.*

In the GJ Type 1d sentences, it is the gender which disagrees and agrees in this manner, as in (9)

(9) *John's mother knew Ben asked about herself.*

The extreme attrition seen in sentences like (8), but not in (9) may be due to S2 being less sensitive to number discrepancies than to differences in gender.

S2's GJ Type 2b sentences exhibit a trend similar to her GJ 1b sentences in that the 2b sentences attrite from an initial average of 100%, to a secondary average of 34%, and a final average of 25%. S2's GJ Type 2d sentences exhibit less attrition, but are very unstable and have low averages from the initial sub-period for which the average is just 17%. This increases to 67% in the secondary sub-period, and drops to 50% in the final period. Taken together then, the scores which combine to create S2's OR averages clearly point to considerable instability and subsequent massive erosion of her knowledge of local binding constraints in English, particularly in constructions using nonfinite embedded clauses such as TV 2a, and having anaphors and antecedents which disagree in number, such as GJ 2b. See data analyses in Appendices J and M. The role of number agreement will not be explored here except to note that, as suggested by Lardiere (1998), the mapping procedures which allow syntactic features such as inflectional morphology to be realized at the level of phonetic form (PF) may fossilize in an L2 learner at a non-target-like level, while other aspects such as case marking are realized at a native-like level. Lardiere suggests that fossilization of the one aspect of grammar, but full development of the other provides evidence that the courses of syntactic and morphological development are independent of each other, and is a potential source of divergence between L1 and L2 final states (p. 1). Were S2's verb morphology fossilized in this way, even at a highly target-like level, we could expect any attrition of the mapping procedures to interfere in her judgements of GJ Type 1b and 2b sentences.

Test subject S3 also exhibits clear signs of attrition over the course of her 12-month data collection period. On her OA scores, S3's averages drop from 97% in the initial sub-period to 65% in the final sub-period. As with S2' results, we can see nonfinite clauses and number agreement contributing the most to this attrition, although in S3's case she also attrites on her finite TV1a sentences. S3's OR scores further support the pattern of attrition seen in S2's results, with S3 exhibiting considerable instability on her GJ1b sentences (67%, 0%, 50%), and heavy attrition on her GJ2b sentences (100%, 50%, 0%). An increase over time on her GJ2d sentences (33%, 50%, 100%) most likely reflects overall instability on judging this form, or may reflect a degree of increasing task awareness. As with S2, test subject S3's OA and OR results provide considerable evidence of a principled loss of reflexive binding constraints in L2 English and support for Hypothesis B.

In the next section I will consider the results of the monoclausal sentence types.

6.3.1.2. *The monoclausal sentence types*

Monoclausal sentences, which include TV Types 3a and 3b sentences and GJ Types 3a, 3b, and 3c sentences, offer less convincing yet still interesting evidence of attrition, in this case of the test subjects' proper antecedent parameter. Unlike the biclausal sentence types, in which the long distance antecedent is grammatically unacceptable in English, monoclausal sentences allow either the subject or the object NP to serve as the antecedent depending on constraints such as pragmatics, and number and gender agreement. Correctness in the TV monoclausal sentences is determined by a preceding brief passage preceding the stimulus sentence. This passage creates a scenario determining which of the two grammatically possible antecedents, either the object NP as in (10), or the subject NP as in (11), is correct.

(10) Richard was a professional photographer. He took a picture of Cliff and Cliff bought it for $50.00.

*Richard sold Cliff a picture of himself.*

(11) Janice was curious about what she was like as a little girl, so she talked to her mother about her childhood.

*Janice asked her mother about herself.*

The monoclausal GJ test items also examine test subjects' understanding of the PAP in English. Test subjects must judge the correctness of sentences containing either subject bound NPs or object bound NPs, based on number agreement as in (12) and (13), or gender agreement as in (14) and (15).

(12) Subject antecedent oriented by number:
*Sally faxed her teachers a report about herself.*

(13) Object antecedent oriented by number:
*Bill and Dan gave Sam a picture of himself.*

(14) Object antecedent oriented by gender
*The policeman gave the traffic accident victim a report about herself.*

(15) Subject antecedent oriented by gender
*Alex never finished the letter to his mother about himself.*

Overall, both the test subjects and the native speakers perform more poorly on the monoclausal sentence types than on the biclausals. This may stem from speakers' (both native and L2 learners') strong preference for subject-only binding in monoclausal sentences (Read and Chou Hare, 1979; Goodluck and Birch, 1988), which causes informants to reject grammatical object antecedents. Thomas (1993) suggests informants may prefer subject antecedents "as a reflex of the generally greater syntactic prominence of subjects" (p 32). In this sense both the test subjects and the controls parallel each other. Most test subjects' early scores on the monoclausal sentence types reflect a strong knowledge of the PAP for English. R1's Type 3 sentences on the TV test, though not 100% accurate, do remain constant throughout her data collection period. Her GJ scores on the monolingual scores are the highest of all informants, including the native speaker control group. This high level of consistent accuracy on the monoclausal sentences, together with her strong results on the biclausal sentences

distinguishes test subject R1 from the other five test subjects as having the most robust knowledge of English reflexive binding.

Although R2's initial test period results show a low TV Type 3a score (50%), her TV Type 3b scores, and especially her monoclausal GJ scores for this period are strong (75%). Furthermore, despite R2's monoclausal scores developing a somewhat random pattern over the course of her data collection period, the difference between her initial and final data collection sub-periods does indicate attrition occurs in her PAP for English. Of most interest is that as R2's knowledge of binding attrites, she comes to reject object NPs in the monoclausals as the correct antecedent more than she rejects the subject NPs. This corresponds to the results of SLA studies of the acquisition of binding by researchers such as Hirakawa (1990), Thomas (1993) and Finer and Broselow (1986) who all find reflexives to be bound to subject NPs more readily than to object NPs. Were R2's binding restrictions to revert to her L1 Japanese, we would expect her to develop this preference to bind the anaphor to the subject NP antecedent in the monoclausal sentences.

R3's monoclausal sentence type results are initially strong, with a preference for subject NPs outweighing object NPs as the proper antecedent. This is true for both her TV and GJ sentences. During the course of her data collection period, this tendency changes as attrition sets in and she surprisingly develops a preference for object NPs. This runs counter to previous research results showing that L2 learners prefer subject NPs in these sentence types. The type of subject oriented monoclausal sentences which seem to suffer the most are grammaticality judgement sentences in which the anaphor and the subject NP agree in number such as (16)

(16) *The consultant e-mailed the customers about herself.*

Sentences in which the anaphor and the subject NP agree in gender such as (17) were less prone to attrition.

(17) *Jack spoke to his sister about himself.*

Turning to the 'S' test subjects, S1's monoclausal sentence types show low initial scores and overall instability. Her TV Type 3a results average 39% during the initial sub-period, and 67% over her secondary (and final for S1) sub-period, and her TV Type 3b sentence results show a slight increase, averaging 72% and 78% for each sub-period. These results are mixed in that, despite increases in her acceptance of object antecedents over time, it appears that S1 also increasingly accepts subject NPs as antecedents on these sentence types. The pattern which emerges from her monoclausal GJ test items is slightly different from that of her monoclausal TV items in that both GJ Type 3a and 3b sentences undergo attrition. Her 3a averages decline from 83% to 67% from the initial to the secondary sub-period, and the 3b scores from 100% to 67%. S1's increasing rejection of the subject NP as an antecedent in monoclausal sentences which correctly favor such an interpretation is in contrast to her TV results which show a slight gain on this sentence type.

Test subject S1's GJ Type 3c sentence results show some instability at first with an average of 67% in the initial sub-period, but gain in the secondary sub-period to a perfect average of 100%. This increase, along the increases observed in her monoclausal TV test sentence results suggest she may have experienced some effect from the test as she became familiar with it.

Taken together, S1's monoclausal sentence results are suggestive of someone who never had the PAP for English firmly set in her mind. Lack of L2 exposure after her return to Japan has little effect on whatever knowledge of these forms she had prior to her return. Despite this, her biclausal results on both the TV and GJ tests point to a fairly robust knowledge of long distance binding restrictions in English. One explanation for this is that the subset principle determining the PAP, and her knowledge of the governing category for English have distinct and separate instantiations in her mind rather than stemming from an overarching, limited knowledge of English binding.

As mentioned above, test subject S2 is the most interesting of the six test subjects as she shows the highest level of attrition on her biclausal sentence types. This is also true of her monoclausal sentences. Although her TV Type 3a results actually increase from 56% in her initial sub-period to 67% in her secondary, they exhibit a severe drop to just 25% in the final sub-period. Though not as severe, her TV Type 3b sentences results also show signs of attrition and then low but stable scores by declining from 83% in the initial sub-period to 67% in both the secondary and final sub-periods. Her Type

3a results are significant in that they reflect an increasingly strong aversion to accepting object NPs as antecedents. At the same time, the decline in her TV Type 3b sentence results indicates she also unexpectedly rejects subject NPs as antecedents, though not as strongly as she rejects the object NPs. Despite the loss in the Type 3a sentences, the results of the 3b sentences point to an uncharacteristic preference for objects antecedents in the monoclausal sentences.

S2's monoclausal GJ sentences also decline over time and exhibit a far greater degree of loss in the Type 3a object oriented sentences than in the Type 3b subject oriented sentences. Her Type 3a sentences decline from an initial average of 100%, to 67% in the secondary sub-period, to 25% in the final sub-period. This level of attrition clearly reflects the change in S2's grammar from initially accepting object antecedents in these monoclausal sentences to strongly rejecting them. S2's GJ Type 3b results exhibit a lesser degree of attrition than do her 3a results. Her averages for the three sub-periods on the GJ 3b sentences drop from an initial 100% to 67% in the secondary period, and then increase slightly to 75% in the final sub-period. Although these results reflect instability in S2's understanding of the principles underlying the acceptance of subject antecedents in monoclausal English sentences, she nevertheless undergoes far less attrition on these forms than she does on the object antecedent Type 3a sentences. S2's PAP appears to have deset to a subset of the original English-like superset, suggesting this is a result transfer from the L1 Japanese subject antecedent orientation.

S2's GJ Type 3c sentence results exhibit the same pattern as test subject S1's. The initial sub-period results average 67%, followed by an average of 100% over the secondary sub-period. S2's final sub-period shows signs of instability returning with an average of 75%. Despite the instability, these results point to a general knowledge of the PAP for English greater than for the specific antecedent-oriented constructions (i.e., Types 3a and 3b).

Test subject S3's monoclausal sentence results reveal an unstable setting of the PAP for English. Her TV Type 3a and 3b results are mostly low and consistent- 3a results are 56% in the initial sub-period, 33% in the secondary sub-period, and 67% in the final sub-period; 3b results are 89%, 83%, and 67%, respectively. This instability can also be seen in her GJ3a and 3b results. Her 3a results reflect an increase in correctly judging local, object antecedents (67%; 50%; 100%), but a decline in accurately judging the subject-oriented Type 3b sentences (100%, 100%, and 0%). S3's

results again reflect an example of a test subject exhibiting a fairly consistent preference for accepting the object NPs as the proper antecedent.

The results of the monoclausal sentence types in both the TV and GJ tests are mixed and indicate the test subjects, as a group, do not develop a strong preference for either object or subject NP antecedents on these test items. This indeterminacy is different from the increasing preference for the long distance bound subject NP antecedent seen in the biclausal sentences. Furthermore, an examination of the ambiguous GJ Type 3c monoclausal sentences shows they are increasingly accepted over time. This increasing level of acceptance shows the test subjects maintain an understanding of the basic principles of reflexives, at least in monoclausal sentences in which either antecedent — the subject NP or the object NP — is acceptable. Why then do the test subjects remain ambivalent about the subject/object NP distinction in the monoclausals? If the native speaker results on monoclausal sentences are considered, we can see that their results too show some indeterminacy in judging these sentence types. The interpretation of monoclausal sentences can therefore be seen as ambiguous by nature, this indeterminacy becoming more pronounced throughout the lengthening periods of no exposure to the target language.

### 6.3.2. *UG constrained attrition*

As discussed in Chapter 4, UG constrains speakers' grammars to bind reflexives according to binding principle A. Wexler and Manzini's (1987) subset principle states that a superset GCP value such as that posited for Japanese should include all lower values in the hierarchy, i.e., including those posited for English. Sharwood Smith and van Buren (1991) suggest that markedness theory may determine the degree to which a parameter setting may be maintained or lost. If a speaker of a language which has a marked GCP setting for instance, moves to an L2 environment which has a less marked GCP setting, that speaker's marked setting will gradually reset to the unmarked L2 setting, supposedly through negative evidence available in classroom instruction. But what of the reverse? L2 speakers of unmarked English moving back to the marked Japanese environment might be expected to increasingly reject the local antecedent restriction in their English by adding the Japanese long distance binding rule to their

grammar for English binding. Were this the case, and were they to ultimately develop an English grammar which completely rejects local binding, we could say that this new grammar is a non-UG sanctioned "rogue" grammar. The GCP does not allow for such a grammar and the L2 English data results presented in this thesis indicate that this is not the case. For the duration of the study reported on here, the data results confirm that as the test subjects' knowledge of reflexive binding attrites, it nonetheless appears to remain within the constraints of UG as predicted by the subset principle. This gradual desetting of the restrictions on binding in English to include the more marked, Japanese superset is in agreement with Principle A of binding theory. The data from the present study support Hypothesis B.

### 6.3.3. *The role of age at initial exposure in the attrition of reflexive binding*

The role of age at initial exposure is addressed in Hypothesis C which claims that a correlation will be exhibited between age at first L2 exposure and degree of retention of L2 reflexive binding principles. What effect, then does age at first exposure to the L2 have on a learner's ability to maintain the L2 in the face of zero input?

The attrition data of six proficient ESL speakers presented in this chapter appear to support Hypothesis C to a certain degree. Three test subjects had childhood exposure to English in an English speaking environment, and three had no exposure until their early 20s. Despite this difference, the initial phase of the study indicates all six had advanced levels of English as evidenced by their high TOEFL scores and the initial results on their TV and GJ tests.

Despite these early similarities, data from the present study show that age at first exposure has some effect on the degree of attrition which sets in following the test subjects' return to Japan. R1, with exposure from birth to almost nine years, has the earliest and longest childhood exposure to English; R2 has with exposure from 6;3 years to just over ten years has four years of exposure. And R3, from 8;3 to 11.1 has just under three years. We would thus predict R1 to have the most robust grammar at the time of her return to Japan and undergo the least attrition. The data support this prediction. R1 consistently accepts local antecedents at native speaker levels throughout her data collection period. Her rejection rate of long distance antecedents drops

somewhat over time, but still remains near-native-like throughout her data collection period. Her results on monoclausal sentences, especially her GJ Type 3 sentences also compare favorably with the native speaker results. R1 is thus an L2 speaker with considerable childhood exposure to the L2 who undergoes little attrition. Although other variables such as attitude and commitment to learning and using English may contribute to the robustness of her reflexive binding, her childhood exposure to English must be considered the primary contributing factor to her ability to resist L2 attrition in the face of zero input.

Test subjects R2, and R3 each have less than half the amount of childhood exposure to the L2 of R1, and that from later ages. We can ask, then, if four years from age six, or three years from age eight are sufficient exposure to stem attrition of the binding principles under consideration here. In the case of R2, her scores over time on her biclausal sentences drop and reflect an increasing tendency to reject local and accept long distance bound antecedents. Her monoclausal sentence results, though not at native speaker level, are comparable to R1's results except for GJ Type 3a sentences which drop by 60%. GJ 3a sentences call for the object NP to be recognized as the correct antecedent and R2's reluctance to accept this sentence type can also be seen in her low, though more consistent TV Type 3a sentences. Overall however, R2's results reflect a knowledge of reflexive binding which remains fairly robust over time despite a much shorter period of childhood exposure to English than that of R1.

Test subject R3, with the shortest amount of childhood exposure to L2 English of the three 'R' test subjects, is also the oldest at first exposure. As we have seen, R3's test results reflect a knowledge of reflexive binding distinctly less robust than R1's or R2's. She moves from rejecting locally bound antecedents only 2% of the time in her initial data collection sub-period, to a rejection rate of more than 30% in her final sub-period. Mirroring these results, R3 increasingly accepts long distance bound antecedents: 17% in the initial sub-period, 37% in the final sub-period. R3's monoclausal results reflect attrition of the PAP as she becomes increasingly unsure of which NP to choose as the correct antecedent. Despite her childhood exposure to the L2, the attrition R3 exhibits is more extreme than that seen in R1, R2 and also S1. This may not only be a result of the relatively limited length of her L2 exposure (2;8 years), but also later exposure. As noted in Chapter 2, Long (1990) suggests that there may be a series of sensitive periods for L2 acquisition, each period corresponding to various aspects of the target language.

Perhaps R3, at age 8;3, was already past a sensitive period beyond which native-like acquisition of binding is impossible.

In addition to L2 exposure, cognitive and affective variables may also have contributed to R3's lower level of binding acquisition and subsequent attrition. R3 had a retiring nature and taciturn attitude toward English and her experience in the States. Additionally, she was a somewhat reluctant test subject which I believe reflected her approach to her ESL experience in general. All these factors combined seem to have limited her intake of the input, consequently leading to distinctly lower results.

The three 'R' test subjects demonstrate that the early childhood exposure to the L2 in the L2 environment can influence the degree of robustness of language over time in the face of zero input. In the cases of R1 and R2, this exposure seems to have allowed them to maintain the binding principles investigated here. In the case of R3, however, there seems to have been little maintaining effect. Her later age at first exposure may mean she had matured beyond the point of reliably acquiring binding, the relative shortness of her L2 exposure, and her negativity about her experience abroad may have contributed to her attrition. Despite these explanations for R3's low scores, it is interesting to note that one of the native speaker controls (NS1) also had test results lower than those of his group-mates.

Let us turn now and consider the 'S' test subjects.

Test subject S1 exhibits almost native-like results on her biclausal sentences despite having no childhood exposure to English. Her rate of acceptance of local antecedents is the least prone to attrition among all six test subjects. Her monoclausal TV scores, though the lowest of the group, actually improve over time. S1's performance represents an adult L2 learner who attains a level of knowledge of English reflexive anaphors which is robust over time and comparable to that of test subjects with childhood exposure to the L2. Her relative success is the strongest evidence in this study for at least partial access to UG by an adult L2 learner. Given the homogeneity of the Japanese education system, though, we must ask why S1 succeeded where S2 and S3 performed more poorly? One simple answer is that her data collection period is the shortest and that she will eventually attrite to levels similar to those of S2 and S3. Yet, S1's OA and OR scores for the two sub-periods available for her do indicate her knowledge of English binding rules is comparable to the 'R' subjects. In particular, S1's scores for her secondary sub-period in remain strong at 93% and 87%, compared to 77%

and 63% for S2, and 70% and 65% for S3. Although limited, S1's longitudinal data describe someone whose ability to maintain her knowledge and use of the binding principles investigated in this study goes beyond deductive learning and has been facilitated by innate processes.

In sharp contrast to S1, test subjects S2 and S3 exhibit quite severe attrition over time as discussed above in section 6.3.1. Their biclausal sentence results initially reflect an overall knowledge of reflexive binding comparable to the other test subjects. However, during the secondary and final data collection sub-periods, this knowledge erodes as attrition sets in and they increasingly come to reject local antecedents and accept long distance antecedents. Is the attrition observed in S2 and S3 solely the result of their lack of childhood exposure to the L2? If this is the case, how can we explain S1's success compared to S2, S3, and even R3 who *did* have childhood exposure to English? Would the results differ significantly with different test subjects? Three variables in addition to childhood exposure which may also influence the degree of attrition obtained include the method of foreign language (FL) instruction in Japan, learner's cognitive ability, and affective variables such as intrinsic motivation.

All test subjects attended public (i.e., state funded) primary and secondary schools in Japan. The 'R' subjects (except R1) prior to and after their childhood stays abroad, and the 'S' subjects for their entire primary and secondary education. These public schools are tightly controlled by the Japanese Ministry of Education which determines most facets of education including curriculum, method of instruction, textbooks to be used, number of hours per week of instruction, etc. Method of instruction in the Japanese L1 environment then, is unlikely to differ significantly among the test subjects. This leaves cognitive development and attitude towards English (intrinsic motivation) as possible variables which may help explain the variety of success and failure in maintaining anaphoric binding found in the data collected for this study. S2 and S3 did have a somewhat lackluster interest in English and their experiences abroad, and they certainly had a less keen attitude towards their studies than did S1. Although speculative, it does seem probable that in addition to their lack of L2 environment exposure, S2 and S3's attrition is compounded by their somewhat indifferent attitudes and less than superior levels of cognitive skills. R3's attrition too, despite her (albeit) later childhood L2 exposure in the L2 environment, may also be a result of attitudinal and cognitive factors.

Although there are other variables not addressed here, such as the type of L2 exposure abroad, the three variables discussed here - childhood exposure, cognitive variables and affective variables - have been shown to conspire to cause various levels of attrition in the test subjects included in this study. Although a connection has been established between age at first L2 environmental exposure and degree of retention of L2 reflexive binding principles, we can see that age at first exposure is not the only factor which will determine ultimate levels of L2 maintenance or loss.

## 6.4. Conclusion

This chapter has examined data collected from both the native speaker control group and the Japanese test subjects. The NS control group results supported the TV and GJ tests as reliable tools with which to collect data on the targeted reflexive binding principles.

The results of the test subject data analysis provide evidence which supports Hypotheses A and B in that test subjects' knowledge of reflexive binding does change and attrite over time in a manner which appears consistent with UG constraints. Furthermore, different principles of reflexive binding do not attrite equally. Principles associated with the GCP appear to attrite in a manner which causes test subjects to increasingly adopt their marked L1 Japanese setting for governing category in their L2 English, allowing long distance subject NPs as antecedents. However, L1 transfer alone does not seem responsible for the change and attrition. Throughout the test subjects' data analyses made in this chapter, sentences with nonfinite verbs in the embedded complement clause have seemed much more prone to attrition than sentences with finite verbs. Reference to the TV test results in Appendix I and to the GJ test results in Appendix K will reflect this trend. This finite-nonfinite distinction is unexpected as finiteness does not effect reflexive binding in either Japanese or English, as it does in some languages, such as Russian. In Russian and Russian-like languages, anaphors can be bound long distance out of complements with non-finite verbs, but not clauses with finite verbs. One interpretation of this unexpected "middle-setting" of the GCP developing in the test subjects' English is that UG may play a role in shaping their L2 grammars, constraining them from developing rogue grammars even while their L2 is

attriting. These UG constraints determine what is instantiated in the mind, while the cognitive constraints referred to above restrict the degree to which UG constraints are realized.

Qualified support for a correlation between age and degree of attrition as stated in Hypothesis C can also be found in the results. Early and prolonged exposure to the target language in the L2 environment does predict a robust knowledge and use of reflexive binding as in the case of test subject R1 and to a certain extent R2. However, R3's poorer results suggest that such exposure must happen prior to the end of a sensitive period for binding which, in her case, appears to have occurred by age 8;3. In addition, S1's acquisition and maintenance of principles of reflexive binding despite her lack of childhood exposure demonstrate that early exposure is not the only predictor of high levels of acquisition and maintenance and that UG may continue to be available to learners after childhood.

# Chapter 7: Conclusion

## 7.0. General remarks

Research into L2 attrition has largely focussed on lexical and morphological loss. No study to my knowledge has addressed the role of UG in L2 loss. This study has thus been unique in that it uses generative-based SLA research tools and methods to investigate L2 attrition. Furthermore, the truth value test and the grammaticality test have provided results which support the hypothesis that principles of reflexive binding attrite in a manner not inconsistent with UG constraints.

The general pattern exhibited by all six test subjects initially shows varying but high levels of knowledge of reflexive binding. These levels then become unstable and knowledge declines, considerably in some cases, over the course of approximately 9 to 16 months. In particular, reflexives in finite subordinate clauses tend to remain bound grammatically to local antecedents to a greater degree than in nonfinite clauses. Reflexive binding in tensed clauses is thus more resistant to attrition. This result correlates with Thomas' (1993) findings in her study of the *acquisition* of L2 English in Japanese and Spanish learners. Thomas' results show that her test subjects acquire the local binding constraint in English non-tensed clauses later than in tensed clauses. These results, considered in light of the attrition data discussed in Chapter 6, lend support to the regression hypothesis proposed by de bot and Weltens (1991) and described here in Chapter 3. If Thomas is right and English binding restrictions in sentences such as Type 2a (biclausal nonfinite) are acquired only after those applying to sentences like 1a (biclausal finite) are in place, then evidence of restrictions applying to Type 2a sentences becoming unstable and attriting while restrictions applying to Type 1a sentences are maintained, supports the "first learned, last lost" claims of the regression hypothesis.

One explanation for this may be that these speakers gradually lose the local governing category restriction in English in the face of no L2 input and come to associate the governing category with their Japanese L1. In this case, the resumption of L1 influence can be held responsible for the loss of the test subjects' English binding restrictions and we can attribute attrition to L1 transfer. In particular, this would explain

the increasing rejection of the locally bound anaphors as Japanese does not allow binding to object antecedents. If changes in the test subjects' grammar are due solely to L1 influence, then the original L2 English settings for reflexive binding may be thought of as having been "deset" to the L1 setting. The attriting grammar would eventually come to take on more and more of the L1 grammar's rules, while continuing to exhibit the more robust qualities of the L2 grammar. This parameter desetting model would not rely on any speaker-internal cognitive knowledge or mechanism further than the already instantiated L1.

If however, the principle-constrained forms discovered during attrition do not exist in either the L1 or the L2, this could provide evidence that L1 transfer was not solely responsible for the new patterns which emerge as attrition sets in. For example, the data in the present study show that, although the increasing distinction made by test subjects such as R3, S2 and S3 between tensed and non-tensed complements plays no role in the interpretation of reflexive anaphors in either Japanese or English, it is consistent with the value for the governing category parameter of a possible language somewhere between them. As Finer and Broselow (1986) suggest, this intermediate setting can be found in a language such as Russian in which the finite/nonfinite distinction does play a role in determining the governing category. Progovac (1992) illustrates this distinction and shows Russian allowing binding outside an infinitival clause, but not outside a finite clause, as in (1).

(1) a. Profesor$_i$ poprosil assistenta$_j$ citat' *svoj*$_{i/j}$ doklad
professor asked assistant-ACC to-read self's report-ACC
'The professor$_i$ asked the assistant to read self's$_{i/j}$ report'

b. Vanja$_i$ znaet citat' Volodja$_j$ ljubit *svoj-u*$*_{i/j}$ zen-u
Vanja knows that Volodja loves self's-ACC wife-ACC
'Vanja$_i$ knows that Volodja$_j$ loves self's$*_{i/j}$ wife'

(Progovac, 1992: 675)

Support for this distinction and the operation of UG in second language acquisition has been discussed in Chapter 4 above. Studies such as Thomas' (1993) and MacLaughlin's (1996) re-analysis of Hirakawa's (1990) data both find evidence of proficient L2 English learners allowing long distance binding out of nonfinite clauses, but not out of finite clauses. In both studies the test subjects' L1s (Thomas = Japanese and Spanish, MacLaughlin/Hirakawa = Japanese) do not make this distinction between finite and nonfinite clauses.

In the present study, test subjects increasingly came to accept long distance antecedents in both the finite and nonfinite biclausal sentences over the course of their data collection periods. Although unacceptable in English, such long distance antecedents are in fact UG-sanctioned. Were the test subjects to only accept long distance antecedents and reject all local antecedents, a case could be made that the test subjects had violated the GCP, and therefore UG constraints. But this did not happen. Despite an increasing acceptance of long distance antecedents, never do any of the test subjects come to uniformly reject all local antecedents, although acceptance of them declines. In particular, the robustness of the test subjects' TV type 1a sentences reflects adherence to this least marked binding pattern. Only 7.5% (13 out of 172) Type 1a sentences misjudgments appear in the data, making the TV Type 1a sentences the sentence type most resistant to attrition over time. This high level of accuracy on the least marked binding pattern indicates that while all six test subjects undergo attrition in other areas, the robust nature of their TV Type 1a sentences supports Hypothesis B, that their grammars remain UG constrained while undergoing attrition. As the governing category requirement attributed to Type 1a sentences (minimal category must have a subject) is the least marked setting of the Subset Principle, it should be allowed in all languages, rendering it the least prone to attrition. The higher degree of rejection of the locally bound anaphors in the TV Type 2a sentences can be attributed to the nonfinite nature of these sentences. As mentioned earlier in Chapter 4, research (Finer and Broselow, 1986, Thomas, 1993) has shown that lower level learners tend to accept sentences of this type. Perhaps the increasing acceptance of these long-distance bound anaphors is a reversion to this earlier stage of interlanguage, and the rejection of locally

bound object anaphors is caused by increasing pressure from the L1 Japanese grammar which does not allow binding to object antecedents. Neither acceptance of the long-distance bound anaphors, nor rejection of the locally bound anaphors, however, violates UG constraints as both conditions are found in natural languages (Japanese = long-distance, English = local). Furthermore acceptance of long-distance bound anaphors dependant on a finite/nonfinite distinction as was found in this study, is also sanctioned by UG as discussed above in the example from Russian.

Hypothesis C addresses the role of age in attrition. As noted in the conclusion of Chapter 6, the sensitive period for the acquisition of reflexive binding may end by around age eight. It could be added that this sensitive period has possibly not yet begun to end at age six, as R2's reflexive binding undergoes little attrition compared to R3's. We can then think in terms of a continuum. Let us assume Long (1990) is right and that various aspects of language such as phonology, morphology, and syntax are subject to differing sensitive periods after which native-level acquisition is not possible. Such a scenario would predict that learners with similar backgrounds (learning context, age, etc) would all tend to fail to fully acquire similar aspects of the target language beyond a given sensitive period. For example, native-level phonology could be attainable if exposure starts by around six years, syntax after seven or eight, perhaps morphology after nine or ten. By puberty, enough sensitive periods would have ended so as to render a native-like acquisition of an L2 almost impossible. Considered in terms of attrition, failure to fully acquire (i.e., to a native level) a certain aspect of language by a certain age would render it and the phenomena associated with that aspect subject to later instability were target language input to become unavailable. We could ask then, by what age must L2 acquisition begin for an individual to reach what Neisser (1984: 33) refers to as a "critical threshold" of language proficiency in order to maintain various aspects of language without undergoing the types of attrition seen here? Considered as a group, all six test subjects in the present study are women of similar ages, their English learning experiences in Japan are similar, and they all had comparable levels of English upon their return from their tertiary studies in the United States. A distinguishing difference between them is age at first exposure to English, namely the three 'R' subjects had childhood exposure and the 'S' subjects did not. It seems likely, then, that R1's and R2's

relative maintenance of anaphoric reflexives can be attributed primarily to their exposure to English in an English speaking environment prior to the sensitive period for anaphoric binding, when parameters for the governing category and the proper antecedent can still be reset. The attrition observed in R3 may be attributed to her somewhat later L2 environmental exposure to English when she may have passed the sensitive period for the setting of binding parameters. S2 and S3's lack of any childhood exposure to English in the L2 environment would also mean their exposure to the L2 occurred after having passed the sensitive period for setting binding parameters. The data from these three test subjects – R3, S1 and S2 – may be indicative of individuals who have not firmly fixed the new settings of the GCP and PAP for English in their minds and quickly lose the settings in the face of zero input upon moving back to Japan, ultimately leading to their attrition of L2 English binding. S1 however, despite her lack of childhood exposure to English, outperforms not only S2 and S3 on the TV and GJ tests, but R3 as well, throughout the nine months of her data collection. S1's maintenance of her knowledge of English binding principles during this time thus suggests that she acquired those principles in a manner more akin to R1, R2 and the native speakers than to R3, S2, and S3, namely through access to UG.

### 7.1. Recommendations

L2 attrition studies emerged as a coordinated field of inquiry in the early 1980s. Anderson (1982), Lambert and Freed (1982), Preston (1982), Sharwood Smith (1983b, 1983c), Weltens et al. (1986) and others conducted controlled attrition experiments and began establishing a framework for further L2 attrition investigations. In particular, Anderson (1982) and Preston (1982) focussed on which linguistic features might be susceptible to attrition, what role UG might play in attrition, and how attrition would ultimately effect the target language. Since then, however, no study of L2 attrition to my knowledge has investigated the effects of lack of exposure on a UG constrained feature of grammar such as reflexive binding.

Further research into this field, then, is open to exploration. Remembering Schachter's (1989) comment that "(n)ot all principles are appropriate candidates for such testing since not all principles are accessible to surface level analysis" (p. 76), we can also see that not all principles are appropriate for attrition studies. But this still leaves a range of attrition phenomena to investigate. Studies into head-position, subjacency, pro-drop, and other "surface level" phenomena constrained by UG are all fertile ground for L2 attrition studies.

As noted above, the present study also incorporated tests designed to tap the subjects' L2 competence. Further test development will broaden our understanding of an attriter's diminishing competence, and confirm or challenge the present study's claim that L2 attrition of reflexive binding is UG constrained.

Appendix A: Abbreviations and terms particular to this thesis

**Data collection period:** The length of time, calculated in months and tenths of months, from the day a test subject returned to Japan, to her last test session.

**Data collection sub-period:** (initial, secondary, and final). One of three approximately four-month sub-divisions of a test subject's data collection period. S2's data collection period lasted only 9 months and thus ends with the secondary sub-period.

GJ: Grammaticality Judgment (test and test items). A data collection instrument which asks informants to judge the grammatical correctness of a set of stimulus sentences.

OA: Overall Acceptance (of local binding). An OA result measures a test subject's overall accuracy of maintenance of local binding. It is calculated by combining the test subject's results of all locally oriented biclausal sentence types on both tests that are correctly judged during a particular data collection sub-period. Sentence types included in an OR result are TV Types 1a and 2a, and GJ Types 1a, 1c, 2a, and 2c. (see Appendix L)

OR: Overall Rejection (of long distance binding). An OR result measures a test subject's overall accuracy in maintaining a rejection of long-distance binding. It is calculated by combining the test subject's results of all long-distance oriented biclausal sentence types on both tests that are correctly judged during a particular data collection sub-period. Sentence types included in an OR result are TV Types 1b and 2b, and GJ Types 1b, 1d, 2b, and 2d. (see Appendix L)

'R': Returnee Test Subjects (R1, R2, R3). These three test subjects all had childhood experience living in the L2 English environment.

'S': Student Test Subjects (S1, S2, S3). These three test subjects had no childhood experience living in the L2 English environment.

TSR: Time Since Return to the L1 Japanese environment. The TSR is calculated from the day a test subject returned to Japan, to a particular test session date. For example, test subject R2 had her fourth test session approximately three months and 24 days after returning to Japan. Her corresponding TSR number is 3.8.

TV: Truth Value Judgment (test and test items). Data collection instrument which presents an informant with a short two to three sentence scenario. Informants must then judge whether the corresponding stimulus sentence accurately summarizes the content of the scenario.

Appendix B: Truth Value Judgment Master List Key.

This table allows the reader to find where specific TV test items appeared in the TV test and to locate each item in the master list of TV test items in Appendix C. Appendix C is divided into six sections, one for each sentence type. The column on the left indicates both the item number on the tests the informants took, and the number of that item in Appendix C. The top row indicates the TV sentence type, and in parentheses, the corresponding section of Appendix C in which sentences of that type are included. For example, on TV Test 5, item numbers 1, 16, and 11 are TV Type 1b sentences and can be found in Appendix C2, item numbers 1, 13, and 14.

| Type: | 1a (C1) | 1b (C2) | 2a (C3) | 2b (C4) | 3a (C5) | 3b (C6) |
|---|---|---|---|---|---|---|
| Test & Master No./ | | | | | | |
| TV Test 1 No: | 11, 5,13 | 17, 1, 12 | 14, 4, 3 | 10, 2, 16 | 18, 6, 8 | 9, 15, 7 |
| Master No: | 1, 2, 3 | 1, 2, 3 | 1, 2, 3 | 1, 2, 3 | 1, 2, 3 | 1, 2, 3 |
| TV Test 2 No: | 7, 11, 5 | 13, 17, 1 | 12, 14, 4 | 3, 10, 2 | 16, 18, 9 | 8, 6, 15 |
| Master No: | 4, 5, 6 | 4, 5, 6 | 4, 5, 6 | 4, 5, 6 | 4, 5, 6 | 4, 5, 6 |
| TV Test 3 No: | 13, 9, 1 | 12, 11, 4 | 3, 17, 2 | 16, 14, 6 | 8, 10, 15 | 7, 18, 5 |
| Master No: | 7, 8, 9 | 7, 8, 9 | 7, 8, 9 | 7, 8, 9 | 7, 8, 9 | 7, 8, 9 |
| TV Test 4 No: | 1, 12, 11 | 4, 3, 17 | 2, 16, 14 | 6, 8, 10 | 15, 13, 18 | 5, 7, 9 |
| Master No: | 10, 11, 12 | 10, 11, 12 | 10, 11, 12 | 10, 11, 12 | 10, 11, 12 | 10, 11, 12 |
| TV Test 5 No: | 4, 13, 17 | 1, 16, 11 | 6, 3, 10 | 2, 7, 14 | 5, 8, 9 | 15, 12, 18 |
| Master No: | 1, 13, 14 | 1, 13, 14 | 1, 13, 14 | 1, 13, 14 | 1, 13, 14 | 1, 13, 14 |
| TV Test 6 No: | 15, 16, 18 | 6, 13, 10 | 1, 7, 11 | 5, 3, 9 | 2, 12, 14 | 4, 8, 17 |
| Master No: | 2, 15, 16 | 2, 15, 16 | 2, 15, 16 | 2, 15, 16 | 2, 15, 16 | 2, 15, 16 |
| TV Test 7 No: | 8, 6, 15 | 10, 18, 11 | 13, 5, 1 | 9, 7, 12 | 3, 4, 2 | 17, 14, 16 |
| Master No: | 3, 17, 18 | 3, 17, 18 | 3, 17, 18 | 3, 17, 18 | 3, 17, 18 | 3, 17, 18 |
| TV Test 8 No: | 2, 7, 14 | 5, 8, 9 | 15, 12, 18 | 4, 13, 17 | 1, 16, 11 | 6, 3, 10 |
| Master No: | 4, 19, 20 | 4, 19, 20 | 4, 19, 20 | 4, 19, 20 | 4, 19, 20 | 4, 19, 20 |
| TV Test 9 No: | 14, 10, 8 | 7, 15, 5 | 18, 9, 13 | 12, 1, 4 | 11, 17, 3 | 16, 2, 6 |
| Master No: | 5, 21, 22 | 5, 21, 22 | 5, 21, 22 | 5, 21, 22 | 5, 21, 22 | 5, 21, 22 |
| TV Test 10 No: | 16, 15, 6 | 18, 10, 13 | 7, 1, 5 | 11, 9, 3 | 12, 2, 4 | 14, 17, 8 |
| Master No: | 6, 23, 24 | 6, 23, 24 | 6, 23, 24 | 6, 23, 24 | 6, 23, 24 | 6, 23, 24 |
| TV Test 11 No: | 4, 13, 17 | 1, 16, 11 | 6, 3, 10 | 2, 7, 14 | 5, 8, 9 | 15, 12, 18 |
| Master No: | 7, 25, 26 | 7, 25, 26 | 7, 25, 26 | 7, 25, 26 | 7, 25, 26 | 7, 25, 26 |
| TV Test 12 No: | 12, 1, 4 | 11, 17, 3 | 16, 2, 6 | 14, 10, 8 | 7, 15, 5 | 18, 9, 13 |
| Master No: | 8, 27, 28 | 8, 27, 28 | 8, 27, 28 | 8, 27, 28 | 8, 27, 28 | 8, 27, 28 |
| TV Test 13 No: | 11, 9, 3 | 12, 2, 4 | 14, 17, 8 | 16, 15, 6 | 18, 10, 13 | 7, 1, 5 |
| Master No: | 9, 29, 30 | 9, 29, 30 | 9, 29, 30 | 9, 29, 30 | 9, 29, 30 | 9, 29, 30 |

TV1a: Biclausal sentences with finite verb in the embedded clause. Pragmatically favor local antecedent interpretation. Bare infinitive distractors marked with asterisk.

(1) Whales are very good swimmers, but during the summer, several whales swam onto the beach and died. It was a mystery why they did it.
*No one could understand why the whales killed themselves.*

(2) The wolf tried to catch the goat, but the goat ran up a steep mountain. So the wolf got tired of running and stopped.
*The wolf gave up because the goat could save itself.*

(3) The young woman was very determined to go to university. Her father had died and her mother didn't have much money. So the young woman worked hard and saved enough money to enter a famous university and finally graduate.
*The mother knew her daughter put herself through university.*

(4) Alan and his friend were athletes at different universities. One day, they were scheduled to race against each other, but Alan's friend didn't want to enter the race.
*Alan was surprised his friend removed himself from the race*

(5) One day, Susan went to a park by herself. She was a little scared. Suddenly she saw a girl fall off her bicycle and cut her leg.
*Susan saw a girl hurt herself.**

(6) One evening, John's wife went to see a movie and John took care of their baby son. The baby cried for a long time, but finally became quiet and slept.
*The father knew his son had cried himself to sleep.*

(7) Alfred took a computer course. The teacher was very good because he didn't do the work for the students, he helped them learn things on their own.
*The computer instructor helped Alfred teach himself the new software**

(8) Bill had a party for his friends. Sam came and was very hungry. Bill watched Sam eat all the cake.
*Bill knew Sam helped himself to a lot of food.*

(9) Alexandra went to Rhona's birthday party and met a lot of new people.
*Rhona watched Alexandra introduce herself to her friends.**

(10) Stew and Tony were soldiers in the jungle. They were looking for the enemy. Suddenly a hand grenade was thrown at them. Tony lay on the grenade to save Stew. The grenade exploded and Tony was killed.
*Stew saw Tony sacrifice himself. **

(11) Amy lost her wedding ring and was very angry. Her mother Alice told her not to worry, but Amy was still upset and banged her head against the wall.
*Alice watched Amy hurt herself.**

(12) Mike's father was very old and forgetful. He had trouble remembering appointments and so on. So he typed out reminders and taped them on his bedroom door.
*Mike knew his father wrote himself messages.*

(13) A hunter was in the forest trying to catch a monkey. The monkey kept watching the hunter and the hunter couldn't get him. Finally the hunter got so depressed he jumped in the river and died.
*The monkey saw the hunter kill himself.* *

(14) The old writer was very famous for his short stories. But recently he couldn't think of interesting things to write about. So he told his publisher he had become a bad writer.
*The publisher heard the writer criticize himself.**

(15) During the election many politicians gave speeches on the TV and radio. They described all the good things they would do for the country.
*The people heard the politicians praise themselves.**

(16) The CIA officer captured a foreign spy who knew many secrets. The spy drank some poison and died.
*The officer knew the spy killed himself.*

(17) During the night a fire broke out in Alex's friend Bill's house. Alex wondered if Bill was OK.
*Alex wondered if Bill had saved himself.*

(18) The telephone company had grown too big and the government wanted to divide it into several smaller companies. The telephone company didn't like the idea and tried to fight the government in court.
*The government listened to the company defend itself.**

(19) Protestors tried to prevent the nuclear plant from opening. The police saw them use handcuffs to attach themselves to the gate.
*The police knew the protestors handcuffed themselves to the power plant.*

(20) During a recent NASA test, a space rocket went out of control. It automatically blew up.
*The NASA scientists knew the rocket destroyed itself.*

(21) The crazy woman was in the hospital. She was always talking, even when no one was listening. The nurse tried to help her, but couldn't.
*The nurse often heard the patient talk to herself.**

(22) Last week at the zoo, the keepers were surprised to see one of the lions bite itself. This is a very unusual thing for lions to do.
*Zookeepers know lions don't usually bite themselves.*

(23) An alien spaceship flew low over a village. The villagers were afraid and watched it fly by. Suddenly it crashed into a mountain and exploded. All the aliens on board were killed.
*The villagers were relieved that the aliens destroyed themselves.*

(24) The computer hacker designed a virus to erase hard disks automatically. He tested it on an old computer and it worked well. He was happy.
*The hacker realized the computer's hard disk erased itself.*

(25) John and Bill were protesting against the new nuclear power plant. Bill thew some rocks at the police. The police tried to catch him and finally he gave up and let the police arrest him.
*John saw Bill surrender himself to the police.**

(26) The busy young father wanted to spend time with his son, so he let his son watch him shave every morning.
*The boy watched his father shave himself.**

(27) Jason's son had a credit card and bought many expensive things. Jason knew because he got the credit card bill.
*Jason knew his son treated himself to expensive things.*

(28) Jane was a famous musician. Only her friend Susan knew that Jane never went to music school and had learned to play at home without a teacher.
*Susan knew Jane had taught herself to play music.*

(29) Julia was a high school student. Rose was her math teacher. Rose thought she was a great teacher and often told her students about all the great things she did.
*Julia often heard her teacher praise herself.**

(30) The poor farmers couldn't afford expensive equipment. The rich landowners, however, had enough money to buy a tractor for their own use.
*The farmers realized the landowners bought themselves a new tractor.*

TV1b: **Biclausal sentences with finite verb in the embedded clause. Pragmatically favor long-distance antecedent interpretation. Bare infinitive distractors marked with asterisk.**

(1) Mo was taking pictures. He wanted to take Curly's picture. So he tried to get Curly's attention.
*Mo made Curly look at himself.* *

(2) The lion ran to catch the gazelle, but the gazelle was able to run away. But then the gazelle tripped and fell so the lion was able to get closer.
*The gazelle realized that the lion would kill itself.*

(3) The old couple didn't have much money. So their children helped them start an internet company. They gave advice to young people about how to be happy. They became wealthy and enjoyed life.
*The old couple was happy because their children helped themselves become rich.*

(4) Jim went to a big public school There were lots of drug users and students with guns and knives. But he had lots of friends and enjoyed the school. Jim's father, though, was worried and finally made Jim change schools.
*Jim was surprised because his father removed himself from the school.*

(5) One day, John was delivering letters when a big dog ran at him. John tried to stop the dog, but it was too late. The dog grabbed John's leg with his teeth.
*John knew the dog bit himself on the leg.*

(6) Trevor was in a hurry. He drove very fast down the highway. Suddenly he saw a police car in his mirror with its lights flashing.
*Trevor knew the police officer saw himself speeding.*

(7) During the war, one computer controlled many guns. So the enemy aimed a missile directly at the computer to blow it up. The computer was very smart, and knew the missile was coming, but couldn't stop it.
*The computer knew the enemy would destroy itself.*

(8) Tim's math teacher knew Tim couldn't understand the homework very well. So he asked Tim to stay after class and reviewed the main points of the exercises.
*Tim was glad the teacher helped himself with the homework.*

(9) Mr. Smith owns a beer store in the city. One day a man with a gun came in and made Mr. Smith give him all the money in the register. Then the robber threw a can of beer at Mr. Smith's head.
*Mr. Smith realized the thief hit himself with the beer.*

(10) Frank took piano lessons every day. Sometimes he practiced hard, other times he was lazy and didn't practice. But his teacher always tried to make Frank play well. Finally Frank performed in a big concert and became famous.
*Frank knew his teacher had trained himself to be a great piano player.*

(11) Anne and her sister were driving in the rain. Suddenly the car went out of control and turned over. Anne was badly hurt, but luckily her sister was OK and pulled Anne to safety.
*Anne was saved because her sister pulled herself from the car.*

(12) Martin and George were brothers. They often made up stories about each other. Once, Martin told their mother that George had destroyed the neighbor's garden shed. George was shocked at the story.
*George couldn't believe what Martin said about himself.*

(13) Ben was trying to get a job at John's company. John asked some of Ben's friends if Ben was a good worker. Ben heard about this.
*Ben knew John asked about himself.*

(14) When Richard was a student he stole some information from his classmate. His classmate found out and got angry at Richard.
*Richard wondered if his classmate trusted himself anymore.*

(15) Jules worked as a salesman at a farm equipment showroom. One day a man came in and wanted to buy a tractor. Jean didn't know about the tractor and made up a story about how great it was. The man wasn't convinced and told Jules.
*Jules knew the customer didn't trust himself.*

(16) Ben was an old dog. He couldn't hear and he was very weak. His owner was very sad, but knew the dog's life had to end. He took a gun and pointed it at Ben.
*The dog knew his owner would kill himself.*

(17) Bill had a lot of money hidden in a book. One day the money disappeared and Bill asked his roommate John if he knew anything about the money.
*John listened to Bill ask himself about the money.* *

(18) Ralph was very sick. Sometimes he went crazy and attacked people for no reason. He went to the hospital for help and told the doctor about his sickness.
*Ralph hoped the doctor would cure himself.*

(19) Recently, there has been a growing number of gangs in the city and the city mayor ordered the police to stop them. So the police made a secret plan to raid the gangs' headquarters, but the gangs found out about the plan.
*The gangs knew the police would attack themselves.*

(20) A group of researchers planned to make a new laboratory, but they needed the government to agree to the plan. So they made a presentation to a group of local politicians about their plan.
*The researchers hoped the politicians would agree with themselves.*

(21) One day, Alice went to school with Jim. While they were walking, Alice's mother drove by. She looked at Alice and Jim.
*Alice knew her mother saw herself with Jim.*

(22) Rex and Susan are famous. They appear on TV quite often. One day Rex and Susan's children were watching TV and saw their parents. They told Rex and Susan about it.
*The parents knew their children saw themselves on TV.*

(23) There were many sick people in the hospital waiting for a special medicine. The doctors were trying very hard to give the medicine to all the patients so they would get better.
*The patients hoped the doctors would give themselves the medicine*

(24) Jack was walking in the forest during hunting season. He was worried a hunter might think he was a deer and shoot him. Suddenly he heard a shot and a bullet hit him in the shoulder.
*Jack wondered if the hunter shot himself on purpose.*

(25) Rita took swimming lessons every day. Sometimes she practiced hard, other times he was lazy and didn't practice. But her teacher always tried to make Rita swim well. Finally Rita swam in a big race and became famous.
*Rita knew the teacher trained herself to be a great swimmer.*

(26) Rachel had an important interview one morning, but she didn't hear her alarm clock. Her sister Becky woke her up just in time to get to the interview.
*Rachel was glad her sister woke herself up.*

(27) Richard worked for the government. Henry was a newspaper reporter. Henry wrote an article that said Richard was a liar and a cheat.
*Richard couldn't believe what Henry wrote about himself.*

(28) Juan thought Jeff was talking to him, but actually Juan was talking to Peter.
*Juan thought that Jeff was talking to himself.*

(29) Sean owned a small law firm with a partner. Sean stole some money from the firm and his partner found out. The partner was very angry.
*Sean wondered if his partner trusted himself anymore.*

(30) It was Christmas and Alice wanted a faster computer. She asked her mother to buy her a new one.
*Alice hoped her mother would buy herself a new computer.*

**TV2a: Biclausal sentences with nonfinite verb in the embedded clause. Pragmatically favor local antecedent interpretation.**

(1) Arnold always asked George what the time was. George got tired of this and helped Arnold buy a watch.
*George wanted Arnold to buy himself a watch.*

(2) The police arrested three terrorists. The terrorists said they wouldn't eat or drink until the government stopped bombing their country. The police hoped terrorists would die.
*The police wanted the terrorists to starve themselves to death.*

(3) Alice and Louise went to the beach. Louise didn't want to get a sunburn, so Alice told her to put the beach blanket over her legs.
*Alice told Louise to cover herself with the blanket.*

(4) Al usually drove Jim to school. But sometimes Al rode his bicycle and then Jim had to walk. Jim often told Al he should use his car every day.
*Jim wanted Al to drive himself to school every day.*

(5) Last year a dangerous criminal escaped from prison. The police went to every house in the neighborhood and warned people to be careful.
*The police told the people to lock themselves in their houses.*

(6) Johnny was only 16, but he looked much older so it was easy for him to buy alcohol. One day Johnny went to buy a quart of whiskey, but as he was paying for it his father came into the shop.
*His father caught Johnny buying himself some whiskey.*

(7) John went to buy a new TV. The salesman connected a video camera to a TV and pointed the camera at John. John could see his face on the TV.
*The salesman wanted John to see himself on TV*

(8) Alan usually drove to work in his Lexus. His friend Rex usually took the train. One day Alan was surprised to see Rex behind the wheel of a red Mercedes Benz.
*Alan was surprised to see Rex driving himself to work.*

(9) A young woman went to see a counselor because her husband drank too much and hit her sometimes. The counselor advised the woman that she should stay in a closet when her husband was drunk.
*The counselor told the woman to hide herself in a closet.*

(10) Three year old Alice May was so hungry she tried to eat her arm! Her mother thought it was funny, but when Alice May started bleeding, her mother finally gave her some food.
*The mother saw the child biting herself.*

(11) Gustav was a spy. He was hiding from the enemy agents in a farmer's barn. The agents were getting closer. The farmer told Gustav to leave at once or he would be shot.
*The farmer ordered the spy to save himself by running away.*

(12) Jack lived in a dangerous neighborhood. Once his son was robbed, so Jack bought him a gun and carefully showed him how to use it.
*Jack trained his son to protect himself with a gun.*

(13) Maurice was in bad shape. He didn't believe he could do anything right. So he went to therapy sessions. The counselor was very helpful and positive.
*The counselor told Maurice to trust himself.*

(14) Because of all the pollution, the ozone layer was destroyed and the suns harmful rays gave many people cancer. The government issued warnings that people should wear hats and use sunblock cream.
*The government warned the people to protect themselves from the sunshine.*

(15) The magician could twist his body into many strange shapes. He trained his son to imitate whatever shape he made himself into. Once he looked like a pretzel.
*The magician trained his son to twist himself into a pretzel shape.*

(16) A group a soldiers led 3 prisoners into the forest. They gave them shovels and told them to dig holes in the ground. The prisoners knew the soldiers would shoot them after they dug the holes.
*The soldiers ordered the prisoners to dig themselves graves in the forest.*

(17) Latisha and Evan were teachers at a local school. The students were not behaving well and wouldn't line up for lunch. Finally Latisha and Evan got angry and yelled at them.
*The teachers screamed at the students to form themselves into straight lines.*

(18) The captain's ship was crossing the ocean in the winter. The seaman steering the ship had been awake for 20 hours- he was cold and tired.
*The captain told the seaman to get himself some coffee.*

(19) After the football team lost many games, the coaches decided the team needed to do more training.
*The coaches asked the team members to train themselves more thoroughly.*

(20) When classes began in September, none of the students had a notebook. The teachers told them to each go to the bookstore and buy one.
*The teachers wanted the students to buy themselves notebooks.*

(21) When Jack and Ruby came home, their front door was open. When they went inside, all their money and jewelry were gone. Suddenly they saw two men go out of their house and hide under a car.
*Jack and Ruby saw the robbers trying to hide themselves.*

(22) Billy often hit other students and said bad words. One day he was very bad and the teacher told him to go see the principal.
*The teacher told the student to report himself to the principal's office.*

(23) Alex had a problem with his car- it often caught fire. Last Thursday he had quite a bad fire and took the car to the garage. The mechanic said Alex's car was too dangerous to drive and he had to get another car.
*The mechanic advised Alex to buy himself a new car.*

(24) Lydia and Suzanne each had a baby boy. One evening they visited each other and the two babies cried a lot, but the mothers didn't feed their babies because they wanted them to fall asleep.
*The mothers wanted their children to cry themselves to sleep.*

(25) John worked at the stock exchange. One day there was a big crash and a lot of people lost money. One man became so depressed he tried to jump out of a window, but the window was locked.
*John saw the worker trying to kill himself.*

(26) During the terrible storm the lighthouse keeper saw a man in a small boat. The man was trying to prevent his boat from crashing onto the rocks.
*The lighthouse keeper watched the man trying to save himself.*

(27) The computer controlled the new robot. One day the robot went out of control and began drinking whiskey. The computer issued the self-destruct command to the robot.
*The computer ordered the robot to destroy itself.*

(28) Wendy noticed her friend Olivia drove very slowly and was afraid to drive on the highway. Wendy tried to encourage her to be a more confident driver.
*Wendy wanted Olivia to trust herself more.*

(29) The crime rate recently rose in the city. The police advised people to carry a gun when they went out during the day and to stay indoors at night.
*The police warned the people to protect themselves against crime.*

(30) The children went to the circus. There were many clowns doing silly things and laughing at their own mistakes.
*The children saw the clowns making fun of themselves*

**TV2b: Bioclausal sentences with non-finite verb in the embedded clause. Pragmatically favor long-distance antecedent interpretation.**

(1) Ed was a customer at a fancy restaurant. He waited a long time for the waiter to take his order. Finally he got up and asked the waiter to take his order.
*Ed asked the waiter to serve himself.*

(2) A boy got lost in the big city. He didn't know how to get back home. Finally he saw a police officer and asked him for directions.
*The child asked the police officer to help himself.*

(3) Tony went hiking in the mountains. Suddenly, he fell down and broke his leg. He lay on the ground for a long time. He got cold. Finally his friend Jeff found him. Tony asked for a blanket.
*Tony told Jeff to cover himself with a blanket.*

(4) Kim's teacher Ms. Wright was famous because she helped sick animals. Once a TV program wanted to interview Ms. Wright about her work. Ms. Wright agreed to appear on TV.
*Ms. Wright told Kim to watch herself on TV.*

(5) Helen was very sick and was in a hospital. She was too weak to take a shower or bath, so she asked the nurse to help her.
*Helen asked the nurse to wash herself.*

(6) Jill couldn't understand her homework assignment so she asked her friend to review the lesson with her.
*Jill asked her friend to help herself.*

(7) Karen wanted to see her hands, but it was too dark. So she asked Sharry to point her flashlight at her.
*Karen told Sharry to shine the light on herself.*

(8) A reporter asked a young politician named Larry for an interview. Larry agreed, but only if the interview appeared on TV.
*Larry wanted the reporter to film himself.*

(9) Frank needed a photograph for his passport, so he went to a photo studio to get his picture taken.
*Frank asked the studio owner to take a picture of himself.*

(10) Rick was very sick, so he went to a hospital for treatment.
*Rick expected the doctor to take care of himself.*

(11) During the war, an army captain ordered a soldier to attack the enemy, but the soldier was afraid and didn't' want to do it. The captain told him he had to do what he said.
*The captain required the soldier to obey himself*

(12) Jonathan was failing math, so he took some notes to a math test and hid them in his shirt. The student next to him saw Jonathan look at his notes during the test. Jonathan got angry at the student.
*Jonathan warned the student not to accuse himself of cheating.*

(13) After the plane took off, a hijacker attacked the pilot. The hijacker had a gun and pointed it at the pilot.
*The pilot asked the hijacker not to shoot himself.*

(14) The garbagecollector wasn't very happy about his job- it was very dirty and he didn't get paid much. So he went to city hall to see the mayor. The worker wanted the mayor to give him a new job.
*The cityworker wanted the mayor to find himself a new job.*

(15) Jeff and Rasheem were old friends. But now Rasheem was president of a big company and Jeff had no job. So Jeff went to see Rasheem about getting a job at his company.
*Jeffwanted Rasheem to hire himself*

(16) The waves were very high at the beach. One swimmer became tired and got swept out to sea. He called to the lifeguard to help him.
*The swimmer screamed at the lifeguard to rescue himself.*

(17) Mary went to a restaurant for dinner. The restaurant was very busy and Mary waited a long time, but still the waitress didn't come to her table. Finally Mary called to the waitress and asked if she could order.
*Mary asked the waitress to serve herself.*

(18) Vanessa was very busy and had a lot of stress. When she went to get her hair done, she decided to get a massage there too.
*Vanessa asked her hairdresser to massage herself.*

(19) The chorus group was giving their first performance in a new city. Not many people knew them, so they requested that the organizers introduce them.
*The singers told the organizers to introduce themselves.*

(20) Don was very rich, but he had many enemies. So he decided to hire a full time guard.
*Don wanted the guard to protect himself.*

(21) Wally was tired of robbing people- he was sorry for all the bad things he had done. So he told a police officer who he was and asked to be put in jail.
*The criminal wanted the police officer to arrest himself.*

(22) Yesterday was Bobbie's birthday. He thought his roommate Jack would have organized a surprise birthday party for him, but there was no party. Bobbie was disappointed.
*Bobbie had expectedJack to surprise himself.*

(23) Sally escaped from jail and went to see her old friend Becky. She knew Becky could keep her secret.
*Sally trusted Becky not to tell anyone about herself.*

(24) Winston was Abdul's body guard. But Winston didn't have a gun to protect Abdul, so Abdul gave him one.
*Abdul gave Winston a gun to protect himself.*

(25) Jack went to visit his friend Bruno. But Bruno was dead- he had been shot. Just then the police arrived. It appeared that Jack had killed Bruno.
*Jack expected a police officer to question himself.*

(26) One night there was no dinner for the rich lady because the cook had run away. So the rich lady made the house servant cook her meal for her.
*The woman ordered the servant to serve herself.*

(27) Anthony forced local shop owners to pay him money every month. One day the owner of the bakery said he would tell the police about Anthony. Anthony told him not to.
*Anthony warned the baker not to reveal himself to the police.*

(28) Terry worked at a bank and lost a lot of the customers' money. He went crazy and decided to rob the bank. The bank guard pointed his gun at Terry and told him to stop, but Terry said he would rather die.
*Terry wanted the guard to shoot himself.*

(29) Jill got married and had a baby boy. But Jill's mother didn't like Jill's husband and never talked about Jill's family. This made Jill sad.
*Jill wanted her mother to talk to herself about her family.*

(30) Tracy loaned Hillary $1,000 to buy an old car. When Tracy needed the money back to pay her bills she asked Hillary to return it.
*Tracy wanted Hillary to give herself the money.*

TV3a: **Monoclausal sentences with subject and object NPs. Pragmatically favor object antecedent interpretation.**

(1) On the first day of school the teachers wanted the students to relax, so they had an open meeting. The students could ask any question they wanted.
*The students asked the teachers about themselves.*

(2) Arnie was sick so he went to see a doctor. The doctor asked Arnie many questions about what he ate, how much he exercised, and how much beer he drank.
*The doctor asked Arnie about himself.*

(3) Rachel was an old woman. She thought she might die soon and wanted her daughter Jenny to know that Jenny had been adopted. So one day she talked to Jenny and told her. Rachel told Jenny about Jenny's real parents and where she had come from. Jenny was shocked.
*Rachel told Jenny about herself.*

(4) Mr. Lynch owned a vegetable shop. One day he saw a boy stealing some carrots. Mr. Lynch caught the boy. He made the boy tell him why he stole the carrots, where he lived, and who his parents were.
*Mr. Lynch questioned the boy about himself.*

(5) Larry was a witness in a murder case. He thought he had seen the killer. In court a Lawyer asked him questions about his eyesight, and how much alcohol he had drunk that night.
*The lawyer cross-examined Larry about himself.*

(6) Jerry was a lawyer in court. He wasn't very good and often made mistakes. At the end of the day the judge asked Jerry where he had studied law and what experience he had.
*The judge asked Jerry about himself.*

(7) Jorge went to a party and met an interesting man. Jorge asked him many questions.
*Jorge asked the man about himself*

(8) Jennifer wanted her picture taken standing next to the big bull on Wall Street in New York. But her camera was broken. So she asked her friend Susan to take the picture. Susan took the picture with her own camera and said she would send the picture to Jennifer.
*Susan promised Jennifer the picture of herself.*

(9) Naomi was very sick when she was a child. Now she was a doctor and wanted to know more about her childhood sickness. She asked her mother to send her a letter about it. Her mother responded right away.
*Naomi's mother wrote to her about herself.*

(10) Richard was a professional photographer. He took a picture of Cliff and Cliff bought it for $50.00.
*Richard sold Cliff a picture of himself.*

(11) Robbie was quite sick and went to the hospital. He didn't trust the doctor who saw him, so Robbie asked him about his training and medical experience.
*Robbie questioned the doctor about himself.*

(12) A group of people was visiting town. They parked their van in a neighborhood garden. The neighbors got upset and asked them why they parked in the garden and when they would leave.
*The neighbors questioned the visitors about themselves.*

(13) Sam was a famous old blues singer. He was looking for a copy of the first recording he ever made. His friend Harold had a copy of the recording and sold it to Sam for one dollar.
*Harold sold Sam a recording of himself playing the blues.*

(14) The children hired a famous painter to paint their parents' picture from a photograph. They gave the picture to their parents on their wedding anniversary.
*The kids gave their parents a painting of themselves for their anniversary.*

(15) After they lost the football game, the team had a meeting with their coaches. The coaches explained what the players did wrong and why they lost the game.
*The coaches told the players about themselves.*

(16) The bankers didn't want to lend the farmers money to buy land. They told the farmers that they were too old, and that they didn't have enough farming experience to successfully run the farm.
*The bankers spoke to the farmers about themselves.*

(17) Hannah's sister was out all night and didn't come home the next day either. Hannah was worried. When her sister finally came home, she didn't explain where she'd been, so Hannah asked her.
*Hannah questioned her sister about herself.*

(18) During the trial, the witness gave confusing answers. So the judge became upset and began directly asking her questions about where she had been and whom she had been with.
*The judge cross-examined the witness about herself.*

(19) Dick got into trouble because he didn't pay his taxes. So he asked a lawyer for a report about his situation.
*Dick's lawyer provided him with a report about himself.*

(20) Donald didn't believe the president and requested a detailed report about where he had been, who he had seen, and how much money he had spent.
*Donald petitioned the president for a report about himself.*

(21) Several women wanted to find out how to care for their new babies. So they asked their grandmothers questions about their experiences of taking care of babies.
*The new mothers asked the grandmothers some questions about themselves.*

(22) Harrison was having some mental problems and went to see a psychologist many times. At the end of the treatment, the psychologist wrote Harrison a report about his condition.
*The psychologist wrote Harrison a description of himself.*

(23) When June was young, Alice took a picture of June and gave it to her. After June died, Alice asked June's husband for the old photograph.
*Alice wanted June's photograph of herself.*

(24) Marcus was having trouble studying so he asked the school counselor for some advice. The counselor asked some questions then gave Marcus some advice about how to improve his study habits.
*The counselor spoke to Marcus about himself.*

(25) Rod wanted a statue of his head and shoulders, so he hired a famous sculptor to make it.
*The sculptor made Rod a statue of himself.*

(26) Maurice the crook was about to be released from prison. Before he was released, he was given a summary of all his crimes and how much time he had spent in prison.
*A prison guard handed Maurice a report about himself*

(27) Jimmy hated school. He didn't get good grades and often caused trouble. The principal got angry and sent him an e-mail and described all the bad things he had done.
*The principal e-mailed the student about himself.*

(28) Harriet liked to take pictures of her younger sister when they were children. When Harriet died, she stated in her will that her sister should have the pictures.
*Harriet left her sister some old photos of herself as a child.*

(29) Richard heard that his good friend Trevor had won a lot of money. Richard sent Trevor an e-mail about this.
*Richard sent Trevor some good news about himself.*

(30) The three astronauts visited a school and encouraged the children to study hard, obey their parents and teachers, and to always try to do the right thing.
*The astronauts talked to the children about themselves.*

**TV3b: Monoclausal sentences with subject and object NPs. Pragmatically favor subject antecedent interpretation.**

(1) Mo wanted a job at the drugstore, so he told the owner about his background and skills. He told the drugstore owner many things.
*Mo told the drugstore owner about himself.*

(2) Dawn lived alone after her husband died. For many years, Dawn regularly wrote letters to her friend Jean, telling her what she did every day, where she went, and how she felt.
*Dawn wrote to Jean about herself.*

(3) When Hugh's father was a young man, he murdered someone. He never told anyone and no one ever found out. But now Hugh's father was an old man and thought he might die soon. He wanted Hugh to know what had done. Hugh was shocked.
*Hugh's father told Hugh about himself.*

(4) Jack and his friend Sam were skiing in the mountains. Jack lost control and crashed into a tree. He was knocked out. When he was conscious again he couldn't remember anything about what had happened, so he asked Sam where he was, how he got there and what had happened.
*Jack questioned Sam about himself.*

(5) Jim wanted to know what his father was like as a child, so he asked him. He found out that his father had been a bad boy and made life very hard for his parents.
*Jim's father told Jim about himself.*

(6) Terry was a mad scientist. He wanted computers and monkeys to communicate with each other. Terry tried to make a monkey understand a computer. The computer asked the monkey many questions about how fast the computer was, what color it was, and where it was.
*The computer asked the monkey questions about itself.*

(7) Marvin was a famous musician and he loved his audiences. After one concert, he told a fan he would send him a copy of his most recent CD.
*Marvin promised the fan a recording of himself.*

(8) Judy was a tourist in Greece. At the airport she was asked many questions by the immigration official. She answered all the questions.
*Judy provided the immigration officialwith information about herself.*

(9) Tim and Tom were twins. They didn't like the presents their parents usually gave them on their birthday. So one year, they bought their own presents and gave them to their parents to give them.
*The kids gave their parents presents for themselves.*

(10) Rebecca was a famous artist, but she was quite selfish. Once she painted a beautiful picture of her sister Jane, but she never let Jane see the picture. Jane got angry and took it.
*Jane stole Rebecca's painting of herself*

(11) Trevor finished his first year at his new job. His boss wanted a report about the work Trevor had done during the year. When Trevor finished the report he gave it to his boss.
*Trevor handed his boss a report about himself*

(12) Victor's usual psychologist killed himself, so Victor had to find a new psychologist. At the first interview, the new psychologist wanted to listen to Victor tell him about his background.
*Victor told the psychologist about himself.*

(13) Janice was curious about what she was like as a little girl, so she talked to her mother about her childhood.
*Janice asked her mother about herself.*

(14) Jimmy was in prison again. He wrote to his lawyer and told him why he was in trouble this time.
*The prisoner wrote to his lawyer about himself.*

(15) June and Sue were sisters. Sue always forgot June's birthday. So one year June bought some flowers for Sue to give to her.
*June gave Sue flowers for herself.*

(16) At the end of every school year, teachers should give out student reports. But one year Ronald's teacher forgot, so he asked the teacher for the report.
*Ronald asked the teacher for the report about himself.*

(17) Harry was an amateur guitarist- he played with his friends. Once he made a recording of his band and gave it to his brother Jimmy.
*Harry gave Jimmy a recording of himself.*

(18) At the start of the new year, the worker expected to make a new contract with his company and wanted to request a bonus. He talked to his supervisor about it.
*The worker asked his supervisor about a bonus for himself.*

(19) Tony felt guilty because he never went to synagogue anymore. He went to see a rabbi and tried to explain why he didn't go.
*Tony spoke to the rabbi about himself.*

(20) Susan and Diana were friends and worked together. Some people at work didn't like Susan and said bad things about her. Diana heard some of the bad things, and Susan wanted to know what they said, so she asked Diana.
*Susan asked Diana about herself*

(21) Jack and Betty were just married. Betty's parents said they wanted a picture of Jack and Betty on their honeymoon, so Jack and Betty said they would send one.
*Jack and Betty promised her parents a photograph of themselves.*

(22) When he was young, a famous artist made a drawing of James and gave it to him. Later, when James was old and needed money, he sold the sketch to an art dealer.
*James sold the art dealer an old drawing of himself.*

(23) Teddy was a criminal. The police wanted him and hung posters of him all over the city. His friend Alex stole one of the posters and Teddy wanted it as a souvenir.
*Teddy wanted Alex's poster of himself.*

(24) Phillip and Rodney were design artists. They hoped to get a contract at a new company. At the interview the company representatives listened to Phillip and Rodney talk about their skills and experiences.
*Phillip and Rodney spoke to the representatives about themselves.*

(25) Veronica painted a beautiful picture of Judy and gave it to her. Judy loved the painting, but eventually needed money and had to sell it.
*Judy sold Veronica's portrait of herself.*

(26) Steve and Greg were friends. Steve was a little crazy and Greg was a psychologist. Greg decided to write a book about Steve and Steve bought a copy when it was published.
*Steve bought Greg's book about himself.*

(27) Igor loved his friend Vladimir and wrote a poem about him. Vladimir liked the poem so much he memorized it and would recite it whenever he could.
*Vladimir recited Igor's poem about himself many times.*

(28) Rosie wasn't married yet, but she was expecting a child. She wanted her mother to know about her situation so she could get some advice.
*Rosie told her mother about herself.*

(29) Sam wanted to attend Everclear College, so he wrote to the admissions officer and mentioned his school background, interests and reasons why he wanted to attend Everclear.
*Sam wrote to the admissions officer about himself.*

(30) William wrote a famous poem about Dan's life, but Dan never liked the poem very much.
*Dan disliked William's poem about himself.*

Appendix D: Grammaticality Judgment Master List Key.

This table allows the reader to find where specific GJ test items appeared in the GJ test and to locate each item in the master list of GJ test items in Appendix E. Appendix E is divided into twelve sections, one for each GJ sentence type. The column on the left indicates both the item number on the tests the informants took, and the number of that item in Appendix E. The top row indicates the GJ sentence type and in parentheses, the corresponding section of Appendix E in which sentences of that type are included. For example, on GJ Test 5, item number 11 is a GJ Type 1c sentence and can be found in Appendix E3, item number 5.

| Type: | 1a(E1) | 1b(E2) | 1c (E3) | 1d (E4) | 2a (E5) | 2b (E6) | 2c (E7) | 2d (E8) | 3a (E9) | 3b (E10) | 3c (E12) |
|---|---|---|---|---|---|---|---|---|---|---|---|
| Test No./ Master No. | | | | | | | | | | | |
| GJ Test 1 No: | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 |
| Master No: | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| GJ Test 2 No: | 2 | 4 | 6 | 8 | 10 | 7 | 3 | 5 | 1 | 9 | 11 |
| Master No: | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 |
| GJ Test 3 No: | 3 | 9 | 1 | 4 | 7 | 10 | 2 | 5 | 8 | 11 | 6 |
| Master No: | 3 | 3 | 3 | 3 | 3 | 3 | 3 | 3 | 3 | 3 | 3 |
| GJ Test 4 No: | 5 | 9 | 2 | 6 | 10 | 3 | 7 | 11 | 4 | 8 | 1 |
| Master No: | 4 | 4 | 4 | 4 | 4 | 4 | 4 | 4 | 4 | 4 | 4 |
| GJ Test 5 No: | 1 | 6 | 11 | 5 | 10 | 4 | 9 | 3 | 8 | 2 | 7 |
| Master No: | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 |
| GJ Test 6 No: | 7 | 2 | 8 | 3 | 9 | 4 | 10 | 5 | 11 | 6 | 1 |
| Master No: | 6 | 6 | 6 | 6 | 6 | 6 | 6 | 6 | 6 | 6 | 6 |
| GJ Test 7 No: | 1 | 3 | 5 | 7 | 9 | 11 | 2 | 4 | 6 | 8 | 10 |
| Master No: | 7 | 7 | 7 | 7 | 7 | 7 | 7 | 7 | 7 | 7 | 7 |
| GJ Test 8 No: | 2 | 10 | 8 | 11 | 3 | 6 | 9 | 1 | 4 | 7 | 5 |
| Master No: | 8 | 8 | 8 | 8 | 8 | 8 | 8 | 8 | 8 | 8 | 8 |
| GJ Test 9 No: | 5 | 11 | 6 | 1 | 7 | 2 | 8 | 3 | 9 | 4 | 10 |
| Master No: | 9 | 9 | 9 | 9 | 9 | 9 | 9 | 9 | 9 | 9 | 9 |
| GJ Test 10 No: | 3 | 6 | 9 | 1 | 4 | 7 | 10 | 2 | 5 | 8 | 11 |
| Master No: | 10 | 10 | 10 | 10 | 10 | 10 | 10 | 10 | 10 | 10 | 10 |
| GJ Test 11 No: | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 |
| Master No: | 11 | 11 | 11 | 11 | 11 | 11 | 11 | 11 | 11 | 11 | 11 |
| GJ Test 12 No: | 7 | 3 | 10 | 6 | 2 | 9 | 11 | 1 | 8 | 4 | 5 |
| Master No: | 12 | 12 | 12 | 12 | 12 | 12 | 12 | 12 | 12 | 12 | 12 |
| GJ Test 13 No: | 10 | 7 | 8 | 9 | 6 | 11 | 4 | 3 | 5 | 1 | 2 |
| Master No: | 13 | 13 | 13 | 13 | 13 | 13 | 13 | 13 | 13 | 13 | 13 |

GJ 1a: Biclausal sentences, finite verbs in embedded clauses. Number agrees. Bare infinitive distractors marked with asterisk.

(1) His parents knew Ben had asked about himself.

(2) The farmers realized the people couldn't feed themselves.

(3) The police officers saw the murderer kill himself.*

(4) The teacher hoped the students would take care of themselves.

(5) The hunter knew the foxes hid themselves in the tree.

(6) The detective realized the terrorists had shot themselves.

(7) Terry knew his children made themselves rich.

(8) Alex couldn't understand why the politicians praised themselves.

(9) The scientists knew the rocket blew itself up.

(10) The official demanded that the workers organize themselves into a union.

(11) The old owl knew that mice couldn't hide themselves very well.

(12) The spy saw the enemy soldiers position themselves under the bridge.*

(13) The jury members couldn't understand why the criminal didn't defend himself.

GJ 1b: Biclausal sentences, finite verbs in embedded clauses. Number disagrees.

(1) The coach knew the players trusted himself.

(2) The students knew the teacher questioned themselves about the new school rules.

(3) The people knew the country could protect themselves.

(4) The victim knew the police officers wanted information about himself.

(5) The oil producing countries thought Europe would monitor themselves.

(6) The bus driver knew the passengers trusted himself.

(7) The swimmers were afraid the shark would bite themselves.

(8) The performers were afraid the audience would attack themselves.

(9) The customers listened to the banker give themselves financial advice.

(10) The government officials hoped the army would defend themselves.

(11) The teacher knew his students didn't trust himself.

(12) The new workers knew the teacher would train themselves.

(13) The terrorist knew the government wasn't afraid of himself.

GJ 1c: Biclausal sentences, finite verbs in embedded clauses. Gender agrees. Bare infinitive distractors marked with asterisk.

(1) Adam thought his daughter had served herself.

(2) The doctor thought his female patient might hurt herself.

(3) The nurse hoped the old man would dress himself.

(4) Bill wondered if Alice would hurt herself.

(5) Sally knew Jim bought himself a new TV.

(6) The old man couldn't hear his daughter sing to herself.

(7) Everyone knew the ship could never save itself.

(8) The mother was glad after her son finally cried himself to sleep.

(9) Jim's mother helped him teach himself German.*

(10) Jim's uncle helped the woman set herself up in the new country.*

(11) Jill realized Henry accidentally locked himself in the room.

(12) Richard heard his daughter cry herself to sleep.*

(13) John couldn't understand why Susan stopped herself from getting married.

GJ 1d: Biclausal sentences, finite verbs in embedded clauses. Gender disagrees
Bare infinitive distractors marked with asterisk.

(1) John's mother knew Ben asked about herself.

(2) Mike wondered if Mary would ask himself about going on a date.

(3) The little boy hoped his mother would help himself with his homework.

(4) Susan knew Mike told a lie about herself.

(5) The mother was surprised because her son bit herself on her leg.

(6) The young woman noticed the man look at herself.*

(7) Fred hoped Mary would help himself with the homework.

(8) Jack couldn't understand why his mother blamed himself.

(9) Aunt Anne demanded that Billy help herself paint the house.

(10) Freddie wondered how long his wife would lie to himself about the car accident.

(11) Jack's sister knew he wouldn't help herself find a new job.

(12) Richard felt Linda slap himself on his back.*

(13) The averageman has no idea what a woman might think about himself.

GJ 2a: Biclausal sentences, non-finite verbs in embedded clauses. Number agrees.

(1) John's parents told him to wash himself.

(2) The doctor advised the patients to rest themselves.

(3) The company owner told the workers to train themselves more thoroughly.

(4) The doctor wanted the sick patients to cover themselves.

(5) The boys asked the movie star to introduce herself.

(6) The police officers didn't want the captured terrorist to starve himself.

(7) The hunter saw the elephants washing themselves in the river.

(8) The government wanted the people to lock themselves in their basements during the storm.

(9) Most companies don't design machines that can fix themselves.

(10) The scientists expected the space ship to force itself away from the earth.

(11) Parents expect a child's cut to usually heal itself in a few days.

(12) The teacher asked the students to test themselves.

(13) The doctor told all his patients not to expose themselves to the sun.

GJ2b: Biclausal sentences, non-finite verbs in embedded clauses. Number disagrees.

(1) The doctors expected the patient to see themselves.

(2) Citizens want the government to protect themselves.

(3) The captain wanted the crew members to know about himself.

(4) The fox heard the dogs trying to find itself.

(5) The parents wanted the child to stay with themselves.

(6) Richard asked the people to question himself.

(7) The politician wanted the people to trust himself.

(8) The church requires its followers to believe in itself.

(9) The computer wanted the astronauts to restart itself.

(10) Jack the thief wanted his victims to sympathize with himself.

(11) The students asked the teacher to write about themselves.

(12) The best man heard his friends making fun of himself.

(13) The customers wanted the waitress to serve themselves some coffee.

GJ 2c: Biclausal sentences, non-finite verbs in embedded clauses. Gender agrees.

(1) The policeman warned the woman to behave herself.

(2) Jack's mother wanted him to write a book about himself.

(3) The fire fighter told the woman to lower herself onto the roof of the building.

(4) Sam's wife wanted him to find out more about himself.

(5) Mary told her husband to cook for himself.

(6) The woman told the policeman to present himself to the chief.

(7) Alice May advised Tommy to hide himself in the basement.

(8) The mother wanted her bad son to redeem himself by apologizing.

(9) Rick asked his daughter to look at herself in the mirror.

(10) John was surprised to see his wife talking about herself on TV.

(11) Judy heard Fred singing to himself in the shower.

(12) Joanna caught her boyfriend buying himself some cigarettes.

(13) Gordon wanted Ellen to see a picture of herself.

GJ 2d: Biclausal sentences, non-finite verbs in embedded clauses. Gender disagrees.

(1) Mary's mother asked Fred to help herself fix her car.

(2) Jane asked Mike to protect herself.

(3) .Max paid Eva to paint a picture of himself.

(4) Bobby saw his sister spying on himself.

(5) Trevor wanted his wife to listen to himself.

(6) The actress told the boy to buy some roses for herself.

(7) Robert wanted Susie to listen to himself.

(8) The queen wanted her husband to support herself during the war.

(9) The mother trained her son to obey herself.

(10) Agnes thanked her brother for volunteering herself for the committee.

(11) Suzie noticed her father smiling at herself.

(12) John asked Mary to take a picture of himself.

(13) Elaine asked her boyfriend to marry herself.

**GJ 3a:** Monoclausal sentences with 2 NPs. Pragmatically favor object antecedents.

(1) Bill and June gave Sam a picture of himself.

(2) The doctors gave the patient a report about herself.

(3) The politicians wrote a report for the government about itself.

(4) John questioned the failing student about herself.

(5) The judge gave the criminals a report about themselves.

(6) The teachers questioned the student about himself.

(7) The aid worker gave the refugees food for themselves.

(8) The policeman gave the traffic accident victim a report about herself.

(9) The teachers asked the boy's father about himself.

(10) The photographer gave Bill & Hillary a picture of themselves.

(11) The newspaper didn't give the politicians a good report about themselves.

(12) The doctors questioned the new patient about herself.

(13) The citizen gave a warning to the officials about themselves.

**GJ 3b:** Monoclausal sentences with 2 NPs. Linguistically favor subject antecedents.

(1) Sally faxed her teachers a report about herself.

(2) The football player gave the fans a picture of himself.

(3) The new workers gave their boss information about themselves.

(4) The couple told the marriage counselor about themselves.

(5) The babysitter gave the parents a recommendation about herself.

(6) The prisoner wrote to his family about himself.

(7) The worried students asked their teacher about themselves.

(8) The actor didn't tell his managers everything about himself.

(9) The artist painted her parents a portrait of herself.

(10) Jack e-mailed his new employers about himself.

(11) The soldier refused to tell the enemy soldiers about himself.

(12) Alex never finished the letter to his mother about himself.

(13) The resigning ministers gave the president a final report about themselves.

**GJ 3c:** Monoclausal sentences with 2 NPs. Pragmatically ambiguous antecedents.

(1) The bad boy told his father a lie about himself.

(2) Richard sang Jimmy a song about himself.

(3) The managers told the new employees about themselves.

(4) The mother wrote to her daughter about herself.

(5) Bill's sick mother questioned her nurse about herself.

(6) The new government gave the country a report about itself.

(7) John e-mailed his son about himself.

(8) John asked Ben about himself.

(9) John painted Billy a picture of himself.

(10) The comedian told Sam a joke about himself.

(11) Merlin read Arthur a poem about himself.

(12) The cancer patient questioned the doctor about himself.

(13) Jake finally told Dan the truth about himself.

Appendix F: Native speaker control group truth value judgment test results. Test item numbers listed across top, test numbers listed down left side. Sentence type indicated above answer. Red letters indicate misjudged answers. Blue cells indicate distractor questions.

**NS1**

| Item/ Test | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | 16 | 17 | 18 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| TV1 | 1b F | 2b F | 2a T | 2a T | 1a T | 3a T | 3b T | 3a F | 3b T | 2b F | 1a T | 1b F | 1a T | 2a T | 3b T | 2b F | 3b F | 3a T |
| TV2 | 1b F | 2b F | 2b F | 2a T | 1a T | 3b T | 1a T | 3b F | 3a T | 2b F | 1a T | 2a F | 1b F | 2a T | 3b T | 3a T | 1b F | 3a T |
| TV3 | 3a T | 2a T | 2a T | 1b F | 3b F | 2b F | 3b T | 3a T | 1a T | 3a T | 1b F | 1b F | 3a T | 2b F | 3a F | 2b F | 2a T | 3b T |
| TV4 | 3a T | 2a T | 1b F | 1b F | 3b F | 2b F | 3b T | 2b F | 3b T | 2b F | 1a T | 2a T | 3a T | 2a T | 3a T | 2a T | 1b F | 3a F |
| TV5 | 3b F | 2b F | 2a F | 1a T | 3a T | 2a T | 2b F | 3a F | 3a T | 2a T | 1b F | 3b F | 3a T | 2b F | 3b T | 1b F | 2a T | 3b T |
| TV6 | 2a T | 3a T | 2b F | 3b T | 2b F | 1b F | 2a T | 3b T | 2b F | 1b F | 2a T | 3a T | 1b F | 3a F | 1a T | 3a T | 3b F | 1a T |
| TV7 | 2a T | 3a T | 3a F | 3a T | 2a T | 1a T | 2b F | 1a T | 2b T | 1b F | 1b F | 2b F | 2a T | 3b F | 3a T | 3b T | 3b F | 1b F |
| TV8 | 3a T | 1a T | 3b T | 2b F | 1b F | 3b F | 1a T | 1b F | 1b F | 3b F | 3a T | 2a T | 2b F | 1a T | 2a F | 3a F | 2b F | 2a T |
| TV9 | 2b F | 3b T | 3a F | 2b F | 1b F | 3b T | 1b F | 1a T | 2a T | 3a T | 3a F | 2b F | 2a T | 1a T | 1b F | 3b F | 3a F | 2a T |
| TV10 | 2a T | 3a F | 2b F | 3a T | 2a T | 1a T | 2a T | 3b T | 2b F | 1b F | 2b F | 3a T | 1b F | 3b F | 1a T | 1a T | 3b T | 1b F |
| TV11 | 1b F | 2b F | 2a T | 3a T | 3a T | 2a T | 2b F | 3a T | 3a F | 2a T | 1b F | 3b F | 3a T | 2b F | 3b F | 1b F | 3a T | 3b F |
| TV12 | 1a T | 2a T | 1b F | 1a T | 3a F | 2a T | 3a T | 2b F | 3b F | 2b F | 1b F | 1a T | 3b F | 2b F | 3a T | 2a T | 1b F | 3b T |
| TV13 | 3b F | 1b F | 1a T | 1b F | 3b F | 2b F | 3b T | 2a T | 3a T | 3a F | 1b F | 3a T | 2a T | 2b F | 2b F | 2a T | 3a F | 3a F |

**NS2**

| Item/ Test | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | 16 | 17 | 18 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| TV1 | 1b F | 2b F | 2a T | 2a T | 1a T | 3a T | 3b T | 3a T | 3b T | 2b F | 1a T | 1b F | 1a T | 2a T | 3b T | 2b F | 3b F | 3a T |
| TV2 | 1b F | 2b F | 2b F | 2a T | 1a T | 3b F | 1a T | 3b T | 3a T | 2b F | 1a T | 2a F | 1b F | 2a T | 3b T | 3a T | 1b F | 3a T |
| TV3 | 3a F | 2a T | 2a T | 1b F | 3b T | 2b F | 3b T | 3a T | 1a T | 3a F | 1b F | 1b F | 3a T | 2b F | 3a F | 2b F | 2a T | 3b T |
| TV4 | 3a T | 2a T | 1b F | 1b F | 3b T | 2b F | 3b T | 2b F | 3b T | 2b F | 1a T | 3a T | 3a T | 2a T | 3a T | 2a T | 1b F | 3a T |
| TV5 | 3b F | 2b F | 2a T | 1a T | 3a T | 2a T | 2b F | 3a F | 3a T | 2a T | 1b F | 3b F | 3a T | 2b F | 3b T | 1b F | 3a T | 3b T |
| TV6 | 2a T | 3a T | 2b F | 3b F | 2b F | 1b F | 2a T | 3b T | 2b F | 1b F | 2a T | 3a F | 1b F | 3a T | 1a T | 3a T | 3b F | 1a T |
| TV7 | 2a T | 3a F | 3a T | 3a T | 2a T | 1a T | 2b F | 1a F | 2b F | 1b F | 1b F | 2b F | 2a T | 3b T | 3a T | 3b T | 3b F | 3b F |
| TV8 | 3a T | 1a T | 3b F | 2b F | 1b F | 3b F | 1a T | 1b F | 1b F | 3b T | 3a F | 2a T | 2b F | 1a T | 2a F | 3a T | 2b F | 2a T |
| TV9 | 2b F | 3b F | 3a T | 2b F | 1b F | 3b T | 1b F | 1a T | 2a T | 3a T | 3a F | 2b F | 2a T | 1a T | 1b F | 3b F | 3a T | 2a T |
| TV10 | 2a T | 3a T | 2b F | 3a F | 2a T | 1a T | 2a T | 3b T | 2b F | 1b F | 2b F | 3a T | 1b F | 3b T | 1a T | 1a T | 3b T | 1b F |
| TV11 | 1b F | 2b F | 2a T | 3a T | 3a T | 2a T | 2b F | 3a T | 3a T | 2a T | 1b F | 3b F | 3a T | 2b F | 3b F | 1b F | 3a T | 3b T |
| TV12 | 1a T | 2a T | 1b F | 1a T | 3a F | 2a T | 3a T | 2b F | 3b F | 2b F | 1b F | 1a T | 3b T | 2b F | 3a T | 2a T | 1b F | 3b T |
| TV13 | 3b T | 1b F | 1a T | 1b F | 3b F | 2b F | 3b T | 2a F | 3a T | 3a F | 1b F | 3a T | 1a F | 2a T | 2b F | 2b F | 2a T | 3a T |

Appendix F (continued): Native speaker control group truth value judgment test results. Test item numbers listed across top, test numbers listed down left side. Sentence type indicated above answer. Red letters indicate misjudged answers. Blue cells indicate distractors.

**NS3**

| Item/ Test | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | 16 | 17 | 18 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| TV1 | 1b F | 2b F | 2a T | 2a T | 1a T | 3a T | 3b T | 3a T | 3b T | 2b F | 1a T | 1b F | 1a T | 2a T | 3b T | 2b F | F | 3a F |
| TV2 | 1b F | 2b F | 2b F | 2a T | 1a T | 3b F | 1a T | 3b F | 3a T | 2b F | 1a T | 2a T | 1b F | 2a T | 3b T | 3a T | 1b F | 3a T |
| TV3 | T | 2a T | 2a T | 1b F | 3b F | 2b F | 3b T | 3a T | 1a T | 3a F | 1b F | 1b F | T | 2b F | 3a T | 2b F | 2a T | 3b T |
| TV4 | T | 2a T | 1b F | 1b F | 3b T | 2b F | 3b T | 2b F | 3b T | 2b F | 1a T | T | 3a T | 2a T | 3a T | 2a T | 1b F | 3a T |
| TV5 | F | 2b F | 2a T | 1a T | 3a T | 2a T | 2b F | 3a F | 3a T | 2a T | 1b F | 3b T | T | 2b F | 3b T | 1b F | T | 3b T |
| TV6 | 2a T | 3a T | 2b F | 3b T | 2b F | 1b F | 2a T | 3b F | 2b F | 1b F | 2a T | 3a F | 1b F | 3a T | 1a T | T | 3b T | 1a T |
| TV7 | 2a T | 3a T | 3a T | 3a T | 2a T | 1a T | 2b F | 1a T | 2b F | 1b F | 1b F | 2b F | 2a T | 3b T | T | 3b T | 3b F | F |
| TV8 | 3a T | 1a T | 3b T | 2b F | 1b F | 3b T | 1a T | 1b F | 1b F | 3b T | 3a F | 2a T | 2b F | 1a T | 2a F | 3a T | 2b F | 2a T |
| TV9 | 2b F | 3b T | 3a T | 2b F | 1b F | 3b T | 1b F | 1a T | 2a T | T | 3a T | 2b F | 2a T | 1a T | 1b F | 3b F | 3a T | 2a T |
| TV10 | 2a T | 3a T | 2b F | 3a T | 2a T | 1a T | 2a T | 3b T | 2b F | 1b F | 2b F | 3a T | 1b F | 3b T | 1a T | 1a T | 3b T | 1b F |
| TV11 | 1b F | 2b F | 2a T | T | 3a T | 2a T | 2b F | 3a T | 3a T | 2a T | 1b F | 3b T | 1a T | F | 3b T | 1b F | T | 3b T |
| TV12 | 1a T | 2a T | 1b F | 1a T | 3a T | 2a T | 3a F | 2b F | 3b T | 2b F | 1b F | 1a T | 3b T | 2b F | 3a T | 2a T | 1b F | 3b T |
| TV13 | 3b T | 1b F | 1a T | 1b F | 3b T | 2b F | 3b F | 2a F | T | 3a F | T | 1b F | 3a T | 2a T | 2b F | 2b F | 2a T | 3a T |

**NS4**

| Item/ Test | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | 16 | 17 | 18 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| TV1 | 1b F | 2b F | 2a T | 2a T | 1a T | 3a T | 3b T | 3a T | 3b T | 2b F | 1a T | 1b F | 1a F | 2a T | 3b T | 2b T | F | 3a T |
| TV2 | 1b F | 2b F | 2b F | 2a T | 1a T | 3b T | 1a T | 3b F | 3a T | 2b F | 1a T | 2a F | 1b F | 2a T | 3b T | 3a T | 1b F | 3a T |
| TV3 | F | 2a T | 2a T | 1b F | 3b F | 2b F | 3b T | 3a T | 1a T | 3a F | 1b F | 1b F | T | 2b F | 3a T | 2b F | 2a T | 3b T |
| TV4 | T | 2a T | 1b F | 1b F | 3b T | 2b F | 3b T | 2b F | 3b T | 2b F | 1a T | T | 3a T | 2a T | 3a T | 2a T | 1b F | 3a F |
| TV5 | F | 2b F | 2a T | 1a T | 3a T | T | 2b F | T | T | 2a T | 1b F | 3b T | T | 2b F | 3b T | 1b F | T | 3b T |
| TV6 | 2a T | 3a T | 2b F | 3b T | 2b F | 1b F | 2a T | 3b F | 2b F | 1b F | 2a T | 3a T | 1b F | 3a F | 1a T | T | 3b T | 1a T |
| TV7 | 2a T | 3a F | 3a F | 3a T | 2a T | 1a T | 2b F | T | 2b F | 1b F | 1b F | 2b F | 2a T | 3b T | T | 3b T | 3b T | F |
| TV8 | 3a T | 1a T | 3b T | 2b F | 1b F | 3b T | 1a T | 1b F | 1b F | 3b T | 3a F | 2a T | 2b F | 1a T | 2a F | 3a F | 2b F | 2a T |
| TV9 | 2b F | 3b F | 3a F | 2b F | 1b F | 3b T | 1b F | 1a T | 2a T | T | 3a T | 2b F | 2a T | 1a T | 1b F | 3b F | 3a F | 2a T |
| TV10 | 2a T | 3a T | 2b F | 3a T | 2a T | 1a T | 2a T | 3b T | 2b F | 1b F | 2b F | 3a T | 1b F | 3b F | 1a T | 1a T | 3b T | 1b F |
| TV11 | 1b F | 2b F | 2a T | T | 3a T | 2a T | 2b F | 3a T | 3a F | 2a T | 1b F | 3b T | T | 2b F | 3b F | 1b F | T | 3b F |
| TV12 | 1a T | 2a T | 1b F | 1a T | 3a T | 2a T | 3a F | 2b F | 3b T | 2b F | 1b F | 1a T | 3b T | 2b F | 3a T | 2a T | 1b F | 3b T |
| TV13 | 3b T | 1b F | 1a T | 1b F | 3b F | 2b F | 3b T | 2a F | T | 3a F | T | 1b F | 3a T | 2a T | 2b F | 2b F | 2a T | 3a T |

Appendix F (continued): Native speaker control group truth value judgment test results. Test item numbers listed across top, test numbers listed down left side. Sentence type indicated above answer. Red letters indicate misjudged answers. Blue cells indicate distractors.

**NS5**

| Item/ Test | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | 16 | 17 | 18 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| TV1 | 1b F | 2b F | 2a T | 2a T | 1a T | 3a T | 3b F | 3a T | 3b T | 2b F | 1a T | 1b F | 1a T | 2a T | 3b T | 2b F | F | 3a T |
| TV2 | 1b F | 2b F | 2b F | 2a T | 1a T | 3b F | 1a T | 3b T | 3a T | 2b F | 1a T | 2a T | 1b F | 2a T | 3b T | 3a T | 1b F | 3a T |
| TV3 | T | 2a T | 2a T | 1b F | 3b T | 2b F | 3b T | 3a T | 1a T | 3a T | 1b F | 1b F | T | 2b F | 3a F | 2b F | 2a T | 3b T |
| TV4 | T | 2a T | 1b F | 1b F | 3b T | 2b F | 3b T | 2b F | 3b T | 2b F | 1a T | T | 3a T | 2a T | 3a T | 2a T | 1b F | 3a T |
| TV5 | F | 2b T | 2a T | 1a T | 3a T | 2a T | 2b F | 3a F | 3a T | 2a T | 1b F | 3b T | T | 2b F | 3b T | 1b F | T | 3b T |
| TV6 | 2a T | 3a T | 2b F | 3b F | 2b F | 1b F | 2a T | 3b T | 2b F | 1b F | 2a T | 3a F | 1b F | 3a F | 1a T | T | 3b T | 1a T |
| TV7 | 2a T | 3a T | 3a T | 3a T | 2a T | 1a T | 2b F | 1a T | 2b F | 1b F | 1b F | 2b F | 2a T | 3b T | T | 3b T | 3b F | F |
| TV8 | 3a T | 1a T | 3b T | 2b F | 1b F | 3b T | 1a T | 1b F | 1b F | 3b T | 3a F | 2a T | 2b F | 1a T | 2a F | T | 2b F | 2a T |
| TV9 | 2b F | 3b T | 3a T | 2b F | 1b F | 3b T | 1b F | 1a T | 2a T | T | 3a T | 2b F | 2a T | 1a T | 1b F | 3b T | 3a T | 2a T |
| TV10 | 2a T | 3a T | 2b F | 3a T | 2a T | 1a T | 2a T | 3b T | 2b F | 1b F | 2b F | 3a F | 1b T | 3b T | 1a T | 1a T | 3b T | 1b F |
| TV11 | 1b F | 2b F | 2a T | T | 3a T | 2a T | 2b F | 3a T | 3a T | 2a T | 1b F | 3b T | T | 2b F | 3b T | F | 3b T | 3b T |
| TV12 | 1a T | 2a T | 1b F | 1a T | 3a T | 2a T | 3a T | 2b F | 3b F | 2b F | 1b F | 1a T | 3b T | 2b F | 3a T | 2a T | 1b F | 3b T |
| TV13 | 3b T | 1b F | 1a T | 1b F | 3b T | 2b F | 3b T | 2a F | F | 3a T | T | 1b F | 3a F | 2a T | 2b F | 2b F | 2a T | 3a T |

**NS6**

| Item/ Test | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | 16 | 17 | 18 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| TV1 | 1b F | 2b F | 2a T | 2a T | 1a T | 3a T | 3b T | 3a T | 3b T | 2b F | 1a T | 1b F | 1a F | 2a T | 3b T | 2b F | F | 3a F |
| TV2 | 1b F | 2b F | 2b F | 2a T | 1a T | 3b T | 1a T | 3b F | 3a T | 2b F | 1a T | 2a T | 1b F | 2a T | 3b T | 3a T | 1b F | 3a T |
| TV3 | T | 2a T | 2a T | 1b F | 3b F | 2b F | 3b T | 3a T | 1a T | 3a T | 1b F | 1b F | T | 2b F | 3a T | 2b F | 2a T | 3b T |
| TV4 | T | 2a T | 1b F | 1b F | 3b T | 2b F | 3b T | 2b F | 3b T | 2b F | 1a T | T | 3a T | 2a F | 3a F | 2a T | 1b F | 3a F |
| TV5 | F | 2b F | 2a T | 1a T | 3a F | 2a T | 2b F | 3a T | 3a T | 2a T | 1b F | 3b T | T | 2b T | 3b T | 1b F | T | 3b T |
| TV6 | 2a T | 3a T | 2b F | 3b F | 2b F | 1b F | 2a T | 3b F | 2b F | 1b F | 2a T | 3a T | 1b F | 3a F | 1a T | T | 3b T | 1a T |
| TV7 | 2a T | 3a F | 3a T | 3a T | 2a T | 1a T | 2b F | 1a F | 2b F | 1b F | 1b F | 2b F | 2a T | 3b T | T | 3b T | 3b T | F |
| TV8 | 3a T | 1a T | 3b T | 2b F | 1b F | 3b F | 1a T | 1b F | 1b F | 3b T | 3a T | 2a T | 2b F | 1a T | 2a T | 3a T | 2b F | 2a T |
| TV9 | 2b F | 3b T | 3a F | 2b F | 1b F | 3b T | 1b F | 1a T | 2a T | F | 3a T | 2b F | 2a F | 1a T | 1b F | F | 3a T | 2a T |
| TV10 | 2a T | 3a T | 2b F | 3a F | 2a T | 1a T | 2a T | 3b T | 2b F | 1b F | 2b F | 3a T | 1b F | 3b T | 1a T | 1a T | 3b F | 1b F |
| TV11 | 1b F | 2b F | 2a T | T | 3a T | 2a T | 2b F | 3a T | 3a F | 2a T | 1b F | 3b T | T | 2b F | 3b T | 1b F | T | 3b F |
| TV12 | 1a T | 2a T | 1b F | 1a T | 3a T | 2a T | 3a T | 2b F | 3b F | 2b F | 1b F | 1a T | 3b T | 2b F | 3a T | 2a F | 1b F | 3b T |
| TV13 | 3b T | 1b F | 1a T | 1b F | 3b T | 2b F | 3b T | 2a F | T | 3a F | T | 1b F | 3a T | 2a T | 2b F | 2b F | 2a T | 3a T |

Appendix G: Native speaker control group grammaticality judgment test results. Test item numbers listed across top, test numbers listed down left side. Sentence type indicated above answer. Red letters indicate misjudged answers. Blue cells indicate distractors.

**NS1**

| Item/ Test | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 |
|---|---|---|---|---|---|---|---|---|---|---|---|
| GJ1 | 1a Y | 1b N | 1c Y | 1d N | 2a Y | 2b N | 2c Y | 2d N | 3a Y | 3b Y | 3c N |
| GJ2 | 3a Y | 1a Y | 2c Y | 1b N | 2d Y | 1c Y | 2b Y | 1d N | 3b Y | 2a N | 3c Y |
| GJ3 | 1c Y | 2c Y | 1a Y | 1d N | 2d N | 3c Y | 2a Y | 3a N | 1b N | 2b N | 3b Y |
| GJ4 | 3c N | 1c Y | 2b N | 3a Y | 1a Y | 1d N | 2c Y | 3b Y | 1b N | 2a Y | 2d N |
| GJ5 | 1a Y | 3b Y | 2d N | 2b N | 1d N | 1b N | 3c Y | 3a Y | 2c Y | 2a Y | 1c Y |
| GJ6 | 3c Y | 1b N | 1a N | 2b N | 2d N | 3b Y | 1a Y | 1c Y | 2a Y | 2c Y | 3a Y |
| GJ7 | 1a Y | 2c N | 1b N | 2d N | 1c N | 3a Y | 1d Y | 3b Y | 2a Y | 3c Y | 2b Y |
| GJ8 | 2d N | 1a Y | 2a Y | 3a Y | 3c Y | 2b N | 3b Y | 1c Y | 2c Y | 1b N | 1d N |
| GJ9 | 1d N | 2b N | 2d N | 3b Y | 1a Y | 1c Y | 2a Y | 2c Y | 3a Y | 3c Y | 1b N |
| GJ10 | 1d N | 2d N | 1a Y | 2a Y | 3a Y | 1b Y | 2b N | 3b N | 1c Y | 2c Y | 3c Y |
| GJ11 | 1a Y | 1b N | 1c Y | 1d N | 2a Y | 2b N | 2c Y | 2d N | 3a N | 3b Y | 3c Y |
| GJ12 | 2d N | 2a Y | 1b N | 3b Y | 3c Y | 2d N | 1c Y | 3a Y | 2b N | 1c Y | 2c Y |
| GJ13 | 3b Y | 3c Y | 2d N | 2c Y | 3a Y | 2a Y | 1b N | 1c Y | 1d N | 1a Y | 2b N |

**NS2**

| Item/ Test | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 |
|---|---|---|---|---|---|---|---|---|---|---|---|
| GJ1 | 1a Y | 1b N | 1c Y | 1d N | 2a Y | 2b N | 2c Y | 2d N | 3a Y | 3b Y | 3c Y |
| GJ2 | 3a Y | 1a Y | 2c Y | 1b N | 2d N | 1c Y | 2b N | 1d N | 3b Y | 2a Y | 3c Y |
| GJ3 | 1c Y | 2c Y | 1a Y | 1d N | 2d N | 3c Y | 2a Y | 3a N | 1b N | 2b N | 3b Y |
| GJ4 | 3c Y | 1c Y | 2b N | 3a Y | 1a Y | 1d N | 2c Y | 3b Y | 1b N | 2a Y | 2d N |
| GJ5 | 1a Y | 3b Y | 2d N | 2b N | 1d N | 1b N | 3c Y | 3a Y | 2c Y | 2a Y | 1c Y |
| GJ6 | 3c Y | 1b N | 1a N | 2b N | 2d N | 3b Y | 1a Y | 1c Y | 2a Y | 2c Y | 3a Y |
| GJ7 | 1a Y | 2c Y | 1b N | 2d N | 1c Y | 3a Y | 1d N | 3b N | 2a Y | 3c Y | 2b N |
| GJ8 | 2d N | 1a Y | 2a Y | 3a Y | 3c Y | 2b N | 3b Y | 1c Y | 2c Y | 1b N | 1d N |
| GJ9 | 1d N | 2b N | 2d N | 3b Y | 1a Y | 1c Y | 2a Y | 2c Y | 3a N | 3c Y | 1b N |
| GJ10 | 1d N | 2d N | 1a Y | 2a Y | 3a Y | 1b N | 2b N | 3b Y | 1c Y | 2c Y | 3c Y |
| GJ11 | 1a Y | 1b N | 1c Y | 1d N | 2a Y | 2b N | 2c Y | 2d N | 3a Y | 3b Y | 3c Y |
| GJ12 | 2d N | 2a Y | 1b N | 3b Y | 3c Y | 1d N | 1c Y | 3a Y | 2b N | 1c Y | 2c Y |
| GJ13 | 3b Y | 3c Y | 2d N | 2c Y | 3a Y | 2a Y | 1b N | 1c Y | 1d N | 1a Y | 2b N |

**NS3**

| Item/ Test | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 |
|---|---|---|---|---|---|---|---|---|---|---|---|
| GJ1 | 1a Y | 1b N | 1c Y | 1d N | 2a Y | 2b N | 2c Y | 2d N | 3a Y | 3b Y | 3c Y |
| GJ2 | 3a Y | 1a Y | 2c Y | 1b N | 2d N | 1c Y | 2b N | 1d N | 3b Y | 2a Y | 3c Y |
| GJ3 | 1c Y | 2c Y | 1a Y | 1d N | 2d N | 3c Y | 2a Y | 3a Y | 1b N | 2b N | 3b Y |
| GJ4 | 3c Y | 1c Y | 2b N | 3a Y | 1a Y | 1d N | 2c Y | 3b Y | 1b N | 2a Y | 2d N |
| GJ5 | 1a Y | 3b Y | 2d N | 2b N | 1d N | 1b N | 3c N | 3a Y | 2c Y | 2a Y | 1c Y |
| GJ6 | 3c Y | 1b N | 1a N | 2b N | 2d N | 3b Y | 1a Y | 1c Y | 2a Y | 2c Y | 3a Y |
| GJ7 | 1a Y | 2c Y | 1b N | 2d N | 1c Y | 3a Y | 1d N | 3b Y | 2a Y | 3c Y | 2b N |
| GJ8 | 2d N | 1a Y | 2a Y | 3a Y | 3c Y | 2b N | 3b Y | 1c N | 2c Y | 1b N | 1d N |
| GJ9 | 1d N | 2b N | 2d N | 3b Y | 1a Y | 1c Y | 2a Y | 2c Y | 3a Y | 3c Y | 1b N |
| GJ10 | 1d N | 2d N | 1a Y | 2a Y | 3a Y | 1b N | 2b N | 3b Y | 1c Y | 2c Y | 3c Y |
| GJ11 | 1a Y | 1b N | 1c Y | 1d N | 2a Y | 2b N | 2c Y | 2d N | 3a N | 3b Y | 3c Y |
| GJ12 | 2d N | 2a Y | 1b N | 3b Y | 3c Y | 1d N | 1c Y | 3a Y | 2b N | 1c Y | 2c Y |
| GJ13 | 3b Y | 3c Y | 2d N | 2c Y | 3a Y | 2a Y | 1b N | 1c Y | 1d N | 1a Y | 2b N |

Appendix G (continued): Native speaker control group grammaticality judgment test results. Test item numbers listed across top, test numbers listed down left side. Sentence type indicated above answer. Red letters indicate misjudged answers. Blue cells indicate distractors.

**NS4**

| Item/ Test | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 |
|---|---|---|---|---|---|---|---|---|---|---|---|
| GJ1 | 1a Y | 1b N | 1c Y | 1d N | 2a Y | 2b N | 2c Y | 2d N | 3a Y | 3b Y | 3c N |
| GJ2 | 3a Y | 1a Y | 2c Y | 1b N | 2d N | 1c Y | 2b N | 1d N | 3b Y | 2a Y | 3c Y |
| GJ3 | 1c Y | 2c Y | [3a] Y | 1d N | 2d N | 3c Y | 2a Y | 3a Y | 1b N | 2b N | 3b Y |
| GJ4 | 3c Y | 1c Y | 2b N | 3a Y | 1a Y | 1d N | 2c Y | 3b Y | 1b N | 2a Y | 2d N |
| GJ5 | 1a Y | 3b Y | 2d N | 2b N | 1d N | 1b N | 3c Y | 3a Y | 2c Y | 2a Y | 1c Y |
| GJ6 | 3c Y | 1b N | [1d] N | 2b N | 2d N | 3b Y | 1a Y | 1c Y | 2a Y | 2c Y | 3a Y |
| GJ7 | 1a Y | 2c Y | 1b N | 2d N | 1c Y | 3a Y | 1d N | 3b Y | 2a Y | 3c Y | 2b N |
| GJ8 | 2d N | 1a Y | 2a Y | 3a Y | 3c Y | 2b N | 3b Y | 1c Y | 2c Y | 1b N | 1d N |
| GJ9 | 1d N | 2b N | 2d N | 3b Y | 1a Y | [2c] Y | 2a Y | 2c Y | 3a N | 3c Y | 1b N |
| GJ10 | 1d N | 2d N | 1a Y | 2a Y | 3a Y | 1b N | 2b N | 3b Y | [1c] Y | 2c Y | 3c Y |
| GJ11 | 1a Y | 1b N | 1c Y | 1d N | 2a Y | 2b N | 2c Y | 2d N | 3a N | 3b Y | 3c Y |
| GJ12 | 2d N | 2a Y | 1b N | 3b Y | 3c Y | [1d] N | [2b] Y | 3a Y | 2b N | [1a] Y | 2c Y |
| GJ13 | 3b Y | 3c Y | 2d N | 2c Y | 3a Y | 2a Y | 1b N | 1c Y | 1d N | 1a Y | 2b N |

**NS5**

| Item/ Test | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 |
|---|---|---|---|---|---|---|---|---|---|---|---|
| GJ1 | 1a Y | 1b N | 1c Y | 1d N | 2a Y | 2b N | 2c Y | 2d N | 3a Y | 3b Y | 3c Y |
| GJ2 | 3a Y | 1a Y | 2c Y | 1b N | 2d N | 1c Y | 2b Y | 1d N | 3b Y | 2a Y | 3c Y |
| GJ3 | 1c Y | 2c Y | [3a] Y | 1d N | 2d N | 3c Y | 2a Y | 3a Y | 1b N | 2b N | 3b Y |
| GJ4 | 3c N | 1c Y | 2b N | 3a Y | 1a Y | 1d N | 2c Y | 3b Y | 1b N | 2a Y | 2d N |
| GJ5 | 1a Y | 3b Y | 2d N | 2b N | 1d N | 1b N | 3c Y | 3a Y | 2c Y | 2a Y | 1c Y |
| GJ6 | 3c Y | 1b N | [1d] N | 2b N | 2d N | 3b Y | 1a Y | 1c Y | 2a Y | 2c Y | 3a Y |
| GJ7 | 1a Y | 2c Y | 1b N | 2d N | 1c Y | 3a Y | 1d N | 3b Y | 2a Y | 3c Y | 2b N |
| GJ8 | 2d N | 1a Y | 2a Y | 3a Y | 3c Y | 2b N | 3b Y | 1c Y | 2c Y | 1b N | 1d N |
| GJ9 | 1d N | 2b N | 2d N | 3b Y | 1a Y | [2c] Y | 2a Y | 2c Y | 3a Y | 3c Y | 1b N |
| GJ10 | 1d N | 2d N | 1a Y | 2a Y | 3a Y | 1b N | 2b N | 3b Y | [1c] Y | 2c Y | 3c Y |
| GJ11 | 1a Y | 1b N | 1c Y | 1d N | 2a Y | 2b N | 2c Y | 2d N | 3a N | 3b Y | 3c Y |
| GJ12 | 2d N | 2a Y | 1b N | 3b Y | 3c Y | [1d] N | [2b] Y | 3a Y | 2b N | [1a] Y | 2c Y |
| GJ13 | 3b Y | 3c Y | 2d N | 2c Y | 3a Y | 2a Y | 1b N | 1c Y | 1d N | 1a Y | 2b N |

**NS6**

| Item/ Test | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 |
|---|---|---|---|---|---|---|---|---|---|---|---|
| GJ1 | 1a Y | 1b N | 1c Y | 1d N | 2a Y | 2b N | 2c Y | 2d N | 3a Y | 3b Y | 3c Y |
| GJ2 | 3a Y | 1a Y | 2c Y | 1b N | 2d N | 1c Y | 2b N | 1d N | 3b Y | 2a Y | 3c Y |
| GJ3 | 1c Y | 2c Y | [3a] Y | 1d N | 2d N | 3c Y | 2a Y | 3a Y | 1b N | 2b N | 3b Y |
| GJ4 | 3c Y | 1c Y | 2b N | 3a Y | 1a Y | 1d N | 2c Y | 3b Y | 1b N | 2a Y | 2d N |
| GJ5 | 1a Y | 3b Y | 2d N | 2b N | 1d N | 1b N | 3c Y | 3a Y | 2c Y | 2a Y | 1c Y |
| GJ6 | 3c Y | 1b N | [1d] N | 2b N | 2d N | 3b Y | 1a Y | 1c Y | 2a Y | 2c Y | 3a Y |
| GJ7 | 1a Y | 2c Y | 1b N | 2d N | 1c Y | 3a Y | 1d N | 3b Y | 2a Y | 3c Y | 2b N |
| GJ8 | 2d N | 1a Y | 2a Y | 3a Y | 3c Y | 2b N | 3b Y | 1c N | 2c Y | 1b N | 1d N |
| GJ9 | 1d N | 2b N | 2d N | 3b Y | 1a Y | [2c] Y | 2a Y | 2c Y | 3a Y | 3c Y | 1b N |
| GJ10 | 1d N | 2d N | 1a Y | 2a Y | 3a Y | 1b Y | 2b N | 3b Y | [1c] Y | 2c Y | 3c Y |
| GJ11 | 1a Y | 1b N | 1c Y | 1d N | 2a Y | 2b N | 2c Y | 2d N | 3a Y | 3b Y | 3c Y |
| GJ12 | 2d N | 2a Y | 1b N | 3b Y | 3c Y | [1d] N | [2b] Y | 3a Y | 2b N | [1a] Y | 2c Y |
| GJ13 | 3b Y | 3c Y | 2d N | 2c Y | 3a Y | 2a Y | 1b N | 1c Y | 1d N | 1a Y | 2b N |

Appendix H: Test subject truth value judgment test results. Test item numbers listed across top, test numbers listed down left side. Sentence type indicated above answer. Red letters indicate misjudged answers. Blue cells indicate distractors.

**R1**

| Item/ Test | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | 16 | 17 | 18 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| TV1 | 1b F | 2b F | 2a T | 2a T | 1a T | 3a T | 3b T | 3a T | 3b T | 2b F | 1a T | 1b F | 1a T | 2a T | 3b T | 2b F | 1b F | 3a F |
| TV2 | 1b F | 2b F | 2b F | 2a T | 1a T | 3b T | 1a T | 3b F | 3a T | 2b F | 1a T | 2a F | 1b F | 2a T | 3b T | 3a T | 1b F | 3a T |
| TV3 | 1b T | 2a T | 2a T | 1b F | 3b T | 2b F | 3b F | 3a T | 1a T | 3a F | 1b F | 1b F | 3a T | 2b F | 3a F | 2b F | 2a T | 3b T |
| TV4 | 1b T | 2a T | 1b F | 1b F | 3b T | 2b F | 3b T | 2b T | 3b T | 2b F | 1a T | 3a T | 3a T | 2a T | 3a T | 2a T | 1b F | 3a F |
| TV5 | 1b F | 2b F | 2a F | 1a T | 3a T | 2a T | 2b F | 3a F | 3a T | 2a T | 1b F | 3b T | 3a T | 2b F | 3b F | 1b F | 1b T | 3b T |
| TV6 | 2a T | 3a T | 2b F | 3b F | 2b F | 1b F | 2a T | 3b T | 2b F | 1b F | 2a T | 3a F | 1b F | 3a F | 1a T | 1a T | 3b T | 1a T |
| TV7 | 2a T | 3a T | 3a F | 3a F | 2a T | 1a T | 2b T | 1a T | 2b F | 1b F | 1b F | 2b F | 2a T | 3b T | 1a T | 3b T | 3b F | 1b F |
| TV8 | 3a F | 1a T | 3b T | 2b T | 1b F | 3b F | 1a T | 1b F | 1b F | 3b T | 3a T | 2a T | 2b F | 1a T | 2a F | 3a T | 2b F | 2a T |
| TV9 | 2b F | 3b F | 3a T | 2b F | 1b F | 3b T | 1b F | 1a T | 2a T | 1a T | 3a T | 2b T | 2a T | 1a T | 1b F | 3b F | 3a F | 2a T |
| TV10 | 2a T | 3a F | 2b T | 3a T | 2a T | 1a T | 2a T | 3b T | 2b T | 1b F | 2b F | 3a T | 1b F | 3b T | 1a T | 1a T | 3b T | 1b F |
| TV11 | 1b T | 2b T | 2a T | 1a T | 3a F | 2a T | 2b F | 3a F | 3a T | 2a T | 1b F | 3b F | 1a T | 2b F | 3b T | 1b F | 3a T | 3b T |
| TV12 | 1a T | 2a T | 1b F | 1a T | 3a T | 2a T | 3a T | 2b F | 3b T | 2b F | 1b F | 1a T | 3b T | 2b F | 3a T | 2a T | 1b F | 3b F |

**R2**

| Item/ Test | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | 16 | 17 | 18 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| TV1 | 1b F | 2b F | 2a T | 2a T | 1a T | 3a T | 3b T | 3a F | 3b T | 2b F | 1a T | 1b F | 1a T | 2a T | 3b T | 2b F | 1b T | 3a F |
| TV2 | 1b F | 2b F | 2b F | 2a T | 1a T | 3b F | 1a T | 3b F | 3a T | 2b F | 1a T | 2a T | 1b F | 2a T | 3b T | 3a T | 1b F | 3a T |
| TV3 | 1b T | 2a T | 2a T | 1b F | 3b F | 2b F | 3b T | 3a T | 1a T | 3a F | 1b F | 1b F | 3a T | 2b T | 3a F | 2b F | 2a T | 3b T |
| TV4 | 1b T | 2a T | 1b F | 1b F | 3b T | 2b T | 3b T | 2b F | 3b T | 2b F | 1a T | 3a T | 3a F | 2a T | 3a T | 2a T | 1b F | 3a F |
| TV5 | 1b F | 2b F | 2a T | 1a T | 3a F | 2a T | 2b F | 3a F | 3a T | 2a T | 1b F | 3b F | 3a T | 2b F | 3b T | 1b F | 1b T | 3b T |
| TV6 | 2a F | 3a T | 2b F | 3b T | 2b F | 1b F | 2a T | 3b F | 2b F | 1b F | 2a T | 3a F | 1b F | 3a T | 1a T | 1a T | 3b F | 1a T |
| TV7 | 2a F | 3a F | 3a T | 3a F | 2a T | 1a T | 2b F | 1a T | 2b T | 1b F | 1b F | 2b F | 2a T | 3b T | 1a T | 3b T | 3b T | 1b F |
| TV8 | 3a F | 1a T | 3b T | 2b T | 1b F | 3b F | 1a T | 1b F | 1b F | 3b T | 3a F | 2a T | 2b F | 1a T | 2a T | 3a F | 2b F | 2a T |
| TV9 | 2b F | 3b F | 3a F | 2b F | 1b F | 3b T | 1b F | 1a T | 2a T | 1a T | 3a T | 2b T | 2a T | 1a T | 1b F | 3b T | 3a F | 2a T |
| TV10 | 2a T | 3a F | 2b F | 3a T | 2a T | 1a T | 2a T | 3b T | 2b T | 1b F | 2b F | 3a T | 1b F | 3b F | 1a T | 1a T | 3b T | 1b F |

Appendix H (continued): Test subject truth value judgment test results. Test item numbers listed across top, test numbers listed down left side. Sentence type indicated above answer. Red letters indicate misjudged answers. Blue cells indicate distractors.

**R3**

| Item/ Test | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | 16 | 17 | 18 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| TV1 | 1b F | 2b F | 2a T | 2a T | 1a T | 3a T | 3b T | 3a F | 3b T | 2b F | 1a T | 1b T | 1a T | 2a T | 3b T | 2b F | 1b F | 3a T |
| TV2 | 1b F | 2b F | 2b F | 2a T | 1a T | 3b F | 1a T | 3b F | 3a T | 2b F | 1a T | 2a T | 1b F | 2a T | 3b F | 3a T | 1b F | 3a F |
| TV3 | 1a T | 2a T | 2a T | 1b F | 3b T | 2b F | 3b T | 3a T | 1a T | 3a F | 1b F | 1b T | 1a T | 2b F | 3a F | 2b F | 2a T | 3b T |
| TV4 | 2a T | 2a F | 1b F | 1b F | 3b T | 2b T | 3b T | 2b F | 3b T | 2b F | 1a T | 3a T | 3a F | 2a T | 3a T | 2a T | 1b T | 3a T |
| TV5 | 1b F | 2b T | 2a T | 1a T | 3a F | 2a F | 2b F | 3a F | 3a T | 2a T | 1b F | 3b F | 3a T | 2b T | 3b T | 1b F | 1a T | 3b T |
| TV6 | 2a F | 3a T | 2b F | 3b T | 2b F | 1b F | 2a F | 3b T | 2b T | 1b T | 2a T | 3a F | 1b F | 3a F | 1a T | 3a T | 3b T | 1a T |
| TV7 | 2a T | 3a T | 3a F | 3a F | 2a T | 1a T | 2b T | 1a T | 2b F | 1b T | 1b F | 2b F | 2a F | 3b T | 3a F | 3b F | 3b F | 1b F |
| TV8 | 3a F | 1a T | 3b T | 2b F | 1b T | 3b F | 1a T | 1b F | 1b F | 3b T | 3a T | 2a T | 2b F | 1a T | 2a F | 3a F | 2b F | 2a F |
| TV9 | 2b F | 3b F | 3a T | 2b T | 1b T | 3b T | 1b F | 1a T | 2a T | 3a T | 3a T | 2b T | 2a F | 1a F | 1b F | 3b F | 3a F | 2a T |
| TV10 | 2a T | 3a F | 2b F | 3a T | 2a F | 1a T | 2a T | 3b T | 2b T | 1b F | 2b F | 3a T | 1b T | 3b F | 1a F | 1a T | 3b T | 1b T |
| TV11 | 1b T | 2b T | 2a F | 3a T | 3a T | 2a T | 2b T | 3a F | 3a F | 2a T | 1b F | 3b F | 3a T | 2b F | 3b T | 1b F | 3a T | 3b T |

**S1**

| Item/ Test | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | 16 | 17 | 18 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| TV1 | 1b F | 2b F | 2a T | 2a T | 1a T | 3a T | 3b F | 3a F | 3b T | 2b F | 1a T | 1b F | 1a T | 2a T | 3b T | 2b F | 3a F | 3a F |
| TV2 | 1b F | 2b F | 2b F | 2a T | 1a T | 3b F | 1a T | 3b T | 3a T | 2b F | 1a T | 2a T | 1b F | 2a T | 3b T | 3a T | 1b F | 3a F |
| TV3 | 3a T | 2a T | 2a T | 1b F | 3b F | 2b F | 3b T | 3a T | 1a T | 3a F | 1b F | 1b F | 3a T | 2b T | 3a F | 2b F | 2a F | 3b T |
| TV4 | 3a T | 2a T | 1b F | 1b F | 3b F | 2b F | 3b T | 2b F | 3b T | 2b F | 1a T | 3a T | 3a F | 2a T | 3a T | 2a T | 1b F | 3a F |
| TV5 | 2b T | 2b F | 2a T | 1a T | 3a T | 2a F | 2b T | 3a F | 3a F | 2a T | 1b F | 3b T | 3a T | 2b F | 3b T | 1b F | 3a T | 3b T |
| TV6 | 2a T | 3a T | 2b F | 3b T | 2b F | 1b F | 2a T | 3b F | 2b F | 1b F | 2a F | 3a F | 1b F | 3a F | 1a T | 3a T | 3b T | 1a T |
| TV7 | 2a T | 3a T | 3a T | 3a F | 2a T | 1a T | 2b F | 1a T | 2b T | 1b T | 1b F | 2b F | 2a T | 3b T | 3a T | 3b T | 3b F | 1b F |
| TV8 | 3a T | 1a T | 3b T | 2b F | 1b F | 3b T | 1a T | 1b F | 1b F | 3b T | 3a F | 2a T | 2b F | 1a T | 2a F | 3a T | 2b F | 2a T |
| TV9 | 2b F | 3b T | 3a F | 2b F | 1b F | 3b T | 1b F | 1a T | 2a F | 3a T | 3a T | 2b T | 2a T | 1a T | 1b F | 3b F | 3a T | 2a T |

Appendix H (continued): Test subject truth value judgment test results. Test item numbers listed across top, test numbers listed down left side. Sentence type indicated above answer. Red letters indicate misjudged answers. Blue cells indicate distractors.

**S2**

S2 Truth Value Test Results

| Item/ Test | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | 16 | 17 | 18 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| TV1 | 1b F | 2b F | 2a T | 2a T | 1a T | 3a T | 3b T | 3a F | 3b T | 2b F | 1a T | 1b F | 1a T | 2a T | 3b T | 2b F | F | 3a F |
| TV2 | 1b F | 2b F | 2b F | 2a T | 1a T | 3b F | 1a T | 3b F | 3a T | 2b F | 1a T | 2a T | 1b F | 2a T | 3b T | 3a T | 1b F | 3a F |
| TV3 | T | 2a T | 2a T | 1b T | 3b T | 2b F | 3b T | 3a T | 1a T | 3a F | 1b F | 1b F | T | 2b T | 3a F | 2b F | 2a T | 3b T |
| TV4 | T | 2a T | 1b F | 1b F | 3b T | 2b F | 3b T | 2b F | 3b T | 2b F | 1a T | T | 3a T | 2a F | 3a F | 2a T | 1b F | 3a T |
| TV5 | F | 2b T | 2a T | 1a T | 3a T | 2a T | 2b F | 3a F | 3a T | 2a T | 1b F | 3b F | T | 2b F | 3b T | 1b F | T | 3b T |
| TV6 | 2a T | 3a T | 2b T | 3b T | 2b F | 1b F | 2a T | 3b T | 2b F | 1b F | 2a T | 3a F | 1b T | 3a T | 1a T | T | 3b T | 1a F |
| TV7 | 2a F | 3a T | 3a T | 3a F | 2a T | 1a T | 2b F | 1a T | 2b T | 1b F | 1b T | 2b F | 2a T | 3b T | T | 3b T | 3b T | F |
| TV8 | 3a T | 1a T | 3b T | 2b T | 1b F | 3b F | 1a F | 1b F | 1b F | 3b T | 3a T | 2a F | 2b F | 1a T | 2a F | 3a T | 2b F | 2a T |
| TV9 | 2b F | 3b F | 3a F | 2b T | 1b T | 3b T | 1b F | 1a T | 2a T | T | 3a T | 2b T | 2a T | 1a F | 1b F | 3b F | 3a F | 2a T |
| TV10 | 2a T | 3a F | 2b T | 3a F | 2a T | 1a T | 2a T | 3b T | 2b T | 1b F | 2b F | 3a T | 1b T | 3b F | 1a T | 1a T | 3b T | 1b F |
| TV11 | 1b T | 2b T | 2a T | T | 3a T | 2a T | 2b F | 3a F | 3a F | 2a T | 1b T | 3b F | T | 2b T | 3b F | 1b F | F | 3b T |
| TV12 | 1a T | 2a F | 1b T | 1a T | 3a F | 2a T | 3a T | 2b T | 3b T | 2b F | 1b T | 1a T | 3b T | 2b T | 3a F | 2a F | 1b F | 3b T |
| TV13 | 3b T | 1b T | 1a T | 1b F | 3b T | 2b T | 3b F | 2a T | F | 3a F | T | 1b T | 3a F | 2a T | 2b F | 2b F | 2a T | 3a F |

**S3**

S3 Truth Value Test Results

| Item/ Test | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | 16 | 17 | 18 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| TV1 | 1b F | 2b T | 2a T | 2a T | 1a T | 3a T | 3b T | 3a F | 3b T | 2b F | 1a T | 1b F | 1a T | 2a T | 3b T | 2b F | T | 3a T |
| TV2 | 1b F | 2b T | 2b F | 2a T | 1a T | 3b F | 1a T | 3b T | 3a F | 2b F | 1a T | 2a F | 1b T | 2a T | 3b T | 3a T | 1b F | 3a T |
| TV3 | T | 2a T | 2a T | 1b F | 3b T | 2b T | 3b T | 3a F | 1a T | 3a F | 1b F | 1b F | T | 2b T | 3a T | 2b F | 2a T | 3b T |
| TV4 | F | 2a T | 1b F | 1b F | 3b T | 2b F | 3b T | 2b T | 3b T | 2b F | 1a T | T | 3a T | 2a T | 3a F | 2a T | 1b T | 3a F |
| TV5 | T | 2b F | 2a F | 1a T | 3a T | 2a F | 2b F | 3a F | 3a F | 2a T | 1b F | 3b T | F | 2b F | 3b T | 1b F | F | 3b T |
| TV6 | 2a F | 3a T | 2b F | 3b T | 2b T | 1b F | 2a T | 3b F | 2b F | 1b F | 2a T | 3a T | 1b F | 3a F | 1a T | 1a F | 3b F | 1a T |
| TV7 | 2a F | 3a T | 3a F | 3a T | 2a T | 1a T | 2b F | 1a T | 2b T | 1b T | 1b T | 2b F | 2a F | 3b T | 1a F | 3b T | T | 1b F |

Appendix I: Truth Value judgement test results by sentence Type. Right-most three columns include percentages for the initial, secondary, and final data collection sub-periods. Sub-periods are color-coded according to key below the table.

### TV TYPE 1a

| Test Number/ Test Subject | TV1 | TV2 | TV3 | TV4 | TV5 | TV6 | TV7 | TV8 | TV9 | TV10 | TV11 | TV12 | TV13 | Initial | Secondary | Final |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| R1 | 3 | 3 | 3 | 3 | 3 | 3 | 3 | 3 | 3 | 3 | 3 | 3 | | 100% | 100% | 100% |
| R2 | 3 | 3 | 3 | 3 | 3 | 3 | 3 | 3 | 3 | 3 | | | | 100% | 100% | 100% |
| R3 | 3 | 3 | 3 | 3 | 3 | 3 | 2 | 3 | 2 | 2 | 3 | | | 100% | 92% | 78% |
| S1 | 3 | 3 | 3 | 3 | 3 | 3 | 3 | 3 | 3 | | | | | 100% | 100% | n/a |
| S2 | 3 | 3 | 3 | 3 | 3 | 2 | 3 | 2 | 2 | 3 | 2 | 3 | 2 | 94% | 67% | 78% |
| S3 | 3 | 3 | 3 | 2 | 1 | 2 | 2 | | | | | | | 100% | 50% | 67% |

### TV TYPE 1b

| Test Number/ Test Subject | TV1 | TV2 | TV3 | TV4 | TV5 | TV6 | TV7 | TV8 | TV9 | TV10 | TV11 | TV12 | TV13 | Initial | Secondary | Final |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| R1 | 3 | 3 | 3 | 3 | 3 | 3 | 3 | 3 | 3 | 3 | 2 | 3 | | 100% | 100% | 93% |
| R2 | 2 | 3 | 3 | 3 | 3 | 3 | 3 | 3 | 3 | 3 | | | | 92% | 100% | 100% |
| R3 | 2 | 3 | 2 | 2 | 3 | 2 | 2 | 2 | 2 | 1 | 2 | | | 75% | 75% | 56% |
| S1 | 3 | 3 | 3 | 3 | 2 | 3 | 2 | 3 | 3 | | | | | 94% | 89% | n/a |
| S2 | 3 | 3 | 2 | 3 | 3 | 2 | 2 | 2 | 2 | 2 | 1 | 1 | 1 | 88% | 67% | 42% |
| S3 | 2 | 2 | 3 | 2 | 2 | 2 | 1 | | | | | | | 78% | 67% | 50% |

### TV TYPE 2a

| Test Number/ Test Subject | TV1 | TV2 | TV3 | TV4 | TV5 | TV6 | TV7 | TV8 | TV9 | TV10 | TV11 | TV12 | TV13 | Initial | Secondary | Final |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| R1 | 3 | 2 | 3 | 3 | 2 | 3 | 3 | 2 | 3 | 3 | 3 | 3 | | 92% | 89% | 93% |
| R2 | 3 | 3 | 3 | 3 | 3 | 2 | 2 | 3 | 3 | 3 | | | | 100% | 83% | 100% |
| R3 | 3 | 3 | 3 | 2 | 2 | 1 | 3 | 1 | 2 | 2 | 2 | | | 92% | 58% | 67% |
| S1 | 3 | 3 | 2 | 3 | 2 | 2 | 3 | 2 | 2 | | | | | 83% | 78% | n/a |
| S2 | 3 | 3 | 3 | 2 | 3 | 3 | 2 | 1 | 3 | 3 | 3 | 1 | 3 | 94% | 67% | 83% |
| S3 | 3 | 2 | 3 | 3 | 1 | 2 | 1 | | | | | | | 89% | 67% | 50% |

### TV TYPE 2b

| Test Number/ Test Subject | TV1 | TV2 | TV3 | TV4 | TV5 | TV6 | TV7 | TV8 | TV9 | TV10 | TV11 | TV12 | TV13 | Initial | Secondary | Final |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| R1 | 3 | 3 | 3 | 2 | 3 | 3 | 2 | 2 | 2 | 1 | 2 | 3 | | 92% | 89% | 66% |
| R2 | 3 | 3 | 2 | 2 | 3 | 3 | 2 | 2 | 2 | 2 | | | | 83% | 83% | 67% |
| R3 | 3 | 3 | 3 | 2 | 1 | 2 | 1 | 3 | 1 | 2 | 1 | | | 92% | 58% | 44% |
| S1 | 3 | 3 | 2 | 3 | 2 | 3 | 2 | 3 | 2 | | | | | 89% | 78% | n/a |
| S2 | 3 | 3 | 2 | 3 | 2 | 2 | 2 | 2 | 1 | 1 | 1 | 1 | 2 | 83% | 50% | 42% |
| S3 | 2 | 2 | 1 | 2 | 3 | 2 | 2 | | | | | | | 55% | 83% | 67% |

### TV TYPE 3a

| Test Number/ Test Subject | TV1 | TV2 | TV3 | TV4 | TV5 | TV6 | TV7 | TV8 | TV9 | TV10 | TV11 | TV12 | TV13 | Initial | Secondary | Final |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| R1 | 2 | 3 | 1 | 2 | 2 | 1 | 1 | 2 | 2 | 2 | 1 | 3 | | 67% | 50% | 67% |
| R2 | 1 | 3 | 1 | 1 | 1 | 2 | 1 | 0 | 1 | 2 | | | | 50% | 33% | 50% |
| R3 | 2 | 2 | 1 | 2 | 1 | 1 | 1 | 1 | 2 | 2 | 1 | | | 58% | 33% | 56% |
| S1 | 1 | 2 | 1 | 1 | 1 | 1 | 2 | 2 | 2 | | | | | 39% | 67% | n/a |
| S2 | 1 | 2 | 1 | 2 | 2 | 2 | 2 | 3 | 1 | 1 | 1 | 1 | 0 | 56% | 67% | 25% |
| S3 | 2 | 2 | 1 | 1 | 1 | 2 | 2 | | | | | | | 55% | 33% | 67% |

### TV TYPE 3b

| Test Number/ Test Subject | TV1 | TV2 | TV3 | TV4 | TV5 | TV6 | TV7 | TV8 | TV9 | TV10 | TV11 | TV12 | TV13 | Initial | Secondary | Final |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| R1 | 3 | 2 | 2 | 3 | 2 | 2 | 2 | 2 | 1 | 3 | 2 | 2 | | 83% | 67% | 67% |
| R2 | 3 | 1 | 2 | 3 | 2 | 1 | 3 | 2 | 2 | 2 | | | | 75% | 67% | 67% |
| R3 | 3 | 0 | 3 | 3 | 2 | 3 | 1 | 2 | 1 | 2 | 2 | | | 75% | 67% | 56% |
| S1 | 2 | 2 | 2 | 2 | 3 | 2 | 2 | 3 | 2 | | | | | 72% | 78% | n/a |
| S2 | 3 | 1 | 3 | 3 | 2 | 3 | 3 | 2 | 1 | 2 | 1 | 3 | 2 | 83% | 67% | 67% |
| S3 | 3 | 2 | 3 | 3 | 2 | 1 | 3 | | | | | | | 89% | 83% | 67% |

Initial sub-period  Secondary sub-period  Final sub-period

Appendix J: Item Analysis for Truth Value test. Blue Master list number = Distractor; Red = misjudged 1x; Green = misjudged 2x

TV 1a Item Analysis

| | Master list number | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Informant | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | 16 | 17 | 18 | 19 | 20 | 21 | 22 | 23 | 24 | 25 | 26 | 27 | 28 | 29 | 30 |
| NS1 | T | T | T | T | T | T | T | T | T | T | T | T | T | T | T | T | T | T | T | T | T | T | T | T | T | T | T | T | T | T |
| NS2 | T | T | T | T | T | T | T | T | T | T | T | T | T | T | T | T | T | T | T | T | T | T | T | T | T | T | T | T | T | T |
| NS3 | T | T | T | T | T | T | T | T | T | T | T | T | T | T | T | T | T | T | T | T | T | T | T | T | T | T | T | T | T | T |
| NS4 | T | T | F | T | T | T | T | T | F | T | T | T | T | T | T | T | T | T | T | T | T | T | T | T | T | T | T | T | T | T |
| NS5 | T | T | T | T | T | T | T | T | T | T | T | T | T | T | T | T | T | T | T | T | T | T | T | T | T | T | T | T | T | T |
| NS6 | T | T | F | T | T | T | T | T | T | T | T | T | T | T | T | T | T | T | T | T | T | T | T | T | T | T | T | T | T | T |
| R1 | T | T | T | T | T | T | T | T | T | T | T | T | T | T | T | T | T | T | T | T | T | T | T | T | T | T | T | T | na | na |
| R2 | T | T | T | T | T | T | T | T | T | T | T | T | T | T | T | T | T | T | T | T | T | T | T | T | na | na | na | na | na | na |
| R3 | T | T | T | T | F | T | T | T | T | T | T | T | T | T | T | T | T | F | T | T | T | T | F | T | T | T | na | na | na | na |
| S1 | T | T | T | T | T | T | T | T | T | T | T | T | T | T | T | T | T | T | T | T | T | T | na | na | na | na | na | na | na | na |
| S2 | T | T | T | T | F | T | T | T | T | T | T | T | T | T | T | F | T | T | F | T | T | T | T | T | T | F | T | T | F | T |
| S3 | T | T | T | T | T | T | T | T | T | F | T | T | F | F | F | T | T | F | na | na | na | na | na | na | na | na | na | na | na | na |

TV 1b Item Analysis

| | Master list number | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Informant | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | 16 | 17 | 18 | 19 | 20 | 21 | 22 | 23 | 24 | 25 | 26 | 27 | 28 | 29 | 30 |
| NS1 | F | F | F | F | F | F | F | F | F | F | F | F | F | F | F | F | F | F | F | F | F | F | F | F | F | F | F | F | F | F |
| NS2 | F | F | F | F | F | F | F | F | F | F | F | F | F | F | F | F | F | F | F | F | F | F | F | F | F | F | F | F | F | F |
| NS3 | F | F | F | F | F | F | F | F | F | F | F | F | F | F | F | F | F | F | F | F | F | F | F | F | F | F | F | F | F | F |
| NS4 | F | F | F | F | F | F | F | F | F | F | F | F | F | F | F | F | F | F | F | F | F | F | F | F | F | F | F | F | F | F |
| NS5 | F | F | F | F | F | F | F | F | F | F | F | F | F | F | F | F | F | F | F | F | F | F | F | F | F | F | F | F | F | F |
| NS6 | F | F | F | F | F | F | F | F | F | F | F | F | F | F | F | F | F | F | F | F | F | F | F | F | F | F | F | F | F | F |
| R1 | F | F | F | F | F | F | T | F | F | F | F | F | F | F | F | F | F | F | F | F | F | F | F | F | F | F | F | F | na | na |
| R2 | T | F | F | F | F | F | F | F | F | F | F | F | F | F | F | F | F | F | F | F | F | F | F | F | na | na | na | na | na | na |
| R3 | F | F | T | T | F | T | T | F | F | F | F | T | F | F | F | T | F | F | F | F | F | T | F | T | F | F | na | na | na | na |
| S1 | T | F | T | F | F | F | F | F | F | F | F | F | F | F | F | F | F | F | F | F | F | F | na | na | na | na | na | na | na | na |
| S2 | F | F | F | F | F | F | T | T | T | F | F | F | F | F | T | F | F | T | F | T | F | T | F | T | F | T | F | T | T | F |
| S3 | T | F | T | T | F | F | F | F | F | F | F | T | F | F | F | F | F | T | na | na | na | na | na | na | na | na | na | na | na | na |

Appendix J (continured): Item Analysis for Truth Value test. Red = misjudged 1x; Green = misjudged 2x

TV 2a Item Analysis

| | Master list number | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Informant | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | 16 | 17 | 18 | 19 | 20 | 21 | 22 | 23 | 24 | 25 | 26 | 27 | 28 | 29 | 30 |
| NS1 | T | T | T | F | T | T | T | T | T | T | T | T | T | T | T | T | T | T | T | T | T | T | T | T | T | T | T | T | T | T |
| NS2 | T | T | T | F | T | T | T | T | T | T | T | T | T | T | T | T | T | T | T | T | T | T | T | T | T | T | T | T | T | T |
| NS3 | T | T | T | F | T | T | T | T | T | T | T | T | T | T | T | T | T | T | T | T | T | T | T | T | T | T | T | T | T | T |
| NS4 | T | T | T | F | T | T | T | T | T | T | T | T | T | T | T | T | T | T | T | T | T | T | T | T | T | T | T | T | T | T |
| NS5 | T | T | T | F | T | T | T | T | T | T | T | T | T | T | T | T | T | T | T | T | T | T | T | T | T | T | T | T | T | T |
| NS6 | T | T | T | T | T | T | T | T | T | T | T | F | T | T | T | T | T | T | T | T | T | T | T | T | T | T | T | T | T | T |
| R1 | T | T | T | F | T | T | T | T | T | T | T | T | F | T | T | T | T | T | T | T | T | T | T | T | T | T | T | T | na | na |
| R2 | T | F | T | F | T | T | T | T | T | T | T | T | T | T | T | T | T | F | T | T | T | T | T | T | na | na | na | na | na | na |
| R3 | F | F | T | F | T | T | T | T | T | F | T | T | T | T | F | T | T | T | T | F | T | F | T | F | F | T | na | na | na | na |
| S1 | F | T | T | F | T | T | T | F | T | T | T | T | T | T | T | F | T | T | T | T | F | T | na | na | na | na | na | na | na | na |
| S2 | T | T | T | T | T | T | T | F | T | T | T | F | T | T | T | T | T | F | F | T | T | T | T | T | T | T | F | T | T | T |
| S3 | T | F | F | F | T | T | T | T | T | T | T | T | F | T | T | T | T | F | na | na | na | na | na | na | na | na | na | na | na | na |

TV 2b Item Analysis

| | Master list number | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Informant | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | 16 | 17 | 18 | 19 | 20 | 21 | 22 | 23 | 24 | 25 | 26 | 27 | 28 | 29 | 30 |
| NS1 | F | F | T | F | F | F | F | F | F | F | F | F | F | F | F | F | F | F | F | F | F | F | F | F | F | F | F | F | F | F |
| NS2 | F | F | F | F | F | F | F | F | F | F | F | F | F | F | F | F | F | F | F | F | F | F | F | F | F | F | F | F | F | F |
| NS3 | F | F | F | F | F | F | F | F | F | F | F | F | F | F | F | F | F | F | F | F | F | F | F | F | F | F | F | F | F | F |
| NS4 | F | F | T | F | F | F | F | F | F | F | F | F | F | F | F | F | F | F | F | F | F | F | F | F | F | F | F | F | F | F |
| NS5 | F | F | F | F | F | F | F | F | F | F | F | F | F | F | F | F | F | F | F | F | F | F | F | F | F | F | F | F | F | F |
| NS6 | F | F | F | F | F | F | F | F | F | F | F | F | F | T | F | F | F | F | F | F | F | F | F | F | F | F | F | F | F | F |
| R1 | F | F | F | T | T | F | T | F | F | F | T | F | F | F | F | F | T | F | F | F | F | F | T | T | F | F | F | F | na | na |
| R2 | F | F | T | T | T | F | F | T | F | F | F | F | F | F | F | F | F | F | F | F | F | F | T | F | na | na | na | na | na | na |
| R3 | T | F | F | F | T | F | T | F | F | T | F | F | F | T | F | T | T | F | F | F | F | T | T | F | T | F | na | na | na | na |
| S1 | F | F | T | F | T | F | F | T | F | F | F | F | T | F | F | F | F | F | F | F | F | F | na | na | na | na | na | na | na | na |
| S2 | T | F | T | T | T | F | T | T | F | F | F | F | F | F | T | F | F | F | F | F | F | T | T | T | F | T | F | T | F | T |
| S3 | F | T | T | F | F | T | F | T | T | F | T | F | F | F | F | F | F | F | na | na | na | na | na | na | na | na | na | na | na | na |

Appendix J (continued): Item Analysis for Truth Value test.
Red = misjudged 1x; Green = misjudged 2x

TV 3a Item Analysis

| | Master list number | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Informant | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | 16 | 17 | 18 | 19 | 20 | 21 | 22 | 23 | 24 | 25 | 26 | 27 | 28 | 29 | 30 |
| NS1 | T | T | F | T | F | T | T | T | F | T | T | F | F | T | T | F | T | T | F | T | F | F | F | T | T | F | T | F | F | T |
| NS2 | F | T | T | T | T | T | F | F | T | T | T | T | F | T | F | T | T | T | T | F | T | T | T | T | T | T | T | T | F | T |
| NS3 | F | T | F | T | T | T | F | F | T | T | T | T | F | T | F | T | T | T | T | F | T | T | T | T | T | T | T | T | F | T |
| NS4 | T | T | T | T | T | T | F | F | T | T | T | F | T | T | F | T | T | F | F | F | F | F | T | T | T | F | T | T | F | T |
| NS5 | T | T | T | T | T | T | T | T | F | T | T | T | F | T | F | F | T | T | T | F | T | T | T | T | T | T | T | T | T | F |
| NS6 | F | T | T | T | T | T | T | T | T | T | T | F | T | T | T | F | T | F | T | T | T | F | T | F | T | F | T | T | F | T |
| R1 | F | T | F | F | T | T | F | F | F | T | T | F | F | T | F | F | F | T | T | T | F | T | F | T | F | T | T | T | na | na |
| R2 | F | T | F | F | T | T | T | F | F | T | F | F | F | T | F | T | F | F | F | F | F | F | F | T | na | na | na | na | na | na |
| R3 | F | T | F | F | F | T | T | T | F | T | F | T | F | T | F | F | F | T | F | T | F | T | F | T | F | F | na | na | na | na |
| S1 | F | T | F | T | F | T | T | F | F | T | F | F | F | F | F | F | F | T | T | F | T | F | na | na | na | na | na | na | na | na |
| S2 | F | T | F | T | F | T | T | F | F | F | T | T | F | T | F | T | F | T | T | T | F | F | F | F | F | F | F | F | F | F |
| S3 | T | T | F | T | T | F | F | F | T | F | T | F | F | F | T | F | T | T | na | na | na | na | na | na | na | na | na | na | na | na |

TV 3b Item Analysis

| | Master list number | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Informant | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | 16 | 17 | 18 | 19 | 20 | 21 | 22 | 23 | 24 | 25 | 26 | 27 | 28 | 29 | 30 |
| NS1 | T | T | F | F | F | F | F | T | F | F | T | T | F | T | T | F | F | T | T | T | T | T | T | T | F | F | T | F | F | F |
| NS2 | T | T | F | F | F | T | T | T | F | T | T | T | T | T | F | T | T | T | T | T | T | T | T | T | T | T | T | F | T | T |
| NS3 | T | T | F | F | F | T | T | T | F | T | T | T | F | T | F | T | T | T | T | T | T | T | T | T | T | T | T | T | T | T |
| NS4 | T | T | F | F | F | F | F | T | F | T | T | T | T | T | F | T | F | T | T | T | F | T | T | T | T | F | T | T | T | F |
| NS5 | T | T | F | T | F | T | T | T | T | T | T | T | T | T | T | T | T | T | T | T | T | T | T | T | T | T | F | T | T | T |
| NS6 | T | F | T | F | F | T | T | T | F | T | T | T | T | T | F | F | T | T | T | T | T | T | F | T | T | F | F | T | T | T |
| R1 | F | F | F | F | F | T | F | F | T | T | T | T | T | T | T | T | T | T | T | T | F | T | T | T | F | T | T | T | na | na |
| R2 | T | T | T | F | F | F | T | T | F | T | T | T | F | T | F | F | T | T | T | T | F | T | T | T | na | na | na | na | na | na |
| R3 | T | T | F | F | F | F | T | T | T | T | T | T | F | T | T | T | T | F | T | T | F | T | T | T | F | T | na | na | na | na |
| S1 | T | T | F | T | F | T | T | T | F | F | T | T | T | T | F | T | T | T | T | T | T | T | na | na | na | na | na | na | na | na |
| S2 | T | T | T | F | F | F | F | T | F | T | T | T | F | T | T | T | T | T | T | T | F | T | T | T | F | T | T | T | T | T |
| S3 | T | T | T | T | F | T | T | T | T | T | T | T | F | T | F | F | T | T | na | na | na | na | na | na | na | na | na | na | na | na |

Appendix K: Test subjects R1, R2, and R3 grammaticality judgment test results. Test item numbers listed across top, test numbers listed down left side. Sentence type indicated above answer. Red letters indicate misjudged answers.

**R1**

| Item/ Test | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 |
|---|---|---|---|---|---|---|---|---|---|---|---|
| GJ1 | 1a Y | 1b N | 1c Y | 1d N | 2a Y | 2b N | 2c Y | 2d N | 3a Y | 3b Y | 3c Y |
| GJ2 | 3a Y | 1a Y | 2c Y | 1b N | 2d N | 1c Y | 2b N | 1d N | 3b Y | 2a Y | 3c Y |
| GJ3 | 1c Y | 2c Y | 1a Y | 1d N | 2d N | 3c Y | 2a Y | 3a Y | 1b N | 2b N | 3b Y |
| GJ4 | 3c Y | 1c Y | 2b N | 3a Y | 1a Y | 1d N | 2c N | 3b Y | 1b N | 2a Y | 2d N |
| GJ5 | 1a Y | 3b Y | 2d N | 2b N | 1d N | 1b N | 3c Y | 3a Y | 2c Y | 2a Y | 1c Y |
| GJ6 | 3c Y | 1b N | 1d N | 2b N | 2d N | 3b Y | 1a Y | 1c Y | 2a N | 2c Y | 3a Y |
| GJ7 | 1a Y | 2c Y | 1b Y | 2d N | 1c Y | 3a Y | 1d N | 3b Y | 2a Y | 3c Y | 2b N |
| GJ8 | 2d N | 1a Y | 2a Y | 3a Y | 3c Y | 2b N | 3b Y | 1c Y | 2c Y | 1b N | 1d N |
| GJ9 | 1d N | 2b N | 2d N | 3b Y | 1a Y | 1c Y | 2a Y | 2c N | 3a Y | 3c Y | 1b N |
| GJ10 | 1d N | 2d N | 1a Y | 2a Y | 3a N | 1b Y | 2b N | 3b Y | 1c Y | 2c Y | 3c Y |
| GJ11 | 1a Y | 1b N | 1c Y | 1d N | 2a Y | 2b Y | 2c Y | 2d N | 3a Y | 3b Y | 3c N |
| GJ12 | 2d Y | 2a Y | 1b N | 3b Y | 3c Y | 1d N | 1a Y | 3a Y | 2b N | 1c Y | 2c Y |

**R2**

| Item/ Test | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 |
|---|---|---|---|---|---|---|---|---|---|---|---|
| GJ1 | 1a Y | 1b N | 1c Y | 1d N | 2a Y | 2b N | 2c Y | 2d N | 3a Y | 3b Y | 3c Y |
| GJ2 | 3a Y | 1a Y | 2c Y | 1b N | 2d N | 1c Y | 2b N | 1d N | 3b Y | 2a Y | 3c Y |
| GJ3 | 1c Y | 2c Y | 1a Y | 1d N | 2d Y | 3c Y | 2a Y | 3a Y | 1b N | 2b N | 3b N |
| GJ4 | 3c Y | 1c Y | 2b N | 3a Y | 1a Y | 1d N | 2c Y | 3b Y | 1b N | 2a Y | 2d N |
| GJ5 | 1a Y | 3b Y | 2d N | 2b N | 1d N | 1b N | 3c Y | 3a Y | 2c Y | 2a Y | 1c Y |
| GJ6 | 3c Y | 1b N | 1d N | 2b N | 2d N | 3b Y | 1a Y | 1c Y | 2a Y | 2c Y | 3a N |
| GJ7 | 1a Y | 2c Y | 1b N | 2d N | 1c Y | 3a Y | 1d N | 3b Y | 2a Y | 3c Y | 2b Y |
| GJ8 | 2d N | 1a Y | 2a Y | 3a N | 3c N | 2b N | 3b Y | 1c Y | 2c Y | 1b N | 1d N |
| GJ9 | 1d N | 2b Y | 2d Y | 3b Y | 1a Y | 1c Y | 2a Y | 2c Y | 3a Y | 3c Y | 1b N |
| GJ10 | 1d N | 2d Y | 1a Y | 2a Y | 3a N | 1b N | 2b N | 3b Y | 1c Y | 2c Y | 3c Y |

**R3**

| Item/ Test | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 |
|---|---|---|---|---|---|---|---|---|---|---|---|
| GJ1 | 1a Y | 1b N | 1c Y | 1d N | 2a Y | 2b N | 2c Y | 2d Y | 3a Y | 3b Y | 3c Y |
| GJ2 | 3a Y | 1a Y | 2c Y | 1b N | 2d N | 1c Y | 2b N | 1d N | 3b Y | 2a Y | 3c Y |
| GJ3 | 1c Y | 2c Y | 1a Y | 1d N | 2d Y | 3c Y | 2a Y | 3a N | 1b N | 2b N | 3b Y |
| GJ4 | 3c Y | 1c Y | 2b N | 3a Y | 1a Y | 1d N | 2c Y | 3b Y | 1b Y | 2a Y | 2d N |
| GJ5 | 1a Y | 3b Y | 2d Y | 2b N | 1d N | 1b N | 3c N | 3a Y | 2c Y | 2a Y | 1c Y |
| GJ6 | 3c Y | 1b N | 1d N | 2b N | 2d Y | 3b Y | 1a Y | 1c Y | 2a N | 2c N | 3a Y |
| GJ7 | 1a Y | 2c Y | 1b N | 2d N | 1c Y | 3a N | 1d N | 3b Y | 2a Y | 3c Y | 2b Y |
| GJ8 | 2d Y | 1a Y | 2a Y | 3a Y | 3c Y | 2b N | 3b N | 1c N | 2c N | 1b N | 1d N |
| GJ9 | 1d N | 2b N | 2d N | 3b Y | 1a Y | 1c Y | 2a N | 2c Y | 3a N | 3c Y | 1b N |
| GJ10 | 1d N | 2d Y | 1a Y | 2a Y | 3a N | 1b Y | 2b N | 3b N | 1c Y | 2c N | 3c Y |
| GJ11 | 1a Y | 1b Y | 1c N | 1d N | 2a N | 2b N | 2c Y | 2d N | 3a N | 3b Y | 3c Y |

Appendix K (continued): Test subjects S1, S2, and S3 grammaticality judgment test results. Test item numbers listed across top, test numbers listed down left side. Sentence type indicated above answer. Red letters indicate misjudged answers.

**S1**

| Item/ Test | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 |
|---|---|---|---|---|---|---|---|---|---|---|---|
| GJ1 | 1a Y | 1b N | 1c Y | 1d N | 2a Y | 2b N | 2c Y | 2d N | 3a Y | 3b Y | 3c N |
| GJ2 | 3a Y | 1a Y | 2c Y | 1b N | 2d N | 1c Y | 2b N | 1d N | 3b Y | 2a Y | 3c Y |
| GJ3 | 1c Y | 2c Y | 1a Y | 1d N | 2d N | 3c Y | 2a Y | 3a Y | 1b N | 2b N | 3b Y |
| GJ4 | 3c N | 1c Y | 2b N | 3a Y | 1a Y | 1d N | 2c Y | 3b Y | 1b N | 2a Y | 2d N |
| GJ5 | 1a Y | 3b Y | 2d N | 2b N | 1d N | 1b N | 3c Y | 3a Y | 2c Y | 2a Y | 1c N |
| GJ6 | 3c Y | 1b N | 1d N | 2b N | 2d Y | 3b Y | 1a Y | 1c Y | 2a N | 2c Y | 3a N |
| GJ7 | 1a Y | 2c Y | 1b N | 2d N | 1c Y | 3a Y | 1d N | 3b N | 2a Y | 3c Y | 2b N |
| GJ8 | 2d Y | 1a Y | 2a Y | 3a Y | 3c Y | 2b N | 3b Y | 1c Y | 2c Y | 1b N | 1d N |
| GJ9 | 1d N | 2b N | 2d N | 3b Y | 1a Y | 1c Y | 2a Y | 2c Y | 3a N | 3c Y | 1b N |

**S2**

| Item/ Test | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 |
|---|---|---|---|---|---|---|---|---|---|---|---|
| GJ1 | 1a Y | 1b N | 1c Y | 1d N | 2a Y | 2b N | 2c Y | 2d Y | 3a Y | 3b Y | 3c N |
| GJ2 | 3a Y | 1a Y | 2c Y | 1b N | 2d Y | 1c Y | 2b N | 1d N | 3b Y | 2a Y | 3c Y |
| GJ3 | 1c Y | 2c Y | 1a Y | 1d N | 2d Y | 3c Y | 2a Y | 3a Y | 1b N | 2b N | 3b Y |
| GJ4 | 3c N | 1c Y | 2b N | 3a Y | 1a Y | 1d N | 2c Y | 3b Y | 1b N | 2a Y | 2d N |
| GJ5 | 1a Y | 3b Y | 2d Y | 2b N | 1d N | 1b Y | 3c Y | 3a Y | 2c Y | 2a Y | 1c N |
| GJ6 | 3c Y | 1b N | 1d N | 2b N | 2d Y | 3b Y | 1a Y | 1c N | 2a N | 2c Y | 3a Y |
| GJ7 | 1a Y | 2c Y | 1b Y | 2d N | 1c Y | 3a Y | 1d N | 3b N | 2a Y | 3c Y | 2b N |
| GJ8 | 2d Y | 1a Y | 2a Y | 3a Y | 3c Y | 2b Y | 3b Y | 1c N | 2c Y | 1b N | 1d N |
| GJ9 | 1d N | 2b Y | 2d N | 3b Y | 1a N | 1c Y | 2a Y | 2c Y | 3a N | 3c Y | 1b N |
| GJ10 | 1d N | 2d Y | 1a Y | 2a Y | 3a Y | 1b Y | 2b N | 3b N | 1c N | 2c Y | 3c Y |
| GJ11 | 1a Y | 1b Y | 1c Y | 1d N | 2a N | 2b Y | 2c Y | 2d N | 3a Y | 3b N | 3c N |
| GJ12 | 2d Y | 2a Y | 1b Y | 3b Y | 3c Y | 1d N | 1a Y | 3a N | 2b Y | 1c N | 2c Y |
| GJ13 | 3b Y | 3c Y | 2d N | 2c Y | 3a N | 2a Y | 1b Y | 1c N | 1d Y | 1a Y | 2b Y |

**S3**

| Item/ Test | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 |
|---|---|---|---|---|---|---|---|---|---|---|---|
| GJ1 | 1a Y | 1b N | 1c Y | 1d N | 2a Y | 2b N | 2c Y | 2d N | 3a Y | 3b Y | 3c N |
| GJ2 | 3a Y | 1a Y | 2c Y | 1b N | 2d Y | 1c Y | 2b N | 1d N | 3b Y | 2a Y | 3c Y |
| GJ3 | 1c Y | 2c Y | 1a Y | 1d N | 2d Y | 3c Y | 2a Y | 3a N | 1b N | 2b N | 3b Y |
| GJ4 | 3c Y | 1c Y | 2b N | 3a N | 1a Y | 1d N | 2c Y | 3b Y | 1b Y | 2a N | 2d N |
| GJ5 | 1a Y | 3b Y | 2d Y | 2b Y | 1d N | 1b Y | 3c Y | 3a Y | 2c Y | 2a Y | 1c Y |
| GJ6 | 3c Y | 1b N | 1d N | 2b Y | 2d N | 3b N | 1a N | 1c Y | 2a N | 2c Y | 3a Y |
| GJ7 | 1a Y | 2c Y | 1b Y | 2d N | 1c Y | 3a Y | 1d N | 3b N | 2a Y | 3c Y | 2b Y |

Appendix L: Test subjects' grammaticality judgement test results by sentence type. Right-most three columns include percentages for the first, secondary, and final data collection sub-periods. Sub-periods are color-coded according to key at bottom.

**TYPE 1a**

| Test Number/ Test Subject | GJ1 | GJ2 | GJ3 | GJ4 | GJ5 | GJ6 | GJ7 | GJ8 | GJ9 | GJ10 | GJ11 | GJ12 | GJ13 | Initial | Secondary | Final |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| R1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | | 100% | 100% | 100% |
| R2 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | | | | 100% | 100% | 100% |
| R3 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | | | 100% | 100% | 100% |
| S1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | | | | | 100% | 100% | n/a |
| S2 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 0 | 1 | 1 | 1 | 1 | 100% | 67% | 100% |
| S3 | 1 | 1 | 1 | 1 | 1 | 0 | 1 | | | | | | | 100% | 100% | 50% |

**TYPE 1b**

| Test Number/ Test Subject | GJ1 | GJ2 | GJ3 | GJ4 | GJ5 | GJ6 | GJ7 | GJ8 | GJ9 | GJ10 | GJ11 | GJ12 | GJ13 | Initial | Secondary | Final |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| R1 | 1 | 1 | 1 | 1 | 1 | 1 | 0 | 1 | 1 | 0 | 1 | 1 | | 100% | 75% | 75% |
| R2 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | | | | 100% | 100% | 100% |
| R3 | 1 | 1 | 1 | 0 | 1 | 1 | 1 | 1 | 1 | 0 | 0 | | | 75% | 100% | 50% |
| S1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | | | | | 100% | 100% | n/a |
| S2 | 1 | 1 | 1 | 1 | 0 | 1 | 0 | 1 | 1 | 0 | 0 | 0 | 0 | 83% | 67% | 0% |
| S3 | 1 | 0 | 1 | 0 | 0 | 1 | 0 | | | | | | | | | |

**TYPE 1c**

| Test Number/ Test Subject | GJ1 | GJ2 | GJ3 | GJ4 | GJ5 | GJ6 | GJ7 | GJ8 | GJ9 | GJ10 | GJ11 | GJ12 | GJ13 | Initial | Secondary | Final |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| R1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | | 100% | 100% | 100% |
| R2 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | | | | 100% | 100% | 100% |
| R3 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 0 | 1 | 1 | 0 | | | 100% | 100% | 50% |
| S1 | 1 | 1 | 1 | 1 | 0 | 1 | 1 | 1 | 1 | | | | | 83% | 100% | n/a |
| S2 | 1 | 1 | 1 | 1 | 0 | 0 | 1 | 0 | 1 | 0 | 1 | 0 | 0 | 67% | 67% | 25% |
| S3 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | | | | | | | 100% | 100% | 100% |

**TYPE 1d**

| Test Number/ Test Subject | GJ1 | GJ2 | GJ3 | GJ4 | GJ5 | GJ6 | GJ7 | GJ8 | GJ9 | GJ10 | GJ11 | GJ12 | GJ13 | Initial | Secondary | Final |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| R1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | | 100% | 100% | 100% |
| R2 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | | | | 100% | 100% | 100% |
| R3 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | | | 100% | 100% | 100% |
| S1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | | | | | 100% | 100% | n/a |
| S2 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 0 | 100% | 100% | 75% |
| S3 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | | | | | | | 100% | 100% | 100% |

**TYPE 2a**

| Test Number/ Test Subject | GJ1 | GJ2 | GJ3 | GJ4 | GJ5 | GJ6 | GJ7 | GJ8 | GJ9 | GJ10 | GJ11 | GJ12 | GJ13 | Initial | Secondary | Final |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| R1 | 1 | 1 | 1 | 1 | 1 | 0 | 1 | 1 | 1 | 1 | 1 | 1 | | 100% | 75% | 100% |
| R2 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | | | | 100% | 100% | 100% |
| R3 | 1 | 1 | 1 | 1 | 1 | 0 | 1 | 1 | 0 | 1 | 0 | | | 100% | 67% | 50% |
| S1 | 1 | 1 | 1 | 1 | 1 | 0 | 1 | 1 | 1 | | | | | 83% | 100% | n/a |
| S2 | 1 | 1 | 1 | 1 | 1 | 0 | 1 | 1 | 1 | 1 | 0 | 1 | 1 | 83% | 100% | 75% |
| S3 | 1 | 1 | 1 | 0 | 1 | 0 | 1 | | | | | | | 100% | 50% | 50% |

**TYPE 2b**

| Test Number/ Test Subject | GJ1 | GJ2 | GJ3 | GJ4 | GJ5 | GJ6 | GJ7 | GJ8 | GJ9 | GJ10 | GJ11 | GJ12 | GJ13 | Initial | Secondary | Final |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| R1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 0 | 1 | | 100% | 100% | 75% |
| R2 | 1 | 1 | 1 | 1 | 1 | 1 | 0 | 1 | 0 | 1 | | | | 100% | 75% | 50% |
| R3 | 1 | 1 | 1 | 1 | 1 | 1 | 0 | 1 | 0 | 1 | 1 | | | 100% | 67% | 75% |
| S1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | | | | | 100% | 100% | n/a |
| S2 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 0 | 0 | 1 | 0 | 0 | 0 | 100% | 34% | 25% |
| S3 | 1 | 1 | 1 | 1 | 0 | 0 | 0 | | | | | | | 100% | 50% | 0% |

**TYPE 2c**

| Test Number/ Test Subject | GJ1 | GJ2 | GJ3 | GJ4 | GJ5 | GJ6 | GJ7 | GJ8 | GJ9 | GJ10 | GJ11 | GJ12 | GJ13 | Initial | Secondary | Final |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| R1 | 1 | 1 | 1 | 0 | 1 | 1 | 1 | 1 | 0 | 1 | 1 | 1 | | 75% | 100% | 75% |
| R2 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | | | | 100% | 100% | 100% |
| R3 | 1 | 1 | 1 | 1 | 1 | 0 | 1 | 0 | 1 | 0 | 1 | | | 100% | 67% | 50% |
| S1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | | | | | 100% | 100% | n/a |
| S2 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 100% | 100% | 100% |
| S3 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | | | | | | | 100% | 100% | 100% |

**TYPE 2d**

| Test Number/ Test Subject | GJ1 | GJ2 | GJ3 | GJ4 | GJ5 | GJ6 | GJ7 | GJ8 | GJ9 | GJ10 | GJ11 | GJ12 | GJ13 | Initial | Secondary | Final |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| R1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 0 | | 100% | 100% | 75% |
| R2 | 1 | 1 | 0 | 1 | 1 | 1 | 1 | 1 | 0 | 0 | | | | 75% | 100% | 0% |
| R3 | 0 | 1 | 0 | 1 | 0 | 0 | 1 | 0 | 1 | 0 | 1 | | | 50% | 33% | 50% |
| S1 | 1 | 1 | 1 | 1 | 1 | 0 | 1 | 0 | 1 | | | | | 83% | 67% | n/a |
| S2 | 0 | 0 | 0 | 1 | 0 | 0 | 1 | 0 | 1 | 0 | 1 | 0 | 1 | 17% | 67% | 50% |
| S3 | 1 | 0 | 0 | 1 | 0 | 1 | 1 | | | | | | | 33% | 50% | 100% |

Initial sub-period  Secondary sub-period  Final sub-period

Appendix L (continued): Test subjects' grammaticality judgement test results.

TYPE 3a

| Test Number/ Test Subject | GJ1 | GJ2 | GJ3 | GJ4 | GJ5 | GJ6 | GJ7 | GJ8 | GJ9 | GJ10 | GJ11 | GJ12 | GJ13 | Initial | Secondary | Final |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| R1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 0 | 1 | 1 | | 100% | 100% | 75% |
| R2 | 1 | 1 | 1 | 1 | 1 | 0 | 1 | 0 | 1 | 0 | | | | 100% | 50% | 50% |
| R3 | 1 | 1 | 0 | 1 | 1 | 1 | 0 | 1 | 0 | 0 | 0 | | | 75% | 33% | 25% |
| S1 | 1 | 1 | 1 | 1 | 1 | 0 | 1 | 1 | 0 | | | | | 83% | 67% | n/a |
| S2 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 0 | 0 | 1 | 0 | 0 | 100% | 67% | 25% |
| S3 | 1 | 1 | 0 | 0 | 1 | 1 | 1 | | | | | | | 67% | 50% | 100% |

TYPE 3b

| Test Number/ Test Subject | GJ1 | GJ2 | GJ3 | GJ4 | GJ5 | GJ6 | GJ7 | GJ8 | GJ9 | GJ10 | GJ11 | GJ12 | GJ13 | Initial | Secondary | Final |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| R1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | | 100% | 100% | 100% |
| R2 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | | | | 100% | 100% | 100% |
| R3 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 0 | 1 | 0 | 1 | | | 100% | 100% | 50% |
| S1 | 1 | 1 | 1 | 1 | 1 | 1 | 0 | 1 | 1 | | | | | 100% | 67% | n/a |
| S2 | 1 | 1 | 1 | 1 | 1 | 1 | 0 | 1 | 1 | 1 | 0 | 1 | 1 | 100% | 67% | 75% |
| S3 | 1 | 1 | 1 | 1 | 1 | 0 | 0 | | | | | | | 100% | 100% | 0% |

TYPE 3c

| Test Number/ Test Subject | GJ1 | GJ2 | GJ3 | GJ4 | GJ5 | GJ6 | GJ7 | GJ8 | GJ9 | GJ10 | GJ11 | GJ12 | GJ13 | Initial | Secondary | Final |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| R1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 0 | 1 | | 100% | 100% | 75% |
| R2 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 0 | 1 | 1 | | | | 100% | 75% | 100% |
| R3 | 1 | 1 | 1 | 1 | 0 | 1 | 1 | 1 | 1 | 1 | 1 | | | 100% | 67% | 100% |
| S1 | 0 | 1 | 1 | 0 | 1 | 1 | 1 | 1 | 1 | | | | | 67% | 100% | n/a |
| S2 | 0 | 1 | 1 | 0 | 1 | 1 | 1 | 1 | 1 | 1 | 0 | 1 | 1 | 67% | 100% | 75% |
| S3 | 0 | 1 | 1 | 1 | 1 | 1 | 1 | | | | | | | 67% | 100% | 100% |

Initial sub-period  Secondary sub-period  Final sub-period

Appendix M: Item Analysis for Grammaticality Judgment test. Master list numbers correspond to the master list in Appendix C

Blue = Distractor; Red = misjudged item

GJ 1a Item Analysis (Object oriented, by number)

| | Master list number | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Informant | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 |
| NS1 | Y | Y | Y | Y | Y | Y | Y | Y | Y | Y | Y | Y | Y |
| NS2 | Y | Y | Y | Y | Y | Y | Y | Y | Y | Y | Y | Y | Y |
| NS3 | Y | Y | Y | Y | Y | Y | Y | Y | Y | Y | Y | Y | Y |
| NS4 | Y | Y | Y | Y | Y | Y | Y | Y | Y | Y | Y | Y | Y |
| NS5 | Y | Y | Y | Y | Y | Y | Y | Y | Y | Y | Y | Y | Y |
| NS6 | Y | Y | Y | Y | Y | Y | Y | Y | Y | Y | Y | Y | Y |
| R1 | Y | Y | Y | Y | Y | Y | Y | Y | Y | Y | Y | Y | na |
| R2 | Y | Y | Y | Y | Y | Y | Y | Y | Y | Y | na | na | na |
| R3 | Y | Y | Y | Y | Y | Y | Y | Y | Y | Y | Y | na | na |
| S1 | Y | Y | Y | Y | Y | Y | Y | Y | Y | na | na | na | na |
| S2 | Y | Y | Y | Y | Y | Y | Y | Y | N | Y | Y | Y | Y |
| S3 | Y | Y | Y | Y | Y | N | Y | na | na | na | na | na | na |

GJ 1b Item Analysis (Subject oriented, by number)

| | Master list number | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Informant | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 |
| NS1 | N | N | N | N | N | N | N | N | N | Y | N | N | N |
| NS2 | N | N | N | N | N | N | N | N | N | Y | N | N | N |
| NS3 | N | N | N | N | N | N | N | N | N | N | N | N | N |
| NS4 | N | N | N | N | N | N | N | N | N | N | N | N | N |
| NS5 | N | N | N | N | N | N | N | N | N | N | N | N | N |
| NS6 | N | N | N | N | N | N | N | N | N | Y | N | N | N |
| R1 | N | N | N | N | N | N | Y | N | N | Y | N | N | na |
| R2 | N | N | N | N | N | N | N | N | N | N | na | na | na |
| R3 | N | N | N | Y | N | N | N | N | N | Y | Y | na | na |
| S1 | N | N | N | N | N | N | N | N | N | na | na | na | na |
| S2 | N | N | N | N | Y | N | Y | N | N | Y | Y | Y | Y |
| S3 | N | Y | N | Y | Y | N | Y | na | na | na | na | na | na |

GJ 1c Item Analysis (Object oriented, by gender)

| | Master list number | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Informant | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 |
| NS1 | Y | Y | Y | Y | Y | Y | Y | Y | Y | Y | Y | Y | Y |
| NS2 | Y | Y | Y | Y | Y | Y | Y | Y | Y | Y | Y | Y | Y |
| NS3 | Y | Y | Y | Y | Y | Y | Y | N | Y | Y | Y | Y | Y |
| NS4 | Y | Y | Y | Y | Y | Y | Y | Y | Y | Y | Y | Y | Y |
| NS5 | Y | Y | Y | Y | Y | Y | Y | Y | Y | Y | Y | Y | Y |
| NS6 | Y | Y | Y | Y | Y | Y | Y | N | Y | Y | Y | Y | Y |
| R1 | Y | Y | Y | Y | Y | Y | Y | Y | Y | Y | Y | Y | na |
| R2 | Y | Y | Y | Y | Y | Y | Y | Y | Y | Y | na | na | na |
| R3 | Y | Y | Y | Y | Y | Y | Y | N | Y | Y | N | na | na |
| S1 | Y | Y | Y | Y | Y | Y | Y | Y | Y | na | na | na | na |
| S2 | Y | Y | Y | Y | N | N | Y | N | Y | N | Y | N | N |
| S3 | Y | Y | Y | Y | Y | Y | Y | na | na | na | na | na | na |

Appendix M (continued): Item Analysis for Grammaticality Judgment test. Master list numbers correspond to the master list in Appendix C

Blue = Distractor; Red = misjudged item

GJ 1d Item Analysis (Subject oriented, by gender)

| | Master list number | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Informant | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 |
| NS1 | N | N | N | N | N | N | N | N | N | N | N | N | N |
| NS2 | N | N | N | N | N | N | N | N | N | N | N | N | N |
| NS3 | N | N | N | N | N | N | N | N | N | N | N | N | N |
| NS4 | N | N | N | N | N | N | N | N | N | N | N | N | N |
| NS5 | N | N | N | N | N | N | N | N | N | N | N | N | N |
| NS6 | N | N | N | N | N | N | N | N | N | N | N | N | N |
| R1 | N | N | N | N | N | N | N | N | N | N | N | N | na |
| R2 | N | N | N | N | N | N | N | N | N | N | na | na | na |
| R3 | N | N | N | N | N | N | N | N | N | N | N | na | na |
| S1 | N | N | N | N | N | N | N | N | N | na | na | na | na |
| S2 | N | N | N | N | N | N | N | N | N | N | N | N | Y |
| S3 | N | N | N | N | N | N | N | na | na | na | na | na | na |

GJ 2a Item Analysis (Object oriented, by number)

| | Master list number | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Informant | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 |
| NS1 | Y | N | Y | Y | Y | Y | Y | Y | Y | Y | Y | Y | Y |
| NS2 | Y | Y | Y | Y | Y | Y | Y | Y | Y | Y | Y | Y | Y |
| NS3 | Y | Y | Y | Y | Y | Y | Y | Y | Y | Y | Y | Y | Y |
| NS4 | Y | Y | Y | Y | Y | Y | Y | Y | Y | Y | Y | Y | Y |
| NS5 | Y | Y | Y | Y | Y | Y | Y | Y | Y | Y | Y | Y | Y |
| NS6 | Y | Y | Y | Y | Y | Y | Y | Y | Y | Y | Y | Y | Y |
| R1 | Y | Y | Y | Y | Y | N | Y | Y | Y | Y | Y | Y | na |
| R2 | Y | Y | Y | Y | Y | Y | Y | Y | Y | Y | na | na | na |
| R3 | Y | Y | Y | Y | Y | N | Y | Y | N | Y | N | na | na |
| S1 | Y | Y | Y | Y | Y | N | Y | Y | Y | na | na | na | na |
| S2 | Y | Y | Y | Y | Y | N | Y | Y | Y | Y | N | Y | Y |
| S3 | Y | Y | Y | N | Y | N | Y | na | na | na | na | na | na |

GJ 2b Item Analysis (Subject oriented, by number)

| | Master list number | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Informant | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 |
| NS1 | N | Y | N | N | N | N | N | N | N | N | N | N | N |
| NS2 | N | N | N | N | N | N | N | N | N | N | N | N | N |
| NS3 | N | N | N | N | N | N | N | N | N | N | N | N | N |
| NS4 | N | N | N | N | N | N | N | N | N | N | N | N | N |
| NS5 | N | Y | N | N | N | N | N | N | N | N | N | N | N |
| NS6 | N | N | N | N | N | N | N | N | N | N | N | N | N |
| R1 | N | N | N | N | N | N | N | N | N | N | Y | N | na |
| R2 | N | N | N | N | N | N | Y | N | Y | N | na | na | na |
| R3 | N | N | N | N | N | N | Y | N | Y | N | N | na | na |
| S1 | N | N | N | N | N | N | N | N | N | na | na | na | na |
| S2 | N | N | N | N | N | N | N | Y | Y | N | Y | Y | Y |
| S3 | N | N | N | N | Y | Y | Y | na | na | na | na | na | na |

Appendix M (continued): Item Analysis for Grammaticality Judgment test. Master list numbers correspond to the master list in Appendix C

Blue = Distractor; Red = misjudged item

GJ 2c Item Analysis (Object oriented, by gender)

| | Master list number | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Informant | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 |
| NS1 | Y | Y | Y | Y | Y | Y | Y | Y | Y | Y | Y | Y | Y |
| NS2 | Y | Y | Y | Y | Y | Y | Y | Y | Y | Y | Y | Y | Y |
| NS3 | Y | Y | Y | Y | Y | Y | Y | Y | Y | Y | Y | Y | Y |
| NS4 | Y | Y | Y | Y | Y | Y | Y | Y | Y | Y | Y | Y | Y |
| NS5 | Y | Y | Y | Y | Y | Y | Y | Y | Y | Y | Y | Y | Y |
| NS6 | Y | Y | Y | Y | Y | Y | Y | Y | Y | Y | Y | Y | Y |
| R1 | Y | Y | Y | N | Y | Y | Y | Y | N | Y | Y | Y | na |
| R2 | Y | Y | Y | Y | Y | Y | Y | Y | Y | Y | na | na | na |
| R3 | Y | Y | Y | Y | Y | N | Y | N | Y | N | Y | na | na |
| S1 | Y | Y | Y | Y | Y | Y | Y | Y | Y | na | na | na | na |
| S2 | Y | Y | Y | Y | Y | Y | Y | Y | Y | Y | Y | Y | Y |
| S3 | Y | Y | Y | Y | Y | Y | Y | na | na | na | na | na | na |

GJ 2d Item Analysis (Subject oriented, by gender)

| | Master list number | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Informant | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 |
| NS1 | N | Y | N | N | N | N | N | N | N | N | N | N | N |
| NS2 | N | N | N | N | N | N | N | N | N | N | N | N | N |
| NS3 | N | N | N | N | N | N | N | N | N | N | N | N | N |
| NS4 | N | N | N | N | N | N | N | N | N | N | N | N | N |
| NS5 | N | N | N | N | N | N | N | N | N | N | N | N | N |
| NS6 | N | N | N | N | N | N | N | N | N | N | N | N | N |
| R1 | N | N | N | N | N | N | N | N | N | N | N | Y | na |
| R2 | N | N | Y | N | N | N | N | N | Y | Y | na | na | na |
| R3 | Y | N | Y | N | Y | Y | N | Y | N | Y | N | na | na |
| S1 | N | N | N | N | N | Y | N | Y | N | na | na | na | na |
| S2 | Y | Y | Y | N | Y | Y | N | Y | N | Y | N | Y | N |
| S3 | N | Y | Y | N | Y | N | N | na | na | na | na | na | na |

GJ 3a Item Analysis (Object oriented)

| | Master list number | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Informant | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 |
| NS1 | Y | Y | N | Y | Y | Y | Y | Y | Y | Y | N | Y | Y |
| NS2 | Y | Y | N | Y | Y | Y | Y | Y | N | Y | Y | Y | Y |
| NS3 | Y | Y | Y | Y | Y | Y | Y | Y | Y | Y | N | Y | Y |
| NS4 | Y | Y | Y | Y | Y | Y | Y | Y | N | Y | N | Y | Y |
| NS5 | Y | Y | Y | Y | Y | Y | Y | Y | Y | Y | N | Y | Y |
| NS6 | Y | Y | Y | Y | Y | Y | Y | Y | Y | Y | Y | Y | Y |
| R1 | Y | Y | Y | Y | Y | Y | Y | Y | Y | N | Y | Y | na |
| R2 | Y | Y | Y | Y | Y | N | Y | N | Y | N | na | na | na |
| R3 | Y | Y | N | Y | Y | Y | N | Y | N | N | N | na | na |
| S1 | Y | Y | Y | Y | Y | N | Y | Y | N | na | na | na | na |
| S2 | Y | Y | Y | Y | Y | Y | Y | Y | N | Y | Y | N | N |
| S3 | Y | Y | N | N | Y | Y | Y | na | na | na | na | na | na |

Appendix M (continued): Item Analysis for Grammaticality Judgment test. Master list numbers correspond to the master list in Appendix C

Blue = Distractor; Red = misjudged item

GJ 3b Item Analysis (Subject oriented)

| | Master list number | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Informant | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 |
| NS1 | Y | Y | Y | Y | Y | Y | Y | Y | Y | N | Y | Y | Y |
| NS2 | Y | Y | Y | Y | Y | Y | N | Y | Y | Y | Y | Y | Y |
| NS3 | Y | Y | Y | Y | Y | Y | Y | Y | Y | Y | Y | Y | Y |
| NS4 | Y | Y | Y | Y | Y | Y | Y | Y | Y | Y | Y | Y | Y |
| NS5 | Y | Y | Y | Y | Y | Y | Y | Y | Y | Y | Y | Y | Y |
| NS6 | Y | Y | Y | Y | Y | Y | Y | Y | Y | Y | Y | Y | Y |
| R1 | Y | Y | Y | Y | Y | Y | Y | Y | Y | Y | Y | Y | na |
| R2 | Y | Y | Y | Y | Y | Y | Y | Y | Y | Y | na | na | na |
| R3 | Y | Y | Y | Y | Y | Y | Y | N | Y | N | Y | na | na |
| S1 | Y | Y | Y | Y | Y | Y | N | Y | Y | na | na | na | na |
| S2 | Y | Y | Y | Y | Y | Y | N | Y | Y | N | N | Y | Y |
| S3 | Y | Y | Y | Y | Y | N | N | na | na | na | na | na | na |

GJ 3c Item Analysis (Ambiguious)

| | Master list number | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Informant | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 |
| NS1 | N | Y | Y | N | Y | Y | Y | Y | Y | Y | Y | Y | Y |
| NS2 | Y | Y | Y | Y | Y | Y | Y | Y | Y | Y | Y | Y | Y |
| NS3 | Y | Y | Y | Y | N | Y | Y | Y | Y | Y | Y | Y | Y |
| NS4 | N | Y | Y | Y | Y | Y | Y | Y | Y | Y | Y | Y | Y |
| NS5 | Y | Y | Y | N | Y | Y | Y | Y | Y | Y | Y | Y | Y |
| NS6 | Y | Y | Y | Y | Y | Y | Y | Y | Y | Y | Y | Y | Y |
| R1 | Y | Y | Y | Y | Y | Y | Y | Y | Y | Y | N | Y | na |
| R2 | Y | Y | Y | Y | Y | Y | Y | N | Y | Y | na | na | na |
| R3 | Y | Y | Y | Y | N | Y | Y | Y | Y | Y | Y | na | na |
| S1 | N | Y | Y | N | Y | Y | Y | Y | Y | na | na | na | na |
| S2 | N | Y | Y | N | Y | Y | Y | Y | Y | Y | N | Y | Y |
| S3 | N | Y | Y | Y | Y | Y | Y | na | na | na | na | na | na |

Appendix N: Test subjects' rates of overall maintenance of acceptability (OA) of local binding and rejection (OR) of long distance binding for the initial, secondary, and final data collection sub-periods. (See Appendix A) Sub-periods are color-coded according to the key below the table.

R1 overall acceptance of locally bound antecedants

| Test Number/ Test Subject | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | Initial | Secondary | Final |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| TV1a | 3 | 3 | 3 | 3 | 3 | 3 | 3 | 3 | 3 | 3 | 3 | 3 | | 100% | 100% | 100% |
| TV2a | 3 | 2 | 3 | 3 | 2 | 3 | 3 | 2 | 3 | 3 | 3 | 3 | | 92% | 83% | 100% |
| GJ1a | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | | 100% | 100% | 100% |
| GJ1c | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | | 100% | 100% | 100% |
| GJ2a | 1 | 1 | 1 | 1 | 1 | 0 | 1 | 1 | 1 | 1 | 1 | 1 | | 100% | 75% | 100% |
| GJ2c | 1 | 1 | 1 | 0 | 1 | 1 | 1 | 1 | 0 | 1 | 1 | 1 | | 75% | 100% | 75% |
| TOTAL | 10 | 9 | 10 | 9 | 9 | 9 | 10 | 9 | 9 | 10 | 10 | 10 | | 95% | 93% | 96% |

R1 overall rejection of long distance bound antecedants

| Test Number/ Test Subject | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | Initial | Secondary | Final |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| TV1b | 3 | 3 | 3 | 3 | 3 | 3 | 3 | 3 | 3 | 3 | 2 | 3 | | 100% | 100% | 92% |
| TV2b | 2 | 3 | 3 | 2 | 3 | 3 | 2 | 2 | 2 | 1 | 2 | 3 | | 83% | 83% | 67% |
| GJ1b | 1 | 1 | 1 | 1 | 1 | 1 | 0 | 1 | 1 | 0 | 1 | 1 | | 100% | 75% | 75% |
| GJ1d | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | | 100% | 100% | 100% |
| GJ2b | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 0 | 1 | | 100% | 100% | 75% |
| GJ2d | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 0 | | 100% | 100% | 75% |
| TOTAL | 9 | 10 | 10 | 9 | 10 | 10 | 8 | 9 | 9 | 7 | 7 | 9 | | 97% | 93% | 81% |

R2 overall acceptance of locally bound antecedants

| Test Number/ Test Subject | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | Initial | Secondary | Final |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| TV1a | 3 | 3 | 3 | 3 | 3 | 3 | 3 | 3 | 3 | 3 | | | | 100% | 100% | 100% |
| TV2a | 3 | 3 | 3 | 3 | 3 | 2 | 2 | 3 | 3 | 3 | | | | 100% | 83% | 100% |
| GJ1a | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | | | | 100% | 100% | 100% |
| GJ1c | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | | | | 100% | 100% | 100% |
| GJ2a | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | | | | 100% | 100% | 100% |
| GJ2c | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | | | | 100% | 100% | 100% |
| TOTAL | 10 | 10 | 10 | 10 | 10 | 9 | 9 | 10 | 10 | 10 | | | | 100% | 97% | 100% |

R2 overall rejection of long distance bound antecedants

| Test Number/ Test Subject | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | Initial | Secondary | Final |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| TV1b | 2 | 3 | 3 | 3 | 3 | 3 | 3 | 3 | 3 | 3 | | | | 92% | 100% | 100% |
| TV2b | 3 | 3 | 2 | 2 | 3 | 3 | 2 | 2 | 2 | 2 | | | | 83% | 83% | 67% |
| GJ1b | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | | | | 100% | 100% | 100% |
| GJ1d | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | | | | 100% | 100% | 100% |
| GJ2b | 1 | 1 | 1 | 1 | 1 | 1 | 0 | 1 | 0 | 1 | | | | 100% | 75% | 50% |
| GJ2d | 1 | 1 | 0 | 1 | 1 | 1 | 1 | 1 | 0 | 0 | | | | 75% | 100% | 0% |
| TOTAL | 9 | 10 | 8 | 9 | 10 | 10 | 8 | 9 | 7 | 8 | | | | 91% | 93% | 70% |

R3 overall acceptance of locally bound antecedants

| Test Number/ Test Subject | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | Initial | Secondary | Final |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| TV1a | 3 | 3 | 3 | 3 | 3 | 3 | 2 | 3 | 2 | 2 | 3 | | | 100% | 89% | 83% |
| TV2a | 3 | 3 | 3 | 2 | 2 | 1 | 3 | 1 | 2 | 2 | 2 | | | 92% | 67% | 58% |
| GJ1a | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | | | 100% | 100% | 100% |
| GJ1c | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 0 | 1 | 1 | 0 | | | 100% | 100% | 50% |
| GJ2a | 1 | 1 | 1 | 1 | 1 | 0 | 1 | 1 | 0 | 1 | 0 | | | 100% | 67% | 50% |
| GJ2c | 1 | 1 | 1 | 1 | 1 | 0 | 1 | 0 | 1 | 0 | 1 | | | 100% | 67% | 50% |
| TOTAL | 10 | 10 | 10 | 9 | 9 | 6 | 9 | 6 | 7 | 7 | 7 | | | 99% | 82% | 65% |

R3 overall rejection of long distance bound antecedants

| Test Number/ Test Subject | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | Initial | Secondary | Final |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| TV1b | 2 | 3 | 2 | 2 | 3 | 2 | 2 | 2 | 2 | 1 | 2 | | | 75% | 78% | 58% |
| TV2b | 3 | 3 | 3 | 2 | 1 | 2 | 1 | 3 | 1 | 2 | 1 | | | 92% | 44% | 58% |
| GJ1b | 1 | 1 | 1 | 0 | 1 | 1 | 1 | 1 | 1 | 0 | 0 | | | 75% | 100% | 75% |
| GJ1d | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | | | 100% | 100% | 100% |
| GJ2b | 1 | 1 | 1 | 1 | 1 | 1 | 0 | 1 | 0 | 1 | 1 | | | 100% | 67% | 75% |
| GJ2d | 0 | 1 | 0 | 1 | 0 | 0 | 1 | 0 | 1 | 0 | 1 | | | 50% | 33% | 50% |
| TOTAL | 8 | 10 | 8 | 7 | 7 | 7 | 6 | 8 | 6 | 5 | 6 | | | 82% | 74% | 69% |

Initial sub-period  Secondary sub-period  Final sub-period

Appendix N (continued): Test subjects' rates of overall maintenance of acceptability (OA) of local binding and rejection (OR) of long distance binding for the initial, secondary, and final data collection sub-periods. (See Appendix A) Sub-periods are color-coded according to the key below. the table

**S1 overall acceptance of locally bound antecedants**

| Test Number/ Test Subject | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | Initial | Secondary | Final |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| TV1a | 3 | 3 | 3 | 3 | 3 | 3 | 3 | 3 | 3 | | | | | 100% | 100% | n/a |
| TV2a | 3 | 3 | 2 | 3 | 2 | 2 | 3 | 2 | 2 | | | | | 83% | 78% | n/a |
| GJ1a | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | | | | | 100% | 100% | n/a |
| GJ1c | 1 | 1 | 1 | 1 | 0 | 1 | 1 | 1 | 1 | | | | | 83% | 100% | n/a |
| GJ2a | 1 | 1 | 1 | 1 | 1 | 0 | 1 | 1 | 1 | | | | | 83% | 100% | n/a |
| GJ2c | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | | | | | 100% | 100% | n/a |
| TOTAL | 10 | 10 | 9 | 10 | 8 | 8 | 10 | 9 | 9 | | | | | 92% | 96% | n/a |

**S1 Overall rejection of long distance bound antecedants**

| Test Number/ Test Subject | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | Initial | Secondary | Final |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| TV1b | 3 | 3 | 3 | 3 | 2 | 3 | 2 | 3 | 3 | | | | | 94% | 89% | n/a |
| TV2b | 3 | 3 | 2 | 3 | 2 | 3 | 2 | 3 | 2 | | | | | 83% | 78% | n/a |
| GJ1b | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | | | | | 100% | 100% | n/a |
| GJ1d | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | | | | | 100% | 100% | n/a |
| GJ2b | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | | | | | 100% | 100% | n/a |
| GJ2d | 1 | 1 | 1 | 1 | 1 | 0 | 1 | 0 | 1 | | | | | 83% | 67% | n/a |
| TOTAL | 10 | 10 | 9 | 10 | 8 | 9 | 8 | 9 | 9 | | | | | 93% | 89% | n/a |

**S2 overall acceptance of locally bound antecedants**

| Test Number/ Test Subject | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | Initial | Secondary | Final |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| TV1a | 3 | 3 | 3 | 3 | 3 | 2 | 3 | 2 | 2 | 3 | 2 | 3 | 2 | 94% | 78% | 83% |
| TV2a | 3 | 3 | 3 | 2 | 3 | 3 | 2 | 1 | 3 | 3 | 3 | 1 | 3 | 94% | 67% | 83% |
| GJ1a | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 0 | 1 | 1 | 1 | 1 | 100% | 67% | 100% |
| GJ1c | 1 | 1 | 1 | 1 | 0 | 0 | 1 | 0 | 1 | 0 | 1 | 0 | 0 | 67% | 67% | 25% |
| GJ2a | 1 | 1 | 1 | 1 | 1 | 0 | 1 | 1 | 1 | 1 | 0 | 1 | 1 | 83% | 100% | 75% |
| GJ2c | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 100% | 100% | 100% |
| TOTAL | 10 | 10 | 10 | 9 | 9 | 7 | 9 | 6 | 8 | 9 | 8 | 7 | 8 | 90% | 80% | 78% |

**S2 overall rejection of long distance bound antecedants**

| Test Number/ Test Subject | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | Initial | Secondary | Final |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| TV1b | 3 | 3 | 2 | 3 | 3 | 2 | 2 | 2 | 2 | 2 | 1 | 1 | 1 | 83% | 67% | 42% |
| TV2b | 3 | 3 | 2 | 3 | 2 | 2 | 2 | 2 | 1 | 1 | 1 | 1 | 2 | 83% | 56% | 42% |
| GJ1b | 1 | 1 | 1 | 1 | 0 | 1 | 0 | 1 | 1 | 0 | 0 | 0 | 0 | 83% | 67% | 0% |
| GJ1d | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 0 | 100% | 100% | 75% |
| GJ2b | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 0 | 0 | 1 | 0 | 0 | 0 | 100% | 33% | 25% |
| GJ2d | 0 | 0 | 0 | 1 | 0 | 0 | 1 | 0 | 1 | 0 | 1 | 0 | 1 | 17% | 67% | 50% |
| TOTAL | 9 | 9 | 7 | 10 | 7 | 7 | 7 | 6 | 6 | 5 | 4 | 3 | 4 | 78% | 65% | 39% |

**S3 overall acceptance of locally bound antecedants**

| Test Number/ Test Subject | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | Initial | Secondary | Final |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| TV1a | 3 | 3 | 3 | 2 | 1 | 2 | 2 | | | | | | | 100% | 50% | 67% |
| TV2a | 3 | 2 | 3 | 3 | 1 | 2 | 1 | | | | | | | 89% | 67% | 50% |
| GJ1a | 1 | 1 | 1 | 1 | 1 | 0 | 1 | | | | | | | 100% | 100% | 50% |
| GJ1c | 1 | 1 | 1 | 1 | 1 | 1 | 1 | | | | | | | 100% | 100% | 100% |
| GJ2a | 1 | 1 | 1 | 0 | 1 | 0 | 1 | | | | | | | 100% | 50% | 50% |
| GJ2c | 1 | 1 | 1 | 1 | 1 | 1 | 1 | | | | | | | 100% | 100% | 100% |
| TOTAL | 10 | 9 | 10 | 8 | 6 | 6 | 7 | | | | | | | 98% | 78% | 70% |

**S3 overall rejection of long distance bound antecedants**

| Test Number/ Test Subject | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | Initial | Secondary | Final |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| TV1b | 2 | 2 | 3 | 2 | 2 | 2 | 1 | | | | | | | 78% | 67% | 50% |
| TV2b | 2 | 2 | 1 | 2 | 3 | 2 | 2 | | | | | | | 56% | 83% | 67% |
| GJ1b | 1 | 0 | 1 | 0 | 0 | 1 | 0 | | | | | | | 67% | 0% | 50% |
| GJ1d | 1 | 1 | 1 | 1 | 1 | 1 | 1 | | | | | | | 100% | 100% | 100% |
| GJ2b | 1 | 1 | 1 | 1 | 0 | 0 | 0 | | | | | | | 100% | 50% | 0% |
| GJ2d | 1 | 0 | 0 | 1 | 0 | 1 | 1 | | | | | | | 33% | 50% | 100% |
| TOTAL | 8 | 6 | 7 | 7 | 6 | 7 | 5 | | | | | | | 72% | 58% | 61% |

Initial sub-period  Secondary sub-period  Final sub-period

## References


Aikawa, T. (1999) Reflexives. In Tsujima, N. (ed.), 154-190.

Al-Kasey, T. and Perez-Leroux, A. (1998) Second language acquisition of Spanish null subjects. In Flynn, S. et al. (eds.), 161-186.

Andersen, R. (1982) Determining the linguistic attributes of language attrition. In Lambert, R. and Freed, B. (eds.), 83-118.

Andersen, R. (ed.) (1984) *Second Language: A Crosslinguistic Perspective*. Rowley, MA: Newbury House.

Aoun, J. (1985) *A Grammar of Anaphora*. Cambridge, MA: MIT Press.

Aoun, J. and Sportiche, D. (1983) On the formal theory of government. *Linguistic Review* 2: 211-236.

Asher, J. and Garcia, R. (1969) The optimal age to learn a foreign language. *Modern Language Journal* 53: 334-341.

Atkinson, M. (1992) *Children's Syntax: An Introduction to Principles and Parameters Theory*. Oxford: Blackwell.

Bahrick, H. (1984a) Fifty years of second language attrition: Implications for programmatic research. *Modern Language Journal* 68: 105-118.

Bahrick, H.. (1984b) Semantic memory content in permastore: Fifty years of memory for Spanish learned in school. *Journal of Experimental Psychology* 113: 1-30.

Bailey, N., Madden, C. and Krashen, S. (1974) Is there a "natural sequence" in adult second language learning? *Language Learning* 21: 235-43.

Bardovi-Harlig, K. (1987) Markedness and salience in second language acquisition. *Language Learning* 37: 385-407.

Bennett, S. (1994) The acquisition of English reflexives by adolescent speakers of Serbo-Croatian. *Second Language Research* 10: 125-156.

Bennett, S. and Progovac, L. (1998) Morphological status of reflexives in second language acquisition. In Flynn, S. et al. (eds.), 187-214.

Berman, R. and Olshtain, E. (1983) Features of first language transfer in second language attrition. *Applied Linguistics* 4: 222-234.

Berwick, R. (1985) *The Acquisition of Syntactic Knowledge*. Cambridge, MA: MIT Press.

Bialystok, E. (1979) Explicit and implicit judgements of L2 grammaticality. *Language Learning* 29: 81-104.

Bialystok, E. and Sharwood Smith, M. (1985) Interlanguage is not a state of mind: An evaluation of the construct for second language acquisition. *Applied Linguistics* 6: 101-117.

Birdsong, D. (ed.) (1999) *Second language acquisition and the critical period hypothesis*. Mahwah, NJ: Lawrence Erlbaum.

Bley-Vroman, R. (1989) The logical problem of second language learning. In Gass, S. and Schachter, J. (eds.), 41-68.

Bley-Vroman, R. (1990) The logical problem of foreign language learning. *Linguistic Analysis* 20: 3-49

Bley-Vroman, R., Felix, S. and Ioup, G. (1988) The accessibility of Universal Grammar in adult language learning. *Second Language Research* 4: 1-32.

Bloomfield, L. (1933) *Language*. New York: Holt, Rinehart and Winston.

Blum-Kulka, S., House, J. and Kasper, G. (eds.) (1989) *Cross-Cultural Pragmatics*. Norwood, NJ: Ablex.

Braine, M. (1971) The acquisition of language in infant and child. In Reed, C. (ed.), 7-95.

Broselow, E. and Finer, D. (1991) Parameter setting in second language phonology and syntax. *Second Language Research* 7: 35-59.

Brown, H., Yorio, C. and Crymes, R. (eds.) (1977) *On TESOL '77*. Washington D.C.: Teaching of English to Speakers of Other Languages (TESOL).

Brown, R. (1973) *A First Language: The Early Stages*. Cambridge, MA: Harvard University Press.

Cancino, H., Rosansky, E. and Schumann, J. (1978) The acquisition of English negative and interrogatives by native Spanish speakers. In Hatch, E. (ed.), 270-287.

Chien, Y.-C. and Wexler, K. (1987) Children's acquisition of the locality condition for reflexives and pronouns. *Papers and Reports on Child Language Development* 26: 30-39.

Chien, Y.-C. and Wexler, K. (1990) Children's knowledge of locality conditions in binding as evidence for the modularity of syntax and pragmatics. *Language Acquisition* 1: 225-295.

Chomsky, N. (1957) *Syntactic Structures*. The Hague: Mouton.

Chomsky, N. (1959) Review of B.F. Skinner, *Verbal Behavior*. *Language* 35: 26-57.

Chomsky, N. (1965) *Aspects of the Theory of Syntax*. Cambridge, MA: MIT Press.

Chomsky, N. (1980a) On binding. *Linguistic Inquiry* 11: 1-46.

Chomsky, N. (1980b) *Rules and Representations*. New York: Columbia University Press.

Chomsky, N. (1981) *Lectures on Government and Binding*. Dordrecht: Foris.

Chomsky, N. (1982) *Some concepts and consequences of the theory of government and binding*. Cambridge, MA: MIT Press.

Chomsky, N. (1986a) *Barriers*. Cambridge, MA: MIT Press.

Chomsky, N. (1986b) *Knowledge of Language: Its Nature, Origin and Use*. Westport, CT: Praeger.

Christie, K. (1992) *Universal Grammar in the Second Language: An Experimental Study of the Cross-Linguistic Properties of Reflexives in English, Chinese and Spanish.* Unpublished Doctoral dissertation, University of Delaware, Newark.

Christie, K. and Lantolf, J. (1998) Bind me up bind me down: Reflexives in L2. In Flynn, S. et al. (eds.), 239-260.

Clahsen, H. (1982) *Spracherwerb in der Kindheit: Eine Untersuchung zur Entwicklung der Syntax bei Kleinkindern*. Tübingen: Gunter Narr.

Clahsen, H. (1984) The acquisition of German word order: A test case for cognitive approaches to L2 development. In Andersen, R. (ed.), 219-242.

Clahsen, H. (1988) Parameterised grammatical theory and language acquisition. In Flynn, S. and O'Neil, W. (eds.), 47-75.

Clahsen, H., Eisenbeiß, S. and Vainikka, A. (1994) The seeds of structure: A syntactic analysis of the acquisition of case marking. In Hoekstra, T. and Schwartz, B.D. (eds.), 85-118.

Clahsen, H., Meisel, J. and Pienemann, M. (1983) *Deutsch als Zweitsprache: der Spracherwerb auslandischer Arbeiter*. Tübingen: Gunter Narr.

Clahsen, H. and Muysken, P. (1986) The availability of Universal Grammar to adult and child learners - a study of the acquisition of German word order. *Second Language Research* 2: 93-119.

Clahsen, H. and Muysken, P. (1989) The UG paradox in L2 acquisition, *Second Language Research* 5: 1-29.

Clyne, M. (1967) *Transference and Triggering*. The Hague: Martinus Nijhoff.

Cohen, A. (1975) Error correction and the training of language teachers. *Modern Language Journal* 59: 414-22

Cohen, A. (1986) Forgetting foreign language vocabulary. In Weltens, B. et al. (eds.), 143-158.

Comrie, B. (1981) *Language Universals and Linguistic Typology*. Oxford: Blackwell.

Cook, V. (1985) Chomsky's Universal Grammar and second language learning. *Applied Linguistics* 6: 1-8.

Cook, V. (ed.) (1986) *Experimental Approaches to Second Language Acquisition*. Oxford: Pergamon.

Cook, V. (1990) Timed comprehension of binding in advanced learners of English. *Language Learning* 40: 557-599.

Cook, V. (1993) *Linguistics and Second Language Acquisition*. Basingstoke: Macmillan.

Cook, V. and Newson, M. (1996) *Chomsky's Universal Grammar*. Cambridge: Blackwell.

Coppieters, R. (1987) Competence differences between native and near-native speakers. *Language* 63: 544-573.

Corder, P. (1967) The significance of learner's errors. *IRAL* 5: 161-170.

Crain, S. and McKee, C. (1986) Acquisition of structural restrictions on anaphora. *Proceedings of the Sixteenth Annual Meeting of the North Eastern Linguistic Society* 16: 94-110.

Crain, S. and McKee, C. (1987) Structure dependence in grammar formation. *Language* 63: 522-543.

Crain, S. and Nakayama, M. (1987) Structure Dependence in Grammar Formation. *Language* 63: 522-543.

de Bot, K. and Clyne, M. (1989) Language reversion revisited. *Studies in Second Language Acquisition* 11: 167-177.

de Bot, K. and Clyne, M. (1994) A 16-year longitudinal study of language attrition in Dutch immigrants in Australia. *Journal of Multilingual and Multicultural Development* 15: 17-28.

de Bot, K., Gommans, P. and Rossing, C. (1991) L1 loss in an L2 environment: Dutch immigrants in France. In Seliger, H. and Vago, R. (eds.), 87-98.

de Bot, K. and Weltens, B. (1991) Regression, recapitulation, and language loss. In Seliger, H and Vago, R. (eds.), 31-52.

de Bot, K. and Weltens, B. (eds.) (1995) Foreign language attrition. *Annual Review of Applied Linguistics* 15: 151-164.

de Groot, A. and Kroll, J. (eds.) (1997) *Tutorials in Bilingualism: Psycholinguistic Perspectives*. Mahwah, NJ: Lawrence Erlbaum.

de Villiers, J. and Roeper, T. (1995) Relative clauses are barriers to Wh-movement for young children. *Journal of Child Language* 22: 389-404.

Diebold, A. (1964) Incipient bilingualism. In Hymes, D. (ed.), 495-511.

Dorian, N. (1973) Grammatical change in a dying dialect. *Language* 49: 411-438.

Dorian, N. (1978) The fate of morphological complexity in language death: Evidence from East Sutherland Gaelic. *Language* 54: 590-609.

Dorian, N. (1981) *Language Death: The Life Cycle of a Scottish Gaelic Dialect*. Philadelphia: University of Pennsylvania Press.

duPlessis, J., Solin, D., Travis, L. and White, L. (1987) UG or not UG, that is the question: A reply to Clahsen and Muysken. *Second Language Research* 3: 56-75.

Eckman, F. (1994) Local and long-distance anaphora in second language acquisition. In Tarone, E. et al. (eds.), 207-226.

Eckman, F. and Hastings, A. (eds.) (1979) *Studies in First and Second Language Acquisition*. Rowley, MA: Newbury House.

Eckman, F., Highland, D., Mileham, J. and Weber, R. (eds.) (1995) *Second Language Acquisition Theory and Pedagogy*. Mahwah, NJ: Lawrence Erlbaum.

Eckman, F., Moravscik, E. and Wirth, J. (eds.) (1986) *Markedness*. New York: Plenum Press.

Eubank, L. (ed.) (1991) *Point Counterpoint: Universal Grammar in the Second Language*. Amsterdam: John Benjamins.

Eubank, L. (1993/1994) On the transfer of parametric values in L2 development. *Language Acquisition* 3: 184-208.

Eubank, L. (1994) Optionality and the initial state in L2 development. In Hoekstra T. and Schwartz, B. D. (eds.), 369-388.

Eubank, L. (1996) Negation in early German-English interlanguage: More valueless features in the L2 initial state. *Studies in Second Language Acquisition* 12: 73-106.

Felix, S. (1985) More evidence on competing cognitive systems. *Second Language Research* 1: 47-72.

Felix, S. and Wode, H. (eds.) (1983) *Language Development at the Crossroads*. Tübingen: Gunter Narr.

Finer, D. (1991) Binding parameters in second language acquisition. In Eubank, L. (ed.), 351-374.

Finer, D. and Broselow, E. (1986) Second language acquisition of reflexive binding. *Proceedings of the Sixteenth Annual Meeting of the North Eastern Linguistic Society* 16: 154-168.

Flaugher, R. and Spencer, R. (1967) College foreign language placement and the intervening years problem. *Modern Language Journal* 51: 394-398.

Flege, J. (1999) Age of leaning and second language speech. In Birdsong, D. (ed.), 101-131.

Flynn, S. (1986) Comprehension vs. production: Differences in underlying competences. *Studies in Second Language Acquisition* 8: 17-46.

Flynn, S. (1987) *A Parameter-setting Model of L2 Acquisition*. Dordrecht: Reidel.

Flynn, S. (1988) Nature of development in L2 acquisition and implications for theories of language acquisition in general. In Flynn, S. and O'Neil, W. (eds.), 76-89.

Flynn, S. (1989) The role of the head-initial/head-final parameter in the acquisition of English relative clauses by adult Spanish and Japanese speakers. In Gass, S. and Schachter, J. (eds.), 89-108.

Flynn, S., Martohardjono, G. and O'Neil, W. (eds) (1998) *The Generative Study of Second Language Acquisition*. Mahwah, NJ: Lawrence Erlbaum.

Flynn, S. and O'Neil, W. (eds.) (1988) *Linguistic Theory in Second Language Acquisition*. Dordrecht: Kluwer.

Freed, B. (1982) Language loss: Current thoughts and future directions. In Lambert, R. and Freed, B. (eds.), 1-5.

Fries, C. (1945) *Teaching and Learning English as a Foreign Language*. Ann Arbor: University of Michigan Press.

Gal, S. (1979) *Language Shift: Social Determinants of Linguistic Change in Bilingual Austria*. New York: Academic Press.

Gardner, R. (1982) Social factors in language retention. In Lambert, R. and Freed, B. (eds.), 24-43.

Gardner, R., Lalonde, R. and MacPherson, J. (1985) Social factors in second language attrition. *Language Leaning* 35: 519-540.

Gardner, R., Lalonde, R., Moorcroft, R. and Evers, F. (1987) Second language attrition: The role of motivation and use. *Journal of Language and Social Psychology* 6: 29-47.

Gass, S. (1979) Language transfer and universal grammatical relations. *Language Learning* 29: 327-344.

Gass, S. (1983) The development of L1 intuitions. *TESOL Quarterly* 17: 273-91.

Gass, S., Cohen, A. and Tarone, E. (eds.) (1994) *Research methodology in second language acquisition*. Hillsdale, NJ: Lawrence Erlbaum.

Gass, S. and Madden, C. (eds.) (1985) *Input in second language acquisition*. Rowley, MA: Newbury House.

Gass, S. and Schachter, J. (eds.) (1989) *Linguistic Perspectives on Second Language Acquisition*. Cambridge: Cambridge University Press.

Gass, S. and Selinker, L. (eds.) (1983) *Language Transfer in Language Learning*. Rowley, MA: Newbury House.

Geoghegan, B. (1950) *The Retention of Certain Secondary School Subjects over the Period of Summer Vacation*. Unpublished Doctoral dissertation, Fordham University.

Gleitman, L. (1981) Maturational determinants of language growth. *Cognition* 10: 103-114.

Gleitman, L., Newport, E. and Gleitman, H. (1984) The current status of the motherese hypothesis. *Journal of Child Language* 11: 43-79.

Goodluck, H. and Birch, B. (1988) Late-learned rules in first and second language acquisition. In Pankhurst, J. et al. (eds.), 94-115.

Green, G. and Morgan, J. (1996) *Practical Guide to Syntactic Analysis*. Stanford: CSLI Publications.

Grodzinsky, J. and Reinhart, T. (1993) The innateness of binding and coreference. *Linguistic Inquiry* 24: 69-102

Haeckel, E. (1910) *The Evolution of Man: A Popular Scientific Study*. New York: Putnam.

Hale, K. (1983) Warlpiri and the grammar of non-configurational languages. *Natural Language and Linguistic Theory* 1:5-48.

Hansen, L. (1999a) Not a total loss: The attrition of Japanese negation over three decades. In Hansen, L. (ed.), 142-153.

Hansen, L. (ed.) (1999b) *Second Language Attrition in Japanese Contexts*. New York: Oxford Uiversity Press.

Hatch, E. (ed.) (1978) *Second Language Acquisition*. Rowley, MA: Newbury House.

Hermon, G. (1992) *Binding Theory and Parameter Setting*. Unpublished manuscript, University of Delaware.

Hill, J. and Hill, K. (1980) Mixed grammar, purist grammar, and language attitudes in Modern Nahuarl. *Language and Society* 9: 321-348.

Hilles, S. (1986) Interlanguage and the pro-drop parameter. *Second Language Research* 2: 33-52.

Hirakawa, M. (1990) L2 acquisition of English reflexives by speakers of Japanese. *Second Language Research* 6: 60-85.

Hirsh-Pasek, K., Treiman, R. and Schneiderman, M. (1984) Brown and Hanlon revisited: Mother's sensitivity to ungrammatical forms. *Journal of Child Language* 11: 81-88.

Hockett, C. (1958) *A Course in Modern Lingusitics*. New York: Macmillan.

Hoekstra, T. and Schwartz, B.D. (eds.) (1994) *Language Acquisition Studies in Generative Grammar: Papers in Honor of Kenneth Wexler from the 1991 GLOW Workshops*. Amsterdam: John Benjamins.

Hornstein, N. and Lightfoot, D. (eds.) (1981) *Explanation in Linguistics*. London: Longman.

Hornstein, N. and Lightfoot, D. (1981) Introduction. In Hornstein, N. and Lightfoot, D. (eds.), 9-31.

Hyams, N. (1986) *Language Acquisition and the Theory of Parameters*. Dordrecht: Reidel.

Hyams, N. (1994) V2, null arguments and COMP projections. In Hoekstra, T. and Schwartz, B. (eds.), 21-55.

Hyltenstam, K. and Obler, L. (eds.) (1989) *Bilingualism Across the Lifespan*. Cambridge: Cambridge University Press.

Hyltenstam, K. and Stroud, C. (1989) Bilingualism in Alzheimer's dementia. In Hyltenstam, K. and Obler, L. (eds.), 202-226.

Hymes, D. (ed.) (1964) *Language in Culture and Society*. New York: Harper and Row.

Ingram, D. (1989) *First Language Acquisition*. Cambridge: Cambridge University Press

Ioup, G. (1994) *The Logical Problem of Language Attrition: Theoretical Directions for L1 Attrition Research*. Unpublished manuscript, University of New Orleans.

Ioup, G., Boustagui, E., Tigi, M. and Moselle, M. (1994) Reexamining the critical period hypothesis: A case of successful adult SLA in a naturalistic environment. *Studies in Second Language Acquisition* 16: 73-98.

Ioup, G. and Kruse, A. (1977) Interference versus structural complexity in second language acquisition: language universals as a basis for natural sequencing. In Brown, H. et al. (eds.), 159-171.

Jaeggli, O. (1982) *Topics in Romance Syntax*. Dordrecht: Foris.

Jakobson, R. (1941) *Kindersprache, Aphasie und Allgemeine Lautgesetze*. Uppsala Universitets Aarsskrift. (Translated as (1968) *Child Language, Aphasia and Phonological Universals*. The Hague: Mouton.)

Johnson, J. (1988) *Critical Period Effects on Universal Properties of Language: The Status of Subjacency in the Acquisition of a Second Language*. Unpublished Doctoral dissertation, University of Illinois, Urbana-Champaign.

Johnson, L. and Newport, E. (1989) Critical period effects in second language learning: The influence of maturational state on the acquisition of English as a second language. *Cognitive Psychology* 21: 60-99.

Jordens, P. (1986) The cognitive function of case marking in German as a native and a foreign language. In Kellerman, E. and Sharwood Smith, M. (eds.), 91-109.

Jordens, P. (1988) The acquisition of word order in Dutch and German as L1 and L2. *Second Language Research*, 4: 41-65.

Jordens, P., de Bot, K. and Trapman, H. (1989) Linguistic aspects of regression in German case marking. *Studies in Second Language Acquisition*, 11: 179-204.

Jordens, P., de Bot, K., van Os, C. and Schumans, J. (1986) Regression in German case marking. In Weltens, B. et al. (eds.), 159-176.

Kanagy, R. (1994) Developmental sequence in learning Japanese: A look at negation. *Issues in Applied Linguistics* 5: 255-277.

Kanno, K. (1996) The status of a non-parameterized principle in the L2 initial state. *Language Acquisition* 5: 317-335.

Kanno, K. (1998b) Consistency and variation in second language acquisition. *Second Language Research* 14: 376-388.

Kanno, K. (1998a) The stability of UG principles in second-language acquisition: Evidence from Japanese. *Linguistics* 35: 1125-1146.

Kellerman, E. (1979) The problem with difficulty. *Interlanguage Studies Bulletin* 4: 27-48.

Kellerman, E. and Sharwood Smith, M. (eds.) (1986) *Cross-linguistic Influence in Second Language Acquisition*. Oxford: Pergamon.

Kellerman, E. and Yoshioka, K. (1999) Inter- and intrapopulation consistency: A comment on Kanno (1998). *Second Language Research* 15: 101-109.

Kennedy, L. (1932) The retention of certain Latin syntactical principles by first and second year Latin students after various time intervals. *Journal of Educational Psychology* 23, 132-146.

Koster, J. and Reuland, E. (1991) *Long-distance anaphor*. Cambridge: Cambridge University Press.

Krashen, S. (1981) *Second Language Acquisition and Second Language Learning*. Oxford: Pergamon.

Kuhberg, H. (1988) The acquisition of German temporality by two eleven-year-old children with Turkish and Polish as their first languages: A longitudinal study. *Delaware Papers in Linguistics* 1: 32-37.

Kuhberg, H. (1992) Longitudinal L2-attrition versus L2-acquisition in three Turkish children - Empirical findings. *Second Language Research* 8: 138-154.

Lado, R (1957) *Linguistics Across Cultures: Applied Linguistics for Language Teachers*. Ann Arbor: University of Michigan Press.

Lakshmanan, U. and Teranishi, K. (1994) Preferences versus grammaticality judgements: Some methodological issues concerning the governing category parameter in second language acquisition. In Gass, S. et al. (eds.), 185-206.

Lambert, R. and Freed, B. (eds.) (1982) *The loss of language skills*. Rowley, MA: Newbury House.

Lardiere, D. (1998) Case and tense in the 'fossilized' steady state. *Second Language Research* 14, 1-26.

Larmouth, W. (1974) Differential interference in American Finnish cases. *Language* 50: 356-366.

Lasnik, H. and Crain, S. (1985) On the acquisition of pronominal reference. *Lingua* 95: 135-154.

Lenneberg, E. (1967) *Biological Foundations of Language*. New York: Wiley & Sons.

Liceras, J. (1985) The role of intake in the determination of learners' competence. In Gass, S. and Madden, C. (eds.), 35-73.

Liceras, J. (1986) *Linguistic Theory and Second Language Acquisition*. Tubingen: Narr

Liceras, J. (1989) On some properties of the "pro-drop" parameter: Looking for missing subjects in non-native Spanish. In Gass, S. and Schachter, J. (eds.), 109-133.

Long, M. (1990) Maturational constraints on language development. *Studies in Second Language Acquisition* 12: 251-286

Lust, B. (1983) On the notion of 'Principle Branching direction': A parameter of universal grammar. In Otsu, Y. et al. (eds.), 137-151.

MacLaughlin, D. (1995) Language acquisition and the Subset Principle. *The Linguistic Review* 12: 143-191.

Maher, J. (1991) A crosslinguistic study of language contact and language attrition. In Seliger, H. and Vago, R. (eds.), 67-86.

Mandell, P. (1999) On the reliability of grammaticality judgement tests in second language acquisition research. *Second Language Research* 15: 73-99.

Manzini, R. and Wexler, K. (1987) Parameters, binding theory, and leanability. *Linguistic Inquiry* 18: 413-44.

Martohardjono, G. (1998) Measuring competence in L2 acquisition: Commentary on Part II. In Flynn, S. et al. (eds.), 151-157.

Mazurkewich, I. (1984) The acquisition of the dative alternation by second language leaners and linguistic theory. *Language Leaning* 34: 91-109.

McConvell, P. (1991) Understanding language shift: A step towards language maintenance. In Romaine, S. (ed.), 143-56.

McMahon, N. (1946) *The Effects of Summer Vacation on Retention of Latin Vocabulary*. Unpublished Master of Arts thesis, Fordham University.

Meisel, J. (1991) Principles of Universal Grammar and strategies and differences between first and second language acquisition. In Eubank, L. (ed.), 231-275.

Miller, W. (1971) The death of language or serendipity among the Shoshoni. *Anthropological Linguistics* 13: 114-120.

Moorcroft, R. and Gardner, R. (1987) Linguistic factors in second-language loss. *Language Leaning* 37: 327-340.

Moore, T. (ed.) (1973) *Cognitive development and the acquisition of language*. New York: Academic Press.

Neisser, U. (1984) Interpreting Harry Bahrick's discovery: What confers immunity against forgetting? *Journal of Experimental Psychology* 113: 32-35.

Newport, E., Gleitman, H. and Gleitman, L. (1977) Mother, I'd rather do it myself: Some effects and non-effects of maternal speech styles. In Snow, C. and Ferguson, C. (eds.), 109-149.

Obler, L. (1989) The boustrophedal brain: Laterality and dyslexia in bi-directional readers. In Hyltenstam, K. and Obler, L. (eds.), 155-172.

Obler, L. and Mahecha, N. (1991) First language loss in bilingual and polyglot aphasics. In Seliger, H. and Vago, R. (eds.), 53-66.

Olshtain, E. (1986) The attrition of English as a second language with speakers of Hebrew. In Weltens, B. et al. (eds.), 187-204.

Olshtain, E. (1989) Is second language attrition the reversal of second language acquisition? *Studies in Second Language Acquisition* 11: 151-165.

Otsu, Y. (1981) *Universal Grammar and Syntactic Development in Children*. Unpublished Doctoral dissertation, Massachusetts Institute of Technology.

Otsu, Y. and Naoi, K. (1986) Structure-dependence in L2 acquisition. *Paper presented at Japan Association of College English Teachers (JACET)*, Keio University, Tokyo, Sept. 1986.

Otsu, Y., van Riemsdijk, H., Inoue, K., Kamio, A. and Kawasaki, N. (1983) (eds.) *Studies in Generative Grammar and Language Acquisition*. Tokyo: International Christian University.

Oyama, S. (1976) A sensitive period in the acquisition of a non-native phonological system. *Journal of Psycholinguistic Research* 5: 261-85.

Pan, B. and Berko-Gleason, J. (1986) The study of language loss: Models and hypotheses for an emerging discipline. *Applied Psycholinguistics* 7: 193-206.

Pankhurst, J., Sharwood Smith, M. and Van Buren, P. (eds.) (1988) *Learnability and Second Languages: A Book of Readings*. Dordrecht: Foris.

Paradis, M. (1977) Bilingualism and aphasia. In Whitaker, H. and Whitaker, H. (eds.), 65-121.

Paradis, M. (1983) (eds.) *Readings on Aphasia in Bilinguals and Polyglots*. Montreal: Dedier.

Paradis, M. (1987) *The Assessment of Bilingual Aphasia*. Hillsdale, NJ: Lawrence Erlbaum.

Paradis, M. (1997) The cognitive neuropsychology of bilingualism. In de Groot, A. and Kroll, J. (eds.), 331-354.

Patkowski, M. (1980) The sensitive period for the acquisition of syntax in a second language. *Language Learning* 30: 449-472.

Phinney, M. (1987) The pro-drop parameter in second language acquisition. In Roeper, T. and Williams, E. (eds.), 221-238.

Pica, P. (1987) On the nature of the reflexivization cycle. *Proceedings of the Seventeenth North Eastern Linguistic Society* 17: 483-499.

Pinker, S. (1984) *Language Learnability and Language Development*. Cambridge, MA: Harvard University Press.

Pitres, A. (1895) Etude sur l'aphasie chez les polyglottes. *Revue de Médecine* 15: 873-899.

Postal, P. (1971/1967) *Cross-Over Phenomena*. New York: Holt, Rinehart and Winston

Pratella, W. (1969) *The Retention of First and Second Year Spanish over the Period of the Summer Vacation*. Unpublished Doctoral dissertation, Fordham University.

Preston, D. (1982) How to lose a language. *Interlanguage Studies Bulletin* 6: 64-87.

Preston, W. (1948) Review of de Goeje (1946). *International Journal of American Linguistic*s 14: 131-134.

Progovac, L. (1992) Relativised SUBJECT: Long-distance reflexives without movement. *Linguistic Inquiry* 23: 671-680.

Radford, A. (1988) *Transformational Grammar: A First Course*. Cambridge: Cambridge University Press.

Radford, A. (1990) *Syntactic Theory and the Acquisition of English Syntax: The Nature of Early Child Grammars of English*. Oxford: Basil Blackwell.

Read, C. and Chou Hare, V. (1979) Children's interpretations of reflexive pronouns in English. In Eckman, F. and Hastings, A. (eds.), 98-116.

Reed, C. (ed.) (1971) *The Learning of Language*. New York: Appleton-Century-Crofts.

Reetz-Kurashige, A. (1999) Japanese returnees' retention of English-speaking skills: Changes in verb usage over time. In Hansen, L. (ed.), 21-58.

Reinhart, T. (1976) *The Syntactic Domain of Anaphora*. Unpublished Doctoral dissertation, Massachusetts Institute of Technology.

Reuland, E. and Koster, J. (1991) Long-Distance Anaphora: An overview. In Koster, J. and Reuland, E., 1-25.

Ribot, T. (1882) *Diseases of Memory: An Essay in The Positive Psychology*. London: Paul.

Rizzi, L. (ed.) (1982a) *Issues in Italian syntax*. Dordrecht: Foris.

Rizzi, L. (1982b) Violations of the WH-Island Constraint and the Subjacency Condition. In Rizzi, L. (ed.), 49-76.

Rizzi, L. (1986) Null objects in Italian and the theory of *pro*. *Linguistic Inquiry* 17: 501-557.

Roeper, T. (1973) Connecting children's language and linguistic theory. In Moore, T (ed.), 187-196.

Roeper, T. and Williams, E. (eds.) (1987) *Parameter Setting*. Dordrecht: Reidel.

Romaine, S. (ed.) (1991) *Language in Australia*. Cambridge: Cambridge University Press.

Romaine, S. (1995) *Bilingualism*. Oxford: Blackwell.

Safir, K. (1987) Comments on Wexler and Manzini. In Roeper, T. and Williams, E. (eds.), 77-89.

Saleemi, A. (1992) *Universal Grammar and Language Learnability*. Cambridge: Cambridge University Press.

Schachter, J. (1974) An error in error analysis. *Language Learning* 24: 205-14.

Schachter, J. (1988) Second language acquisition and its relationship to Universal Grammar. *Applied Linguistics* 9: 219-235.

Schachter, J. (1989) Testing a proposed universal. In Gass, S. and Schachter, J. (eds.), 73-88.

Schachter, J. (1990) On the issue of completeness in second language acquisition. *Second Language Research* 6: 93-124.

Scherer, G. (1957) The forgetting rate in learning German. *German Quarterly* 30: 275-277.

Schmidt, A. (1985) *Young people 's Dyirbal*. Cambridge: Cambridge University Press.

Schwartz, B.D. (1998) On two hypothesis of "Transfer" in L2A: Minimal Trees and Absolute L1 Influence. In Flynn, S. et al. (eds.), 35-59.

Schwartz, B.D. and Sprouse, R. (1996) L2 cognitive states and the Full Transfer/Full Access model. *Second Language Research* 12: 40-72.

Seliger, H. and Vago, R. (eds.) (1991) *First language attrition*. Cambridge: Cambridge University Press.

Selinker L (1972) Interlanguage. *International Review of Applied Linguistics*. 10: 209-30.

Sharwood Smith, M. (1983a) Crosslinguistic aspects of second language acquisition. *Applied Linguistics* 4: 192-31.

Sharwood Smith, M. (1983b) On explaining loss. In Felix, S. and Wode, H. (eds.), 49-69.

Sharwood Smith, M. (1983c) On first language loss in the second language acquirer: Problems of transfer. In Gass, S. and Selinker, L. (eds.), 222-231.

Sharwood Smith, M. (1988) On the role of linguistic theory in explanations of second language developmental grammars. In Flynn, S. and O'Neil, W. (eds.), 173-198.

Sharwood Smith, M. (1989) Crosslinguistic influence in language loss. In Hyltenstam, K. and Obler, L. (eds.), 185-201.

Sharwood Smith, M. and van Buren, P. (1991) First language attrition and the parameter setting model. In Seliger, H. and Vago, R. (eds.), 17-30.

Silva-Corvalan, C. (1991) Spanish language attrition in a contact situation with English. In Seliger, H. and Vago, R. (eds.), 151-171.

Skinner, B. (1957) *Verbal Behavior*. New York: Appleton-Century-Crofts.

Snow, C. and Ferguson, C. (eds.) (1977) *Talking to Children*. Cambridge: Cambridge University Press.

Solan, L. (1983) *Pronominal Reference: Child language and the theory of grammar*. Dordrecht: Reidel.

Solan, L. (1987) Parameter setting and the development of pronouns and reflexives. In Roeper, T. and Williams, E. (eds.), 189-210.

Sorace, A. (1988) Linguistic intuitions in interlanguage development: The problem of indeterminacy. In Pankhurst, J. et al. (eds.), 167-190.

Sportiche, D. (1981) Bounding nodes in French. *Linguistic Review* 1: 219-246.

Tarone, E., Gass, S. and Cohen, A. (eds.) (1994) *Research Methodology in Second Language Acquisition*. Hillsdale, NJ: Lawrence Erlbaum.

Tavakolian, S (1981) *Lauguage Acquisition and Linguistic Theory*. Cambridge, MA: MIT Press.

Thomas, M. (1989) The interpretation of english reflexive pronouns by non-native speakers. *Studies in Second Language Acquisition* 11: 281-303.

Thomas, M. (1991) Universal Grammar and the interpretation of reflexives in a second language. *Language* 67: 211-239.

Thomas, M. (1993) *Knowledge of Reflexives in a Second Language*. Amsterdam: John Benjamins.

Thomas, M. (1995) Acquisition of the Japanese reflexive *zibun* and movement of anaphors in Logical Form. *Second Language Research* 11: 206-234.

Thomas, M. (1998) Binding and related issues in L2 acquisition: Commentary on Part III. In Flynn, S. et al. (eds.), 261-276.

Tsujima, N. (ed.) (1999) *The Handbook of Japanese Linguistics.* Malden, MA: Blackwell.

Vago, R. (1991) Paradigmatic regularity in first language attrition. In Seliger, H. and Vago, R. (eds.), 241-251.

Vainikka, A. and Young-Scholten, M. (1994) Direct access to X'-theory: Evidence from Korean and Turkish adults learning German. In Hoekstra, T. and Schwartz, B.D. (eds.), 265-316.

Vainikka, A. and Young-Scholten, M. (1996) Gradual development of L2 phrase structure. *Second Language Research*, 12: 7-39.

Vainikka, A. and Young-Scholten, M. (1998) The initial state in the L2 acquisition of phrase structure. In Flynn, S. et al. (eds.), 17-34.

Van Buren, P. and Sharwood-Smith, M. (1985) The acquisition of preposition stranding by second language learners and parametric variation. *Second Language Research* 1: 18-46.

van Vlerken, M. (1980) *Adverbial Placement in English: A study of First Language Loss.* Unpublished Masters dissertation, Utrecht University.

Waas, M. (1993) *Language Attrition among German Speakers in Australia: A Sociolinguistic Inquiry.* Unpublished Doctoral dissertation, Macquarie University, Sydney.

Wakabayashi, S. (1996) The nature of interlanguage: SLA of reflexives. *Second Language Research* 12: 266-303.

*Webster's Ninth New Collegiate Dictionary* (1985) Springfield, MA: Merriam Webster.

Weltens, B. (1989) *The Attrition of French as a Foreign Language.* Dordrecht: Foris.

Weltens, B., de Bot, K. and van Els, T. (eds.) (1986) *Language Attrition in Progress.* Dordrecht: Foris.

Wexler, K. and Chien, Y.-C. (1985) The development of lexical anaphors and pronouns. *Papers and Reports on Child Language Development* 24: 138-149.

Wexler, K. and Manzini, R. (1987) Parameters and learnability. In Roeper, T. and Williams, E. (eds.), 41-76.

Whitaker, H. and Whitaker, H. (eds.) (1977) *Studies in Neurolinguistics, Volume 3.* New York: Academic Press.

White, L. (1985) The pro-drop parameter in adult second language acquisition. *Language Learning* 35: 47-62.

White, L. (1986a) Implications of parametric variation for adult second language acquisition: An investigation of the pro-drop parameter. In Cook, V. (ed.), 55-72.

White, L. (1986b) Markedness and parameter-setting: Some implications for a theory of second language acquisition. In Eckman, F., et al. (eds.), 309-327.

White, L. (1989a) The adjacency condition on case assignment: Do L2 learners observe the subset principle? In Gass, S. and Schachter, J. (eds.), 134-158.

White, L. (1989b) *Universal Grammar and Second Language Acquisition*. Amsterdam: John Benjamins.

White, L. (1990) Another look at the logical problem of foreign language learning: A reply to Bley-Vroman. *Linguistic Analysis* 20: 50-63.

White, L. (1995) Input, triggers, and second language acquisition: Can binding be taught? In Eckman, F. et al. (eds.), 63-78.

White, L., Bruhn-Garavito, J., Kawasaki, T., Pater, J. and Prevost, P. (1997) The researcher gave the subject a test about himself: Problems of ambiguity and preference in the investigation of reflexive binding. *Language Learning* 47: 145-172.

Yang, D.-W. (1983) The extended binding theory of anaphors. *Language Research* 19: 169-192.

Yoshitomi, A. (1992) Towards a model of language attrition: Neurobiological and psychological contributions. *Issues in Applied Linguistics* 3: 293-318.

Yoshitomi, A. (1994) *The Attrition of English as a Second Language of Japanese Returnee Children*. Unpublished Doctoral dissertation, University of California, Los Angeles.

Yoshitomi, A. (1999) On the loss of English as a second language by Japanese returnee children. In Hansen, L. (ed.), 80-111.

Yusa, N. (1998) A Minimalist approach to second language acquisition. In Flynn, S. et al. (eds.), 215-238.

Zobl, H. (1988) Configurationality and the subset principle: The acquisition of V1 by Japanese learners of English. In Pankhurst, J. et al. (eds.), 116-131.