SSLA, **24**, 27–54. Printed in the United States of America.

JAPANESE ADULT LEARNERS' DEVELOPMENT OF THE LOCALITY CONDITION ON ENGLISH REFLEXIVES

Yasuhiro Akiyama
Toho Junior-Senior High School, Japan

This study explores the developmental pattern observed when Japanese adult learners acquire the locality condition on English reflexives. Experimental tasks were designed specifically to deal with the methodological problems of earlier research and then administered to Japanese learners of English at five proficiency levels ($n = 411$) as well as English and Japanese control groups ($n = 40$). Results from the learner groups indicate that the locality condition is acquired significantly better with sentences containing embedded *that*-clauses (type E-1) than with sentences containing embedded infinitival clauses (type E-2). This asymmetry exists even at beginning stages of learning and persists through later stages. For type E-2 clauses, there is an appreciable percentage of advanced learners (about 35% in this study) who failed to acquire the locality condition, which, I argue, is extremely difficult to account for within the UG models proposed thus far.

It is widely known that English and Japanese reflexives differ in that they take different local domains for antecedents. For example, consider the sentences in (1) and (2).

(1) a. *John*$_i$ *thought that Tom*$_j$ *was blaming himself*$_{(*i/j)}$.
b. *John*$_i$ *wanted Tom*$_j$ *to know himself*$_{(*i/j)}$ *better.*

This is a revised version of the paper that I submitted to the University of Hawai'i for my Master's degree. I appreciate a lot of insightful comments and constant encouragement from Prof. Bley-Vroman, many other professors of the university, and the anonymous *SSLA* reviewers. Any flaws or inadequacies in the paper, however, should be interpreted as my own.

Address correspondence to: Yasuhiro Akiyama, Toho Gakuen, 3-1-10 Naka, Kunitachi, Tokyo 186-0004, Japan; e-mail: jmyasu@a1.mbn.or.jp.

(2) a. *John*$_i$ *-ga* *Tom*$_j$*-ga* *zibun*$_{(i/j)}$*-o semeteiru-to omotta (koto).*
John-NOM Tom-NOM self-ACC was blaming thought (fact)
b. *John*$_i$*-ga* *Tom*$_j$*-ni* *motto* *zibun*$_{(i/j)}$*-o shitte moraitakatta (koto).*
John-NOM Tom-DAT better self-ACC knew wanted (fact)

As shown in (1a) and (1b), *himself* only refers to the local antecedent *Tom* (local binding) but not to the long-distance (LD) antecedent *John* (LD binding). In the corresponding Japanese sentences, (2a) and (2b), in which the most commonly used reflexive form *zibun* is employed,[1] the reflexive can refer to either the local or the LD antecedent. The local domain of English reflexives is therefore smaller and more restrictive than that of Japanese reflexives.

An interesting question arises in the case of Japanese learners of English acquiring this locality condition. For one thing, their knowledge of Japanese reflexives does not help, because of the difference in locality between Japanese and English reflexives, and for another, it is very unlikely that they are explicitly told or taught about such a restriction.[2] If so, how do Japanese learners acquire the locality condition on English reflexives without receiving any instruction or any information from their L1? Or, more fundamentally, is it possible for them to acquire such a restriction at all?

Although the acquisition of this locality condition has been extensively studied (e.g., Cook, 1990; Eckman, 1994; Finer, 1991; Finer & Broselow, 1986; Lakshmanan & Teranishi, 1994; Shimura, 1990; Thomas, 1989, 1991, 1993; Wakabayashi, 1996), there have been only a few studies (e.g., Hirakawa, 1990; Matsumura, 1994) that have investigated it from a developmental perspective. Furthermore, the conclusions from those few studies are contradictory.

Hirakawa (1990), for example, did a multiple-choice antecedent identification task with Japanese learners from Grades 10 through 13 (first year of college) and found no significant difference in performance across the four levels. In contrast, Matsumura (1994) found statistically significant development between two proficiency levels. Matsumura conducted a similar task using the same levels of Japanese students as Hirakawa but regrouped his participants into two levels according to their scores on a cloze test that was independently administered prior to the task. He found that the group with higher scores performed significantly better on both embedded *that*-clauses and embedded infinitival clauses (cf. [1a] and [1b], respectively) than the group with lower scores.

It should be pointed out, however, that neither study is completely adequate as a developmental study. In Hirakawa (1990), the number of participants for each level is small (13, 14, 18, and 20 participants for Grades 10, 11, 12, and 13, respectively) and therefore one can question how faithfully this small number represents each level. Likewise, the two levels Matsumura (1994) set up would be insufficient to reveal any so-called pattern of development, though they might show a difference between the two. Additionally, these two studies suffer from serious methodological problems that also exist in other studies on the acquisition of the locality condition, as will be pointed out later. Owing to these methodological problems, it has been quite difficult

for the field to draw any definitive conclusions about the development of the locality condition on English reflexives. Therefore, in this study, I would first like to trace the progress of the previous studies carefully and attempt to establish the best possible methodology for investigating the development of the locality condition. Then, conducting a study that is based on that methodology, I would like to provide an answer to the following old but as yet unanswered question: What is the developmental pattern observed when Japanese adult learners acquire the locality condition on English reflexives?[3]

METHODOLOGICAL PROBLEMS IN PREVIOUS RESEARCH

The methodological problems in previous L2 studies of the acquisition of the locality condition on English reflexives can be summarized by the following three points: (a) a failure to elicit ambiguity from participants, (b) an inappropriate choice of verbs that take infinitival clauses, and (c) limitations of the analyses of aggregate data. I address each of these points in turn.

Elicitation of Ambiguity

One of the key points in designing an experiment for the study at issue here is that, although English sentences like (1a) are unambiguous and should be interpreted as such, L2 learners may misinterpret them ambiguously in the same way that Japanese native speakers interpret the corresponding Japanese sentence (2a). Thus, an experimental device must be created carefully enough to elicit this potential ambiguity from participants. Most studies, however, have been unsuccessful in this respect. The task employed by Cook (1990), for example, would be the most inappropriate in eliciting ambiguity from participants because it did not allow them to have any ambiguous interpretations. In his experiment, the participants, presented with a stimulus sentence like *Peter said that John voted for himself*, were asked to say as quickly as possible which out of two possible persons, namely *Peter* or *John*, was referred to by the reflexive *himself*. Thus, in his design, the participants were not allowed to have the ambiguous interpretation—that is, to interpret *himself* as "either Peter or John." It should also be noted that this task does not guarantee that an English reflexive refers exclusively to its local antecedent, because even those participants who chose *John* (who were considered to possess correct English grammar in Cook's analysis) did not necessarily know that the reflexive cannot also refer to *Peter*. They might have chosen *John* just because they preferred it to *Peter* even though their grammar allowed both interpretations. Thus, the so-called error rate he used in his analysis—the rate of incorrect antecedents chosen per sentence type—would be underestimated as compared to the true rate of incorrect judgments that should have been sought in the study.

The same criticism applies to almost all other previous studies, although to different degrees. The point is that, in tasks employed in these studies,

https://doi.org/10.1017/S027226310200102X Published online by Cambridge University Press

such as picture identification (Finer, 1991; Finer & Broselow, 1986; Eckman, 1994) or multiple-choice antecedent identification (Hirakawa, 1990; Matsumura, 1994; Thomas, 1989, 1991), there is always a danger of underestimating the ambiguous interpretation. This is evidenced in the percentages of ambiguous interpretation when the same task was conducted with native control groups whose L1 allows both local and LD binding like Japanese. Finer reports that only 25% of his Japanese control group opted for the ambiguous interpretation when he conducted the same task with Japanese translations. Likewise, Thomas (1991), Hirakawa, Eckman, and Lakshmanan and Teranishi (1994) report 10%, 9–10%, 25%, and 58%, respectively, for the percentage of ambiguous interpretations made by the Japanese control group. Such percentages prove that those experimental devices were not appropriate for eliciting ambiguity from their participants. (Finer & Broselow, Cook [1990], Matsumura, and Wakabayashi [1996] do not even provide this percentage.) That said, several attempts were made to improve the rate of the ambiguous interpretation. For example, Thomas (1991) had a special training session for the ambiguous interpretation prior to the main study; Matsumura seems to have done similar training prior to his main study, though it was not as explicit as Thomas's; Eckman cautioned his participants to always think of the possibility of two interpretations; Wakabayashi asked his participants to put numbers in preference order rather than forcing them to choose one or two antecedents. However, it seems that none of them were quite successful, as just discussed.

Of all the tasks employed, the task in Lakshmanan and Teranishi (1994) was the most successful in eliciting ambiguity, as evidenced by the relatively high percentage of the performance of their Japanese control group (58%). They asked their participants about the impossibility of local binding and the impossibility of LD binding separately for each sentence. Although their task was quite complicated, as would be imagined from asking about impossibility rather than possibility, and required their participants to use a high degree of metalinguistic knowledge, their method of treating local and LD binding separately seems to enjoy an advantage over other methods and, more importantly, would be theoretically motivated as a proper way of assessing L2 learners' knowledge about the locality condition for the following two reasons. First, this method avoids the previously mentioned problems with eliciting ambiguity from participants because they are asked about only one interpretation at a time. Second, note that, to acquire the correct binding domain of English reflexives, L2 learners have to know two properties: One is that local binding is possible, and the other is that LD binding is impossible.[4] These two properties must be assessed separately because they would be acquired in different learning environments. There should be no problem when L2 learners discover that English reflexives allow local binding because they are constantly exposed to English sentences that show that this is the case: English reflexives only refer to local antecedents. They would be able to learn this property with ease simply by noticing it through these data, namely, by means of positive evidence. However, matters would not be so simple when Japanese

learners acquire the second property—that is, that LD binding is impossible. What should be noted here is that, as contrasted with the acquisition of the first property, there seems to be no evidence available to them indicating that this restriction exists in English. As discussed with respect to (1), the L1 (Japanese) does not provide any relevant data because it simply has no such restriction.[5] Data from the L2 (English) may provide ample evidence showing that local binding is possible, but this by no means suggests that LD binding is impossible in English. Furthermore, as already pointed out, it is quite rare that learners are explicitly told or taught about such a restriction. Hence, the acquisition of the second property is a true case of learning without (negative) evidence and, therefore, should have been treated separately from the case of learning the first property in the previous studies.

Choice of Verbs

It is conceptually problematic to use (object) control verbs, such as *tell*, *ask*, and *order*, as verbs that take infinitival clauses as their complements because the structure in which these verbs appear involves control as well as binding. Consider the sentence in (3a) and its linguistic analysis in (3b), where the subject of the embedded clause (PRO) is controlled by *Jane*, and the reflexive *herself* is bound by this PRO, thus being interpreted as *Jane*, as indicated by the indices.[6]

(3) a. *Mary told Jane to look at herself.*
b. *Mary told Jane*$_i$ [PRO$_i$ *to look at herself*$_i$].

It has been tacitly assumed that, if a participant chooses *Jane* as an antecedent of the reflexive *herself* in sentences like (3a), he has found the antecedent within a clause (local binding); that is, he is considered to have identified PRO, which is a clause mate of *herself* as its antecedent. However, it would also be conceivable that the participant has directly interpreted *herself* as *Jane*, which is the object of the main clause, beyond the clause boundary, in which case he should be judged to have opted for LD binding. Put differently, it is impossible to investigate the locality condition for binding using control verbs because there is always the risk of confounding control and binding. Therefore, it would be fair to claim that the studies in which (object) control verbs are used in an experiment aiming to investigate the local domain of reflexives (Cook, 1990; Eckman, 1994; Finer, 1991; Hirakawa, 1990; Matsumura, 1994; Wakabayashi, 1996) face a serious conceptual problem.[7]

Methods of Analysis

There have been two methods employed to analyze data in the previous L2 binding studies. The first method is to aggregate data collected from a group and analyze them without any consideration of the consistency or the varia-

tion that exists within each individual. The second method is to analyze data by the consistency each individual exhibits for a particular rule.

When participants form a homogeneous group, aggregate data collected from such a group exhibit a certain pattern of performance that would be regarded as a characteristic of that group. Most of the previous studies, however, simply analyzed aggregate data without observing this homogeneity requirement. For example, participants in Finer's (1991) study consisted of three L1 groups that varied in English proficiency. Participants in Thomas's (1989) study were composed of 97 native speakers of 20 different languages whose English proficiency also varied. As mentioned earlier, Hirakawa (1990), whose study was conducted with participants in Grades 10–13, concluded that there was no difference in performance between the grades, which was rather surprising and questionable even to Hirakawa herself (p. 73). Matsumura (1994), as pointed out before, conducted a similar developmental study with participants also in Grades 10–13, but he reclassified them into two proficiency groups. He found that there was a significant difference between the two. Therefore, if Matsumura's conclusion was right, Hirakawa's surprising result might be attributed simply to inappropriate classification of her participants: Each group might have lacked homogeneity in her study.

The limitations of the analysis of aggregate data have been pointed out in the literature (Eckman, 1994; Wakabayashi, 1996). The point is that such data do not indicate how systematic each individual's grammatical knowledge is. In other words, they would not reveal to what extent each individual's performance is rule governed (Thomas, 1991) or what kind of systematic knowledge, or "interlanguage grammar," each individual has in his or her process of acquiring a second language. Finer (1991), for example, concluded that the participants in his study set the Governing Category Parameter (Wexler & Manzini, 1987) at the value *C*, which is neither the value of their L1 (Korean) nor the L2 (English), based solely on the tendency to perform more like native speakers of English on sentences with embedded *that*-clauses than on sentences with embedded infinitival clauses. However, to demonstrate that his participants actually had the parameter set at the value *C*, it would be necessary to show that the individual participant(s), rather than the aggregate data from the group, consistently performed well on sentences with embedded *that*-clauses and at the same time consistently performed poorly on sentences with embedded infinitival clauses. The aggregate data Finer presented did not show the existence of such participants and should therefore be regarded as inappropriate for drawing such a conclusion.

One strong advantage of an analysis based on consistency is that it is possible to discover a systematicity that is observed neither in an L1 nor an L2. Unlike the aggregate data analysis, where data are analyzed mostly with respect to the norm set by a target language, an analysis based on consistency is completely free from the grammar of an L1 or an L2, which makes it possible to explore the nature and the forms of interlanguage grammar more accurately (see Bley-Vroman, 1983; Gass, 1983).

Table 1. Number of participants in each group

Group	*n*	Participants excluded	Participants used
Level 1 (Grade 8)	83	23	60
Level 2 (Grade 9)	91	36	55
Level 3 (Grade 10)	95	38	57
Level 4 (Grade 11)	91	27	64
Level 5 (advanced)	51	2	49
Subtotal	411	126 (30.7%)	285 (69.3%)
Control (English)	20	0	20
Control (Japanese)	20	0	20
Grand total	451	126	325

METHOD

In an attempt to solve the problems raised in the previous section, I incorporated the following three improvements into the design of the present study:

1. A story-based truth-value judgment task was employed.
2. The verb that takes an infinitival clause as its complement was restricted to the verb *want*.
3. The data obtained were processed using two methods: an analysis of aggregate data and an analysis of consistency for individual participants.

Participants

Organization of Participants. The overall organization of the participants in this study is summarized in Table 1. Five levels were set up depending on the length of time spent learning English. The first four levels were composed of students at a private junior and senior high school in Japan (a six-year school for students in Grades 7–12). Students in Grades 8, 9, 10, and 11 were assigned to levels 1, 2, 3, and 4, respectively. Most students started learning English seriously after entering this school. Because the school year in Japan starts in April and the experiment was done at the end of January, students in Grade 8, for example, had studied English for approximately 2 years at the start of the experiment. Furthermore, the transition between grades (especially between junior and senior high school) was assumed to be continuous because the school adopts a six-year system that allows all students at the junior high school to continue studying at the same senior high school.

Level 5 (advanced) comprised 51 native speakers of Japanese who were studying full-time at the University of Hawai‘i at Manoa. The majority of the participants were graduate students, and their level of English proficiency was estimated to be advanced.[8] Most of them had started learning English at the age of 13 (Grade 7) in Japan, as did the participants at levels 1–4. The ages of the participants in level 5 ranged from 20 to 38 years, the mean age being 28

https://doi.org/10.1017/S027226310200102X Published online by Cambridge University Press

years and 2 months. The mean length of their residence in English-speaking countries was 3 years and 10 months.[9]

In addition to the experimental group, 20 English and 20 Japanese native speakers participated as the two control groups. The participants in the English control group were all American students studying at the University of Hawai'i, and the participants in the Japanese control group lived in Japan and did not use English in their daily lives. The experiment was conducted both in Japan and in the United States. The participants in levels 1–4 as well as the Japanese control group did the experimental task in Japan; the participants in level 5 and the English control group completed the task in the United States.

Screening. For a cross-sectional developmental study, it is crucial to maintain homogeneity among participants at each level, as discussed earlier. For this purpose, the following two screening procedures were employed with levels 1–4. First, participants who had spent more than three consecutive months in any English-speaking country were excluded from the total pool of participants.[10] Second, a series of syntax tests, details of which will be discussed in the next section, were administered to ensure that the participants had the knowledge of basic structures of English that were presumed to be possessed by all the participants of this study. Those who failed to answer correctly any one item on the tests were excluded. The participants in level 5 did not take these syntax tests for reasons that will be explained later. Instead, in a debriefing session after the task, they were asked whether they had learned or been explicitly taught about the locality condition on English reflexives.[11] Two participants reported that they had learned the relevant rule in a linguistics class at college and were consequently removed from the data analyses. As a result of these procedures, about 31% of the participants were excluded, as shown in Table 1. Only the data from the remaining participants were processed in the analyses of this study.

Materials

Story-Based Truth-Value Judgment Task. To avoid the ambiguity problem pointed out earlier, I adopted a story-based truth-value judgment task (see Crain & McKee, 1986; Thomas, 1995), in which participants were provided with a context in the form of a story. The context was created to lead them to only one interpretation—either the local or the LD interpretation—for a potentially ambiguous sentence. Thus, the ambiguity (or lack thereof) of one stimulus sentence was determined by two judgments made in two different contexts. For instance, for the stimulus sentence in (4), there were two different stories provided, one for LD binding and one for local binding, as in (5) and (6), respectively.

(4) X thought that Y hated himself/herself.
(5) Story for LD binding:

https://doi.org/10.1017/S027226310200102X Published online by Cambridge University Press

John and Taro were best friends. John fell in love with Taro's girlfriend. Taro's girlfriend became interested in John too, and she broke up with Taro. Since then Taro has never even tried to see John.
Stimulus sentence:
John thought that Taro hated himself. Yes/No

(6) Story for local binding:
Yesterday Susan saw Tomoko sitting alone on a bench. Tomoko looked very depressed. Susan asked Tomoko, "What's the matter?" Tomoko answered quietly, "I am not doing well in my classes. I am not good at sports. I seem to fail at everything I try. I am not attractive, and I am not popular among boys!"
Stimulus sentence:
Susan thought that Tomoko hated herself. Yes/No

Given a story as in (5) or (6), participants were asked to judge whether the content of the stimulus sentence presented could be naturally inferred from the story provided. Thus, it was assumed that if participants allowed the ambiguous interpretation for a given stimulus sentence, they would answer "Yes" to the sentences for both stories.

The effect of this story-based truth-value judgment task for a binding study has been attested to in White, Bruhn de Garavito, Kawasaki, Pater, and Prévost (1997). One of the advantages of this method is that it does not require participants to consider two possible meanings for an ambiguous sentence at the same time, which is often said to be extremely difficult for those who have not received special linguistic training. Furthermore, unlike a multiple-choice antecedent identification task that explicitly asks about grammaticality in a particular interpretation, the task itself does not necessitate participants' metalinguistic knowledge. They are asked, in principle, only about the truth value of each given sentence.

Stimulus Sentences. Two types of stimulus sentences were used in this experiment: sentences with embedded *that*-clauses (type E-1) and sentences with embedded infinitival clauses (type E-2). Three tokens were prepared for each type, as specified in (7). Two distracter sentences were added to the total of six stimulus sentences in order to distract participants' attention from the target structure, as well as to make sure that participants relied on the hierarchical rather than linear structure of sentences when they interpreted reflexives. The result showed that almost all participants pointed out correct antecedents for these two distracters.

(7) Stimulus sentences (including two distracters)
   a. Type E-1
      X thought that Y hated himself/herself.
      X thought that Y was blaming himself/herself.
      X said that Y wrote about himself/herself in the letter.
   b. Type E-2
      X wanted Y to talk about himself/herself in class.
      X wanted Y to think about himself/herself (again).
      X wanted Y to know himself/herself better.

https://doi.org/10.1017/S027226310200102X Published online by Cambridge University Press

c. Distracters
A fan of Takanohana hit himself.
The girl beside Susan saw herself in the mirror and laughed.

X and Y in the sentences in (7) stand for the names of persons who appeared in the respective stories. As discussed earlier, each of the six stimulus sentences appeared twice, following two different stories. To avoid unwanted interferences that might be caused by using the same sentence twice, different names were used in each of the two experimental sentences made from one stimulus sentence. (For a complete set of materials, see the Appendix.)

These six stimulus sentences were carefully chosen. The six English sentences and the corresponding Japanese translations were piloted on both English and Japanese native speakers, respectively, and refined several times until expected interpretations were consistently obtained from both groups. Though neglected in most previous studies, it should be noted that it is not always easy to obtain the ambiguous interpretation that potentially exists in biclausal Japanese sentences corresponding to the English sentences of types E-1 or E-2. It depends not only on the property of *zibun* but also on other factors, such as the choice of particular verbs (Lakshmanan & Teranishi, 1994) and the choice of case particles like *wa*, *ga*, *ni*, and so forth. Thus, to investigate the issue raised in this study, one must first guarantee that, whereas these six sentences are unambiguous in English, they should be interpreted as ambiguous by Japanese people when translated into Japanese.

From a total of 14 stories (12 for the stimulus sentences and 2 for the distracters), two forms of the test materials were prepared in which the orders of the 14 stories were counterbalanced. Half of the participants took one form of the task and the other half took the other form.

Syntax Tests. To check whether the participants had the knowledge of basic structures of English necessary for the task, three types of syntax tests were created. They were designed to detect whether the participants knew (a) that *himself* is an anaphor and distinct from a pronoun (Anaphor Test), (b) the structure of a *that*-clause as a sentential complement (*That*-Clause Test), and (c) the structure of an infinitival clause as a sentential complement (Infinitival Clause Test). (For a complete set of the syntax tests, see the Appendix.)

Transfer Test. The three stimulus sentences of type E-1 with appropriate names inserted were prepared for the transfer test. In this test, the participants translated these sentences into Japanese. The test was formulated with the intent of assessing which Japanese reflexive word(s) each participant used and how this affected his or her performance in the main study. (See the Appendix.)

Procedures

The participants were given the materials in booklets. They were first given a story that was written in Japanese on one page in the booklet and then, flip-

ping the page, were presented with an English sentence for the story, thereby ensuring that they were first given the context and then the sentence (for the importance of keeping this order, see White et al., 1997).[12] The participants were allowed to read the story as many times as they liked after reading the sentence. They were also allowed to ask about the meaning of a word they did not understand if there was one. Prior to the task, the experimenter carefully explained the procedure of the task to the participants using a sample item.[13] The participants in levels 1–4 performed the task in class during a regular class period and the participants in level 5 did so either individually or two at a time. All the participants were carefully observed and supervised by the experimenter while completing the task. The task was not timed, so each participant worked at his or her own pace.

Following the truth-value judgment task, the syntax tests and then transfer test were given to the participants in levels 1–4. These two types of tests, however, were not administered to the participants in level 5 because their proficiency level was presumed to be advanced enough to pass the tests or not to be affected by translation effects. After the task (and the syntax and transfer tests for the participants in levels 1–4), all the participants were asked to fill out a questionnaire that inquired about any previous experience learning English. It took approximately 30–40 minutes for each participant in levels 1–4 and about 10–15 minutes for each participant in level 5 to complete the whole experiment.

RESULTS

For the reasons discussed earlier, I will present the results of this study using two methods of analysis: the analysis of aggregate data and the analysis of data based on the consistency each participant exhibited. An alpha level of .05 was used for all statistical tests.

Analysis of Aggregate Data

Table 2 shows the mean scores and standard deviations of each sentence type depending on binding locality for each level. One point was given when a participant answered "Yes" to a sentence placed in the local context or "No" to a sentence placed in the LD context. Because there were three tokens for each category, the maximum score each participant could obtain in any one category was 3. As shown in Table 2, the English control group performed almost perfectly on all four categories. (The data for type E-2 at level 1 have been excluded because it was discovered, at the start of the experiment, that the participants at this level had not yet learned the V + NP + *to*-VP construction.) A three-way ANOVA (2 [sentence] Types × 2 Locality [types] × 4 [proficiency] Levels) was performed on the data from levels 2–5. Results are summarized in Table 3, where one finds a significant main effect for all three independent

https://doi.org/10.1017/S027226310200102X Published online by Cambridge University Press

Table 2. Mean scores and standard deviations of five proficiency levels

	Type E-1		Type E-2	
Group	Local "Yes"	LD "No"	Local "Yes"	LD "No"
Level 1				
M	2.65	2.07	—	—
SD	0.68	1.01	—	—
Level 2				
M	2.84	2.25	2.69	1.23
SD	0.50	0.97	0.57	0.98
Level 3				
M	2.79	2.50	2.60	1.49
SD	0.49	0.83	0.84	1.14
Level 4				
M	2.72	2.39	2.48	1.23
SD	0.57	0.83	0.77	1.08
Level 5				
M	2.78	2.73	2.80	1.57
SD	0.46	0.53	0.50	1.14
Control (E)				
M	2.85	3.00	2.90	2.95
SD	0.37	0	0.31	0.22

Note. Dashes indicate excluded data.

Table 3. ANOVA for levels 2–5

Source	*df*	*SS*	*MS*	*F*
Between groups				
Level	3	8.67	2.89	3.12*
Error (level)	221	205.72	0.93	
Within groups				
Type	1	83.91	83.91	143.69**
Type × Level	3	0.57	0.19	0.80
Error (type)	221	129.06	0.58	
Locality	1	136.78	136.78	213.40**
Locality × Level	3	4.61	1.54	2.40
Error (locality)	221	141.65	0.64	
Type × Locality	1	50.38	50.38	117.73**
Type × Locality × Level	3	0.98	0.33	0.77
Error (type × locality)	221	94.58	0.43	
Total	899	856.91	278.60	

*$p < .05$. **$p < .01$.

https://doi.org/10.1017/S027226310200102X Published online by Cambridge University Press

Table 4. Performance of Japanese control group in comparison to Hirakawa (1990)

Clause type	Hirakawa (1990) (n = 110)	Present study (n = 60)
Type E-1		
LD	62.73% (69)	25.00% (15)
Local	26.36% (29)	3.33% (2)
Either	9.10% (11)	68.33% (41)
Neither	0 (0)	3.33% (2)
Type E-2		
LD	70.91% (78)	8.33% (5)
Local	19.09% (21)	0 (0)
Either	10.00% (11)	91.67% (55)
Neither	0 (0)	0 (0)

variables: Level, $F(3, 221) = 3.12$, $p < .05$; Type, $F(1, 221) = 143.69$, $p < .01$; Locality, $F(1, 221) = 213.40$, $p < .01$. There was also a significant interaction between Type and Locality, $F(1, 221) = 117.73$, $p < .01$.

Because there was no significant interaction between Type and Level, the significant effect of Type implies that performance on type E-1 is better than performance on type E-2 consistently across all four levels. This suggests that there existed a similar difference in performance between type E-1 and type E-2 at every level.[14] The claim made earlier that local and LD binding would involve different acquisition processes turned out to be statistically supported due to the existence of the significant main effect of Locality.

Because a significant effect for Level was obtained, post hoc Scheffé tests were performed. The tests indicated that a significant difference obtained only in performance on Type E-1 (LD "No") between Level 2 and Level 5, $p = .03$. This suggests that performance improved with an increase in proficiency level in the case of Type E-1 (LD "No") but not in the other three cases. Thus, it would be fair to say that the participants exhibited different developmental patterns concerning LD "No" between Type E-1 and Type E-2. Gradual improvement in Type E-1 (LD "No") is more clearly observed when the data are seen from Level 1 through Level 5 in Table 2, whereas no such development is observed in Type E-2 (LD "No").

Table 4 summarizes the performance of the Japanese control group in comparison with Hirakawa (1990). Responses that were categorized as "Either" in this study—for example, are those in which a participant answered "Yes" to both the story for local binding and the story for LD binding. As is clear from Table 4, the rate of the ambiguous interpretation dramatically improves in this study when compared to Hirakawa's results (from 9.10% to 68.33% in type E-1 and from 10.00% to 91.67% in type E-2),[15] which suggests that the experimental device used for this study was superior to the methodology used by Hirakawa.

https://doi.org/10.1017/S027226310200102X Published online by Cambridge University Press

	X (English)	Y (Japanese)	Z	W
Local binding	Yes	Yes	No	No
Long-distance binding	No	Yes	Yes	No

Figure 1. Four types of grammar.

Analysis Based on Consistency

One of the strengths of the experimental design of this study is that, for each stimulus sentence, it is possible to determine what type of grammar a participant is considered to possess. There are four conceivable types of grammar, depending on the judgments made on the two stories for each stimulus sentence, as shown in Figure 1. Type X stands for the grammar in which only local binding is allowed, as in English (*himself*). Likewise, type Y stands for the grammar in which both local and LD binding are allowed, as in Japanese (*zibun*). Hyams and Sigurjónsdóttir (1990), Reinhart and Reuland (1991, 1993), and Sigurjónsdóttir and Hyams (1992) report that Norwegian *seg* and Icelandic *sig* take type Z, in which only LD binding is permitted. Type W, though logically possible, should not occur because it would indicate that an anaphor does not have an antecedent within a sentence, which contradicts its definition.

For the analysis of the data based on the consistency each participant exhibited, I set the criterion for consistency in the following way: Those who showed the same type of grammar for 2 or 3 out of 3 stimulus sentences were regarded as possessing that type of grammar consistently (the "2 out of 3 criterion"). By this criterion, the participants were classified for each sentence type as either having a particular type of grammar or being inconsistent in their judgments. The validity of criteria for consistency has been a controversial issue in the literature. Hamilton (1996) argued that 6 out of 8 is "the minimum standard for experimental studies in the field of psychology" (p. 440); Wakabayashi (1996) criticized Thomas (1991, 1993) and argued for 100% correct performance for consistency. Therefore, in addition to the results analyzed by the 2 out of 3 criterion described above, I also analyzed the data by the 3 out of 3 criterion. The two sets of results are summarized in Table 5, where *I* stands for "inconsistent judgments." Table 5 indicates that the results by the 3 out of 3 criterion basically showed the same tendencies as those observed by the 2 out of 3 criterion, though inconsistency was more frequent in the former. Furthermore, the performance of the two native control groups seems to suggest that the 3 out of 3 criterion was too strict for this study because a significant percentage of the native speakers, who were presumed to exhibit consistent judgments, turned out to be inconsistent in their judgments when classified on that criterion. This would imply that a criterion for consistency should not be determined absolutely but rather empirically de-

Table 5. Type of grammar for each sentence type judged by consistency

	2 out of 3 Criterion					3 out of 3 Criterion				
Type	X(E)[a]	Y(J)[b]	Z	W	I[c]	X(E)	Y(J)	Z	W	I
Level 1 (n = 60)										
E-1	38	12	1	4	5	21	3	1	0	35
P	63.33	20.00	1.67	6.67	8.33	35.00	5.00	1.67	0	58.33
Level 2 (n = 55)										
E-1	44	9	1	0	1	26	4	0	0	25
P	80.00	16.36	1.81	0	1.81	47.27	7.27	0	0	45.45
E-2	17	28	1	0	9	7	9	1	0	38
P	30.90	50.90	1.81	0	16.36	12.72	16.36	1.81	0	69.09
Level 3 (n = 57)										
E-1	46	6	0	1	4	35	2	0	0	20
P	80.70	10.53	0	1.75	7.02	61.40	3.51	0	0	35.09
E-2	26	22	3	1	5	11	11	2	1	32
P	45.61	38.60	5.26	1.75	8.77	19.30	19.30	3.51	1.75	56.14
Level 4 (n = 64)										
E-1	51	7	0	3	3	28	3	0	0	33
P	79.69	10.94	0	4.69	4.69	43.75	4.69	0	0	51.56
E-2	23	28	5	0	8	9	12	2	0	41
P	35.94	43.75	7.81	0	12.50	14.06	18.75	3.13	0	64.06
Level 5 (n = 49)										
E-1	46	2	0	0	1	30	0	0	0	19
P	93.88	4.08	0	0	2.04	61.22	0	0	0	38.78
E-2	26	20	1	0	2	13	17	0	0	29
P	53.06	40.82	2.04	0	4.08	26.53	14.29	0	0	59.18
Control (English) (n = 20)										
E-1	19	0	0	0	1	17	0	0	0	3
P	95.00	0	0	0	5.00	85.00	0	0	0	15.00
E-2	20	0	0	0	0	17	0	0	0	3
P	100	0	0	0	0	85.00	0	0	0	15.00
Control (Japanese) (n = 20)										
E-1	0	16	3	0	1	0	7	1	0	12
P	0	80.00	15.00	0	5.00	0	35.00	5.00	0	60.00
E-2	0	20	0	0	0	0	15	0	0	5
P	0	100	0	0	0	0	75.00	0	0	25.00

[a]E = English. [b]J = Japanese. [c]I = Inconsistent judgments

pending on the nature of the task employed for each study. For these reasons, I will primarily discuss the results obtained by the 2 out of 3 criterion.

Overall, as shown in Table 5, the rate of inconsistent judgments was quite low across all levels. For example, even at level 1, there were only 5 participants (out of 60) who showed inconsistency in their judgments. This would suggest that L2 learners' performance is not arbitrary but rule governed, as has been claimed in terms of interlanguage since the birth of SLA studies

https://doi.org/10.1017/S027226310200102X Published online by Cambridge University Press

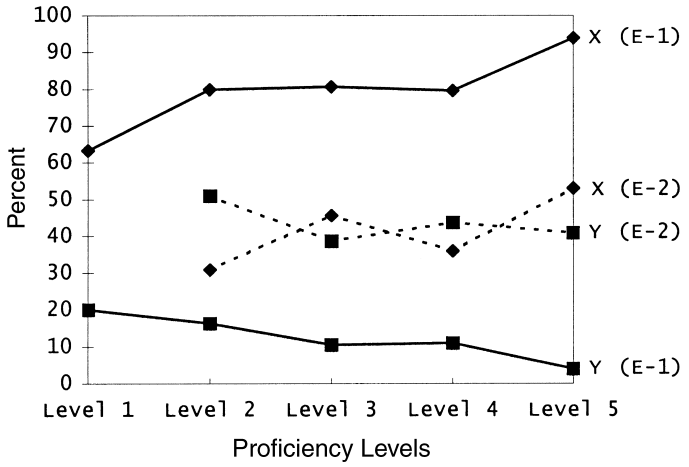

Figure 2. Development of each type of grammar using the 2 out of 3 criterion.

(Corder, 1967; Selinker, 1972). Figure 2 shows how X and Y types of grammar develop from level 1 through level 5 with respect to the two sentence types. It appears that the two types of grammar for type E-1 develop in certain determined directions: X is progressing toward 100% whereas Y is regressing toward zero. The development for type E-2, on the other hand, is more complex, with the two types of grammar appearing to develop in a reciprocal fashion, with no seeming progression or regression. Thus, Figure 2 clearly shows that the two sentence types exhibit different developmental patterns. The next analysis, focusing on two types of grammar, X (English type) and Y (Japanese type), aimed to see what combination of the two grammar types was consistently possessed for the two sentence types within one participant. Table 6 shows the results of this analysis. For example, X-X in the first column stands for the case in which a participant consistently accepted only local binding for both type E-1 and type E-2. Likewise, placement in X-Y indicates that the participant consistently accepted only local binding for type E-1 and both local and LD binding for type E-2. As evidenced in the results specified under the X-Y combination, there was an almost invariable ratio (around 30%) of participants at every level who allowed both local and LD binding for type E-2 but only local binding for type E-1. This would imply that some Japanese learners cannot acquire the locality condition on English reflexives at all by positive evidence alone. However, it should be noted that this does not mean that all Japanese learners cannot acquire the locality condition. As shown under the X-X combination in Table 6, about half of the participants in level 5 had seemingly acquired the locality condition for English, both for type E-1 and type E-2. This fact, however, was obscured in the aggregate data (cf. Table

https://doi.org/10.1017/S027226310200102X Published online by Cambridge University Press

Table 6. Interaction of grammar types for two sentence types

	2 out of 3 Criterion				3 out of 3 Criterion			
Group	X-X	Y-X	X-Y	Y-Y	X-X	Y-X	X-Y	Y-Y
Level 2	16	1	20	8	4	0	1	3
P	29.09	1.81	36.36	14.55	7.27	0	1.81	5.45
Level 3	24	0	15	6	13	0	5	2
P	42.11	0	26.31	10.53	22.80	0	8.77	3.51
Level 4	20	0	21	5	8	0	7	1
P	31.25	0	32.81	7.81	12.50	0	10.93	1.56
Level 5	26	0	17	2	12	0	3	0
P	53.06	0	34.69	4.08	24.49	0	6.12	0
Control (E)	19	0	0	0	15	0	0	0
P	95.00	0	0	0	75.00	0	0	0
Control (J)	0	0	0	16	0	0	0	7
P	0	0	0	80.00	0	0	0	35.00

2), which suggests again the importance of analyses based on consistency for a study of this kind.

DISCUSSION

The Tensed-Infinitive Asymmetry

It has been reported that L2 learners, when identifying correct antecedents for English reflexives, perform better on sentences with embedded *that*-clauses (type E-1 in this study) than on sentences with embedded infinitival clauses (type E-2). Consider (8), for example (from Finer & Broselow, 1986).

(8) a. *Mr. Fat*$_i$ *thinks that Mr. Thin*$_j$ *will paint himself*$_{(*i/j)}$.
b. *Mr. Fat*$_i$ *wants Mr. Thin*$_j$ *to paint himself*$_{(*i/j)}$.

This means that there are learners who interpret *himself* correctly in (8a) but incorrectly in (8b), taking the antecedent to be *Mr. Fat*. Yuan (1994) called this phenomenon the tensed-infinitive asymmetry. Some studies (e.g., Cook, 1990; Finer, 1991; Finer & Broselow, 1986; Hirakawa, 1990; Wakabayashi, 1996) reported that the asymmetry exists, whereas other studies (e.g., Eckman, 1994) reported conflicting results indicating that there is no such asymmetry.[16] Results from this study clearly show that the tensed-infinitive asymmetry exists throughout all proficiency levels in the case of Japanese learners. The results also show that the locality condition on English reflexives develops differently depending on clause type. Furthermore, it was demonstrated that the acquisition of the locality condition on English reflexives involves two separate issues: (a) that local binding is possible, and (b) that LD binding is impossible. It is the unsuccessful acquisition of the latter that causes the asymmetry.

https://doi.org/10.1017/S027226310200102X Published online by Cambridge University Press

Limitations of UG Proposed Thus Far

Most studies on the acquisition of the locality condition on English reflexives have centered around the question of whether Universal Grammar (UG) and learning principles, which are considered to be operative as the language acquisition device (LAD) in L1 acquisition (Chomsky, 1965, 1981, 1986), are accessible to L2 learners. What light can the results of this study shed on this question? Broadly speaking, two models of UG have been proposed regarding the locality condition on reflexives (anaphors). One is based on parameters (Manzini & Wexler, 1987; Wexler & Manzini, 1987), and the other is based on the morphological complexity of reflexives (Cole, Hermon, & Sung, 1990; Cole & Sung, 1994; Progovac, 1992, 1993; Reinhart & Reuland, 1991, 1993).

The parameterized binding theory claims that world languages allow five local domains for anaphors. Principles in the theory contain a parameter for choosing among these options. Thus, for each anaphor that learners encounter in the process of acquisition, they need to set the parameter at the value appropriate for that anaphor. Furthermore, the five local domains corresponding to five values of the parameter are nested within one another such that any two of the five form a subset-superset relationship. Because L1 learning is considered to occur without negative evidence (Braine, 1971; Brown & Hanlon, 1970), setting the parameter for an anaphor is guided by a learning principle called the Subset Principle (Berwick, 1985) which leads learners to choose the smallest possible domain compatible with the input data.

The results of this study clearly suggest that the Subset Principle is not operative in L2 learning, which replicates other studies (Cook, 1990; Hirakawa, 1990; Lakshmanan & Teranishi, 1994). As pointed out at the beginning of this paper, the local domain of English reflexives is narrower than the local domain of Japanese *zibun*. If L2 learning were guided by the Subset Principle, it would be expected that L2 learners take the narrowest possible domain compatible with the input from English, which is the domain for English reflexives, not the domain for Japanese *zibun*. However, as shown in Table 5, a considerable number of participants consistently chose Y (the local domain for Japanese *zibun*) for English reflexives.

As Wexler and Manzini (1987) argued, UG and learning principles are independent (although both are included in the LAD), and thus UG can still hold true even when learning principles collapse. This means that learners may make use of the binding principles even though they do not follow the Subset Principle in the process of acquisition. Results of this study, as seen in Table 5, show that all types of learner grammars seem to be within the sanction of UG because few participants had a grammar of type Z (the "rogue grammar").[17] Furthermore, the X-Y type specified in Figure 2 is also compatible with the theory, and, in fact, Finer and Broselow (1986) insisted that this constituted evidence for UG access in L2 learning (because learner grammars are still within the sanction of UG, although they may be neither the grammar of the L1 nor the grammar of the L2).[18] However, the parameterized binding theory

has been criticized (e.g., Cole & Sung, 1994; MacLaughlin, 1995) for being inappropriate as a theory of language and effectively collapsed as a result, being replaced with the nonparameterized binding theory.[19] Thus, Finer and Broselow's argument does not hold any longer.

The nonparameterized binding theory, in which morphological complexity is assumed to be a determinant of whether a reflexive allows LD binding, basically maintains that polymorphemic anaphors such as *himself* allow only local binding whereas monomorphemic anaphors such as *zibun* allow LD binding (as well as local binding). Thus, by claiming that L2 learners misanalyze English reflexives as monomorphemic, the theory explains why L2 learners allow LD binding for English reflexives.

Progovac's (1992, 1993) approach seems to provide the most straightforward account for the tensed-infinitive asymmetry observed in this study. She argued that a morphologically simplex anaphor (X^o reflexive) is only bound to an X^o category—that is, Agr(eement). She further argued that Agr is anaphorically linked to the Agr of a higher clause when it lacks an overt referential agreement, as in infinitival clauses in English. This means that a morphologically simplex anaphor allows LD binding when it occurs in an embedded infinitival clause but does not do so when it occurs in an embedded *that*-clause. In this way, Progovac's account explains why some learners, identified under X-Y in Table 6, allow LD binding for sentences with embedded infinitival clauses but reject it for sentences with embedded *that*-clauses.

However, it should be pointed out that Progovac's (1992, 1993) approach crucially depends on the assumption that those who exhibit the tensed-infinitive asymmetry take English reflexives to be monomorphemic (i.e., X^o reflexives). This study showed that 35% of the participants in level 5 still exhibited the asymmetry in question. This would mean that those 35%, who are assumed to have reached an advanced level, still misinterpret *himself* as monomorphemic even after they have learned English and stayed in English-speaking countries for a considerable length of time (3 years and 10 months, on average), which is very hard to imagine. Therefore, the nonparameterized binding theory also does not give a coherent account for the facts obtained in this study.

In summary, from the discussion above, it should be clear that the UG models proposed thus far fail to account for the L2 learning manifested by the findings of this study. This might suggest that L2 learners do not have (at least full-fledged) access to UG or that UG (or the LAD) that has been proposed thus far in linguistics is inappropriate as a model that explains language acquisition. The issue, however, is still left open.

X-*Zisin*

Earlier in this paper I noted that, when Japanese learners acquire the locality condition on English reflexives, data from the Japanese language are not available to them. However, some researchers working on the relationship between morphological complexity of a reflexive and its locality (e.g., Yuan,

Table 7. Choice of Japanese reflexives for all participants

Group	*zibun*	X-*zisin*[a]	N-*zisin*[b]	Others	Total
Level 1	32	105	29	5	171
Level 2	31	88	26	17	162
Level 3	28	95	28	14	165
Level 4	58	58	36	27	179

[a]X-*zisin* = pronoun (anaphor)-*zisin* (e.g., *kare [kanojo]-zisin*, *zibun-zisin*). [b]N-*zisin* = person's name-*zisin* (e.g., *Taro-zisin*, *Tomoko-zisin*).

1994), argue that the locality restriction on phrasal (polymorphemic) reflexives such as *kare-zisin* and *zibun-zisin* can be generalized to reflexives in English. Because these phrasal reflexives are considered to behave like English reflexives in locality, it is theoretically possible for Japanese learners to acquire the condition at issue. This amounts to claiming that the locality condition for English is already available to Japanese learners of English when they start to learn that language. However, there seem to be at least five problems to be solved before this claim can be confirmed.

First, it is true that, in addition to *zibun*, there are phrasal reflexives in Japanese. It should also be pointed out that English and Japanese still differ in terms of an inventory of reflexives: English has only one kind of reflexive whereas Japanese has several kinds (viz., *zibun*, *zisin*, *zibun-zisin*, and *kare-zisin*). Furthermore, these reflexives all differ in locality, subject orientation, and logophoricity (Li, 1994). Thus, even if one may claim that Japanese learners can make use of their knowledge of Japanese reflexives, how can they determine from several different kinds of Japanese reflexives, each of which has a different behavior, the one that exactly fits the properties of English reflexives?

Second, it is doubtful whether Japanese actually has a reflexive that corresponds exactly to the properties of English reflexives. A brief inquiry I conducted with Japanese native speakers showed, for example, that the locality of *zibun-zisin* is weaker than that of English reflexives, presumably because of the interference from *zibun*: Some Japanese speakers accept LD antecedents for *zibun-zisin*.

Third, it is questionable how strongly X-*zisin* affects Japanese learners. As Yusa (1998) pointed out, *kare-zisin* is a literal translation of English *himself* and is not used very often in normal conversation. He argued that this word is (unnaturally) introduced to Japanese learners when they learn English *himself* in classes at junior high school. He then went on to say that Japanese learners may overuse *kare-zisin* at the junior high school level but cease to use it later, realizing that it is an unnatural, translation-favored Japanese word. Interviews I conducted with teachers of the participants in levels 1–4 revealed that students are actually taught *kare-zisin* as the equivalent of *himself* when the word is first introduced. Results from the translation test, as summarized in Table 7, also show that Yusa's claim is partially supported.

Table 8. Choice of Japanese reflexives for participants who had X grammar consistently

Group	*zibun*	X-*zisin*[a]	N-*zisin*[b]	Others	Total
Level 1	15	35 (58.33%)	10	0	60
Level 2	8	47 (60.26%)	14	9	78
Level 3	12	60 (58.82%)	23	7	102
Level 4	19	26 (33.33%)	20	13	78

[a]X-zisin = pronoun (anaphor)-*zisin* (e.g., *kare [kanojo]-zisin*, *zibun-zisin*). [b]N-zisin = person's name-*zisin* (e.g., *Taro-zisin*, *Tomoko-zisin*).

The participants in level 4 (second year of senior high school) tended to avoid using X-*zisin*: The percentage of those who used *X-zisin* at level 4 was 32.4%, whereas the percentages for levels 1, 2, and 3 were 61.4%, 54.3%, and 57.6%, respectively. If that is a general tendency, how can such an uncommon word affect the Japanese when learning English?

Fourth, the data from Japanese learners do not support a direct influence from X-*zisin*. If X-*zisin* does indeed directly influence the interpretation of English reflexives, it is expected that Japanese learners who know the locality condition of English perfectly well will dominantly rely on the word when asked about the locality of English reflexives. Table 8 shows the kinds of Japanese words used for *himself* by the participants who were judged to have adopted X grammar (English type) by the stricter 3 out of 3 criterion. The results, however, do not indicate the dominant use of X-*zisin*: about 60% of responses were X-*zisin* at levels 1–3 and only 30% at level 4.

Fifth and finally, if knowledge of X-*zisin* affects Japanese learners, why do some use that knowledge for embedded *that*-clauses but not for embedded infinitival clauses? As shown in Table 6, this study manifested the existence of a significant number of such learners (i.e., those who were classified as possessing X-Y type). The account based on the influence from X-*zisin* would not provide any convincing answer to this question.

CONCLUSION

What is the developmental pattern observed when Japanese adult learners acquire the locality condition on English reflexives? The present study revealed that the tensed-infinitive asymmetry exists even at beginning stages of learning and persists through later stages. Furthermore, it was demonstrated that the locality condition exhibits different developmental patterns depending on sentence type. In acquiring the restriction on LD binding, the participants of this study showed statistically significant development across proficiency levels for sentences with embedded *that*-clauses (type E-1), whereas they did not show such development for sentences with embedded infinitival clauses (type E-2). Analyses based on the consistency each individual exhibited also showed

https://doi.org/10.1017/S027226310200102X Published online by Cambridge University Press

that the locality condition with these two sentence types clearly developed in a different way: The development of type E-1 appeared to proceed unidirectionally, whereas the development of type E-2 was more complex and proceeded as if there were no ultimate endpoint. It was further found that about 40% of the participants at an advanced level failed to acquire the locality condition for type E-2.

Although sentences of type E-2 were assumed to be sentences with embedded infinitival clauses in this study, it would be premature to conclude that all verbs that take embedded infinitival clauses exhibit the tensed-infinitive asymmetry. This study, restricting the verb for type E-2 stimulus sentences to the verb *want*, focused on the question of whether a clause boundary formed by infinitives functions as a binding domain of English reflexives, and results showed that it does not, at least for some Japanese learners. Thus, the next question to be asked would be how those Japanese learners who were judged to lack the locality condition would perform on sentences where the verb *want* is replaced by other (kinds of) verbs—for example, (object) control verbs such as *tell*, *persuade*, and *order*. As argued in the previous section, an answer to this question will provide an important insight into the process of L2 learners' acquisition of the clause structures of English and the question of how the asymmetry arises in their grammar.[20]

Reflecting on the methodologies employed in the previous studies, this study attempted to explore the best possible way of assessing L2 learners' knowledge concerning reflexives. However, it did not provide sufficient explanation of why such an asymmetry arises, although it briefly pointed out that the present forms of UG models are not adequate for explaining this knowledge. More elaborate theoretical consideration is needed to answer this question. It should also be recalled that this study was conducted with learners who were situated in rather restricted contexts: Japanese learners who started to learn English seriously after puberty (or a sensitive period) in an EFL context. It would be interesting to ask how far the results obtained here can be generalized to learners in other contexts. For example, does the asymmetry exist in Spanish learners whose L1 has the same locality condition on reflexives as English? Does learning in an ESL situation affect the acquisition of the locality condition? Or, is the asymmetry not observed if one starts to learn English before a sensitive period comes to an end? Chien and Wexler (1990) report that children acquiring English as their L1 do not exhibit the tensed-infinitive asymmetry. No studies, however, have tested L2 learning in these respects.

(*Received 12 April 2001*)

NOTES

1. The influence from other reflexives such as *kare-zisin* and *zibun-zisin*, which are often claimed to have the same local domain as English reflexives (Yuan, 1994), are discussed in the section entitled "Discussion."

2. I inquired about the instruction of the relevant condition by examining textbooks used in class, asking English teachers in Japan, and asking participants directly about their learning experiences. I found, in fact, that almost all participants in this study had not been taught such a restriction.

3. Here, adult learners are defined as those who start learning the L2 seriously after a sensitive period. Almost all participants of this study began learning English at the age of 13.

4. An anonymous reviewer points out that the former is automatically deduced from the latter given the universal properties of reflexives (i.e., a reflexive must have its antecedent within the sentence that contains it). This is true and, in fact, the participants in this study who answered "No" for LD binding almost all answered "Yes" for local binding (see Table 5). This might lead one to conclude that it would be sufficient to ask only about the impossibility of LD binding in the case of English reflexives, but that may not be true. This is because, even when participants have correct knowledge of English reflexives, it is possible for them to answer "No" for local binding as well because of various unexpected flaws in experimental devices.

5. Yuan (1994) argued that underdetermination (Hamilton, 1996, 1998; White, 1989) does not hold in Japanese learners acquiring English reflexives, because Japanese has other reflexives (X-*zisin*) that are considered to have a restriction on LD binding as English reflexives do. This point is discussed extensively in the section entitled "Discussion."

6. An anonymous reviewer points out that, depending on the status of PRO in L2 grammars (e.g., no PRO at a certain stage of development), the clause boundary might be placed in another position (e.g., between *told* and *Jane*). Control verbs are problematic in this sense, too.

7. The verb *ask* is more problematic, as Matsumura (1994) pointed out. It is an object control verb in most cases, but it can also behave as a subject control verb, as the following sentence indicates:

(i) *The student*$_i$ *suddenly got to his feet and asked the teacher* [PRO$_i$ *to go to the restroom*].

8. The university requires nonnative speakers of English to earn 600 on the TOEFL in order to be admitted to graduate programs. Furthermore, many of the participants had already spent a considerable amount of time in English-speaking countries when the experiment was conducted.

9. An anonymous reviewer comments that the participants in level 5 seem to be very different both in terms of age and experience. It may be true that this level appears to lack homogeneity as a group, compared with the other levels, and therefore it might be hard to characterize it truly as "advanced." However, it seems fair to say that, at least with respect to the structure at issue in this study (viz., the relationship between a reflexive and its antecedent), the participants at this level have sufficient learning experience. In this sense, it will be legitimate to regard them as a group of advanced learners.

10. Some participants were found to have some experience learning English before entering junior high school. These students were not excluded because such pre–junior high school learning is not seriously undertaken in most cases (usually only once a week) and it is often said that the effect of such learning is not maintained through Grade 8.

11. Interviews with the teachers of the participants in levels 1–4 revealed that they had not been taught about the restriction at issue. Furthermore, the grammar books used at the school did not mention anything about it.

12. The stories were given in Japanese rather than in English for the following reasons: (a) the proficiency of the participants at lower levels was not high enough to fully understand stories written in English, so it was expected that reading stories all in English might be more cognitively taxing; and (b) I needed to ensure that all participants understood the stories completely before presenting them with the stimulus sentences. The booklet-style presentation was employed to avoid potential interferences from Japanese. The control group of Japanese native speakers was presented with both stories and stimulus sentences in Japanese. Likewise, the control group of English native speakers was presented with both stories and stimulus sentences in English.

13. The only purpose of this sample item was to familiarize the participants with the procedure of the experiment and make it possible for them to do the subsequent experimental task smoothly. To avoid a response bias, the answer to the sample item was not provided, although the participants were allowed to ask about the procedure itself. A few participants had some difficulty interpreting the phrase "naturally inferred" in the directions, but its interpretation was left to each participant and, in fact, no serious problem seemed to occur with this treatment.

14. As will be discussed in the next section, this was termed the "tensed-infinitive asymmetry" by Yuan (1994). The phenomenon itself was already reported in early studies, such as Finer and Bro-

selow (1986), though their conclusion was obscured by methodological problems, as discussed earlier.

15. One possible reason why this percentage was lower than expected is that some participants found the sequence of NP-*ga* NP-*ga* awkward and so presumably placed primacy on the first NP.

16. Possible reasons for why Eckman (1994) found no asymmetry would be that (a) he used control verbs for type E-2, (b) only four native Japanese speakers participated in his study, and (c) the picture identification task he used failed to elicit ambiguity from his participants.

17. See Finer (1991) and Thomas (1991). In early studies, many cases of rogue grammars were observed, which constituted a weakness for the UG-access argument.

18. I will not get into details of the theory here but instead will simply point out that a value of the parameter for this type of grammar (or anaphor) was considered to be neither the value of Japanese *zibun* nor the value of English reflexives but the one between the two.

19. One piece of evidence against the theory came from the existence of reflexives that only allow LD binding, as found in Greek, Norwegian, Icelandic, and so forth.

20. The results from Matsumura (1994) are interesting in this respect. He showed that the participants at his higher level chose a correct antecedent for the verb *tell* at the rate of 92.86%, whereas for the verb *want* they did so at a rate of 14.29%. This might suggest that there is a stage in the process of acquiring English clausal structure at which the clause formed by an infinitive is not perceived (or PRO is missing or somehow invisible) and that, for object control verbs like *tell*, many participants in the previous studies interpreted the antecedent of a reflexive within an infinitival clause as the object of the verb of the matrix clause rather than as the PRO of the infinitival clause.

REFERENCES

Berwick, R. (1985). *The acquisition of syntactic knowledge*. Cambridge, MA: MIT Press.

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

Braine, M. (1971). On two types of models of the internalization of grammars. In D. Slobin (Ed.), *The ontogenesis of grammar* (pp. 153–188). San Diego, CA: Academic Press.

Brown, R., & Hanlon, C. (1970). Derivational complexity and the order of acquisition in child speech. In J. R. Hayes (Ed.), *Cognition and the development of language* (pp. 155–207). New York: Wiley and Sons.

Chien, Y., & 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. (1965). *Aspects of the theory of syntax*. Cambridge, MA: MIT Press.

Chomsky, N. (1981). *Lectures on government and binding*. Dordrecht: Foris.

Chomsky, N. (1986). *Knowledge of language: Its nature, origin, and use*. New York: Praeger.

Cole, P., Hermon, G., & Sung, L.-M. (1990). Principles and parameters of long-distance reflexives. *Linguistic Inquiry, 21*, 1–22.

Cole, P., & Sung, L.-M. (1994). Head movement and long-distance reflexives. *Linguistic Inquiry, 25*, 355–406.

Cook, V. J. (1990). Timed comprehension of binding in advanced L2 learners of English. *Language Learning, 40*, 557–599.

Corder, S. P. (1967). The significance of learners' errors. *International Review of Applied Linguistics, 5*, 161–170.

Crain, S., & McKee, C. (1986). Acquisition of structural restrictions on anaphora. *Proceedings of the North Eastern Linguistic Society, 21*, 1–22.

Eckman, R. R. (1994). Local and long-distance anaphora in second-language acquisition. In A. Cohen, S. Gass, & E. Tarone (Eds.), *Research methodology in second-language acquisition* (pp. 207–225). Mahwah, NJ: Erlbaum.

Finer, D. (1991). Binding parameters in second language acquisition. In L. Eubank (Ed.), *Point counterpoint: Universal Grammar in the second language* (pp. 351–374). Amsterdam: Benjamins.

Finer, D., & Broselow, E. (1986). Second language acquisition of reflexive binding. *Proceedings of the North Eastern Linguistic Society, 16*, 154–168.

Gass, S. (1983). The development of L2 intuition. *TESOL Quarterly, 17*, 273–291.

Hamilton, R. (1996). Against underdetermined reflexive binding. *Second Language Research, 12*, 420–446.

Hamilton, R. (1998). Underdetermined binding of reflexives by adult Japanese-speaking learners of English. *Second Language Research, 14*, 292–320.

Hirakawa, M. (1990). A study of the L2 acquisition of English reflexives. *Second Language Research, 6*, 60–85.

Hyams, N., & Sigurjónsdóttir, S. (1990). The development of "long-distance anaphora": A cross-linguistic comparison with special reference to Icelandic. *Language Acquisition, 1*, 57–94.

Lakshmanan, U., & Teranishi, K. (1994). Preferences versus grammaticality judgments: Some methodological issues concerning the governing category parameter in second-language acquisition. In A. Cohen, S. Gass, & E. Tarone (Eds.), *Research methodology in second-language acquisition* (pp. 185–206). Mahwah, NJ: Erlbaum.

Li, Y. (1994). The Japanese dialectal monomorphemic reflexive *zisin* and its theoretical implications: A contrast with Chinese *ziji*. In B. Lust, G. Hermon, & J. Kornfilt (Eds.), *Syntactic theory and first language acquisition: Crosslinguistic perspectives: Vol. 2. Binding, dependencies, and learnability* (pp. 83–91). Mahwah, NJ: Erlbaum.

MacLaughlin, D. (1995). Language acquisition and the Subset Principle. *Linguistic Review, 12*, 143–191.

Manzini, M. R., & Wexler, K. (1987). Parameters, binding theory, and learnability. *Linguistic Inquiry, 18*, 413–444.

Matsumura, M. (1994). Japanese learners' acquisition of the locality requirement of English reflexives: Evidence for retreat from overgeneralization. *Studies in Second Language Acquisition, 16*, 19–42.

Progovac, L. (1992). Relativized SUBJECT: Long-distance reflexives without movement. *Linguistic Inquiry, 23*, 671–680.

Progovac, L. (1993). Long-distance reflexives: Movement-to-Infl versus relativized subject. *Linguistic Inquiry, 24*, 755–772.

Reinhart, T., & Reuland, E. (1991). Anaphors and logophors: An argument structure perspective. In J. Koster, & E. Reuland (Eds.), *Long-distance anaphora* (pp. 283–321). New York: Cambridge University Press.

Reinhart, T., & Reuland, E. (1993). Reflexivity. *Linguistic Inquiry, 24*, 657–720.

Selinker, L. (1972). Interlanguage. *International Review of Applied Linguistics, 10*, 209–231.

Sigurjónsdóttir, S., & Hyams, N. (1992). Reflexivization and logophoricity: Evidence from the acquisition of Icelandic. *Language Acquisition, 2*, 359–413.

Shimura, A. (1990). *Testing the availability of Universal Grammar in second language acquisition: The governing category parameter and the proper antecedent parameter in Japanese as a second language*. Unpublished Master's thesis, University of Hawai'i at Manoa.

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: Benjamins.

Thomas, M. (1995). Acquisition of the Japanese reflexive *zibun* and movement of anaphors in Logical Form. *Second Language Research, 11*, 206–234.

Wakabayashi, S. (1996). The nature of interlanguage: SLA of English reflexives. *Second Language Research, 12*, 266–303.

Wexler, K., & Manzini, R. (1987). Parameters and learnability in binding theory. In T. Roeper & E. Williams (Eds.), *Parameter setting* (pp. 41–76). Dordrecht: Reidel.

White, L. (1989). *Universal Grammar and second language acquisition*. Amsterdam: Benjamins.

White, L., Bruhn de Garavito, J., Kawasaki, T., Pater, J., & Prévost, 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.

Yuan, B. (1994). Second language acquisition of reflexives revisited. *Language, 70*, 539–545.

Yusa, N. (1998). A minimalist approach to second language acquisition. In S. Flynn, G. Martohardjono, & W. O'Neil (Eds.), *The generative study of second language acquisition* (pp. 215–238). Mahwah, NJ: Erlbaum.

APPENDIX

EXPERIMENTAL MATERIALS

Truth-Value Judgment Task (English-Translated Version)

Example context:

Mary's grandmother has never been to Disneyland. Mary had planned to take her today, but she couldn't. She had work to finish as soon as possible. Then, Mary asked John, "Will you please take her to Disneyland?" "Of course; no problem," answered John. Mary's grandmother was very happy because she had a good time in Disneyland.

Example stimulus sentence:

Today John and Mary's grandmother went to Disneyland.

Type E-1 (embedded finite clause) contexts and sentences.

1. A thought that B hated himself/herself. (LD, local)
   LD context:
   John and Taro were best friends. John fell in love with Taro's girlfriend. Taro's girlfriend became interested in John too, and she broke up with Taro. Since then Taro has never even tried to see John.
   Stimulus sentence:
   John thought that Taro hated himself.
   Local context:
   Yesterday Susan saw Tomoko sitting alone on a bench. Tomoko looked very depressed. Susan asked Tomoko, "What's the matter?" Tomoko answered quietly, "I am not doing well in my classes. I am not good at sports. I seem to fail at everything I try. I am not attractive, and I am not popular among boys!"
   Stimulus sentence:
   Susan thought that Tomoko hated herself.
2. A thought that B was blaming himself/herself. (LD, local)
   LD context:
   Today after school, Tomoko came to Susan and said, "Why did you tell the teacher on me?" Tomoko looked very angry. Then, Tomoko said, "I was scolded really badly by the teacher because you told on me!"
   Stimulus sentence:
   Susan thought that Tomoko was blaming herself.
   Local context:
   Taro and John are on the same baseball team. Their team lost a game yesterday. Taro said to John, "We lost the game because I made a lot of errors." He added, "I shouldn't have played in the game."
   Stimulus sentence:
   John thought that Taro was blaming himself.
3. A said that B wrote about himself/herself in the letter. (LD, local)
   LD context:
   Taro's mother saw him and said, "What are you thinking about so seriously?" Taro replied, "John sent me a letter and said that I should act more carefully and be more considerate of others."

https://doi.org/10.1017/S027226310200102X Published online by Cambridge University Press

Stimulus sentence:
Taro said that John wrote about himself in the letter.
Local context:
Mrs. Sato was Tomoko's teacher when she was at elementary school. The other day, Mrs. Sato received a letter from Tomoko for the first time since graduation. Mrs. Sato read the letter and said to her husband, "The letter says that Tomoko graduated from college and she is now working as a nurse. Letters like this always make me happy."
Stimulus sentence:
Mrs. Sato said that Tomoko wrote about herself in the letter.

Type E-2 (embedded nonfinite clause) contexts and sentences.

1. A wanted B to talk about himself/herself in class. (LD, local)
LD context:
Susan is now studying with Tomoko at a school in Japan. Susan became friends with Tomoko while Tomoko was staying in America. Yesterday, in the social studies class, Tomoko talked a lot about her experiences in America, but she did not talk at all about her encounters with Susan. Susan was a little disappointed.
Stimulus sentence:
Susan wanted Tomoko to talk about herself in class.
Local context:
Taro began to study with John at a school in America. In the first class, the teacher asked Taro to make a short speech in front of his classmates. Taro talked only about Japanese politics and culture. John wanted to know about Taro rather than about Japan.
Stimulus sentence:
John wanted Taro to talk about himself in class.
2. A wanted B to think about himself/herself (again). (LD, local)
LD context:
Kazu is the captain of John's soccer team. Yesterday Kazu announced the starting players for the next game. Unfortunately, John's name was not included in the list. John said to Kazu, "I wish I could be on the starting team, too."
Stimulus sentence:
John wanted Kazu to think about himself again.
Local context:
Taro is Mr. Tanaka's student. Taro is mean and violent, so he always gets into trouble with his friends. Yesterday Mr. Tanaka said to Taro harshly, "Since you are like this, everybody hates you now." Mr. Tanaka simply wanted Taro to be nicer.
Stimulus sentence:
Mr. Tanaka wanted Taro to think about himself.
3. A wanted B to know himself/herself better.
LD context:
Mr. Ito is the director of the drama club, and he was looking for an actor to play the main character in the next play. Since Taro had acting experience from when he was in America, Taro thought, "I will be perfect for the role!" So, he sent a long letter to Mr. Ito, telling him about his acting experiences in America.
Stimulus sentence:
Taro wanted Mr. Ito to know himself better.
Local context:
Tomoko wants to be an actress some day. Tomoko always says to Susan, "I am pretty and look clever. Once I get into show business, I will be given the role of a heroine immediately." Unfortunately, however, nobody around Tomoko thinks that this will happen. Susan is a little worried about Tomoko.
Stimulus sentence:
Susan wanted Tomoko to know herself better.

Distracters.

1. A fan of Takanohana hit himself.
   Context:
   Yesterday Taro and John went to watch sumo wrestling. Takanohana was entering the sumo ring. Suddenly, a man got out of his seat, hit Takanohana on the back, and said, "You have a nice body." Taro and John thought the man was very rude to Takanohana.
   Stimulus sentence:
   A fan of Takanohana hit himself. (No)
2. The girl beside Susan saw herself in the mirror and laughed.
   Context:
   Susan's school had a costume parade today. There were many girls preparing for the parade in the ladies room. Susan began to put on make-up in front of a big mirror to look like a witch. A girl putting on make-up next to Susan looked at her and said laughing, "Wow, you are really scary!!" Susan thought, "Maybe I overdid the make-up."
   Stimulus sentence:
   The girl beside Susan saw herself in the mirror and laughed. (No)

Syntax Tests

Anaphor test.

1. John met Bill on the way home. Suddenly John hit himself. Q: Who did John hit?
2. John met Bill on the way home. Suddenly Bill hit him. Q: Who did Bill hit?
3. John and Bill stood in front of the mirror. Bill saw himself there. Q: Who did Bill see in the mirror?
4. John and Bill stood in front of the mirror. John saw him there. Q: Who did John see in the mirror?

That-*clause/infinitival clause test.* Make a sentence with the indicated meaning using the words given.

1. (thinks, Tom, loves, Bill, Mary, that, love) very much.
2. (wrote, Bill, a letter, write, knows, Tom) last night.
3. (him, a doctor, be, to, wants, John's father, is) in the future.
4. (Mary, the party, to, to, came, come, wanted, John) last night.

Transfer test. Put into Japanese.

1. Bill thought that Kenji hated himself.
2. Lisa said that Yumiko wrote about herself in the letter.
3. Ichiro thought that Tomoki was blaming himself.