SYNTACTIC FEATURE TRANSFER AND REFLEXIVE BINDING IN INTERLANGUAGE

NAO ISHINO
Kwansei Gakuin University

The purpose of this paper is not simply to give an empirical description, but rather to give a theoretical explanation of the syntactic mechanism for reflexive binding in interlanguage grammar by proposing Feature Transfer and Feature Learning Hypothesis, with the idea of φ-feature (de)composition. Our hypothesis demands that formal features in an L1 item, if specified featurally, should be transferred to interlanguage grammar, while underspecified features in an L1 item should not. We will demonstrate that the reflexives in the interlanguage grammar of Japanese learners of English are φ-defective and their syntactic properties are explained with our hypothesis coherently under the current minimalist binding theory through Agree.*

Keywords: Feature Transfer, Feature Learning, φ-feature decomposition, Agree theory of reflexive binding

1. Introduction
1.1. Aims and Data
In the literature on the UG-based studies of SLA, the L2 acquisition of the English reflexives by Japanese learners of English (JLsE) has been investigated extensively. To be specific, it has widely been observed (e.g. Finer and Broselow (1986) and Hirakawa (1990)) that JLsE are apt to misunderstand the binding dependency of the English reflexives in their intermediate acquisition stage (which is called interlanguage (IL) grammar (cf. White (2003))). It seems, however, that most of the previous studies have

---

* Portions of this paper were presented at Gifu University, Keio University, and Kobe University. I wish to thank the audience at those meetings. I would like to express my sincere gratitude to Hiroyuki Ura and Kiyomi Kusumoto for their invaluable suggestions and encouraging comments on the materials presented herein. Thanks also go to members of Generative Lyceum, held at Kwansei Gakuin University, for judgments and helpful comments. Particular thanks also go to two anonymous EL reviewers. All remaining inadequacies are my own.
succeeded in drawing descriptive generalizations about their empirical ob-
servations/facts on the basis of their experimental surveys; that is, they have
inductively shown how the reflexives in the JLsE’s IL grammar behave in
syntactic respects, but they have failed to explicate the theoretical rationale
as to why JLsE’s IL reflexives syntactically behave as such.

The purpose of this paper, therefore, is not simply to give an empiri-
cal description, but rather to give a theoretically consistent explanation, by
proposing a hypothesis with the idea of the φ-feature (de)composition of
a reflexive, to the syntactic mechanism for the L2 acquisition of reflexive
binding. We will hypothesize that the syntactic mechanism of the L2 ac-
quision should be recast within the minimalist syntax; more specifically,
we will propose that formal features within an L2 lexical item, but not
lexical items per se nor syntactic properties inherent in a particular construc-
tion (such as syntactic parameters), are to be acquired in the course of the
L2 acquisition. Syntactic formal features include person-, gender-, and
number-features. Following, basically, the idea about the decomposition of
φ-features proposed by Bouchard (1984) and Burzio (1991), this paper will
reveal that the English reflexives in JLsE’s IL grammar are φ-defective; that
is, they are not fully specified featurally.

In order to achieve our goal, we will first report an empirically new dis-
covery concerning the specification of each φ-feature within the reflexive
item in JLsE’s IL grammar: Through conducting several experimental sur-
veys, we will reveal (I) that the person-feature within the reflexive item in
JLsE’s IL grammar is strictly specified; as a result, JLsE at their IL stage
do not have the person-neutral interpretation for the English reflexives; (II)
that the gender-feature within the reflexive item in JLsE’s IL grammar is
strictly specified; as a result, they do not have the gender-neutral interpreta-
tion for the English reflexives; and (III) that the number-feature within the
reflexive item in JLsE’s IL grammar is partially specified; as a result, they
are apt, mistakenly, to regard himself/herself as having the distributive read-
ing. These facts are very significant on empirical grounds, because each of
the φ-features within the English reflexives in JLsE’s IL grammar has sel-
dom been revealed in the literature on SLA.

It is important to notice, here, that our main goal is to provide a theoreti-
cally adequate explanation to the issue as to why the IL reflexives show the
syntactic properties which have been noticed through the major experiments
conducted in some previous literature on the relevant topic. While arguing
that two leading approaches to JLsE’s IL grammar of reflexive binding can-
not provide any satisfactory explanation to the empirical/experimental dis-
covery, we will propose a new hypothesis, which states that some but not all of the formal features within an L1 item are transferred to IL grammar (the hypothesis which we call Feature Transfer and Feature Learning Hypothesis (FTFL)). FTFL founds itself conceptually on the feature-based derivational syntax under the Minimalist framework (cf. Chomsky (1995) and subsequent work). More specifically, FTFL stipulates that each of the formal features within an L1 lexical item, unless underspecified featurally, should be transferred to IL grammar in the course of the L2 acquisition. Employing FTFL, we will reveal in a deductive manner that the English reflexives in JLsE’s IL grammar are φ-defective. Then, under the current minimalist assumptions of binding through Agree (see Reuland (2005, 2011), Heinat (2008), Quicolli (2008), Gallego (2010), and among others), we will give a theoretical explanation to the experimentally obtained fact, oft-reported in the literature on JLsE’s acquisition of English reflexives, that JLsE are apt to allow an IL reflexive to be bound by an antecedent even when the binding dependency between them is separated by a non-tensed clause boundary (though disallowing the binding dependency between them when they are separated by a tensed clause boundary), and JLsE’s IL reflexives show subject orientation.

One of the theoretically important consequences of our theory to be proposed in this paper is that FTFL, if reinforced with the idea utilizing the φ-feature decomposition of reflexives, leads to a lucid solution of the theoretically interesting issue as to what triggers the erroneous usage of reflexives by JLsE in their IL grammar: We will propose, endorsing the concept of markedness in the field of L2 studies, that learning of specified features within an L2 item is easy in a case where their equivalent L1 feature is underspecified because specified features within an L2 item are regarded as unmarked, whereas learning of underspecified features within an L2 item is rather difficult in a case where their equivalent L1 feature is specified because underspecified features within an L2 item are regarded as marked. Moreover, it will also be proposed that Feature Transfer takes precedence over Feature Learning where the specification of the L1 feature at issue exactly parallels that of the corresponding L2 item. FTFL will, then, be shown to be theoretically more adequate than major previous approaches to JLsE’s acquisition of reflexives.

1.2. Organization

This paper is organized as follows: In §2, we will argue that two major approaches to the L2 acquisition of reflexive binding have serious problems
on theoretical grounds. In §3, we will present the results of our experimental surveys on syntactic properties of JLsE’s IL reflexives, which include their neutral interpretation and their distributive reading. Our empirical observation indicates that the previous approaches have problems on empirical grounds. In §4, through highlighting the number-feature specification of a reflexive and its distributive reading, we will express how the specification of the φ-features within a lexical item is concerned with its syntactic properties. Then, we will elucidate the φ-feature composition of JLsE’s IL reflexives and conclude that JLsE’s IL reflexives are φ-defective. In §5, we will propose our new hypothesis (i.e. FTFL), which enables us to explain why some of the syntactic properties of a certain L1 reflexive are transferred and materialized in JLsE’s IL grammar, but its other syntactic properties are not. In §6, with the help of some independently motivated assumptions under the minimalist framework, we will explicate the question as to why JLsE’s IL reflexives behave as what we have discovered through our experimental surveys. Finally, §7 will conclude the present paper.

2. Previous Approaches and Their Conceptual Problems

Among the UG-based studies on SLA, the Parameter Resetting Approach (PRA) (cf. Finer and Broselow (1986), Hirakawa (1990), and Watanabe et al. (2008), among many others), and the (Lexical) Transfer Approach (LTA) (cf. Yuan (1994), and Ishino and Ura (2009), among others) are widely recognized as a foremost approach to L2 acquisition of reflexive binding. According to PRA, L2 learners’ IL grammar is formed through resetting the parametric values of their L1; on the other hand, according to LTA, they adopt the parametric values of a lexical item in their L1 grammar to form their IL grammar (see White (2003) for comparative discussion). Both of them, nevertheless, agree that UG is available somehow to L2 learners.

2.1. PRA

According to the Binding Theory under the GB framework (cf. Chomsky (1981) and Chomsky and Lasnik (1993)), an anaphor must be bound in its syntactically local domain (what is called “binding domain”). Advocates for PRA (e.g. Hirakawa (1990), Thomas (1995), and MacLaughlin (1998)) have built their rationale upon the conception of Governing Category Parameter (GCP), which states that the binding domain of an anaphor is alleged to be parametrically determined (cf. Yang (1984) and Wexler and
The five parametric values determine five different possible binding domains, which range from the most restrictive English type (i.e. the value (a)) to the least restrictive Japanese *zibun*-type (i.e. the value (e)). Then, some of the advocates for PRA (e.g. Hirakawa (1990) and Watanabe et al. (2008)) adopt the Subset Principle of Wexler and Manzini (1987) (cf. also, Berwick (1985)), and try to explain the experimental result that JLsE in their IL stage adopt neither their own L1 parameter value for the binding domain (i.e. the value (e)), nor the English value (i.e. the value (a)); rather, they reset their L1 parameter into the Russian parameter value (i.e. the value (c)), which is intermediate between (a) and (e). According to PRA, the reason why it is rather hard for JLsE to acquire the correct parameter value for the English reflexives stems from the fact that the Japanese value includes the English one as a proper subset, and the Subset Principle predicts that it is rather hard for JLsE to acquire only with positive evidence.

PRA, if reinforced with the Subset Principle, maintains that the parametric value intermediate between the L1 value and the L2 one is reset in learners’ IL grammar when the relevant parameter value in L2 does not include the one in L1 as a subset. For the PRA advocates, this plays the most important role of their explanation for the binding dependency of JLsE’s IL reflexives, as we argued above. It should be noted, however, that there are two values other than the value (c), which lie between the Japanese value (e) and the English value (a); namely, the value (d) and the value (b), in addition to the value (c), lie between them. Then, why is it that there are so many JLsE who adopt the value (c), instead of adopting the value (d) or the value (b)? It seems that there is no consistent explanation of this issue under PRA and no cogent speculation on it without any proviso has yet been provided as far as we can detect.

2.2. LTA

Advocates for LTA have hypothesized (e.g. Yuan (1994) and Ishino and Ura (2009)) that the syntactic properties of a lexical item in learners’ L1

---

1 GCP (Wexler and Manzini (1987: 53)) states that $\alpha$ is a governing category for $\beta$ if $\alpha$ is the minimal category which contains $\beta$ and (a) has a subject, (b) has an INFL, (c) has a TNS, (d) has an indicative TNS, or (e) has a root TNS.

2 The binding domain of the Russian reflexive *sebja* is alleged to show the value (c) of GCP: It must be bound in a tensed clause, but it can allow the long-distance binding over a clause boundary when it is embedded in a non-tensed clause.
grammar are copied (i.e. transferred) to the corresponding lexical item in their target language. Returning, with this in mind, to JLSE’s acquisition of reflexives, we are led to predict that the syntactic properties of zibun ‘SELF’ in L1 Japanese are copied to the English reflexives in JLSE’s IL grammar when zibun is selected as the representative of the Japanese reflexives; likewise, when either zibun-zisin ‘SELF-self’ or pronoun+zisin (such as kare-zisin ‘himself’) is selected as the representative of the Japanese reflexives, its syntactic properties are copied to the corresponding lexical item in JLSE’s IL grammar.

LTA, too, is not free from any conceptually serious problem: A reflexive in learners’ L1 grammar must be selected as the representative of their L1 reflexives before its syntactic properties are copied (i.e. transferred) to the corresponding lexical item in their target language. Thus, Yuan (1994) proposes that pronoun+zisin should be selected and transferred to JLSE’s IL grammar in the course where they learn the English reflexives, and Ishino and Ura (2009) propose that zibun-zisin, instead, should be selected and transferred. Those proposals, no matter how much descriptive sufficiency they might have on empirical grounds, have a conceptually fatal problem, however: As correctly pointed out by MacLaughlin (1995), there is no theoretically/conceptually evident reason as to which one of the Japanese reflexives is to be selected as the representative of them and transferred to JLSE’s IL grammar. Unless this issue is given a lucid solution, LTA is not sufficient on conceptual grounds.

In this section we have argued that the two major approaches to JLSE’s acquisition of the English reflexives (i.e. PRA and LTA) are somewhat deficit on theoretical grounds. As will be argued in detail in the next section, they have empirical problems, too. In the literature on the SLA of reflexives, extensive attention has been paid almost exclusively to their binding domain in IL grammar and occasionally/sporadically to the existence/absence of their subject orientation in IL grammar. But what about other kinds of syntactic properties that reflexives show in JLSE’s IL grammar? Our experimental survey, which will be introduced in §3, indicates that some properties (such as the availability of the distributive reading) of the IL reflexives are not explainable, without any ad hoc proviso, under PRA or LTA.

3. Syntactic Properties of Interlanguage Reflexives

For our theory to be proposed hereinafter, it is important to explicate the precise nature of the φ-feature composition of the reflexives in JLSE’s L1
grammar (i.e. Japanese) and that of the reflexives in their target language (i.e. English). Let us take a closer look at the Japanese reflexives *zibun* and *zibun-zisin*. As shown in (1), they allow both the person-neutral interpretation (as shown in (1a) and (1b)) and the gender-neutral interpretation (as shown in (1c)) in the grammar of L1 Japanese:

(1) a. Watasiₜₜ-wa kanozyoₜₜ-ni zibun(-zisin)ₜₜ-h/-o
   I-Top her-Dat SELF-self-Acc
   understand-Cause-Past
   Lit. ‘I made her understand SELF(-self).’

b. Anataₜₚ-wa kareₜₚ-ni zibun(-zisin)ₜₚ-h/-o
   You-Top he-Dat SELF-self-Acc
   understand-Cause-Past
   Lit. ‘You made him understand SELF(-self).’

c. Sono otokoₜₜ-wa ano onnaₜₚ-ni zibun(-zisin)ₜₚ-h/-o
   the man-Top that woman-Dat SELF-self-Acc
   understand-Cause-Past
   Lit. ‘The man made that woman understand SELF(-self).’

Next, let us look at (2) below:

(2) [John to Mary]ₜₜ-wa zibun(-zisin)ₜₜ-h/-o hihansi-ta.
   and -Top SELF-self-Acc criticize-Past
   Lit. ‘[John and Mary] criticized SELF(-self).’

The sentence in (2), if it is to be interpreted as “the couple of John and Mary criticized the couple of John and Mary” (as delineated by its indices), is unacceptable in L1 Japanese. This indicates that the singular reflexive form *zibun(-zisin)* is not tolerable with the number-neutral interpretation; that is, the singular reflexive form cannot be bound by the plural DP *[DP John to Mary]*. It is interesting, however, that the sentence in (2) is acceptable if *zibun(-zisin)* is distributively bound by John and Mary; that is, (2) is grammatical if it means “John criticized John, and Mary criticized Mary.” Throughout this paper, we call this type of reading *distributive reading*.

Now that the distributive reading is one of the syntactic properties that the reflexives in L1 Japanese show, it is highly interesting, for our study of the SLA of reflexives, to examine how the syntactic property concerning the distributive reading is materialized in JLsE’s IL grammar. Accordingly, we have conducted the following experimental surveys, with which we will deal
minutely in this section.

3.1. Experiments
3.1.1. Subjects

We first set out the following four surveying tests in order to examine (i) whether JLsE’s IL reflexives allow the person-neutral interpretation (Test 1), (ii) whether they allow the gender-neutral interpretation (Test 2), (iii) whether they allow the number-neutral interpretation (Test 3), and (iv) whether they allow the distributive reading (Test 4). It should be noted that the existing studies on the SLA of reflexives have paid no attention to these four points. In addition, we set out the two follow-up tests to examine (v) whether JLsE’s IL reflexives show the subject orientation (Test 5) and (vi) what is the locality for their binding dependency (Test 6). As mentioned above, these two points have been extensively investigated in the L2 literature (As for these follow-up tests, we will discuss them in detail in §6).

59 university students (26 male students and 33 female students) participated in Test 1, 2, and 3, and 154 university students (69 male students and 85 female students), which included the abovementioned 59 students, participated in Test 4. With respect to the follow-up tests, 369 university students (208 male students and 161 female students) and 288 university students (152 male students and 136 female students), both of which included the abovementioned 154 students, participated in Test 5 and Test 6. All of our experimental subjects were native Japanese and had studied English as a second language for more than seven years, but none of them had lived in English-speaking countries for a short or long period of time. At the time of our experiments, they were all in the freshman year in their university and registered for a course in linguistics. Their average age was 18.7. Their average score of TOEFL(P) (TOEFL Paper-based Test) was approximately 505. A preliminary survey conducted at the beginning of their linguistic class indicates that they have the correct capability both to distinguish an embedded clause from the matrix clause and to recognize that the subject is structurally higher than the object in an active, declarative clause (that is, the subject asymmetrically c-commands the object in an ordinary clause). With respect to clausal distinction, we have confirmed that our JLE subjects correctly recognized that seriously in (i) is adjoined to the matrix TP and seriously in (ii) is adjoined to the embedded TP; for, many of them successfully recognized that the former

3 With respect to clausal distinction, we have confirmed that our JLE subjects correctly recognized that seriously in (i) is adjoined to the matrix TP and seriously in (ii) is adjoined to the embedded TP; for, many of them successfully recognized that the former
in their L2 English acquisition stage. In addition, we had a control group consisting of 20 native speakers of English, all of whom stayed in Japan as exchange students (they were all undergraduate students at college and their average age was 18.9).

3.1.2. Procedures
At the time of our experimental survey, the participants were told about the aim of the experiments and their anonymity. They were given a sheet of questionnaire and were informed that their anonymous answers would be used in a linguistic research. Some couple of test sentences with a few control sentences were presented (see Appendix for all test sentences). At the top of the questionnaire, the following directions were given in Japanese: “Do it yourself without consulting with anyone,” and “Answer each question within a minute. Answer using your first impression, and do not over-think.” All questions were presented in the way that is called Grammatical Judgment Test. The following direction was given in Japanese: “If you regard each sentence as grammatical/acceptable, fill the blank with a ○; if not (that is, if you regard it as ungrammatical/unacceptable), fill the blank with a ×.” Even when our experimental subjects could not determine the acceptability of a sentence, they were required to answer based on their intuition. Our experimental subjects were not allowed to change their answer once they filled a blank in the questionnaire.

3.2. Tests
3.2.1. Test 1: Person-Neutral Interpretation
In order to discern whether JLsE consider the English reflexives to allow the person-neutral interpretation in their IL grammar, we presented our experimental subjects with the following test sentences, both of which are ill-formed in L1 English (see Appendix Test 1 for the other test sentences):

(i) John believes seriously that Mary loves Bill.
(ii) John believes that Mary loves Bill seriously.

Without the knowledge about clausal distinction, they could not have succeeded in the interpretational difference between (i) and (ii).

4 As will be shown in the Appendix, additional test sentences were indeed given in our survey, the number of which seems to be adequate to ensure that our experimental results reflect the IL grammar of JLsE.
(3)  a. *Bill and I hit yourselves.
     b. *Bill likes yourself.

Table 1 delineates the result of this survey:

<table>
<thead>
<tr>
<th></th>
<th>JLSa</th>
<th>Controlb</th>
</tr>
</thead>
<tbody>
<tr>
<td>acceptable</td>
<td>0.0</td>
<td>0.0</td>
</tr>
<tr>
<td>unacceptable</td>
<td>100.0</td>
<td>100.0</td>
</tr>
</tbody>
</table>

\( a_{n=59}. \quad b_{n=20}. \)

As Table 1 shows, 100% of our 59 JLE subjects correctly interpreted the test sentences in (3) as unacceptable. Our experimental survey further revealed that every experimental subject cannot interpret all of the eight variant forms (i.e. myself, ourselves, yourself, yourselves, himself, herself, itself, and themselves) as referring to any other person than the person the morphology of the respective form designates (see Appendix 1).

3.2.2. Test 2: Gender-Neutral Interpretation

Next, in order to discern whether JLS consider the English reflexives to allow the gender-neutral interpretation in their IL grammar, we presented our experimental subjects with the test sentences in (4), both of which are ill-formed in the grammar of L1 English (see Appendix Test 2 for the other test sentences):

     b. *The boy criticized herself.

We intended these test sentences to make it clear whether our JLE subjects can interpret himself in (4a), for example, as being coreferential with the girl. The result of this survey is delineated in Table 2:

<table>
<thead>
<tr>
<th></th>
<th>JLSa</th>
<th>Controlb</th>
</tr>
</thead>
<tbody>
<tr>
<td>acceptable</td>
<td>0.0</td>
<td>0.0</td>
</tr>
<tr>
<td>unacceptable</td>
<td>100.0</td>
<td>100.0</td>
</tr>
</tbody>
</table>

\( a_{n=59}. \quad b_{n=20}. \)

As Table 2 shows, 100% of our 59 JLE subjects correctly interpreted the test sentences in (4) as unacceptable. Our experimental survey further revealed that every experimental subject cannot interpret the IL third-person, non-neuter, singular reflexives (i.e. himself and herself) as referring to the
opposite gender (as reported in Appendix Test 2).

3.2.3. Test 3: Number-Neutral Interpretation

Now, in order to discern whether JLsE consider the English reflexives to allow the number-neutral interpretation, we presented our experimental subjects with the test sentences in (5), both of which are ill-formed in the grammar of L1 English (see Appendix Test 3 for the other test sentences):

(5) a. *Bill likes *themselves.
    b. *We like *myself.

The result of this survey is delineated in Table 3:

<table>
<thead>
<tr>
<th></th>
<th>%</th>
<th>JLsE</th>
<th>Control</th>
</tr>
</thead>
<tbody>
<tr>
<td>acceptable</td>
<td>1.7</td>
<td>1.7</td>
<td>0.0</td>
</tr>
<tr>
<td>unacceptable</td>
<td>98.3</td>
<td>98.3</td>
<td>100.0</td>
</tr>
</tbody>
</table>

\(a_n=59. \quad b_n=20.\)

As Table 3 shows, only 1.7% of our 59 JLE subjects regard the sentences in (5) as acceptable.

3.2.4. Test 4: Distributive Reading

We also contrived another test in order to reveal one of the yet unknown properties of JLsE’s IL reflexives. In (6) and hereinafter, the index \(k\otimes j\) indicates that the so-indexed item is bound distributively by the item with the index \(k\) and the item with the index \(j\).

(6) John\(_k\) to Bill\(_j\)-wa (sore-zore) zibun(-zisin)\(_k\otimes j\)/*kare-zisin\(_k\otimes j\)-o and -Top (each) SELF(-self)/himself-Acc

hihansi-ta.
criticize-Past

Lit. ‘John criticized John, and Bill criticized Bill.’

For the time being, we will provisionally ignore PRONOUN+zisin (such as kare-zisin), which disallows the distributive reading (as shown in (6)), and concentrate our attention only on zibun-zisin. As shown in (6), zibun-zisin allows the distributive reading, irrespective of the existence/absence of the distributive operator sore-zore ‘each.’

On the other hand, Heim, Lasnik, and May (1993) report that the singular reflexives in L1 English, regardless of whether the distributive operator each overtly exists or not, cannot permit the distributive reading, as shown in (7).
(7) a. *John$_k$ and Bill$_j$ (each) criticized himself$_{k\otimes j}$.
   b. *Mary$_k$ and Jane$_j$ (each) criticized herself$_{k\otimes j}$.

Now in order to make it clear whether JLsE consider the English reflexives to allow the distributive reading in their IL grammar, we adopted the test sentence in (8), which is unacceptable in L1 English (see Appendix Test 4 for the other test sentences):

(8) *Mary$_k$ and Sue$_j$ saw herself$_{k\otimes j}$ in the mirror.

The result is shown in Table 4:

<table>
<thead>
<tr>
<th>(%)</th>
<th>JLSEx</th>
<th>Controlb</th>
</tr>
</thead>
<tbody>
<tr>
<td>acceptable</td>
<td>62.2</td>
<td>0.0</td>
</tr>
<tr>
<td>unacceptable</td>
<td>37.8</td>
<td>100.0</td>
</tr>
</tbody>
</table>

n=154. n=20. Notes. There is a statistically significant difference between the JLE group and the control group: t=6.59, p<.01.

Of great importance on empirical grounds is the result that 62.2% of the total JLE subjects mistakenly regarded their IL singular reflexive as allowing the distributive reading.

3.3. Empirical Problems in Predictions with PRA and LTA

So far in this section we have presented the results of our experimental survey concerning some syntactic properties of JLsE’s IL reflexives. In this subsection we will demonstrate that these results raise empirically serious problems for PRA and for LTA, which were outlined in §2 as major approaches to the SLA of reflexives. Either or both of them fail(s) to provide any consistent account of our experimental observation that JLsE’s IL reflexives have the following properties: (A) They disallow the person-neutral interpretation, (B) they disallow the gender-neutral interpretation, and (C) they allow the distributive reading.

3.3.1. Problems for PRA

It is true that PRA is good at explaining the oft-reported observation concerning the locality for the binding dependency of JLsE’s IL reflexives, but it cannot be predicted, under PRA, whether the reflexives in JLsE’s IL grammar have the neutral interpretation in person, gender and number; for, it is obvious that GCP is not relevant to person-/gender-interpretation or to the distributive reading of a given reflexive, and it seems highly implausible to speculate that there exists a parameter for the person-/gender-neutral in-
interpretations or a parameter of the distributive reading in UG. If this kind of speculation should be approved, then UG would admit extremely many parameters, the number of which equates to the number of all the possible interpretations observable in UG, resulting in the collapse of the generative enterprise. To conclude, PRA cannot make a consistent prediction for the reflexive binding in JLsE’s IL grammar with respect to their neutral interpretation in person, gender and number and distributive reading; as a consequence, PRA is insufficient in explaining JLsE’s SLA of reflexive binding regardless of any experimental result concerning JLsE’s reflexive binding.

3.3.2. Problems for LTA

In Chart 1 below, we make a comparison between the syntactic properties of each reflexive form in L1 Japanese, L1 English, and JLsE’s IL reflexives.

<table>
<thead>
<tr>
<th>Chart 1. The Syntactic Properties of L1 Japanese/IL/L1 English Reflexives</th>
</tr>
</thead>
<tbody>
<tr>
<td>L1 Japanese</td>
</tr>
<tr>
<td>zibun(-zisin) / PRON.+zisin</td>
</tr>
<tr>
<td>(A) Person-Neutral Interpretation</td>
</tr>
<tr>
<td>(B) Gender-Neutral Interpretation</td>
</tr>
<tr>
<td>(C) Distributive Reading</td>
</tr>
</tbody>
</table>

Suppose that LTA is on the right track. Then, it is predicted that there must be a reflexive form in L1 Japanese that syntactically behaves the same as JLsE’s IL reflexives; for, LTA postulates that all of the syntactic properties of the selected reflexive form in L1 are transferred and materialized in learners’ IL grammar. Now, LTA will be approved if the above prediction is borne out; otherwise, LTA turns out to be insufficient in explaining JLsE’s SLA of reflexive binding. Contrary to the prediction by LTA, neither of the Japanese reflexive forms disaccords behaviorally with JLsE’s IL reflexives, as can be seen in Chart 1. Therefore, this counts as an empirically serious problem for LTA. To conclude, LTA cannot give a consistent account to all of our experimental results from (A) through (C). This indicates that LTA, too, is insufficient on empirical grounds.

4. Φ-Feature Specification of Reflexives

In the previous section, we pointed out some empirical problems imma-
nant in PRA and LTA, respectively, each of which is too troublesome for PRA or LTA to be maintained as a plausible theory about JLsE’s acquisition of reflexive binding. In what follows, we will argue that the φ-feature specification of a lexical item plays the most significant role in a possible theory for JLsE’s L2 acquisition of reflexives. Before offering an alternative proposal, which is based conceptually on the idea of the φ-feature decomposition within a reflexive, we will first sketch out the specification of each φ-feature within the reflexives in L1 Japanese and L1 English and investigate how the feature specification of the relevant reflexives determines their syntactic properties.

4.1. Three Types of Feature Specification

First, we assume that feature specification can logically be classified into three types: (α) underspecified, (β) strictly specified, and (γ) partially specified. Our definition of φ-feature specification can be illustrated in Chart 2 below:5

<table>
<thead>
<tr>
<th>(α) underspecified</th>
<th>Ø</th>
<th>→ defective</th>
</tr>
</thead>
<tbody>
<tr>
<td>(β) strictly specified</td>
<td>+</td>
<td>→ complete</td>
</tr>
<tr>
<td>(γ) partially specified</td>
<td>−</td>
<td>→ defective</td>
</tr>
</tbody>
</table>

We demarcate the three way distinction for the value of feature specification in precisely the same way as the logical/mathematical theory of relations.6 Now we define the following: (α) A feature F within a lexical item is underspecified iff every morphologically possible variation of F in a language L allows arbitrarily free (or neutral) interpretation; otherwise, F is specified.7 (β) A specified F is strictly specified iff each one of F’s mor-

---

5 Throughout this paper, ‘+’ stands for a strictly specified feature, ‘−’ stands for a partially specified feature, and ‘Ø’ stands for an underspecified feature.

6 For example, the mathematical relation of reflexivity is defined ternarily: A relation R is either reflexive or non-reflexive, and a non-reflexive relation is irreflexive or not, as a consequence, R is necessarily either of reflexive, irreflexive, or non-reflexive, but not anything else (cf. Partee et al. (1990)).

7 The term ‘every morphologically possible variation’ does not mean every variation in all human languages (i.e. universally). For example, some languages, such as German, have a morphologically three-way variation for the gender-feature (i.e. male/female/neuter distinction), whereas English has a morphologically two-way variation for the gender-feature. Hence, for the purpose of discerning whether the gender-feature of the
phologically possible variations in $L$ allows only one specific interpretation; (γ) otherwise, $F$ is *partially specified*. With respect to the defectiveness of a feature, we presume that only a strictly specified feature is defined as *complete* and an underspecified/partially specified feature is defined as *defective*.8

More importantly, we would like to reinterpret the afore-defined defectiveness as identical to the featural defectiveness proposed in Burzio (1991); as a consequence, a lexical item counts as $\phi$-complete if it includes no defective formal feature within it; otherwise, it counts as $\phi$-defective in the sense of Burzio (1991). This is because Burzio’s (1991) $\phi$-completeness requires the featurally complete specification of a given lexical item.

Given these definitions, we will, in the following subsections, provide a piece of evidence to demonstrate how the feature specification of each $\phi$-feature within L1 Japanese/L1 English reflexives is identified through analyzing linguistic data.

### 4.2. Person-Features and Gender-Features

The reflexives in L1 English are specified with respect to their *person*-features, as shown in (9) below:

\[
(9) \quad \begin{align*}
\text{a. I criticized} & \text{ myself}/*\text{yourself}/*\text{himself}/*\text{herself}. \\
\text{b. You criticized} & \text{*myself}/*\text{yourself}/*\text{himself}/*\text{herself}. \\
\text{c. He (She) criticized} & \text{*myself}/*\text{yourself}/*\text{okhimself (herself)}. 
\end{align*}
\]

The ill-formed examples in (9) show that, in the grammar of L1 English, (i) the first-person, singular reflexive *myself* cannot be bound either by the second-person pronoun (i.e. *you*) or by the third-person pronouns (i.e. *he/she*); (ii) the second-person, singular reflexive *yourself* cannot be bound either by the first-person pronoun (i.e. *I*) or by the third-person pronouns; and (iii) the third-person, singular reflexives *himself/herself* cannot be bound either by the first-person pronoun or by the second-person pronoun. As can be seen from these facts, no arbitrarily free interpretation in terms of the *person*-feature is allowed for the L1 English reflexives. It follows that the *person*-feature of the L1 English reflexives is *not* underspecified. It is

---

8 In this paper we presume that underspecified/partially specified features are *defective* because their interpretation is imperfectly specified, whereas strictly specified features are *complete* because they have perfectly fixed interpretation.
obvious from the ill-formed examples in (9) that the first-person reflexives disallow the second-/third-person interpretation, the second-person reflexives disallow the first-/third-person interpretation, and the third-person reflexives disallow the first-/second-person interpretation; that is, their interpretation in *person* is unarily fixed as their morphological variation shows; as a consequence, all of the L1 English reflexives are strictly specified in terms of their *person*-features.

Moreover, the third-person singular reflexives in L1 English are specified with respect to their *gender*-features, as shown in (10) below:

(10) a. The man* criticized *himself/*herself.
    b. The woman criticized *himself/*herself.

The ill-formed examples in (10) show that, in the grammar of L1 English, (i) the third-person, singular, female reflexive *herself* cannot be bound by the third-person, singular, male antecedent; (ii) the third-person, singular, male reflexive *himself* cannot be bound by the third-person, singular, female antecedent. From these facts, it is concluded that no arbitrarily free interpretation in terms of the *gender*-feature is allowed for the L1 English third-person, singular reflexives. It follows that the *gender*-features of the L1 English third-person, singular reflexives are *not* underspecified. The ill-formed examples in (10) show that the third-person, singular, male reflexives only allow the male interpretation, and the third-person, singular, female reflexives only allow the female interpretation; that is, their interpretation in *gender* is unarily fixed as their morphological shapes demonstrate. Because the English reflexives have only two morphological variants in terms of the *gender*-features, which are manifested by *himself* and *herself*, the above fact that these third-person, singular reflexives in English allow a unarily fixed interpretation shows that the L1 English reflexives are strictly specified in terms of their *gender*-features. Thus, we conclude, from these observations, that the *person*-features and the *gender*-features within the L1 English reflexives are strictly specified; that is, they do not allow the *person*-neutral and *gender*-neutral interpretations.

On the other hand, arbitrarily free interpretation in *person* and *gender* is allowed for the Japanese reflexive *zibun(-zisin)*, as confirmed by the well-formedness of (1) in §3; that is, the *person*-feature and *gender*-feature within the L1 Japanese reflexive *zibun(-zisin)* are underspecified.

4.3. **Number-Features and Distributivity**

Now we will consider the specification of the *number*-feature in the L1 Japanese reflexives. Let us look at (11):
SYNTACTIC FEATURE TRANSFER AND REFLEXIVE BINDING IN INTERLANGUAGE

(11) John,-ga zibun(-*tati)/zibun(-*tati)-zisin-o seme-ta.
    -Nom SELF(-PL)/SELF(-PL)-self-Acc blame-Past
Lit. ‘John blamed John.’
The plural forms of the L1 Japanese reflexives (i.e. zibun-tati and zibun-tati-
-zisin) disallow the singular interpretation. This fact leads us to conclude
that the L1 Japanese reflexive zibun-tati(-zisin) is not underspecified in
terms of the number-feature; however, the singular form of the L1 Japanese
reflexive zibun(-zisin) allows the singular interpretation, as shown by the
well-formedness of (11) with the singular reflexive. Let us look at (6) (re-
peted as (12) below):

(12) John,-to Bill,-wa (sore-zore) zibun(-zisin)kΩ/*kare-zisinkΩ-o
and -Top (each) SELF(-self)/himself-Acc
hihansi-ta.
criticize-Past
Lit. ‘John criticized John, and Bill criticized Bill.’

It is natural to presume that, among the φ-features, the number-feature
plays an important role in determining the distributive reading of a reflex-
ive. There are two morphological variants of the Japanese reflexives in
terms of the number-feature (i.e. the singular form zibun(-zisin) and the plu-
ral form zibun-tati(-zisin)), but the singular form can be bound distributively
by the plural antecedents, here [John to Bill], as shown in (12). This fact
indicates that it is not the case that the singular reflexive form zibun(-zisin)
allows only one specific interpretation in number (i.e. the singular inter-
pretation).

According to Heim (2008), a bound pronominal/reflexive enters narrow
syntax without φ-features, and the morphophonological realization of its
φ-features is somehow materialized (at Spell-Out) by copying them from its
syntactic antecedent (namely, the DP that binds (i.e., c-commands and is co-
indexed with) it in narrow syntax).

The Japanese reflexive zibun(-zisin) allow the distributive reading even
when the overt distributive operator sore-zore ‘each’ is not contained in its
clause, as shown in (12) above. Given Heim’s (2008) proposal, the fact
observed in (12) indicates that zibun(-zisin) is not strictly specified in terms

9 It is true that (11) is acceptable when zibun is used logophorically; namely, the group
reading of zibun (i.e. ‘John blamed the group of [John and someone else].’) is available,
but this paper aims to investigate the L2 acquisition of syntactic binding and the logo-
phoric/emphatic use of a reflexive is neglected throughout this paper.
of its number-feature; for, the antecedent that c-commands it in (12) is the plural DP [John to Bill]-ga, so that it unavoidably inherits the plurality from its antecedent. Now that the singular form of zibun(-zisin) can be interpreted either as singular (as in (11)) or as plural (as in (12)), it follows that it allows arbitrary free interpretation; nevertheless, its plural form zibuntati(-zisin) disallows its singular interpretation, as shown in (11). The conclusion is that all of the morphological variations of the Japanese reflexive zibun(-zisin) disallow arbitrary free interpretation nor strictly fixed interpretation; whence, we arrive at the conclusion, on the basis of our three way definition, that the number-feature of zibun(-zisin) is partially specified.

In contrast, the English third-person, singular, masculine reflexive himself does not have the distributive reading even when the overt distributive operator each is involved, as exemplified in (7a) (repeated as (13) below):10

10 One might wonder whether the distributive operator has some influence on the contrast between (12) and (13); however, what is empirically significant here is the fact that the Japanese example allows the distributive reading without the distributive operator, whereas the English counterpart disallows it without the distributive operator. The conclusion we draw from these facts is that the existence of the distributive operator has nothing to do with the distributive reading for a reflexive, as we maintain in the text.

(13) *Johnk and Billj (each) criticized himselfk/j.

The ill-formedness of (13) indicates that (13) does not have the distributive reading; that is, it cannot be interpreted as “John criticized John, and Bill criticized Bill” even with the existence of the overt distributive operator (see Heim, Lasnik and May (1993) for syntactic/semantic discussion on the ill-formedness of (13)).11

Given Heim’s (2008) hypothesis, this fact leads us to conclude that the number-feature of the reflexives in L1 English is strictly specified; for, the

11 In fact, it seems that the following English sentence, entailing “John saw John, and Bill saw Bill,” can be said to have the distributive reading:

(i) [John and Bill] saw themselves in the mirror.

But, as Langendoen (1978) and Langendoen and Magloire (2003) argue, this distributive reading can be generated as a logical implication from the reading “[John and Bill] saw [John and Bill],” the reading which is brought out when the reflexive is bound by the conjoined plural DP [John and Bill] as a whole. In other words, the ostensibly distributive reading in (i) is attributed not to the defectiveness of the number-feature of the reflexive in (i), but to the reflexive’s plurality (see Langendoen and Magloire (2003)). It should be noted, here, that the Japanese sentence in (12), with which we claimed the truly distributive reading is concerned, does not have the plural reading (i.e., it does not mean “[John and Bill] criticized [John and Bill],”); consequently, the distributive reading in (12) comes not from the reflexive’s plurality, but from the defectiveness of the number-feature of the reflexive in (12).
antecedent that c-commands the reflexive in (13) is the plural DP [John and Bill]; resulting in the ill-formedness of the singular reflexive in (13). The conclusion is that the fact shown in (13) indicates that the L1 English reflexives are strictly specified in their number-feature.13

4.4. Summary of Feature Specification

Here, let us recapitulate our conclusion about the φ-feature specifications of the reflexives in L1 Japanese and in L1 English. From the discussion we made through the previous subsections, we have reached the conclusion that the φ-feature composition of a reflexive in L1 Japanese/L1 English is schematized in Chart 3 below:

<table>
<thead>
<tr>
<th>φ-feature composition</th>
<th>Japanese</th>
<th>English</th>
</tr>
</thead>
<tbody>
<tr>
<td>person</td>
<td>Ø</td>
<td>+</td>
</tr>
<tr>
<td>gender</td>
<td>Ø</td>
<td>+</td>
</tr>
<tr>
<td>number</td>
<td>−</td>
<td>+</td>
</tr>
</tbody>
</table>

We have observed that zibun(-zisin) in the grammar of L1 Japanese has the underspecified person- and gender-features and the partially specified number-feature, and all of the person-/gender-/number-features in the L1 English reflexives are strictly specified. The conclusion, thus, is that zibun(-zisin) in L1 Japanese is φ-defective and the L1 English reflexives are φ-complete in the sense of Burzio (1991).

A comment on the other reflexive form in Japanese is in order: What does the featural specification of the φ-features within PRONOUN+zisin look like in the grammar of L1 Japanese? Its person- and gender-features are strictly specified; for, the pronominal prefixed to zisin (e.g. watasi ‘I,’ anata ‘you,’ and kare ‘he’) has a fully specified featural morphology. With respect to its number-feature, PRONOUN+zisin (such as kare-zisin) is not tolerable with

12 Notice, here, that the distributive operator does not have any influence on the availability of the distributive reading, as stated in footnote 10.

13 An anonymous reviewer questioned whether the defectiveness of the number-feature makes any prediction regarding any other aspects of JLsE’s IL grammar, such as their difficulty in acquiring the distinction between countable and uncountable nouns as argued in Inagaki (2009). Indeed, it seems very likely that our conclusion has some relevance to the countability of noun phrases and pursuing this possibility seems to be highly promising, but it is far beyond the topic in this paper and we hope to investigate it in a future research.
the distributive reading, as the ill-formedness of (12) above shows. The conclusion we have reached is that PRONOUN+zin in L1 Japanese is strictly specified in terms of its number-feature; for, it disallows the distributive reading. From our observations that PRONOUN+zin in L1 Japanese is strictly specified with respect to its person-, gender-, and number-features, it follows that PRONOUN+zin in L1 Japanese is a φ-complete reflexive in the sense of Burzio (1991).

Now that we have revealed the φ-feature specification of zibun(-zin) in L1 Japanese and that of the L1 English reflexives. What about the specification of each φ-feature within the reflexive in JLsE’s IL grammar? In §5, we will try to answer this question by making a new proposal.

Before closing this section, we re-summarize our experimental tests and their results concerning the φ-specification of the reflexives in JLsE’s IL grammar. We contrived the experiments and obtained their results:

<Test 1> If JLE subjects regard the examples of the person-neutral interpretation (such as (3)) as acceptable, their IL reflexives are concluded to be underspecified in terms of the person-feature; otherwise, they are specified in terms of the person-feature. The result of Test 1 indicates that their IL reflexives allow only one specific interpretation (i.e., they allow a unarily fixed interpretation); therefrom, it is concluded that the person-feature of JLsE’s IL reflexives is strictly specified.

<Test 2> If JLE subjects can accept the test sentences (such as (4)), their IL reflexive is concluded to be underspecified in terms of the gender-feature; otherwise, it is specified in terms of the gender-feature. The result of Test 2 indicates that their IL reflexives allow only one specific interpretation (i.e., they allow a unarily fixed interpretation); therefrom, it is concluded that the gender-feature of JLsE’s IL reflexives is strictly specified.

<Test 3, Test 4> If JLE subjects regard the test sentences (such as (5)) as acceptable, their IL reflexives are concluded to be underspecified in terms of the number-feature because underspecification in terms of the number-feature leads to arbitrarily any interpretation in number. As shown in the result of Test 3, it follows that it is not the case that the number-feature of JLsE’s IL reflexive allows arbitrarily any interpretation in number; as a result, it is not underspecified in terms of the number-feature. Moreover, as shown in the result of Test 4, JLE subjects are likely to accept the distributive reading of the test sentences (such as (8)); as a result, the number-feature of their IL reflexive is concluded to be not strictly specified, as argued in §4.3. Taking into consideration this result together with the result reported in Table 3 (which we have interpreted as indicating that the number-feature of JLsE’s
IL reflexives is not underspecified), we have now arrived at the conclusion that the number-feature within JLsE’s IL reflexives is partially specified.

We have therefore revealed the φ-feature specification of JLsE’s IL reflexives as follows: They have the strictly specified person-feature (as shown in Table 1), the strictly specified gender-feature (as shown in Table 2), and the partially specified number-feature (as shown in Table 3 and Table 4). From the discussion we made through this section, we now would like to draw the conclusion that the φ-feature composition of JLsE’s IL reflexives is schematized as in Chart 4 below:

<table>
<thead>
<tr>
<th>φ-feature composition</th>
<th>Interlanguage</th>
</tr>
</thead>
<tbody>
<tr>
<td>person</td>
<td>+</td>
</tr>
<tr>
<td>gender</td>
<td>+</td>
</tr>
<tr>
<td>number</td>
<td>−</td>
</tr>
</tbody>
</table>

Given our proposal (made in §4.1 above) that our definition of defectiveness should be identified with the featural defectiveness in the sense of Burzio (1991), it follows that JLsE’s IL reflexives turn out to be φ-defective; for, their φ-features include a defective feature (i.e., their number-feature turns out, through our experiments, to be defective). Following, essentially, Burzio’s (1991) insight, we assume that a lexical item counts as φ-defective unless all of its φ-features are featurally complete.

In what follows, we will demonstrate that our experimental data are all coherently explainable under our theory of the feature (de)composition of reflexives. In the next section, we will try to give a theoretically adequate account to all of the abovementioned facts concerning the specification of each φ-feature within JLsE’s IL reflexives.

5. Proposal: Feature Transfer and Feature Learning Hypothesis (FTFL)

What we would like to propose here is a hypothesis (what we will call FTFL), which is so stipulated as to explain which of the formal features within a given lexical item in learners’ L1 should be transferred into the corresponding lexical item in their IL grammar.

The rationale behind this hypothesis comes from the conception about the learnability of formal features with respect to their φ-feature specification: It is very plausible that strictly or partially specified features are regarded as unmarked (see White (1986) and Gair (1988) for ‘markedness’ in SLA), and
underspecified features are regarded as marked. Of particular significance in our proposal is that we do not adopt the subset relation to define the (un)markedness of features; that is, the subset relation of the feature specification is not concerned in FTFL. Instead, we naturally assume that learners start to assume an unmarked value in the course of their language acquisition (cf. Chomsky (1981) and Gair (1988)). To put it differently, unmarked features, unless provided sufficient positive evidence, are harder for learners to overwrite once they have been learned in the course of their L1 acquisition; as a result, we demand that an L1 feature, if it has an unmarked value, should be transferred to IL grammar, and that marked features are easier to discard due to their perceptible transformation from the pristine state (unmarkedness), and very easy to reconvert to their unmarked state. Thus, we propose to hypothesize that FTFL requires that a given feature of an L1 lexical item is to be transferred to IL grammar (FEATURE TRANSFER) if it is strictly or partially specified in the L1 grammar (i.e., if it is unmarked) regardless of whether the L2 corresponding feature of the lexical item in learners’ target language is strictly or partially specified or underspecified in the grammar of the target language, and a given feature of an L1 lexical item, if it is underspecified (i.e., if it is marked), is to be discarded and the corresponding feature in IL grammar is to be learned from the grammar of the target language (FEATURE LEARNING) regardless of whether the L2 corresponding feature of the lexical item in the target language is strictly or partially specified or underspecified in the grammar of the target language.

(14) illustrates the core idea of our hypothesis:

(14) **Feature Transfer and Feature Learning Hypothesis**

<table>
<thead>
<tr>
<th></th>
<th>Native language</th>
<th>IL</th>
<th>Target language</th>
</tr>
</thead>
<tbody>
<tr>
<td>(i) <strong>FEATURE TRANSFER</strong></td>
<td>− − − − − − − −</td>
<td>−</td>
<td>+/Ø</td>
</tr>
<tr>
<td>(ii) <strong>FEATURE LEARNING</strong></td>
<td>Ø</td>
<td>+/-</td>
<td>+/-</td>
</tr>
</tbody>
</table>

Under FTFL, we specifically predict the following: (i) When a given feature of an L1 item is partially specified, the corresponding feature of the IL item is set as partially specified, no matter how the corresponding feature of the L2 item may be specified (FEATURE TRANSFER); and (ii) When a given feature of an L1 item is underspecified, the corresponding feature of the IL item is set as the same specification as the corresponding feature of the L2
Returning to our main topic in this paper, let us consider, under FTFL, how the feature transfer/feature learning takes place in the case of JLsE’s acquisition of reflexive binding. Take a look at (15) below:

(15) FTFL in the L2 acquisition of the English reflexives by JLsE

<table>
<thead>
<tr>
<th>language</th>
<th>Native language</th>
<th>Interlanguage</th>
<th>Target language</th>
</tr>
</thead>
<tbody>
<tr>
<td>lexical item</td>
<td>zibun(-zisin)</td>
<td>himself</td>
<td>himself</td>
</tr>
<tr>
<td>φ-feature defectiveness</td>
<td>φ-defective</td>
<td>φ-defective</td>
<td>φ-complete</td>
</tr>
<tr>
<td>φ-feature composition</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>person</td>
<td>Ø</td>
<td>+</td>
<td>←</td>
</tr>
<tr>
<td>gender</td>
<td>Ø</td>
<td>+</td>
<td>←</td>
</tr>
<tr>
<td>number</td>
<td>−</td>
<td>−</td>
<td>−</td>
</tr>
</tbody>
</table>

Notes. *FEATURE TRANSFER: (→→) FEATURE LEARNING: (←←)

First, it is evident, as Burzio (1991) argues, that the reflexives in the grammar of L1 English are φ-complete: As we have observed in §3, the native speakers of English cannot interpret the third-person, masculine reflexive (i.e. *himself*) as first-/second-person, feminine, or plural, in any context. Consequently, all of its φ-features are concluded to be strictly specified; whereupon, all of its φ-features in (15) are marked as “+.” Similarly, we also observed in §3 that the person- and gender-features of zibun(-zisin) are underspecified and its number-feature is partially specified; whence the former are marked as “Ø” and the latter as “−” in (15) above.

Now, let us make a provisional stipulation that the feature composition of zibun(-zisin) is selected as the target of FTFL when JLsE are learning the usage of the English reflexives. Given this stipulation, we can infer that FTFL provides that JLsE’s IL reflexives should have the strictly specified person- and gender-features through FEATURE LEARNING and have the

---

14 Needless to say, there are other possible options for the pair of the feature specification in L1/L2, the study of which we wish to pursue in a future research (see Ishino (2012) for some relevant discussion).

15 It is true that this stipulation critically degrades our theory on conceptual grounds (just like LTA is accused of being degraded for the same reason, as we argued in section 2), but we will soon deduce this stipulation from a more general principle in the UG-based theory of language acquisition.
partially specified number-feature through Feature Transfer, because the former types of features are underspecified (i.e. marked) in L1 Japanese but strictly specified (i.e. unmarked) in L1 English whereas the latter is partially specified (i.e. unmarked) in L1 Japanese but strictly specified (i.e. unmarked) in L1 English. Put differently, FTFL leads us to predict that JLsE learn, from English, the strictly specified person- and gender-features for JLsE’s IL reflexives and they transfer the partially specified number-feature of their L1 reflexives onto their IL reflexives. As the result of these Transfer/Learning, we conclude that the φ-feature composition of JLsE’s IL reflexives is depicted as in (15) above; namely, JLsE’s IL reflexives have the strictly specified person- and gender-features and the partially specified number-feature, and they are φ-defective in the sense of Burzio (1991) due to their partially specified number-feature. Given this conclusion, we can naturally explain our experimental observation (which was reported in §3) that many JLsE are apt to disallow the person- and gender-neutral interpretation for their IL reflexives, but they allow the distributive reading.

The remaining problem is that we have to deduce our provisional stipulation that zibun(-zisin), instead of PRONOUN+zisin (such as kare-zisin), is regarded as the default reflexive form for their L1 reflexives and is selected as the target of FTFL when JLsE are learning the usage of the English reflexives: If we fail to derive this stipulation from a more general principle, FTFL is not free from a conceptually serious deficit, which is very similar to the one that we claimed in §2.2 to be immanent in LTA. It has occasionally been claimed (cf. Bennett and Progovac (1998)) that the L1 acquisition of morphologically complex reflexives with LC-property is more difficult for children than the L1 acquisition of morphologically simplex reflexives with the LD-property. Given this observation that the LD-property manifests itself in an earlier stage of L1 acquisition, we would like to speculate that zibun, the morphologically simplex reflexive with the LD-property, is acquired earlier than PRONOUN+zisin, the speculation which

16 “LC,” which means “local,” is used herein to refer to the binding dependency between a reflexive and its antecedent, both of which are within a single clause, irrespective of whether the clause is tensed or non-tensed.

17 “LD,” which means “long-distance,” is used herein to refer to the binding dependency between a reflexive and its antecedent astride a clause boundary. Notice that the nature of syntactic binding demands that, if an anaphor shows LD, it also shows LC (cf. Wexler and Manzini’s (1987) idea about Governing Category Parameter). See footnote 1 above.
conforms to the major consensus among some L2 researchers (Yusa (1998), Shirahata (2007), among others), who claim that PRONOUN+zisin is the least common (i.e. marked) and rarely used among the three reflexive forms in Japanese. This leads us to the plausible hypothesis that the feature composition of zibun(-zisin), which is the default element (i.e. unmarked), is recognized by JLsE as the target of FTFL when JLsE are learning the usage of the English reflexives. Given this hypothesis, we can naturally deduce the theoretical rationale as to why zibun(-zisin) is selected as the target of FTFL when JLsE are learning the usage of the English reflexives.

In this section, we have explored FTFL under the more general theory of L2 acquisition through considering how our feature-(de)composition analysis of reflexives can be implemented under FTFL, and demonstrated that FTFL is an adequate theory for JLsE’s acquisition of reflexive binding on conceptual/empirical grounds. In order to show that the theory proposed in this paper is theoretically more adequate than the previous approaches to the SLA of reflexives, we will, in the next section, explore some implication of our theory within the UG-based approach to SLA.

6. Syntactic Binding of Φ-defective Reflexives within the Agree Theory

Our conclusion that JLsE’s IL reflexives are φ-defective brings theoretically important consequences, which we will discuss in what follows. From this conclusion, what else is predictable concerning the syntactic behaviors of JLsE’s IL reflexives? Under the recent minimalist syntax, not a few researchers (cf. Reuland (2005, 2011), Heinat (2008), Quicolli (2008), Gallego (2010), and Ishino and Ura (2011, 2012), to mention only a few) have proposed that a φ-defective reflexive should have its φ-features checked in order for it to be properly interpreted at LF (see Bouchard (1984) and Burzio (1991) for a pre-minimalist idea behind this proposal).

The consequence of this proposal is that the syntactic binding of (φ-defective) reflexives should be recast within the Agree theory under the current minimalist Probe-Goal framework (Chomsky (2004) and subsequent work): A φ-defective reflexive must have its φ-features valued by a Probe with the whole φ-feature amalgam in order to be properly interpreted at LF (the idea which we hereafter call Agree Theory of Reflexive Binding (ATRB) (cf. Uriagereka and Gallego (2006) and Gallego (2010))). Under ATRB, it is assumed that α binds β if they are both Goals of a single relevant Probe; otherwise, α and β are obviative. Accordingly, the binding relation between a reflexive and a DP (i.e. its antecedent) is materialized not through c-com-
mand plus referential co-indexing, as the traditional binding theory assumes, but through *Agree*. To be concrete, when a single Probe with the whole $\varphi$-feature amalgam agrees with two Goals in terms of the $\varphi$-feature checking, such as a DP and a $\varphi$-defective reflexive, the binding relation between the DP and the reflexive is successfully materialized. (See Gallego (2010) for further applications of ATRB to empirically broad data.)

### 6.1. Subject Orientation

As a concrete instantiation of ATRB, take a brief look at the subject orientation of a reflexive, which is found in not a few languages in the world. It is a well-known fact that the Japanese reflexives *zibun* and *zibun-zisin* have subject orientation (cf. Kuroda (1965) and Aikawa (1999)). Given Ishino and Ura’s (2011, 2012) assumption that only T with the whole $\varphi$-feature amalgam can serve as a Probe for $\varphi$-defective reflexives (cf. also, Uriagereka and Gallego (2006)), ATRB leads to the prediction that a reflexive $R$ shows subject orientation if $R$ is $\varphi$-defective. This is because T always bears a probe-goal relation with the subject DP in a clause to value the nominative Case of the DP; as a consequence, both the subject DP and a $\varphi$-defective reflexive are always mediated by T with $\varphi$-complete specifications through *Agree*, resulting in the binding relation between the subject DP and the $\varphi$-defective reflexive; whence, the subject orientation of the $\varphi$-defective reflexive follows.$^{18}$

Interestingly, PRONOUN+$zisin$ in Japanese does not have subject orientation whereas *zibun* and *zibun-zisin* show subject orientation (see Nakamura (1987) and Aikawa (1999) for further discussion). As observed elsewhere in this paper, PRONOUN+$zisin$ is a $\varphi$-complete reflexive and, hence, it needs no agreement with T, but *zibun* and *zibun-zisin* are both $\varphi$-defective. Therefore, we can predict under ATRB that *zibun* and *zibun-zisin*, but not PRONOUN+$zisin$, show subject orientation, and this prediction is borne out empirically.

More importantly, our conclusion that the English reflexives in JLsE’s IL grammar are $\varphi$-defective leads to the prediction that they show subject orientation despite the fact that the reflexives in L1 English do not show subject orientation. By conducting an experimental survey, we have ascertained that this prediction is indeed borne out. In order to examine whether

---

$^{18}$ We omit going into any detail about ATRB, for which the reader is referred to the extensive discussion presented in Gallego (2010).
JLsE’s IL reflexives show subject orientation or not, we have utilized the English test sentences in (16) and (17) (see Appendix Test 5 for all the test sentences):

(16)  The boy told these girls about himself.
(17)  These girls told the boy about himself.

If an experimental subject cannot accept the interpretation under which the reflexive in (17) is coreferential grammatically with the object, he/she is alleged to regard their IL reflexives as showing the subject orientation. The results of the above tests are shown in Table 5:

Table 5. Acceptability of Object Antecedents

<table>
<thead>
<tr>
<th>(%)</th>
<th>JLsE^a</th>
<th>Control^b</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>SUB</td>
<td>OBJ</td>
</tr>
<tr>
<td>acceptable</td>
<td>91.8</td>
<td>22.2</td>
</tr>
<tr>
<td>unacceptable</td>
<td>8.2</td>
<td>77.8</td>
</tr>
</tbody>
</table>

^a n=369.  ^b n=20. Notes. There are statistically significant differences: between the reflexive which is bound by the subject (as shown in (16)) for the JLE group and the reflexive which is bound by the object (as shown in (17)) for the JLE group (t=−19.6, p<.01); between the reflexive which is bound by the object (as shown in (17)) for the JLE group and that for the control group (t=−4.57, p<.01.)

It should be noted that 77.8% of the total JLE subjects mistakenly interpreted himself as being unable to be bound by the object; that is, the English reflexives in their IL grammar show the subject orientation. This is exactly what FTFL predicts, and hence it provides strong support for FTFL.

6.2. Binding Dependency

In the grammar of L1 English, a reflexive cannot be bound astride a clause boundary, irrespective of whether the clause is tensed or non-tensed, as exemplified in (18) below:

(18)  a. *John said [that the students praised himself].
      b. *John made [the students praise himself].

Let us try to explain the binding dependency of a φ-defective reflexive with ATRB. Take a look at the Japanese examples in (19) below:

      mi-ta to]  omot-ta.
      see-Past Comp think-Past
      Lit. ‘Taro thought that Jiro saw SELF-self/himself.’
b. Taro-wa [Jiro-ki/ni zibun-zisin-ki/kare-zisin*-ki-ku-o
-Top -Dat SELF-self/him-self-Acc
mi]-sase]-ta.
see-Cause-Past
Lit. ‘Taro made Jiro see SELF-self/himself.’

The Japanese reflexive form *zibun-zisin*, when embedded within a tensed clause, shows LC (as the ill-formedness of (19a) indicates), but it shows LD when it is embedded within a non-tensed clause (as the well-formedness of (19b) indicates): This observation was first reported in Kuroda (1965) but has long been neglected since then. In contrast, the Japanese reflexive form *PRONOUN+zisin* (such as *kare-zisin*) cannot be bound over a clause boundary, irrespective of whether it is tensed or non-tensed (namely, it shows LC everywhere), as shown by the ill-formedness of both (19a) and (19b).

Given that *zibun-zisin* is φ-defective but *PRONOUN+zisin* is φ-complete, the facts shown in (19) can be explained as in the following way: Following, basically, Kitagawa (1986), the Japanese causative clause, which is non-tensed, takes two types of clausal complement: One corresponds to vP without T, and the other corresponds to TP[−tense] with the whole φ-feature amalgam. When the causative clause is vP, the reflexive within the causative clause must have its φ-features checked by T[+tense] in the superordinate/matrix tensed clause. As required, this results in the binding relation between the φ-defective reflexive in the embedded non-tensed causative clause and the subject DP in the superordinate tensed clause, which has had its Case valued by T[+tense] in the tensed clause. When, on the other

---

19 It has widely been alleged in the literature on Japanese syntax that *zibun-zisin*, one of the morphologically complex reflexives in Japanese, shows the same binding dependency as *PRONOUN+zisin* does; namely, it shows LC in every context (cf. Nakamura (1987) and Katada (1991)), though Kitagawa (1986) and Manning et al. (1999) endorsed Kuroda’s (1965) observation. Through polling more than 200 native Japanese speakers, including more than 20 linguistic professionals, Ishino and Ura (2009) revealed (i) that more than 90% of them judge that (19a) is unacceptable, and (ii) that 75% of those who judge (19a) as unacceptable judge that *zibun-zisin* in (19b) is far better than (19a) though *kare-zisin* in (19b) is as bad as (19a). This indicates that Kuroda’s (1965) observation that *zibun-zisin* shows the LD-property in a non-tensed clause and shows the LC-property in a tensed clause is correct. For much extensive discussion, see Ishino and Ura (2009, 2012).

20 As we argued above, it is because φ-defective reflexives must be interpreted properly at LF that they must be provided with the full specification of φ-features through Agree by T. Here, we assume that Agree between a φ-defective reflexive and T is executed at
hand, the causative clause has $T[−\text{tense}]$ with the whole $\phi$-feature amalgam, the reflexive within the causative clause has its $\phi$-features checked by the $T[−\text{tense}]$ within the causative clause, because the $T[−\text{tense}]$ within the causative clause is the nearest to the reflexive among the possible Probes for the $\phi$-defective reflexive. This results in the binding relation between the $\phi$-defective reflexive and the causee DP within the non-tensed causative clause; for, the causee DP has had its Case valued by $T[−\text{tense}]$ within the non-tensed causative clause because the $T[−\text{tense}]$ within the non-tensed causative clause has the whole $\phi$-feature amalgam and acts as the Probe for the causee’s Case-valuation in the causative clause. We can, thus, explicate the ambiguity concerning the binder of $\text{zibun-zisin}$ in (19b); that is, we can explain why $\text{zibun-zisin}$ in (19b) can be bound either by the subject DP in the superordinate tensed clause or by the causee DP within the non-tensed causative clause. This explains the LD-property of $\text{zibun-zisin}$ in a non-tensed clause.

On the other hand, when a $\phi$-defective reflexive is embedded within a tensed clause, the reflexive always agrees with $T[+\text{tense}]$ in the embedded tensed clause that includes the reflexive; for, $T[+\text{tense}]$, which has the whole $\phi$-feature amalgam, is the nearest to the $\phi$-defective reflexive among any other possible $T[+\text{tense}]$ within the whole sentence. This explains the LC-property of $\text{zibun-zisin}$ in a tensed clause. Furthermore, we assume that $\text{zibun}$, another $\phi$-defective reflexive in L1 Japanese, can move at LF due to its morphological simplicity (see Katada (1991) for much discussion on this idea), and it can undergo a long-distance LF movement over a tensed clause boundary. Therefore, $\text{zibun}$ in the embedded tensed clause can agree with the subject DP in the matrix clause by the mediation of the matrix T, resulting in the LD-property of $\text{zibun}$ even in a tensed clause. On the other hand, $\text{zibun-zisin}$, due to its morphological complexity, cannot move at LF beyond a tensed clause boundary (see, again Katada (1991) for discussion), resulting in its LC-property in a tensed clause, as required.

As for PRONOUN+$\text{zisin}$, which we already concluded to be $\phi$-complete, it does not need to agree with T under ATRB. Following Aikawa (1993), we propose to assume that PRONOUN+$\text{zisin}$ is a reflexivizer à la Reinhart and Reuland (1993) and the binding relation between PRONOUN+$\text{zisin}$ and its antecedent is materialized through co-argumenthood. Given this, the

LF, where Spell-Out has already been executed. Thus, a phase, which is the domain of Spell-Out, is irrelevant to Agree between a $\phi$-defective reflexive and T.
LC-property of PRONOUN+zisin in any context follows naturally, just like the LC-property of the English reflexives in the grammar of L1 English (see Reinhart and Reuland (1993)).\textsuperscript{21}

Given our conclusion that JLsE’s IL reflexives are φ-defective and morphologically complex, ATRB leads us to predict that its binding dependency is LC in a tensed clause and LD in a non-tensed clause. In what follows, we will show that this prediction is, indeed, borne out and pertinent to experimental results.

For the purpose of discerning the precise binding domain of JLsE’s IL reflexives, the subjects in our experiments were asked whether various types of LD-bound reflexives are acceptable or not. We have utilized test sentences in which an English reflexive is contained either within an embedded tensed clause (as in (20)), or within a causative clause (as in (21)) (see Appendix Test 6 for all the test sentences):

\begin{enumerate}
\item[(20)] *The boy believes [that the girl loves \textit{himself}].
\item[(21)] *The boy made [the girl understand \textit{himself}].
\end{enumerate}

Note that these English test sentences are ill-formed in the grammar of L1 English, because there is no proper antecedent for each of the reflexives within the embedded clause, irrespective of whether the clause is tensed or non-tensed. If an experimental subject judges (20) as acceptable, he/she is alleged to interpret the reflexive as being bound astride a tensed clause boundary; namely, he/she regards a reflexive within a tensed clause as showing LD. If a subject judges (20) as unacceptable, he/she is alleged to interpret the reflexive as not being bound astride a tensed clause boundary; namely, he/she regards a reflexive within a tensed clause as showing LC. Moreover, we examined the binding dependency of JLsE’s IL reflexives within a non-tensed clause: Those who judge (21) as acceptable regard a reflexive within a non-tensed clause as showing LD. The results of our experimental tests are shown in Table 6:

\textsuperscript{21} We purposely neglect the logophoric use of reflexives throughout this paper.
Table 6. Locality for the Binding Dependency

<table>
<thead>
<tr>
<th></th>
<th>JLsE&lt;sup&gt;a&lt;/sup&gt;</th>
<th>Control&lt;sup&gt;b&lt;/sup&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Tensed</td>
<td>Non-tensed</td>
</tr>
<tr>
<td>LC</td>
<td>64.6</td>
<td>24.9</td>
</tr>
<tr>
<td>LD</td>
<td>35.4</td>
<td>75.1</td>
</tr>
</tbody>
</table>

<sup>a</sup><sub>n=288</sub>. <sup>b</sup><sub>n=20</sub>. Notes. There are statistically significant differences: between the reflexive which is contained in the tensed clause (as shown in (20)) for the JLE group and the reflexive which is contained in the non-tensed clause (as shown in (21)) for the JLE group (<sub>t=7.01, p<.01</sub>); between the reflexive which is contained in the non-tensed clause for the JLE group and that for the control group (<sub>t=12.71, p<.01</sub>.)

As the bold numbers (i.e. 64.6% and 75.1%) in Table 6 show, the result of our experimental survey has reconfirmed the well-known observation (see, among many others, Hirakawa (1990), and Watanabe et al. (2008)) that many JLsE in their IL grammar regard an English reflexive within a tensed clause as showing LC and one within a non-tensed clause as showing LD.

Of great importance is the fact that these experimental results (as well as the results of many former experiments in the same vein) are what we predict with FTFL. This, in turn, lends strong support, on both empirical and theoretical grounds, to our claim that FTFL is a more adequate theory for the L2 acquisition of reflexive binding.

6.3. Summary

We have explained our empirical/experimental data about the syntactic properties of reflexive binding in the IL grammar of JLsE: Given that JLsE’s IL reflexives are φ-defective, FTFL, if reinforced with ATRB, enables us to provide a coherent account to the following observations obtained from our experimental survey: (i) JLsE’s IL reflexives show subject orientation; and (ii) JLsE’s IL reflexives show LC in a tensed clause and LD in a non-tensed clause.

7. Concluding Remarks

In this paper, we argued that our newly introduced hypothesis about L2 acquisition (i.e. FTFL) is not only pertinent to the empirical/experimental data, which were reported in the previous literature or obtained through our own experiments, but also able to give a theoretically consistent account for all of the syntactic properties of JLsE’s IL reflexives. In some recent L2 studies, the importance of feature-based theory has been recognized
(cf. Liceras, Zobl, and Goodluck (2008), and García Mayo and Hawkins (2009)). Under FTFL, the specification of each formal feature within a lexical item is crucial in determining the syntactic mechanism for the L2 acquisition of the relevant item. Hence, it is very important and highly interesting to contemplate FTFL’s relevancy to those feature-based L2 studies; however, this paper will not go into this issue any further due to space limitations (see Ishino (2012) for extensive discussion in relevance).

Before closing this paper, we have to comment on the possible stages in IL grammar. In this paper we have implicitly distinguished three stages/levels in IL grammar, and this distinction is, indeed, demanded theoretically by FTFL: (i) For beginners, FTFL demands that only feature transfer should take place; that is, every feature specification in their L1 reflexive should be transferred to their IL reflexives. As a result, for example, beginner JLsE’s IL reflexives should have the same feature specification as L1 zibun-zisin; (ii) For intermediate to advanced learners, FTFL demands that feature learning should take place; that is, the transferred L1 feature should be overwritten by the corresponding L2 feature, when the L1 feature is different from the one in the target language in terms of the markedness of their feature specification. As a result, for example, the person/gender-features of IL reflexives for intermediate to advanced JLsE are overwritten and change to strictly specified; (iii) For even more advanced learners, FTFL demands that no feature learning should take place, when both L1 and L2 features are the same in terms of the markedness; that is, the transferred L1 feature should not be overwritten by the corresponding L2 feature. As a result, for example, the partially specified number-feature in IL reflexives is to be retained for even more advanced JLsE.

In our future research, we wish to confirm that FTFL is further expandable empirically: It is expected to be applicable to the L2 acquisition of reflexive binding in other languages as well: This paper deals with the L2 acquisition of reflexive binding in the case where an L1 item includes underspecified features and a partially specified feature. We need to proceed to a next step to make a more extensive research on the L2 acquisition where L1 and/or L2 items have other possible feature specification for the purpose of investigating whether FTFL can give a theoretically consistent account to these cases. We will, however, leave it to future research to pursue such investigations.
Appendix

The following sentences were presented to the experimental subjects in our surveys. The asterisks, indices and brackets, which were not shown to the subjects, were attached herein for the purpose of providing the grammaticality of each sentence to the reader. In our experimental surveys, these test sentences were randomly ordered.

Test 1. *Person*-Neutral Interpretation:
1. *Bill and I hit yourselves.
2. *Bill likes yourself.
3. I believe that Mary likes herself.
5. *That boy and this girl like ourselves.
6. *You and I hit themselves.
7. *I hit itself, but I did not know what it was.
8. John saw himself in the mirror.

Test 2. *Gender*-Neutral Interpretation:
11. Bill praised himself at the concert.
12. Mary introduced herself in the class.

Test 3. *Number*-Neutral Interpretation:
15. *We like myself.
16. *I do not like ourselves.
17. *Bill and Jane hit himself.
18. The robot fixed itself.

Test 4. Distributive Reading:
19. *Mary_k and Sue_j saw herself_k@j in the mirror.
20. *Bill_k and John_j criticized himself_k@j.
21. *Jane_k and Mary_j blamed herself_k@j.
22. *This boy_k and that boy_j hit himself_k@j.
23. *Bill_k and Tom_j praised himself_k@j in the TV show.

Test 5. Subject Orientation:
24. The boy told these girls about himself.
25. These girls told the boy about himself.
26. Bill did not tell Mary about herself.
27. Mary told Bill about himself.
28. You told me about myself.

Test 6. Locality on Binding Dependency:
29. *The boy believes [that the girl loves himself].
30. *The boy made [the girl understand himself].
31. *John said [that the students praised himself].
32. John made [the students praise themselves].
33. *Bill believes [her to love himself].

Partial Control Sentences:
34. John and Mary saw each other. 35. Mary loves her mother.
36. We believe that Mary loves Bill.
37. He is always criticizing others.

REFERENCES


García Mayo, María del Pilar and Roger Hawkins (2009) Second Language Ac-
35

SYNTACTIC FEATURE TRANSFER AND REFLEXIVE BINDING IN INTERLANGUAGE


Uriagereka, Juan and Ángel J. Gallego (2006) “(Multiple) Agree as Local (Binding and) Obviation,” paper presented at Going Romance XX, Vrije Universiteit Amsterdam.


Martohardjono and Wayne O’Neil, 215–238, Lawrence Erlbaum, Mahwah, NJ.

[received April 30 2011, revised and accepted December 8 2011]

Graduate School of Humanities
Kwansei Gakuin University
1–1–155, Uegahara, Nishinomiya
Hyogo 662–8501
e-mail: nao.ishino@kwansei.ac.jp