Reflexive Pronouns in Second Language Acquisition

William O’Grady
University of Hawai‘i at Manoa

1. Introduction

The study of second language acquisition within formal linguistics has long been based on the assumption that learners acquire a grammar consisting of abstract parameterized principles that apply to intricate tree-like syntactic representations. I call this assumption into question here by exploring an alternative processing-based approach to language learning.

The key idea that I wish to put forward is that the successes and failures that characterize the various stages of first and second language acquisition are best explained in processing terms—an idea that I will explore here by focusing on the interpretation of reflexive pronouns. I will begin in the next section by illustrating how processing works in the sort of theory that I have in mind and how this approach might contribute to an account of pronoun interpretation in English and Japanese (section 3). I’ll then turn my attention to how the analysis I propose can shed light on the emergence of interpretive strategies in the course of first and second language acquisition (sections 4 and 5, respectively). The paper ends with a brief set of concluding remarks.

2. A proposal for processing

As in all work on processing, I assume that the processor operates in an incremental manner—interpreting each word as it is encountered and immediately integrating it into the sentence’s emerging semantic representation. In the system I propose, sentences are processed in a series of steps, each involving the application of an appropriate routine, or automatized procedure. As illustrated in the sample transitive sentence below, some routines look up the referent of a proper name, others call up the predicate associated with a particular lexical item, and still others associate a referent with a position in a predicate’s argument grid. Procedures of this sort are posited in one form or another in all approaches to processing, and there is no controversy over their existence.

* I am grateful to the organizers of the 2012 J-SLA conference for inviting me to give the keynote address and to Takaaki Suzuki, Naoko Yoshinaga, Kevin Gregg, and various members of the audience for their helpful questions and comments.
(1) Richard trusts Mary.

a. The nominal Richard is assigned an interpretation, represented here as the index \( r \).

\[
\text{Richard} \quad r
\]

Based on its position, the referent of Richard is then identified as likely first argument of a yet-to-be-determined predicate.

\[
\text{PRED} \quad <r \ldots >
\]

The referent of Richard corresponds to first argument of an anticipated predicate.

b. The transitive predicate trust is encountered, and its argument structure is integrated into the semantic representation.

\[
\text{Richard trusts} \quad \text{TRUST} \quad <r \_>
\]

c. The nominal Mary is assigned a referent (represented by the index \( m \)), and is immediately interpreted as the predicate’s second argument.

\[
\text{Richard trusts Mary} \quad \text{TRUST} \quad <r m>
\]

The referent of Mary corresponds to the verb’s second argument.

Things work essentially the same way for Japanese, as the example in (2) helps illustrate.

(2) Richard-ga Mary-o mi-ta
Richard-Nom Mary-Acc see-Pst
Richard saw Mary.

a. The nominal Richard is assigned an interpretation.

\[
\text{Richard-ga} \quad r
\]

Based on the nominal’s case, its referent is identified as likely first argument of a yet-to-be-determined predicate.

\[
\text{PRED} \quad <r \ldots >
\]

The referent of Richard corresponds to first argument of an anticipated predicate.
b. The nominal *Mary* is assigned an interpretation.

\[ \text{Mary-o} \]
\[ m \]

Based on the nominal’s case, its referent is identified as likely second argument of the anticipated predicate.

\[ \text{PRED} \]
\[ \langle r \ m \ldots \ \uparrow \rangle \]

The referent of *Mary* corresponds to second argument of the anticipated predicate.

c. The transitive predicate *see* is encountered, and its argument structure is integrated into the semantic representation.

\[ \text{Richard-ga Mary-o mi-ta} \]
\[ \text{SEE} \]
\[ \langle r \ m \rangle \]

In sum, as these examples from English and Japanese illustrate, the processor operates on words of a particular form arranged in a particular way, converting them into a simple semantic representation. The challenge, of course, is to show that such a system can be ‘scaled up’ to deal with phenomena involving a higher order of intricacy. The interpretation of reflexive pronouns is a case in point.

### 3. Reflexive Pronouns

Unlike names, reflexive pronouns are not capable of direct reference. I will represent this property by assigning them the index *x*.

(3)  
\[ \text{herself}_x, \text{himself}_x \]

Informally speaking, we can say that reflexive pronouns introduce a ‘referential dependency,’ which the processor must resolve by finding an appropriate referent.

(4)  
\[ \text{Richard trusts himself}. \]
\[ \text{TRUST} \]
\[ \langle r \ x \rangle \]
\[ \leftarrow ? \]

Two processing-related forces shape this search:

(i) pressure to find the referent as quickly and as locally as possible, so as to reduce the burden on working memory

(ii) sensitivity to the topicality of potential referents and to their contextual appropriateness, so as to maximize coherence
Each of these pressures is associated with a different type of processing.

On the one hand, the pressure for immediate action is typical of **sentence-level (or structural) processing**, which focuses on the lexical properties of a sentence’s component words, their morphological form, and their linear relationship to each other. A classic example of sentence-level processing is the use of case and/or word order to identify the proper noun in *Jerry bought a book* as the verb’s first argument (its ‘subject’).

On the other hand, **pragmatic processing** focuses on what is being talked about, who is talking to whom, the setting in which an utterance is produced, and so forth, thereby connecting the sentence to the larger communicative context. A prototypical example of pragmatic processing involves the interpretation of definite pronouns (*he, she, him, her*), which typically refer to a recently mentioned individual who is prominent in the consciousness of the speaker and addressee (e.g., Foraker & McElree 2007, Song & Fisher 2007).

<table>
<thead>
<tr>
<th>Sentence-level processing</th>
<th>Pragmatic processing</th>
</tr>
</thead>
<tbody>
<tr>
<td>focuses on the lexical properties, form &amp; arrangement of words</td>
<td>focuses on what is being talked about, who is talking to whom, the setting, etc.</td>
</tr>
</tbody>
</table>

Let us consider how these two types of processing influence the interpretation of reflexive pronouns in English and Japanese.

Responsibility for interpreting reflexive pronouns in English falls largely to the sentence-level processor, which seeks an immediate and local opportunity to resolve the referential dependency, as illustrated in (5).¹

(5)  $\text{Richard}_r \text{ trusts } \text{himself}_x$.

a. The nominal *Richard* is assigned an interpretation, represented here as the index $r$, and identified as likely first argument of a predicate that is to follow.

```
<table>
<thead>
<tr>
<th>Richard</th>
</tr>
</thead>
<tbody>
<tr>
<td>PRED</td>
</tr>
<tr>
<td>$&lt;r \ldots$</td>
</tr>
<tr>
<td>↑</td>
</tr>
</tbody>
</table>
```

The referent of *Richard* corresponds to first argument of an anticipated predicate.

¹ Because of space limitations, I restrict my discussion to direct object reflexives in English. O’Grady (2005) considers a much broader range of patterns, including so-called ‘exempt reflexives’ (e.g., John thought that a picture of himself had been published).
b. The transitive verb *trust* is encountered and integrated into the semantic representation.

*Richard trusts*

\[
\text{TRUST} \\
<r> \\
\]

c. The reflexive pronoun *himself* is identified as the verb’s second argument; the referential dependency that it introduces is resolved by the index of *Richard*, which is already in the verb’s argument grid.

*Richard trusts himself*

\[
\text{TRUST} \\
<r_x> \\
\leftarrow r \\
\]

As can be seen here, the sentence-level processor acts in a way that minimizes the burden on working memory, immediately seeking out an antecedent within the portion of the semantic representation that it is constructing. No heed is paid to other possibilities, including the discourse topic *Harry* in examples such as (6).

(6) \(\text{Harry}_h\) suddenly realized something. \(\text{Richard}_r\) trusts \(\text{himself}_r\).

In contrast, the interpretation of the principal reflexive pronoun in Japanese (*zibun*) seems to reflect pragmatic processing, as shown by its well-known sensitivity to factors such as topicality, empathy, perspective, and contextual plausibility (Kuno & Kaburakai 1977, Kuno 1978, Sells 1987), without regard for structural position per se.\(^2\) As a result, we find the range of patterns partly exemplified in (7), the first two involving sentence-internal antecedents and third involving a sentence-external referent.

(7) a. Local co-argument antecedent:

\text{Taroo-ga_t} \text{ zibun-o_t} \text{ seme-ta}.

\text{Taroo-Nom} self-Acc blame-Pst

‘Taroo\(_t\) blamed self\(_t\).’

b. Local non-co-argument antecedent (Abe 1997:27):

\text{[Masao-no_m tegami-ga]} \text{ zibun-no_m \_ie -ni tuita}

\text{Masao-Gen letter-Nom self-Gen house-at arrived}

‘Masao’s letter arrived at self’s house.’

\(^2\) A partially overlapping set of factors appears to be relevant to the interpretation of English plain pronouns (*he*, *she*, etc); see, e.g., Kehler (2002), Wolf, Gibson, & Desmet (2004), Song & Fisher (2007:1961).


‘Taroo went to the police station. It was because the chief had summoned self.’

A variety of factors license the choice of an antecedent in these examples—
topicality in the case of the subject antecedent in (7a), point of view in the case of
the genitive antecedent in (7b), and contextual plausibility in the case of the extra-
sentential antecedent in (7c). None of these factors fall within the domain of the
sentence-level processor—responsibility for resolving the referential dependency
introduced by zibun has been passed to the pragmatic processor. This can be
represented as follows ($P = \text{pragmatic processor}$).

(8) Syotyoo-ga zibun-o ... (yonda)
    chief-Nom self-Acc ... (summoned)
    PRED
    $<s, x>$

    $\Rightarrow$

    $P$

    $\Leftarrow ?$

If these ideas are on the right track, English and Japanese differ in terms of
how they go about resolving the referential dependency introduced by a reflexive
pronoun. In English, the task is handled by the sentence-level processor, acting
immediately and locally in response to the pressure to minimize the burden on
working memory. In Japanese, in contrast, the job is handled by the pragmatic
processor, which can converge on local or distant antecedents, depending on the
context and circumstances.

(9) Sentence-level routine for interpreting $X$-self in English:
    Resolve the referential dependency immediately (hence locally).

(10) Pragmatically oriented routine for interpreting zibun in Japanese:
    Pass the referential dependency to the pragmatic system.

As we will see next, a dual-system approach to processing can also shed light on
the manner in which children and adults learn the facts and contrasts appropriate
to the type of reflexive pronoun found in the language that they are acquiring.
4. First Language Acquisition

It is well known that children are able to correctly interpret English reflexive pronouns by age 3 (e.g., Chien & Wexler 1990, Guasti 2002:285). Thus a child who is presented with the task exemplified in (11) will almost invariably respond in the affirmative, just as adults do.

(11) This is Mama Bear; this is Goldilocks.
Is Mama Bear touching herself?

![Sample picture from Chien & Wexler (1990)](image)

Early success in the interpretation of reflexive pronouns has also been reported for Japanese—Sugisaki and Otsu (2011:309) observe that ‘children as young as three have already acquired the major properties of zibun, such as subject orientation, long-distance binding, and the c-command requirement.’

Such findings are important and valuable, especially in light of the relative infrequency of reflexive pronouns in the input, but they are not sufficient. We need information not just about whether reflexive pronouns are correctly interpreted, but also about how the appropriate interpretations are derived in the course of real-time processing. A study by Clackson, Felser & Clahsen (2011) sheds important light on this question.

Clackson et al.’s experiment focused on 40 English-speaking children aged 6 to 9 who had performed well (mean score of 97%) on a task that required them to look at pictures and then answer questions containing a reflexive pronoun, as in (11) above. The key experiment involved an eye-tracking task, in which participants listened to a short passage as they looked at a set of four pictures, as exemplified in (12).

---

3 A search that I conducted of speech to Adam, Eve and Sarah in the CHILDES data base turned up just 17 instances of himself, compared to 487 instances of him; there was just one instance of themselves, versus 717 of them. A similar asymmetry is reported for myself and me by Bloom, Barss, Nicol & Conway (1994). Anaphoric uses of zibun also appear to be rare in speech to Japanese children—Sugisaki (2011) reports just 2 examples in the Aki corpus (from age 1;05 to 3;00), none in the Ryo corpus (1;04-4;00), and one in the Tai corpus (1;05-3;01).
(12) *Peter* was waiting outside the corner shop. He watched as *Mr. Jones* bought a huge box of popcorn for *himself* over the counter.

![Figure 2](image)

**Figure 2.** Sample picture from Clackson et al. (2011:143)

By examining eye gaze right after the reflexive pronoun is heard, Clackson et al. were able to determine that the children temporarily considered the discourse-prominent *Peter* as antecedent for *himself*, before eventually settling (correctly) on *Mr. Jones*. The effect was stronger in the younger children (ages 6 & 7) than in the older children (ages 8 & 9), and was barely discernable in adults.\(^4\) As Clackson et al. note (p. 140), it is evidently difficult to deactivate a highly prominent referent such as *Peter*, which is introduced as subject of the lead-in sentence and is referred to twice before the reflexive pronoun in encountered.

Taken together, the results of Clackson et al.’s online and offline tasks suggest a two-stage developmental sequence. In an initial stage, probably complete by age 3, the sentence-level structural routine (S) establishes its primacy, guaranteeing selection of a local antecedent. This is why classic acquisition studies report high rates of interpretive success by age 3 (e.g., Chien & Wexler 1990). Crucially, however, development is far from complete at this point. The competing pragmatic routine (P) continues to be activated when a reflexive pronoun is encountered in a sentence where a discourse-prominent referent is also in play. The full suppression of the pragmatic routine in this context takes many years, creating a developmental profile that is characterized by two stages.

\(^4\) Badecker & Straub (2002) and Sturt (2003) report effects of this type in adults in reading tasks, however.
Establishment of a preferred routine
Gradual entrenchment of the preferred routine at the expense of its competitor

Routine S ⇒ Routine S ⇒ Routine S
Routine P ⇒ Routine P ⇒ Routine P

FIGURE 3. Developmental profile for first language acquisition: x-self

This is precisely the developmental profile that one would expect if language acquisition consists essentially of processing amelioration, as I have suggested elsewhere (2005:193ff, 2011, in press). In a first step, the processor seeks out an efficient procedure for interpreting reflexive pronouns that is consistent with the input. In the case of English, that procedure turns out be a sentence-level routine focused on immediacy and locality, rather than on discourse prominence or pragmatic fit. That routine is then strengthened in response to repeated activation until the competing pragmatic routine is fully suppressed.

To my knowledge, there has not yet been experimental work on Japanese comparable to Clackson et al.’s research on English. As noted by Takaaki Suzuki (p.c.), a key question in this regard has to do with whether there is competition in Japanese, at least in children, between the language’s dominant pragmatic routine and a weaker sentence-level routine sensitive to the advantages of immediate local resolution of a referential dependency. I regret having to leave that question open for the time being.

5. Second Language Acquisition

I propose that second language acquisition works essentially the same way as first language acquisition in that development is the product of attempts to improve processing by creating low-cost routines in response to input. The acquisition of English as a second language by native speakers of Japanese is especially revealing in this regard.

5.1 The course of development

When native speakers of Japanese acquire English, they appear to go through a preliminary stage characterized by acceptance of long-distance antecedents in sentences such as John said Richard cut himself (e.g., Thomas 1989, Hirakawa 1990, Matsumura 1994, Wakabayashi 1996; for a critical summary, see Sachs 2010). Eventually, successful learners overcome this propensity and focus on the local antecedent, just as native speakers of English do.

Although interesting, these findings are inadequate for our purposes. As in the case of first language acquisition, we cannot hope to understand development
without information about how learners go about processing reflexive pronouns in real time in the course of comprehension. A study by Felser, Sato & Bertenshaw (2009) offers suggestive findings. (Unfortunately, I know of no comparable study involving the acquisition of Japanese as a second language.)

Felser et al.’s study focused on 22 adult Japanese learners of English as an L2 whose grammar scores on the Oxford Proficiency Test ranged from ‘mid-intermediate’ to ‘very advanced.’ Unlike less advanced learners, Felser et al.’s subjects responded as accurately as native speakers (96%) on a written questionnaire, thereby establishing their awareness of the fact that English reflexive pronouns require a local antecedent. Crucially, however, the L2 learners fared very differently on a reading-time task, on which they took longer to read the reflexive pronoun in (13a) than (13b). (Native speakers of English manifested no such difference.)

(13)

\[\text{longer reading time here...}\]

\[\text{\ldots than here}\]

a. [John said [Richard had cut himself with a very sharp knife]].

b. [Jane said [Richard had cut himself with a very sharp knife]].

Felser et al. suggest that the slowdown in (13a) is due to the fact that L2 learners initially consider the prominent sentence-initial NP John as a possible antecedent for the gender-matching reflexive pronoun, before ultimately settling (correctly) on the local antecedent Richard.

The results of Felser et al.’s study present a revealing picture of L2 development, at least for mid-intermediate and advanced learners. On the one hand, it is evident that the processing routine that favors immediate (local) resolution of the referential dependency is stronger than the routine that favors discourse prominence. That’s why the learners select the correct antecedent 96% of the time on the written questionnaire. On the other hand, it appears that the locally oriented sentence-level routine is not strongly enough entrenched to fully suppress the routine that favors a pragmatically prominent antecedent. That’s why the learners, unlike their native speaker counterparts, are distracted by the gender-matching NP in the matrix clause in sentences such as (13a).

We are thus left with a developmental scenario that is characterized by the same two stages that are found in the case of first language acquisition—but with an interesting difference. As in the case of first language acquisition, the routine that seeks local resolution of referential dependencies has become strong enough to ensure selection of the correct interpretation for English reflexive pronouns. Crucially, however, the competing pragmatic routine continues to be active, identifying the previously encountered and highly topical matrix subject as a potential antecedent. In contrast to what happens in the case of first language
acquisition (figure 3), there is no indication that this routine is ever entirely eliminated, at least for the learner population considered by Felser et al.

Establishment of
a preferred routine

Gradual entrenchment of the preferred
routine at the expense of its competitor

Routine S \Rightarrow \text{Routine S} \\
Routine P \quad \text{Routine P}

FIGURE 4. Developmental profile for second language acquisition: \textit{x-self}

5.2 The role of transfer

Why do Japanese learners of English manifest the ongoing sensitivity to discourse prominence that showed up in Felser et al.’s study? One possibility is transfer, especially in light of the fact that the interpretation of reflexive pronouns in Japanese is highly sensitive to discourse and pragmatic factors, as we have seen (section 3).

I have no objection to transfer-based explanations in general. To the contrary, I believe that transfer—and especially the transfer of processing routines—has a major role to play in understanding second language acquisition (see, e.g., O’Grady 2011, in press). However, a transfer explanation appears inappropriate here for two reasons.

First, the phenomenon documented by Felser et al. does not involve errors in the choice or content of routines. The fact the Japanese-speaking learners come up with the right interpretation for the test sentences strongly suggests that they apply the correct interpretive routine. The problem is that access to that routine is hindered by activation of the pragmatic processor.

Second, there is good evidence that activation of the pragmatic processor happens even when the learner’s first language requires a local antecedent for its reflexives. Felser & Cunnings (2012) studied the interpretation of English reflexive pronouns by 25 adult German-speaking learners, using a reading task similar to the one used by Felser at al. for Japanese-speaking learners. The key finding was that German-speaking learners too are initially attracted to the discourse-prominent referent in patterns such as the following, slowing down when its gender does not match that of the reflexive pronoun.\(^5\)

(14) \textbf{Helen} had worked in the army hospital for years. She noticed [that the \textbf{soldier} had wounded \textbf{himself} while on duty in the Far East]. Life must be difficult when you are in the army.

\(^5\) As Felser & Cunnings note (p. 588), this is not quite the same result as for Japanese-speaking learners, who were slowed down if the discourse-prominent NP was of the \textit{same} gender as the reflexive pronoun.
Once again, we see evidence of an attempt to link the reflexive pronoun to a distant but discourse-prominent antecedent rather than to the local antecedent (*the soldier*) that is favored by the sentence-level processor. The attempt is short-lived, however, and the learners almost instantly converge on the local antecedent. As in the case of the Japanese-speaking learners, the right processing routine is in place; it just isn’t accessed immediately. As Felser & Cunnings suggest (p. 600), the problem seems to be one of timing: the L2 learners’ initial reaction to a reflexive pronoun involves pragmatic rather than sentence-level processing. But why?

One possibility, in the spirit of an idea put forward by Felser & Cunnings (p. 601), is that pragmatic processing requires less computation in the examples at hand, as it simply accesses a referent that is already prominent by virtue of its previous mention, its topicality, and its reactivation by the definite pronoun *she*. In contrast, construction of the semantic representation needed to compute locality requires the use of structure-building routines that are sensitive to lexical choices, morphological properties, and linear order, among other factors. On my view, genuine sentence-level processing *does* take place in the course of both first language and second language acquisition. But as Felser & Cunnings suggest, it cannot always be done quickly enough to pre-emptively suppress the alternative offered by the pragmatic processor, especially in the early years of development.

6. Concluding remarks

This paper focuses on two related ideas about the emergence of the relationship between form and meaning, one having to do with syntax and the other with language acquisition.

First, as outlined in sections 1–3, I have proposed an approach to syntax that offers a processing-based account for how reflexive pronouns are interpreted in English and Japanese. The key idea is that the relationship between form and meaning is regulated by the processor, without the mediation of syntactic structure. This in turn leads to an account of the comprehension of reflexive pronouns that posits fundamentally different processing routines for different languages—a sentence-level structural routine for English and a pragmatically driven routine for Japanese.

Second, this way of looking at syntax and development opens the door for a new approach to understanding second language acquisition. In particular, as illustrated by the case study of reflexive pronouns on which we have been focusing, development is best seen as a two-part phenomenon. In an initial stage, an appropriate processing routine emerges that permits the interpretation of reflexive pronouns in a manner consistent with the input from the language to which the learner is exposed. Then, in a second stage that can unfold over a period of years, that routine is strengthened and reinforced to the point where any
competing interpretive strategy is expunged, at least in the case of first language acquisition.

To date little attention has been paid to the possibility that processing considerations lie at the heart of development. However, if the ideas put forward here are on the right track, we can expect further study of that possibility to contribute to a deeper understanding both of development and of the mechanisms that ultimately mediate the relationship between form and meaning in natural language.

References


