Online Processing of Anaphora by Advanced English Learners

Rong Liu and Janet Nicol
University of Arizona

1. Introduction[1]

It has long been observed that second language (L2) learners have particular difficulty mastering agreement morphology, both in production and in comprehension (e.g., Aaronson & Ferres, 1987; Bailey, Madden & Krashen, 1974; Chen, 2009; Ellis, 1988; Guillelmon & Grosjean, 2001; Hahne Mueller, & Clahsen, 2006; Jiang, 2004, 2007; Johnson & Newport, 1989; Johnson, Shenkman, Newport, & Medin, 1996; Krashen et al., 1977; Krashen & Pon, 1975; Lardiere, 1998; Long, 1997; O'Grady, 2006; Perkins & Larsen–Freeman, 1975; Prevost & White, 2000; Schmidt, 1983; Shapira, 1978; Wei, 2000). Research has examined bound inflectional morphology, such as English plural *–s*, number marking on verbs, and tense marking on verbs, in addition to unbound morphemes, such as gender-marked determiners like French *le, la* and Spanish *el, la*. In most of these cases, grammatical agreement is required between one sentential element and another.

Another linguistic domain in which one finds agreement between two sentential entities is anaphora. In a sentence such as *Jane likes herself*, the reflexive pronoun herself must agree in person, number, and gender with its antecedent *Jane*. Most of the research on anaphora in L2 has focused on the ability of learners to acquire new syntactic constraints—the binding constraints (Chomsky, 1981)—on the coreference possibilities of pronouns and reflexives (Hirakawa, 1990; Lee & Schachter, 1997; Thomas, 1995; Wakabayashi, 1996; White, 1998). These constraints show cross-linguistic variation, and some theories of language acquisition claim that once the constraints are "set" for one language, they cannot be "reset" for another. Typically, these studies use methods that require a judgment from the participants that reveals how they interpreted the pronoun or reflexive. Participants may be presented with a sentence such as *John wanted Tom to know himself better* and asked explicitly who *himself*, or asked to make a judgment about whether the sentence is true of a particular scenario, which requires that participants interpret the sentence as assign an antecedent to the reflexive (e.g., Akiyama, 2002). It is worth noting that in general, these studies do not use simple sentences like *Jane likes herself*, because such sentences are not subject to different binding constraints across languages. It is the more complex structures, containing multiple potential antecedents that give rise to the differences.

Our focus in this paper is on the comprehension of anaphora by L2 learners: native Chinese speakers whose L2 is English. In this work, we focus not on binding as a grammatical constraint that may or may not be reset in the acquisition of L2. Rather, we use anaphora as an interesting case of agreement between entities in the sentence. One way in which pronouns and reflexives differ from other kinds of agreeing elements is in the sheer phonological saliency of the agreeing form (See O'Grady, 2006). Note that in determiner-noun agreement, for example, such as *la bouche* (the$_{\text{-feminine}}$ mouth$_{\text{-feminine}}$), *la* is unstressed (and it so happens, differs minimally from other definite determiners),

[1] This work is part of the dissertation project supported by the Language Learning Dissertation Grant, a dissertation grant from the Social and Behavioral Sciences Research Institute and a grant from the Asian American Faculty, Staff and Alumni Association at the University of Arizona. The second author also acknowledges support for travel and accommodation from the Program in Cognitive Science at the University of Arizona.

whereas in *Jane likes herself*, both elements are stressed. This could make it easier for L2 learners to master agreement, and in experiments, to detect violations in agreement.

Another way in which our work differs from much of the past research on anaphora is our focus on the comprehension of sentences *as they are being read*. Therefore, we use a reading task that yields word-by-word reading times to grammatical vs. ungrammatical sentences. This allows us to determine whether L2 readers are able to comprehend sentences incrementally, as native speakers do. In our view, the primary goal of L2 learning is to be able to use language in everyday situations. To do so successfully requires a degree of automaticity of language processing: language input is fast, and speakers must be able to comprehend sentences within a relatively short span of time. This means that it is not enough merely to "know" vocabulary and rules of grammar; the knowledge must be deployed rapidly and automatically. Likewise, in the assessment of L2 acquisition, it is important to test not only language knowledge through standard paper-and-pencil tests, but also language *processing* through "on-line" methods (N. Ellis, 1993; Hulstijin, 2001; Segalowitz, 2003; Segalowitz, Segalowitz, & Wood, 1998).

Only relatively recently have researchers in the L2 acquisition field begun to study L2 grammatical processing (e.g., Clahsen & Felser, 2006; Dussias, 2001; Felser 2005; Frenck-Mestre & Pynte, 1997; Hahne & Friederici, 2001; Sabourin, 2003; Osterhout & Mobley, 1995) and some of these studies have suggested that L2 learners do not comprehend language in the same manner as native speakers, even though they have learned the relevant information. Such "processing" errors may have a number of different causes. One may be that L2 comprehenders use inappropriate processing strategies, possibly transferred from their L1 (Marinis, 2003). Another reason, proposed by Clahsen and Felser (2006), is that L2 sentence processing is fundamentally different from L1 processing in the sense that the syntactic structure that is computed is "shallower" and that sentence comprehension relies more heavily on lexical and pragmatic information. Finally, L2 comprehenders may suffer from failures of automaticity which affect the speed of lexical retrieval (Ardal, Donald, Meuter, Muldrew, & Luce, 1990) and retrieval of morphosyntactic information (Hahne & Friederici, 2001). Our research addresses this third possibility, that L2 readers may be unable to deploy their grammatical knowledge in the same automatic fashion as native speakers.

Some compelling evidence for this view comes from studies by Jiang (2004, 2007), who tested a population similar to ours on their sensitivity to number marking. In Jiang (2004), participants read sentences like the following (for clarity, we have underlined the region of interest):

(1) The bridges to the island <u>were about</u> ten miles away.
*The bridge to the island <u>were about</u> ten miles away.

Results showed that native speakers slowed down at both "were" and "about" when reading the ungrammatical version. This indicates that readers noticed the number disagreement in the sentence. Nonnative speakers, however, did not take longer to read these words, indicating that they were not sensitive to the failure of agreement. However these participants did show sensitivity to a different kind of violation that was tested as part of the same experiment: violations of verb subcategorization (such as *The teacher insisted the children to mail the letter to the president*). An "off-line," paper-and-pencil task was also administered to non-native participants; performance on this test indicated that they knew the rules of subject-verb agreement. Hence, it was the on-line computation of subject-verb agreement that was especially difficult for this group.

Follow-up work by Jiang (2007) found that advanced L2 learners had difficulty in detecting the anomalous sentences containing the plural morpheme –s during their online reading of texts. His non-native participants showed no reading time differences in the relevant sentence regions of sentences such as "The visitor took several of the rare <u>coins in the</u> cabinet." and "The visitor took several of the rare <u>coin in the</u> cabinet." Native speakers, on the other hand, slowed down after the word "coin" (compared to "coins"), presumably reflecting the difficulty incurred from reading the ungrammatical sequence "several of the coin." Overall, Jiang's results suggest that although non-native speakers can attend to plural marking in off-line situations that are not time-constrained, when they are reading for comprehension, L2 learners either fail to notice it, or fail to retrieve it when they reach a point in the sentence where agreement is relevant.

In our experiments, we address two issues. First, we examine L2 learners' sensitivity to number agreement in the comprehension of reflexives and compare it to their sensitivity to gender agreement.

This provides an interesting contrast because the plural form of a noun in English typically only involves the addition of an –s. In terms of salience, number-marking is far less noticeable than gender-marking. English nouns are not typically marked by way of a morpheme (the -ess/-er distinction does exist, though it is waning), but rather gender is part of the word's meaning, as in *girl*, or *matriarch*. Another difference has to do with the properties of Chinese. In general, gender agreement is more common than number agreement, given the fact that nouns are not marked for number. If this fact influences sensitivity to these grammatical features during language processing in general, then we would expect speakers to be more sensitive to violations of gender agreement than to violations of number agreement.

Second, we examine learners' sensitivity to gender agreement between a subject pronoun and prior antecedent in conditions where pragmatic information either conflicts with gender agreement or does not. Specifically, we examine the effects of so-called "implicit causality" of main verbs (Garnham & Oakhill, 1985; McDonald & MacWhinney, 1995). Research has shown that in a two-clause construction containing a simple transitive sentence followed by a *because*-clause, the main verb may be biased toward the first noun phrase (NP1) or the second noun phrase (NP2) as the antecedent the subject pronoun in the following clause (Stewart, Pickering, & Sanford, 2000). For example, in the sentence fragment *The boy admires the man because he…*, readers naturally interpret *he* as *the man (NP2)* instead of *the boy (NP1)*: the verb *admire* is an "NP2-bias verb." In contrast, in the sentence "*The boy bores the man because he …,*" *he* is more likely to be interpreted as *the boy*. The verb "bore", therefore, is an "NP1-bias verb." At the root of this verb difference is the thematic roles that the two arguments carry: NP2-bias verbs assign Experiencer, Stimulus to the subject and object, respectively. NP2-bias verbs, on the other hand, assign Stimulus, Experiencer roles. In general, a Stimulus makes a better causal agent than an Experiencer, and this is what leads to the interpretation bias. Research by McDonald and MacWhinney (1995) (among many others) showed that when there is a conflict between the verb bias and gender congruence, reading is slowed. For example, people take more time to read a sentence such as *Jane bores Jack because he is repetitious* than *Jane bores Jack because she is repetitious* because *bore* is a Stimulus/Experiencer verb and so it would be expected that the following subject pronoun would corefer with NP1, the Stimulus. When the pronoun forces a different interpretation, based on gender agreement, reading is slowed. In order to understand these sentences appropriately, readers must keep track of the gender features of the nouns and their position in the string (subject or object), have access to the thematic properties of the verb, and make the appropriate thematic role assignments. These sentences were included here to test whether gender agreement goes awry when other variables come into play.

3. The study

3.1. Research questions and hypotheses

The purpose of the experiment is to test L2 learners' ability to use their knowledge of coreference constraints *on-line*, as they read sentences. One set of materials tests sensitivity to featural constraints on coreference, and another set of materials tests sensitivity to so-called "implicit causality" associated with verbs. If advanced L2 learners of English use feature information and verb cues *during* processing, they should show the same slow-downs as native speakers (and at the same points in the sentence) when they read sentences that are unacceptable precisely with respect to these types of information. It is assumed that native speakers (NSs) will, of course, show such slow-downs (relative to the acceptable counterparts) in all cases. It is predicted that NNSs will be more sensitive to gender than to number. With respect to implicit causality, we expect NNSs to pattern with NSs only if they are indeed sensitive to gender, and are able to retrieve verb information and compute sentence meaning in a relatively rapid manner.

3.2. Methodology

The self-paced reading methodology was used. This is a standard psycholinguistic method in which participants press a key to advance word-by-word throughout a sentence. It is assumed that a slowdown in reading an anomalous sequence reflects participants' sensitivity to the grammatical

constraints under investigation. As discussed above, this technique has been used extensively with native-language readers and successfully with L2 learners (e.g., Jiang, 2004, 2007).

A comparison of reading time at the critical positions in grammatical vs. ungrammatical sentences (as in the underlined segments in the example below), is used to determine whether or not participants are sensitive to the relevant information (gender and number features, on one hand, and verb/thematic role information, on the other). For example, consider the pair of sentences below:

(2) The careless pedestrian found themselves covered with mud.
(3) The careless pedestrian found himself covered with mud.

If a reader is able to compute the ungrammaticality of the first sentence, at the reflexive or in subsequent regions, the reader should show longer reading times than for those same regions in the second (grammatical) sentence. Relatively longer reading times are assumed, in this case, to reflect sensitivity.

It is important to mention here that slowdowns in reading do not always occur at the very point of ungrammaticality, nor are they localized to just one word in the sentence. It is quite typical to see reading slowdown not just at a point when the sentence becomes ungrammatical, but in subsequent regions as well. This is referred to as "spillover." It is also common to see a delay: the effect emerges one or two words *after* the sentence has become ungrammatical. This happens because readers advance ahead in the sentence before they have completely processed the ungrammatical word.

In addition to the on-line reading task, traditional paper-and-pencil tests and background questionnaires were administerd to assess grammatical knowledge and overall proficiency in English. This is necessary both to ensure homogeneity among the participants, and to allow comparisons with other studies in the literature.

3.3. Materials

3.3.1. Self-paced reading tasks: Reflexive sentences

Thirty-two grammatical/ungrammatical sentence pairs were created. Each sentence contained a subject, verb, and reflexive, followed by additional material. Sentences were balanced in the following ways. Half of the sentences probed number agreement, and half probed gender agreement. Half of the number-agreement sentences contained a singular subject and half contained a plural subject. Half of the gender-agreement sentences contained a feminine subject and half contained a masculine subject. Examples appear below (A complete list of all the items used in the self-paced reading tasks can be found in Appendix B. Some of the materials were adapted from Osterhout and Mobley (1995).

(4a) The shy girls forced themselves to sing the part at the concert. (Number)
(4b) The shy girls forced herself to sing the part at the concert. (Number)

(5a) The lonely grandfather made himself a cup of tea. (Gender)
(5b) The lonely grandfather made herself a cup of tea. (Gender)

3.3.2. Self-paced reading tasks: "Implicit causality" sentences

This set of materials (adapted from a number of studies, such as McDonald & MacWhinney, 1995; Rigalleau & Caplan, 2000; Stewart et al., 2000) consists of 32 *double-check* sentences containing two clauses, a simple clause containing a subject, verb and object, followed by a *because*-clause containing a subject pronoun. There were two types of sentences those which contained a "NP1-bias" verb such as *amuse*, which assigns a Stimulus role to its subject and Experiencer role to its object (as in (6) below), and those containing a "NP2-bias" verb such as *admire*, which assigns an Experiencer role to its subject, and Stimulus to its object (as in (7) below). Within each type, pairs of sentences were constructed. In one member of each pair, the "verb bias" was consistent with gender agreement (as in (6a) and (7a) below); in the other member of the pair, these were in conflict (as in (6b) and (7b) below). Note that the latter sentences are not *ungrammatical*, they are simply a little strange or implausible, and for the proficient English reader, more difficult to comprehend.

(6a) The mother amused the father because she told funny jokes at dinner. (NP1 biased)
(6b) The mother amused the father because he told funny jokes at dinner. (NP1 biased)

(7a) The boy admired the girl because she was so intelligent. (NP2 biased sentences)
(7b) The boy admired the girl because he was so intelligent. (NP2 biased sentences)

We note that another factor in pronoun resolution is a reader preference for "parallel function": a subject NP will be favored as the antecedent of a subject pronoun. This could add to the NP1 bias and detract from the NP2 bias.

The Reflexive sentences and Implicit-causality sentences were interleaved into one experiment with two counter-balanced lists. Each list contained 16 grammatical "number" sentences, 16 ungrammatical "number" sentences, 16 grammatical "gender" sentences and 16 ungrammatical "gender" sentences, and 32 "implicit causality" sentences. Half of the "implicit causality" sentences resolve in favor of the bias and half against the bias. So the total number of critical sentences was 96, with 16 tokens per type. No participants saw all variants of a sentence. Half of the sentences were followed by yes/no comprehension questions; these were included to ensure that participants were actually reading for general comprehension. For example, following *The careless pedestrians found themselves/himself covered with mud* would be the question *Were the pedestrians covered with water?* Most of the comprehension questions followed grammatical sentences; however, to keep participants paying attention to both grammatical/plausible and ungrammatical/implausible sentences, several ungrammatical/implausible sentences were followed by a comprehension probe. For these, either a Yes or a No answer was treated as correct. There were also 32 filler sentences. In order to ensure that the sentences were understandable to the participants, a larger set of items was pre-tested on several participants drawn from the same target participants group and only those judged to be easy to read by the researcher were included in the experiment. Comprehensibility of the materials was also confirmed by participants' informal report after the experiment.

3.3.3. Offline written test

Materials also included a written test administered to NNSs to assess their explicit knowledge of gender, number constraints and "implicit causality" constraints in the resolution of anaphora.

Because the focus of this study is on the on-line processing of anaphora, it is crucial to establish that participants at least have the explicit knowledge of the constraints under study. It is unlikely (though not impossible) that they will automatically process this information online if they do not demonstrate knowledge of these constraints in untimed tasks. There were 15 items altogether. There were three items targeting the masculine *himself*, three items targeting *herself*, three items targeting a number-matched reflexive, three targeting NP1 biased verbs and three NP2 biased verbs. Items were similar or identical to those used in the online self-paced reading task. Items were presented in pseudo-random order. A complete list of items appears in Appendix C. The following presents the sample stimuli for each category:

Sample item testing the gender cue:
The lonely grandfather made_____(himself/herself) a cup of coffee.
The gracious wife introduced_____(himself/herself) to the guests.

Sample item testing the number cue:
The shy girls forced _____(themselves/herself) to sing the part at the concert.

Sample item testing the NP1 biased verb cue:
The mother amused the father because____(he/she) told funny jokes.

Sample item testing the NP2 biased verb cue:
The widow envied the bachelor because _____ (he/she) won the national lottery.

3.3.4. Language profile

ESL participants' language background information was collected via questionnaire (see Appendix A), which was adapted from Jiang (2007). It was a one page survey with questions about participants' demographic information, including age, age at which a participant began to learn English, length of residence in the United States, language learning experience, and self-ratings of English proficiency in listening, reading, speaking and writing skills. Such information is informative with respect to both homogeneity of the subject group and comparability with participants tested in previously published research.

3.4. Participants

Two groups of participants were tested: English native speaker controls and advanced ESL learners. Forty-one native speakers were tested. Data from nine were excluded from analysis because of high error rates in the written test or in the reading comprehension questions task (more than 20%). All were undergraduate students at the University of Arizona. Most were taking introductory Psychology courses, although some were recruited from introductory Linguistics classes. They received course credit for participating in the experiment.

Forty-one Chinese ESL learners participated in the study. Data from seven participants were excluded from analysis due to high error rates in the written test or in the reading comprehension questions (more than 20%). The ESL learners were graduate students at the University of Arizona and their spouses. All participants had normal vision or correctly to normal vision. They all spoke Chinese as their first language. Most of them were science majors. ESL participants' ages ranged from 19 to 41, with an average of 30. Based on their TOEFL score and self-ratings of their English proficiency, they would be considered advanced ESL learners. Their average TOEFL score was 627, with a range of 560 to 677 (The minimum requirement of TOEFL for admission to most American universities is 550.). Their average self-ratings of English in the four skills (i.e., speaking, listening, reading and writing) were 3.9, 4.1, 4.2, 3.9, respectively (on a scale of one to five with five being most proficient). They were well-educated, with an average of 20 years of formal education. On average, they had been in the United States for 3.7 years and they had been studying in English speaking countries for more than three years. Most participants started to learn English after puberty. (One person reported starting to learn English at the age of 3; however, her TOEFL score was 623, nearly the same as the group average). The following table summarizes the background information of the non-native participants in this experiment:

Table 1. *Nonnative Speakers' Language Profile in the Experiment*

	M	*Min*	*Max*	*SD*	*N*
Age	30	19	41	4.9	34
TOEFL scores	624	560	677	27.7	30
Age starting English	12	3	16	2.4	34
Years of formal education	20	13	25	2.6	34
Year of residence in the USA	3.7	1	10	2.4	34
Years of education in English speaking countries	3.3	1	9	1.8	34
Self rating of English proficiency					
Speaking	3.9	3	5	0.7	34
Listening	4.1	3	5	0.7	34
Reading	4.2	3	5	0.6	34
Writing	3.9	2	5	0.8	34

3.5. Procedures

Participants were tested individually in a quiet test booth. After providing signed consent, participants were seated in front of a computer monitor and button-box for the self-paced "moving-window" reading task. They were told that they would read sentences for general comprehension one word at a time and as soon as they read each word they would press the right-hand button to advance to the next word. The current word disappeared when the next word appeared. Each sentence was first presented with each word consisting of a string of x's, one x for each letter in the word. With a button-press, the current word turned "back" into x's and the next word changed from x's to letters. Spaces between words were preserved. Participants were told to proceed as quickly as they could but to make sure that they fully understood what they were reading. They were also told that occasionally there would be a yes/no comprehension question about what they had just read and they were expected to press either the right-hand button for "Yes" or the left-hand button for "No." They were told to read as quickly as possible and answer the comprehension questions as accurately as possible.

To familiarize the participants with the task, they were asked to read 10 practice items of various types. DMDX, an experimental software developed by Forster and Forster at the University of Arizona (see Forster & Forster, 2003), was used to present the materials and record the participant's reaction time to each word in a sentence, along with responses to each comprehension question. The self-paced reading task took less than 35 minutes. After the online task, participants filled out the language profile questionnaire and the paper-pencil written test.

3.6. Data Analysis and Results

The on-line reading data were analyzed only for participants who performed the task satisfactorily. Exclusion of participants was based on several criteria.

First, nonnative speakers' test scores in the written test were calculated. The written test was administered immediately after the self-paced reading tasks were completed. Those whose error rates were more than 10% were excluded from the analysis. Two nonnative speakers were excluded for the reason, leaving the total number of nonnative speakers at 39. Their error rates fell between 0% and 10% with an average error rate of 7%. The written test results showed that nonnative speakers in the study have mastered the structures investigated in the study as measured by the offline task.

In addition, for all participants, errors on comprehension questions were calculated. Those who had an error rate higher than 20% were excluded from further analysis. Nine native speakers were excluded for this reason, leaving the total number of native speakers at 32. Those native speakers' error rates ranged from 2.5% to 20%, with an average error rate of 10%. Five nonnative speakers were excluded, leaving the total number of nonnative speakers at 34. Those nonnative speakers' error rates fell between 3.8% to 20%, with an average error rate of 14%. The comprehension rates suggested that the two groups of participants followed the instructions to read the sentences for general comprehension and that they understood the sentences reasonably well.

In addition, following standard procedure, reaction times (RTs) that were two *SD*s longer or shorter than a given participant's mean, or higher or lower than the high and low cutoffs set at 2,000 and 200ms, respectively, were excluded. Reaction time Data involving incorrect responses were excluded and only correct RTs were included in the analysis. The trimmed data, along with missing data and display errors, accounted for 5.4% of the data. For the participant analyses, 32 mean RTs were computed for each participant, one for each test position (there are four test positions) in each condition (there are two conditions: grammatical vs. ungrammatical) for each cue (there are four cues: the gender cue, the number cue, the NP1 biased verb cue and the NP2 biased verb cue). These means were subjected to analysis of variance computed by the statistical program SPSS. For the item analyses, sixteen means were calculated for each of the 96 sentences, one for each of the four test positions in each of the four conditions. The participants' RTs for each test position, each condition, and each cue are given in the following tables. The native speaker results appear first, followed by the non-native speaker results.

Table 2. *Native Speakers' Mean RTs (ms) in Four Regions (of an example sentence) for Grammatical and Ungrammatical Sentences Involving Number Cues and Gender Cues*

Cue	Number				Gender			
Test position	1 forced	2 -self	3 to	4 sing	1 made	2 -self	3 a	4 cup
Grammatical	449	436	415	398	420	415	407	392
Ungrammatical	438	436	460	429	428	430	465	444
Difference	11	0	-45*	-31*	-8	-15*	-58*	-52*

*$p < .05$

Table 3. *Nonnative Speakers' Mean RTs (ms) in Four Regions for Grammatical and Ungrammatical Sentences Involving Number Cues and Gender Cues*

Cue	Number				Gender			
Test position	1 forced	2 -self	3 to	4 sing	1 made	2 -self	3 a	4 cup
Grammatical	560	550	484	477	546	542	469	454
Ungrammatical	556	553	502	498	540	558	534	491
Difference	4	-3	-18	-21	6	-16	-65*	-37*

*$p < .05$

Table 4. *Native Speakers' Mean RTs (ms) in Four Regions (of an example sentence) for Plausible and Implausible Sentences Involving "Implicit Causality"*

Cue	NP1 biased verb				NP2 biased verb			
Test position	1 because	2 s/he	3 told	4 funny	1 because	2 s/he	3 was	4 so
Plausible	410	402	399	391	419	405	391	402
Implausible	414	402	414	408	423	420	413	411
Difference	-4	0	-15*	-17*	-4	-15	-22*	-9

*$p < .05$

Table 5. *Nonnative Speakers' Mean RTs (ms) in Four Regions for Plausible and Implausible Sentences Involving "Implicit Causality"*

Cue	NP1 biased verb				NP2 biased verb			
Test position	1 because	2 s/he	3 told	4 funny	1 because	2 s/he	3 was	4 so
Plausible	507	444	450	464	500	445	445	451
Implausible	518	462	480	482	504	459	469	479
Difference	-11	-18*	-30	-18	-4	-14	-24	-28*

*$p < .05$

The data for NNSs and native speakers were analyzed separately. The SPSS Paired-Samples *t*-Test was chosen to compare the participants' mean RTs for grammatical/plausible and ungrammatical/implausible sentences at each of the four different sentence positions and for each of the four cues. For both groups (i.e., the native group and the nonnative group), thirty-two pairs were compared both for subject analysis and item analysis. Because the prediction is directional—ungrammatical/implausible versions are predicted to take longer to read than grammatical/plausible ones—the one-tailed test is used.

The discussion of statistical results is broken down as follows. We discuss the statistical results for number, then gender, then verb bias. In each case, we start with the data at Position one, where we

expect no statistical differences in RT. Next we discuss the point at which the anaphor appears, Position two. Finally, we discuss the subsequent two regions, Positions three and four, where we expect to see RT differences due to either ungrammaticality or implausibility.

3.6.1. Statistical analyses of reading time differences

3.6.1.1. Reflexive sentences: Number

Position one. At this position, there should be no RT differences because participants read the same stimuli in the two experiment conditions, grammatical or ungrammatical. This is exactly what was found. For native speakers (NSs), as in Table 2, RT difference at this position is not statistically significant. For nonnative speakers (NNSs), the same is true. The RT difference is not significant.

Position two. For NSs (See Table 2), there is no effect of number congruence at position two. The same is also true for NNSs (See Table 3). This is the first point of divergence for the sentence pairs; this is where the anaphoric element appears. One might expect to see RT differences begin to emerge here; however, this is not observed. As mentioned above, it is not uncommon to an effect emerging one or two words after the point of divergence, especially if this point contains a high frequency word.

Positions three and four. For NSs, RTs between the grammatical and ungrammatical sentences at the two positions for the number cue were statistically significant both in subject analysis, $t1(31) = 4.8$, $p < .05$ for position three, $t1(31) = 4.1$, $p < .05$ for position four and in item analysis, $t2(31) = 5.7$, $p < .05$ for position three and $t2(31) = 3.6$, $p < .05$ for position four (Following convention, $t1$ refers to the subject analysis and $t2$ refers to the item analysis.). Unlike the NSs, NNSs showed no reliable differences, either on the subject analysis, $t1(33) = 1.6$, $p = .056$, or the item analysis, $t2(31) = 1.2$, $p = .114$ for the third position. For the fourth position, it was found that results were significant in the subject analysis, $t1(33) = 2.1$, $p < .05$ but nonsignificant in the item analysis, $t2(31) = .90$, $p = .189$. Overall, the NNSs group showed no reliable effect of number congruence.

3.6.1.2. Reflexive sentences: Gender

Position one. As shown in Tables 2 and 3, there is no effect of gender congruence at this position for both NSs and NNSs.

Position two. For native speakers (NSs), there is a significant effect of gender congruence at the reflexive position. NNSs, however, do not show such an effect here.

Positions three and four. For NSs, the gender congruence condition showed the same effects as number: $t1(31) = 7.5$, $p < .05$ for position three, $t1(31) = 6.3$, $p < .05$ for position four and in item analysis, $t2(31) = 6.3$, $p < .05$ for position three, $t2(31) = 6.0$, $p < .05$ for position four. NNSs show the same effects: RTs between the grammatical and ungrammatical sentences at the two positions were statistically significant both in the subject analysis, $t1(33) = 5.6$, $p < .05$ for position three, $t1(33) = 3.3$, $p < .05$ for position four and in the item analysis, $t2(31) = 4.4$, $p < .05$ for position three and $t2(31) = 2.2$, $p < .05$ for position four. Overall, both groups showed a robust gender congruence effect.

3.6.1.3. Implicit causality sentences

Position one. There is no effect for either the NP1 biased verb condition and the NP2 biased verb condition. This is true for both NSs and NNSs.

Position two. For NP1 biased verb sentences, NSs show no effect; however the NNSs show a significant effect on both the participant analysis, $t1(33) = 2.0$, $p < .05$, and the item analysis, $t2(31) = 2.2$, $p < .05$, suggesting a clear expectation for the pronoun to corefer with the subject of the main clause. For NP2 biased verb sentences, NSs showed no reliable effect. The RT difference was only significant in the participant analysis, $t1(31) = 1.7$, $p < .05$, $t2(31) = 1.5$, $p = .067$. The effect for NNSs was nonsignificant, $t1(33) = 1.2$, $p = .112$, $t2(31) = 1.5$, $p = .073$.

Positions three and four. In the NP1 biased verb condition, for NSs, the RT difference between plausible and implausible sentences in position three was significant in both the subject analysis, $t1(31) = 1.9$, $p < .05$ and item analysis, $t2(31) = 1.8$, $p < .05$. In position four, results were also statistically significant on both analyses, $t1(31) = 2.6$, $p < .05$ and $t2(31) = 2.1$, $p < .05$. NNSs showed the following pattern: in position three, there was a 30 ms difference that was significant in the subject

analysis, $t1(33) = 2.8$, $p < .05$ and approached statistical significance in the item analysis, $t2(31) = 1.5$, $p = .077$. In position four, the 18ms difference was statistically significant in the subject analysis, $t1(33) = 1.9$, $p < .05$, but not in the item analysis, $t2(31) = 1.1$, $p = .135$. Overall, both groups showed an effect of NP1 biased verb type.

In the NP2 biased verb condition, for NSs, the RTs difference between plausible and implausible sentences in position three was found to be statistically different in both the subject and item analyses: $t1(31) = 3.2$, $p < .05$, $t2(31) = 3.4$, $p < .05$. In position four, the difference in RTs was statistically nonsignificant in both the subject analysis, $t1(31) = 1.3$, $p = .200$ and item analysis, $t2(31) = 1.1$, $p = .271$. For NNSs, in position three, the difference was significant only in the subject analysis, $t1(33) = 2.1$, $p < .05$ but not in the item analysis, $t2(31) = 1.6$, $p = .059$. In position four, the difference was significant in both the subject analysis, $t1(33) = 3.1$, $p < .05$, and the item analysis, $t2(31) = 1.9$, $p < .05$. Overall, both groups showed sensitivity to both NP1 biased verbs and NP2 biased verbs during online processing.

4. Discussion and conclusion

This experiment was designed to explore whether L2 learners process gender-marked pronouns and gender and number-marked reflexives as L1 speakers do. We examined the processing of gender in two different contexts: simple reflexive-antecedent contexts, and contexts in which pronoun-antecedent agreement interacts with pragmatic information.

For the reflexive study, there are several findings of interest. Our results showed that L2 learners noticed violations of gender agreement in sentences like *The new stepmother prepared himself to meet the family*, as evidenced by reading slowdowns after the reflexive compared to the gender-congruent condition. Their reading times patterned with those of the native speakers: reading slowed down at the reflexive and continued into subsequent regions although for the L2 learners, the slowdown was not statistically significant until after the reflexive. The findings suggest that not only are these non-native speakers able to extract gender information during sentence processing, they are able to use this information in computing agreement with a later-occurring element. Further, they do so in roughly the same time-course as native speakers. They are certainly slower in their reading than the native speakers, but the effects show up in the expected sentence regions. As for the sentence condition with number agreement violation such as *The hungry guests helped himself to the delicious meal*, the L2 learners did show reading time slowdowns, but in no region was the slowdown statistically significant. This contrasts with the results for the native speakers, who showed a robust effect after the reflexive. Overall, then, the non-native speakers are fully capable of extracting (and using) agreement information during sentence processing, but they do not do so when the information is about number.

Why would even advanced ESL learners fail to show reliable sensitivity to the number feature? This is not an unusual result. Many researchers have found that the acquisition of the plural morpheme is difficult for L2 learners, and they tend to remain nonnative-like in their marking of plurals (Long, 2003; Schmidt, 1983; White, 2003). It has been reported that L2 learners, even after a 20-year period of exposure in the target language, may still be unable to use the correct plural forms of nouns in their spoken English. A plausible explanation for the current data is that in Chinese, nouns are not marked for number, so native Chinese speakers are not used to paying attention to number marking during sentence comprehension and do not readily do so in a second language that requires it. There is even some evidence that this inattention to number seeps into other kinds of tasks involving picture-sorting (Lucy, 1992) and sentence-picture matching (Jiang et al., 2008). Interestingly, Chinese reflexives *may* be marked for number: “ta” means third person singular while “ta men” is a third person plural form. This means that number surfaces morphologically in the language. So it would not be the absence of plural marking *in general* that would lead to insensitivity to number agreement violations, rather it is the absence of plural marking on the antecedent noun phrase. By contrast, Chinese nouns carry gender information, just as they do in English, and Chinese reflexives (and pronouns) are also gender-marked.[2] Therefore, Chinese readers are used to attending to gender information, and using it for agreement.

[2] Third-person pronouns and reflexives in Chinese are gender-marked if written in Chinese characters. In spoken language and Pinyin script, this distinction is neutralized. Our subjects would have had vast experience with written language in which the marking occurs.

However, it is also possible that something else is at play here. Even for the native speakers, violations of gender agreement appear to produce stronger effects than violations of number agreement. This may have to do with the fact that the plural marking on a noun in English is rather small in form, usually only one character. Gender, on the other hand, is part of the noun's meaning, and may be difficult to ignore. Both of our participant groups showed a larger effect for gender violations; it just so happened that the smaller effect for number violations did not reach statistical significance in the L2 group.

We turn now to the results of the "implicit causality" study. Results revealed the following. First, native speakers showed a significant slowdown when there was a clash between gender-based antecedent assignment and the "bias" inherent in the verb. This was true for both NP1-biased verbs (as in sentences such as *The landlady angered the policeman because he was cruel to animals*), and NP2-biased verbs (as in sentences like *The girl comforted the boy because she didn't have the chance to go to college*). In both cases, the slowdown effect did not reach statistical significance until after the pronoun; it appears that subjects pressed ahead before resolving the conflict. Non-native speakers showed effects for both verb types, but the timing of these effects differed from the timing for the native speakers. For the NP1-bias verbs, the slowdown was statistically significant only at the pronoun itself, and for the NP2-bias verbs, the slowdown was statistically significant only two words after the pronoun. The pattern of reading times is not identical to that of the native speakers, but in all cases, agreement-violation effects show up. Earlier, it was suggested that there could be an additional effect of "parallel function." In general, subject pronouns that are coreferent with an antecedent that is also a subject are more readily processed than subject pronouns that are coreferent with objects. In these causal constructions, we did not find this to be the case: the magnitude of the violation effect was no larger for NP1-bias verbs than NP2-bias verbs. Overall, these results suggest that the L2 readers were able to keep track of gender information associated with the two nouns and use it to compute agreement. At the same time, they were able to extract information associated with the verb about the roles played by the participants and use this information to compute a causal representation. The fact that these various computations were accomplished during the processing of the sentence once again suggests that sentence comprehension for the nonnative participants was relatively rapid and automatic.

Appendices

Appendix A. Language profile/questionnaire

Language Profile

Subject # _____

We are conducting research on second language acquisition. Please fill in the form as truthfully as possible. Information on this form is kept entirely confidential. Only your number code appears with the information you provide. By completing the questionnaire, you consent to our use of the information.

1. Date:_____
2. Age:_____ 3. Gender:_____ 4. Place of birth:_____
5. Do you have any hearing or vision problems? Yes_____ No_____
6. What is your mother tongue?_____
7. What is your major?_____
8. What is your TOEFL score (if you have one)?_____
9. At what age did you start to learn English? _____
10. Years of residence in the US:_____Year(s)_____Month(s)
11. Years of formal education in total:_____
12. Years of formal education in English speaking countries_____
13. Estimate your level of English on a scale of 1 (beginner) to 5 (advanced)

Speaking	1	2	3	4	5	Listening	1	2	3	4	5
Reading	1	2	3	4	5	Writing	1	2	3	4	5

14. List any other languages you speak besides English and your month tongue_____
    Estimate your level for each: a. Beginner b. Intermediate c. Advanced

Appendix B. Sentences to test L2 learners' sensitivity to featural constraints and verb cues

Number agreement (+ indicate a positive answer and – indicate a negative answer):

1. The careless pedestrians found themselves/himself covered with mud.
   -Were the pedestrians covered with water?
2. The nervous actors calmed themselves/himself down after work.
3. The famous actresses prepared themselves/herself to face the crowd.
4. The dirty soldiers cleaned himself/themselves before going to bed.
5. The emotional sisters felt themselves/herself getting sentimental after the speech.
   -Did the sisters feel happy?
6. The shy girls forced themselves/herself to sing the part at the concert.
   +Did the girls attend the concert?
7. The hungry guests helped themselves/himself to the delicious meal.
8. The tired waitresses poured some coffee for themselves/herself after a hard day's work.
   +Did the waitresses have some coffee?
9. The talented tailor made themselves/himself some beautiful clothes.
10. The careless scientist burned themselves/himself with the dangerous chemical.
11. The successful hunter cleaned themselves/himself off after walking through the woods.
    +Did the hunter walk through the woods?
12. The young mechanic considered himself/themselves to be very handsome.
13. The old lady gathered flowers for themselves/herself every morning during the spring.
    -Did the lady plant flowers during the summer?
14. The busy housewife allowed herself/themselves a one-hour nap.
15. The tired nurse administered the injection herself/themselves without telling the doctor.
    -Did the nurse tell the doctor about the injection?
16. The popular midwife taught herself/themselves how to dance.
    +Was the midwife popular?

Gender agreement:

1. The lonely grandfather made himself/herself a cup of tea.
   -Did the grandfather cook some soup?
2. The anxious cowboy prepared himself/herself for the performance.
3. The nervous son checked himself/herself in the large mirror.
   +Did the son look in the mirror?
4. The young husband found himself/herself without a job or money.
   +Was the husband poor?
5. The desperate boyfriend told himself/herself to forgive the girl.
6. The confused brother left himself/herself a note on the wall.
7. The stubborn gentleman did the work himself/herself during the meeting.
   -Did the gentleman assign the work to someone else?
8. The proud king looked at himself/herself in the mirror in the living room.
9. The hungry waitress ordered herself/himself a big meal.
10. The successful woman congratulated herself/himself on the promotion.
    +Was the woman promoted?
11. The wealthy queen built himself/herself a big castle in the forest.
    -Did the queen build a small castle?
12. The new stepmother prepared herself/himself to meet the family.
13. The youthful grandmother planted a garden for herself/himself in the yard.
    +Was the garden for the grandmother?
14. The brave girl told herself/himself not to worry.
15. The ambitious lady flew herself/himself to Paris for business.
    -Did the lady go to Paris by train?
16. The new aunt introduced herself/himself to the guests.

Implicit causality (verb cues):

1. The mother amused the father because he/she told funny jokes at dinner.
2. The landlady angered the policeman because he/she was cruel to animals.
   +Did the landlady make the policeman angry?
3. The aunt scared the uncle because he/she got mad so easily.
4. The wife bored the husband because he/she told the same story over and over again.
   +/-Did the husband tell the same story many times?
5. The congresswoman disappointed the chairman because he/she lost the election.
6. The gentleman frightened the hostess because he/she carried a bomb into the vehicle.
7. The princess apologized to the prince because he/she was sorry for the mistake.
   +Did the princess apologize to the prince because of the mistake?
8. The brother annoyed the sister because he/she kept lying to people.
   +Was the sister annoyed with the brother?
9. The lady inspired the actor because he/she volunteered at a homeless shelter.
10. The boy amazed the mother because he/she ran so fast in the race.
    -Did the mother feel sad about the race?
11. The woman charmed the policeman because he/she was so adorable in the hat.
    +Did the policeman like the hat?
12. The cowboy cheated the cowgirl because he/she wanted to win the game.
13. The uncle confessed to the policewoman because he/she felt guilty about the incident.
    -Was the policewoman happy about the incident?
14. The grandson fascinated the grandmother because he/she was a wonderful singer.
15. The salesman deceived the girl because he/she wanted to sell the car.
    -Was the car red?
16. The guy humiliated the lady because he/she told the story in public.
17. The boy admired the girl because he/she was so intelligent.
    -Was the boy admired?
18. The nephew hated the niece because he/she cheated all the time.
19. The grandfather congratulated the granddaughter because he/she won the race.
20. The stepfather feared the daughter because he/she was such an aggressive person.
21. The widow envied the bachelor because he/she won the national lottery.
    -Did the widow and the bachelor win the lottery?
22. The girl noticed the man because he/she was wearing such bright colors.
    +Were the clothes brightly colored?
23. The actress praised the actor because he/she had given a good presentation yesterday.
    -Was the presentation given last week?
24. The mother blamed the son because he/she missed the target in the game.
25. The granddaughter adored the father because he/she was the nicest person in the world.
    +Was the father adored by the granddaughter?
26. The woman likes the boy because he/she is very honest.
27. The mother thanked the son because he/she had washed the car.
    +Was the car clean?
28. The doorman assisted the lady because he/she was carrying a large suitcase.
29. The girl comforted the boy because he/she didn't have the chance to go to college.
30. The policeman honored the grandmother because he/she deserved the award for saving people.
    +/-Did the policeman deserve the award?
31. The husband trusts the wife because he/she always tells the truth.
32. The stepfather despised the mother because he/she stole money from a church.
    +Was money taken from a church?

Appendix C. A written test to evaluate explicit knowledge about gender agreement, number agreement and implicit causality in English (to be scrambled in the actual test)

Instructions: Complete the following sentences by circling the correct word in the parentheses

1. The lonely grandfather made_____(himself/herself) a cup of coffee.
2. The overweight cowboy denied____(himself/herself) a hamburger.
3. The successful man congratulated_____(herself/himself) on the promotion.
4. The gracious wife introduced_____(himself/herself) to the guests.
5. The infamous princess looked at _____(himself/herself) in the mirror.
6. The landlady worked_____(himself/herself) into a frenzy.
7. The careless pedestrians found _____(themselves/himself) covered with mud.
8. The famous actresses prepared____ (themselves/herself) to face the crowd.
9. The dirty soldiers cleaned_____(themselves/himself) before going to bed.
10. The mother amused the father because____(he/she) told funny jokes.
11. The congressman disappointed the chairwoman because _____(he/she) lost the election.
12. The brother annoyed the sister because _____ (he/she) kept lying to people.
13. The widow envied the bachelor because _____ (he/she) won the national lottery.
14. The actress praised the actor because _____ (he/she) had given a good presentation.
15. The grandfather congratulated the granddaughter because_____ (he/she) won the reading competition.

References

Aaronson, Doris & Ferres, Steven (1987). The impact of language differences on language processing: An example from Chinese-English bilingualism. In Peter Homet, Michael Palij, & Doris Aaronson (Eds.), *Childhood bilingualism: Aspects of linguistic, cognitive, and social development* (pp. 75–119). Hillsdale, NJ: Erlbaum.

Akiyama, Yasuhiro (2002). Japanese adult learners' development of the locality condition on English reflexives. *Studies in Second Language Acquisition, 24*, 27–54.

Ardal, Sten, Donald, Merlin W., Meuter, Renata, Muldrew, Shannon, & Luce, Moira (1990). Brain responses to semantic incongruity in bilinguals. *Brain and Language*, *39*, 187–205.

Bailey, Nathalie, Madden, Carolyn, & Krashen, Stephen D. (1974). Is there a "natural sequence" in adult second language learning? *Language Learning, 24*, 235–243.

Chen, Chang-Ching (2009). *The representation and processing of past tense in Chinese English-language learners*. Available from ProQuest Dissertations and Theses database (UMI No. AAT 3352346).

Chomsky, Noam (1981). *Lectures on Government and Binding*. Dordrecht, Holland: Foris.

Clahsen, Harald & Felser, Claudia (2006). Grammatical processing in language learners. *Applied Psycholinguistics, 27*, 3–42.

Cole, Peter & Sung, Li-May (1994). Head movement and long distance reflexives. *Linguistic Inquiry, 25*, 355–406.

Dussias, Paola E. (2001). Sentence parsing in fluent Spanish-English bilinguals. In Janet Nicol (Ed.), *One mind, two languages: Bilingual language processing.* Maiden, MA: Blackwell.

Ellis, Nick (1993). Rules and instances in foreign language learning: Interactions of explicit and implicit knowledge. *European Journal of Cognitive Psychology*, *5,* 289–318.

Ellis, Rod (1988). The effects of linguistic environment on the second language acquisition of grammatical rules. *Applied Linguistics, 9*, 257–274.

Felser, Claudia (2005). Experimental psycholinguistic approaches to second language acquisition. *Second Language Research, 21*, 95–97.

Felser, Claudia, Roberts, Leah, Gross, Rebecca, & Marinis, Theodore. (2003). The processing of ambiguous sentences by first and second language learners of English. *Applied Psycholinguistics, 24*, 453–489.

Forster, Kenneth I. & Forster, Jonathan C. (2003). DMDX: A windows display program with millisecond accuracy. *Behavior Research Methods, Instruments, & Computers, 35*(1), 116–124.

Frenck-Mestre, Cheryl & Pynte, Joel (1997). Syntactic ambiguity resolution while reading in second and native languages. *Quarterly Journal of Experimental Psychology*, *50A*, 119–148.

Garnham, Alan & Oakhill, Jane (1985). On-line resolution of anaphoric pronouns: effects of inference making and verb semantics. *British Journal of Psychology, 76,* 385–393.

Guillelmon, Delphine & Grosjean, Francois (2001). The gender marking effect in spoken word recognition: The case of bilinguals. *Memory and Cognition, 29*, 503–511.

Hahne, Anja & Friederici, Angela D. (2001). Processing a second language: Late learners' comprehension mechanisms as revealed by event-related brain potentials. *Bilingualism: Language and Cognition, 4*, 123–141.

Hahne, Anja, Mueller, Jutta L., & Clahsen, Herald (2006). Morphological processing in a second language: Behavioral and event-related brain potential evidence for storage and decomposition. *Journal of Cognitive Neuroscience, 18*(1), 121–134

Hawkins, Roger & Chan, Yuet-hung C. (1997). The partial availability of universal grammar in second language acquisition: the failed features hypothesis. *Second Language Research, 13*, 187–226.

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

Hulstijin, Jan H. (2001). Intentional and incidental second language vocabulary learning: A reappraisal of elaboration, rehearsal and automaticity. In Peter Robinson (Ed.), *Cognition and second language instruction* (pp. 258–286). New York: Cambridge University Press.

Jiang, Nan (2004). Morphological insensitivity in second language processing. *Applied Psycholinguistics, 25*, 603–634.

Jiang, Nan (2007). Selective integration of linguistic knowledge in adult second language learning. *Language Learning, 57*, 1–33.

Jiang, Nan, Masuda, Kyoko, & Wang, Xin (2008, March). *Morphological insensitivity in L2 processing: a universal or L1-specific phenomenon?* Paper presented at the American Association for Applied Linguistics 2008 Conference, Washington, D.C.

Johnson, Jacqueline S., Shenkman, Kenneth D., Newport, Elissa L., & Medin, Douglas (1996). Indeterminacy in the grammar of adult language learners. *Journal of Memory and Language, 35*, 335–352.

Johnson, Jacqueline S. & Newport, Elissa L. (1989). Critical period effects in second language learning: The influences of maturational state on the acquisition of English as a second language. *Cognitive Psychology, 21*, 60–99.

Juffs, Alan (2001). Psycholinguistically oriented second language research. *Annual Review of Applied Linguistics*, *21*, 207–220.

Krashen, Stephen D., Houck, Noel, Guunchi, Paola, Bode, Sharon, Birnbaum, Robert, & Strei, Gerald. (1977). Difficulty order for grammatical morphemes for adult second language performers using free speech. *TESOL Quarterly, 11*, 338–341.

Krashen, Stephen D. & Pon, Pauline (1975). An error analysis of an advanced learner of ESL: The importance of the monitor. *Working Papers on Bilingualism, 7*, 125–129.

Lardiere, Donna (1998). Case and tense in the "fossilized" steady state. *Second Language Research, 14*, 1–26.

Lee, Dami & Schachter, Jacqueline (1997). Sensitive period effects in binding theory. *Language Acquisition, 6,* 333–62.

Long, Michael, H. (1997, October). *Fossilization: Rigor mortis in living linguistic systems?* Paper presented at Second Language Research Forum, East Lansing, MI.

Long, Michael, H. (2003). Stabilization and fossilization in interlanguage development. In Catherine J. Doughty & Michael H. Long (Eds.), *Handbook of second language acquisition* (pp. 487–535). Oxford: Blackwell.

Lucy, John A. (1992). *Grammatical categories and cognition: A case study of the linguistic relativity hypothesis.* Cambridge: Cambridge University Press.

McDonald, Janet L. & MacWhinney, Brian. (1995). The time course of pronoun resolution: Effects of implicit verb causality and gender. *Journal of Memory and Language, 34*, 543–566.

MacLaughlin, Dawn (1998). The acquisition of the morphosyntax of English reflexives by non-native spearkers. In Maria L. Beck (Ed.), *Morphology and its interfaces in second language knowledge*. Amsterdam: John Benjamins.

Marinis, Theodore (2003). Psycholinguistic techniques in second language acquisition research. *Second Language Research, 19*, 144–161.

Nicol, Janet L. (1988). *Coreference processing during sentence comprehension*. Unpublished doctoral dissertation, Massachusetts Institute of Technology, Cambridge, MA.

O'Grady, William (2006). The problem of verbal inflection in second language acquisition. *Proceedings of the 11th Conferences of the Pan-Pacific Association of Applied Linguistics*, 169–190.

Osterhout, Lee & Mobley, Linda A. (1995). Event-related brain potentials elicited by failure to agree. *Journal of Memory and Language, 34,* 739–773.

Perkins, Kyle & Larsen–Freeman, Diane (1975). The effect of formal language instruction on the order of morpheme acquisition. *Language Learning, 25*, 237–243.

Prevost, Philippe & White, Lydia (2000). Missing surface inflection or impairment in second language acquisition? Evidence from tense and agreement. *Second Language Research, 16*, 103–133.

Rigalleau, Francois & Caplan, David (2000). Effects of gender marking in pronominal coindexation. *Quarterly Journal of Experimental Psychology: Section A, 53*, 23–52.

Sabourin, Laura (2003). *Grammatical gender and second language processing: An ERP study*. Available from ProQuest Dissertations and Theses database. (UMI No.AAT CA33213)

Schmidt, Richard W. (1983). Interaction, acculturation, and the acquisition of communicative competence. In Nessa Wolfson & Joan Manes (Eds.), *Sociolinguistics and second language acquisition* (pp. 137–174). Rowley, MA: Newbury House.

Segalowitz, Norman (2003). Automaticity and second languages. In Catherine Doughty & Michael H. Long (Eds.), *The handbook of second language acquisition* (pp. 382–408). Malden, MA: Blackwell.

Segalowitz, Sidney J., Segalowitz, Norman S., & Wood, Anthony G. (1998). Assessing the development of automaticity in second language word recognition. *Applied Psycholinguistics, 19,* 53–67.

Shapira, Rina G. (1978). The non-learning of English: Case study of an adult. In Evelyn Hatch (Ed.), *Second language acquisition* (pp. 246–255). Rowley, MA: Newbury House.

Stewart, Andrew J., Pickering, Martin J., & Sanford, Anthony J. (2000). The time course of the influence of causality structures. *Journal of Memory and Language, 42,* 423–443.

Thomas, Margaret (1993). *Knowledge of reflexives in a second language*. Amsterdam: Benjamins.

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

Vainikka, Anne & Young-Scholten, Martha (1996). Gradual development of L2 phrase structure. *Second Language Research, 12*, 7–39.

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

Wei, Longxing (2000). Unequal election of morphemes in adult second language acquisition. *Applied Linguistics, 21*, 106–140.

White, Lydia (1985). The pro-drop parameter in adult second language acquisition. *Language Learning, 35*, 47–62.

White, Lydia (1998). Second language acquisition and binding principle b: Child/adult difference. *Second Language Research, 14,* 425–439.

White, Lydia (2003). Fossilization in steady state L2 grammars: Persistent problems with inflectional morphology. *Bilingualism: Language and Cognition*, *6*, 129–141.

Yip, Virginia & Tang, Gladys (1998). Acquisition of English reflexive binding by Cantonese learners: Testing the positive transfer hypothesis. In Maria L. Beck (Ed.), *Morphology and its interfaces in second language knowledge*. Amsterdam: John Benjamins.

Yuan, Boping (1998). Interpretation of binding and orientation of the Chinese reflexive *ziji* by English and Japanese speakers. *Second Language Research, 14*, 324–340.

Selected Proceedings of the 2008 Second Language Research Forum: Exploring SLA Perspectives, Positions, and Practices

edited by Matthew T. Prior, Yukiko Watanabe, and Sang-Ki Lee

Cascadilla Proceedings Project Somerville, MA 2010

ISBN 978-1-57473-439-3 library binding

Ordering information

Orders for the library binding edition are handled by Cascadilla Press.
To place an order, go to www.lingref.com or contact:

Cascadilla Press, P.O. Box 440355, Somerville, MA 02144, USA
phone: 1-617-776-2370, fax: 1-617-776-2271, e-mail: sales@cascadilla.com

Web access and citation information

This entire proceedings can also be viewed on the web at www.lingref.com. Each paper has a unique document # which can be added to citations to facilitate access. The document # should not replace the full citation.

This paper can be cited as:

Liu, Rong and Janet Nicol. 2010. Online Processing of Anaphora by Advanced English Learners. In *Selected Proceedings of the 2008 Second Language Research Forum*, ed. Matthew T. Prior et al., 150-165. Somerville, MA: Cascadilla Proceedings Project. www.lingref.com, document #2391.