NATIVE AND NON-NATIVE PROCESSING OF SHORT AND LONG AGREEMENT DEPENDENCIES IN FRENCH

Caitlin E. Coughlin & Annie Tremblay
University of Illinois at Urbana-Champaign

1. Introduction

It is well documented in the second/foreign language (L2) acquisition literature that adult L2 learners do not always show sensitivity to agreement morphology in sentence comprehension (e.g., Foote, 2011; Hopp, 2006, 2010; Jiang, 2004, 2007; Keating, 2009; McDonald, 2006; McDonald & Roussel, 2010; Neubauer & Clahsen, 2009; Sagarra & Herschensohn 2010; Sato & Felser, 2010; Silva & Clahsen, 2008). These difficulties can persist at advanced levels of proficiency, and they can be found even when similar grammatical features are instantiated in the native and target languages (e.g., Hopp, 2006, 2010).

Two hypotheses (among others) have been put forward to explain why L2 learners have difficulty comprehending agreement morphology: (i) a representational deficit hypothesis, according to which L2 learners do not have native-like representations of agreement morphology (e.g., Clahsen et al., 2010; Jiang, 2004, 2007; Keating, 2009; Neubauer & Clahsen, 2009; Sato & Felser, 2010; Silva & Clahsen, 2008); and (ii) a computational deficit hypothesis, according to which the internalized L2 grammar is similar to the native language grammar, but other processes (such as computing inflectional affixes and establishing agreement dependencies between non-adjacent words) can interfere with the systematic use of agreement morphology in sentence comprehension (e.g., Foote, 2011; Hopp, 2006, 2010; McDonald, 2006; McDonald & Roussel, 2010; Sagarra & Herschensohn, 2010). The present study aims to tease these two hypotheses apart by examining native and non-native speakers' online processing of short and long number agreement dependencies in French.

2. Previous Research

One methodology employed to study L2 learners' sensitivity to agreement morphology in sentence comprehension is the self-paced reading paradigm. In self-paced reading tasks, sentences are presented on a screen one word (or region) at a time; the previous word (or region) disappears as the participant presses a button to reveal the next word (or region). Participants thus need to retain lexical and morphosyntactic information of the previous words in memory while they read the rest of the sentence. If the sentences to be read contain agreement violations, sensitivity to these violations is evidenced by increased reading times at the word (or region) containing these violations as compared to sentences in which the corresponding word (or region) does not contain such violations.

In his study of intermediate-to-advanced Chinese L2 learners of English, Jiang (2004) used self-paced reading tasks to investigate L2 learners' sensitivity to agreement violations involving the plural marker *–s*. His stimuli included sentences containing agreement dependencies between a subject noun and the

Actes du congrès annuel de l'Association canadienne de linguistique 2011.
Proceedings of the 2011 annual conference of the Canadian Linguistic Association.

plural form of the verb *be*, and between a plural determiner and a noun; agreement violations were created by removing the plural *–s* from the nouns, as illustrated in (1).

(1) a. The bridges/*bridge to the island were about 10 miles away.

b. The visitor took several of the rare coins/*coin in the cabinet

The results showed that the L2 learners did not demonstrate increased reading times when encountering number agreement violations at the verb (1a) or at the noun (1b), unlike native English speakers. Jiang concluded from these results that his Chinese L2 learners of English did not have native-like knowledge of number agreement rules (for similar results, see Jiang, 2007).

One should acknowledge, however, that the agreement dependencies in Jiang's (2004) experimental items were not local (i.e., several words intervened between the subject noun and the verb, and between the determiner and the noun). In such cases, the speaker is required to retain agreement information in memory over a longer period of time, which can create difficulties and result in reduced sensitivity to agreement violations. It could thus be argued that Jiang's results in fact support a computational deficit hypothesis rather than a representational one.

Sagarra and Herschensohn (2010) used a similar self-paced reading task to examine whether intermediate English L2 learners of Spanish were sensitive to both gender and number agreement in sentence processing. Their experimental stimuli were sentences that contained gender or number agreement violations between a noun and an adjective, as shown in (2). Importantly, the test sentences contained agreement dependencies involving adjacent words, thus eliminating the computational difficulties associated with long agreement dependencies.

(2) a. El ingeniero presenta el prototipo famoso/*famosa en la conferencia.
the-masc-sg engineer-masc-sg presents the-masc-sg prototype-masc-sg famous-masc-sg/*famous-fem-sg in the-fem-sg conference-fem-sg

b. El ingeniero presenta el prototipo famoso/*famosos en la conferencia.
the-masc-sg engineer-masc-sg presents the-masc-sg prototype-masc-sg famous-masc-sg/*famous-masc-pl in the-fem-sg conference-fem-sg

'The engineer presents the famous prototype at the conference.'

The results showed that in such sentences, the L2 learners were in fact sensitive to both gender and number agreement violations, suggesting native-like representations of gender and number agreement rules.

Using a similar task with advanced English L2 learners of Spanish, Foote (2011) examined L2 learners' sensitivity to gender and number agreement violations in local and non-local dependencies. Her experimental sentences contained gender agreement violations between a noun and an adjective, and number agreement violations between a subject and a verb. She compared sensitivity to violations in local dependencies (i.e., where the agreeing word was

adjacent to the word it agreed with) and sensitivity to violations non-local dependencies (i.e., where the agreeing word was distant from the word it agreed with), as illustrated in, respectively, (3a) and (3b).

(3) a. Dicen que el libro blanco/*blanca está en esa mesa.
they-say that the book-masc-sg white-masc-sg/*white-fem-sg is on that-fem-sg table-fem-sg
'They say that the white book is on that table.'

b. El pollo del táco está rico/*rica pero picante.
the-masc-sg chicken-masc-sg of-the-masc-sg taco-masc-sg is tasty-masc-sg/*tasty-fem-sg but spicy-sg
'The chicken in the taco is tasty but spicy.'

The results showed that the L2 learners were sensitive to gender and number violations not only in local dependencies, but also in non-local dependencies. This suggests that these L2 learners had native-like representations of gender and number agreement rules, and the computational difficulty associated with non-local dependencies was not sufficiently large to eliminate L2 learners' sensitivity to gender and number agreement in that condition.

Keating (2009) used a more sensitive methodology, eye-tracking during a reading task, to investigate how native English speakers at different proficiencies in Spanish process gender agreement in local and non-local dependencies. His stimuli included sentences containing gender agreement violations between a noun and an adjective, as shown in (4). In the local condition, the noun and adjective were adjacent (4a), whereas in the non-local conditions, the noun was the subject of the sentence and the adjective was located inside a verb phrase (4b) or a subordinate clause (4c).

(4) a. *Una fiesta pequeño es ideal para una persona tímida o introvertida.
a-fem-sg party-fem-sg small-masc-sg is ideal for a-fem-sg person-fem-sg shy-fem-sg or introverted-fem-sg
'A small party is ideal for a shy or introverted person.'

b. *Una casa es bastante pequeño cuando tiene sólo una habitación.
a-fem-sg house-fem-sg is enough small-masc-sg when it-has only one-fem-sg bedroom
'A house is small enough when it has only one bedroom.'

c. *Una biblioteca no tiene computadoras cuando es pequeño y falta dinero.
a-fem-sg library-fem-sg not have computers-fem when it-is small-masc-sg and lacks money-masc-sg
'A library does not have computers when it is small and does not have enough money.'

The results showed that only the L2 learners at an advanced proficiency were sensitive to gender agreement violations, and unlike native Spanish speakers, they were sensitive to these violations only when the noun and adjective were in

adjacent positions. Keating interpreted these findings as evidence for the Shallow Structure Hypothesis (Clahsen & Felser, 2006), according to which L2 learners' online processing is restricted to shallow morphosyntactic structures. For agreement morphology, this hypothesis predicts that native-like computation of agreement dependencies will be limited to local domains such as "between closely adjacent constituents" (p. 111). Thus, according to Keating, the L2 learners in his study did not represent agreement dependencies in the same fashion as native speakers did. Notice, however, that the non-local agreement dependencies in his study also imposed greater memory demands than the local dependencies, thus not ruling out a computational account.

Recent research has also looked at the non-linguistic factors that might interfere with L2 learners' use of morphosyntactic information in sentence processing. McDonald (2006) (see also McDonald & Roussel, 2010) conducted a series of experiments with native English speakers and L2 learners of English from various language backgrounds to determine whether L2 learners' varying sensitivity to agreement morphology could stem from a computational deficit. In a first series of tasks, she examined whether native and non-native speakers' performance on working memory, lexical access, and decoding tasks could predict their performance on a grammaticality judgment task that contained agreement (and other) violations. These included tense and number agreement violations in verbs, and number agreement violations in nouns, as illustrated in (5).

(5) a. Last night my friend walked/*walk home after dark.

b. The boy jumps/*jump whenever he is startled.

c. There are 20 flutes/*flute in our marching band.

The results showed that L2 learners performed significantly more poorly than native speakers on all the tasks. Moreover, L2 learners' performance on the grammaticality judgment task could be predicted from their performance on the working memory and decoding tasks: those who had better working memory capacity and decoding abilities provided more accurate judgments.

Additionally, McDonald was able to simulate the L2 learners' grammaticality judgment results in a second group of native speakers by applying certain stressors to them. She found that native speakers who were placed under a working memory load or under decoding stress (in this case, noise) showed grammaticality judgments that were similar to those of non-native speakers (who were not placed under such stressors). McDonald concluded that L2 learners' reduced sensitivity to morphosyntactic information stems from a computational deficit rather than from a representational one.

Much of the research looking at L2 learners' sensitivity to agreement morphology in sentence comprehension has yielded inconsistent results. These inconsistencies may be attributable to differences in stimuli and tasks, differences in the agreement paradigms exhibited by the target languages, and differences in L2 learner characteristics (e.g., whether the native language has the corresponding agreement feature, and how proficient L2 learners are in the target language). The present study aims to shed further light on the

representational vs. computational deficit debate by examining whether L2 learners' sensitivity to agreement morphology is modulated by proficiency and by the different working memory demands of short vs. long agreement dependencies. It adds to the previous studies on the processing of agreement morphology in that it focuses on a different target language, French, and unlike Keating (2009) and Foote (2011), it carefully controls the structure of sentences with short and long agreement dependencies, thus manipulating the memory demands of the task while holding syntactic computations constant. For this reason, we will refer to the agreement dependencies used in this study as short vs. long rather than as local vs. non-local.

This study focuses specifically on L2 learners' sensitivity to number agreement violations in third-person direct object clitics that are close to or distant from their antecedents. Third-person direct object clitics in French include *le* 'him/it-masc,' *la* 'her/it-fem,' and *les* 'them'. They agree in gender (when singular) and number with their antecedent. Unlike in English, these pronouns occur in preverbal position (*Marie le mangera* 'Marie it will-eat') and they cannot be stressed (**Marie LE mangera; Marie le mangera LUI* 'Mary it will-eat IT'). L2 learners' sensitivity to number agreement is examined in sentences where the object clitic agrees or does not agree with its left-dislocated antecedent, and where the object clitic is close to or distant from its left-dislocated antecedent, as shown in (6).

(6) a. Ce fruit Marie le/*les mangera pour sa collation avant l'entretien.
this fruit Marie it/*them will-eat for her snack before the interview

b. Ces fruits Marie les/*le mangera pour sa collation avant l'entretien.
these fruits Marie them/*it will-eat for her snack before the interview

c. Ce fruit avant l'entretien Marie le/*les mangera pour sa collation.
this fruit before the interview Marie it/*them will-eat for her snack

d. Ces fruits avant l'entretien Marie les/*le mangera pour sa collation.
these fruits before the interview Marie them/*it will-eat for her snack

'This (these) fruit(s) (before the interview) Marie will eat it/them for her snack (before the interview).'

Left-dislocated phrases are common in spoken French, and they make it possible to assess the participants' sensitivity to morphological information in short and long number agreement dependencies without introducing a context prior to the critical sentences.

Since English pronouns also agree in number with their antecedents, the goal of the present study is not to determine whether non-native speakers can learn to use morphosyntactic information not instantiated in the native language, but rather to investigate the effects of proficiency and working memory demands on L2 learners' sensitivity to agreement violations. We predict that less proficient L2 learners will be less sensitive to agreement violations than more proficient L2 learners, and L2 learners will be less sensitive to agreement violations when the clitic is distant from its antecedent than when it is close to it.

3. Present Study

3.1 Participants

The participants included 31 adult English L2 learners of French (experimental group; age: 19-30; 26 females) and nine native French speakers (control group; age: 22-45, 6 females). All L2 learners grew up speaking only English before puberty, and all native French speakers grew up speaking only French before puberty. All participants were either undergraduate or graduate students at the University of Illinois at the time of testing.

The L2 learners all had at least three semesters of French and were enrolled in at least one advanced (300-level or above) French class at the time of testing. Their proficiency in French was measured with the help of a cloze (i.e., fill-in-the-blank) test (Tremblay & Garrison, 2010; Tremblay, 2011). Two proficiency groups were created based on the cloze test scores: a mid-level group (n=16) and a high-level group (n=15). The L2 learners also filled out a language background information questionnaire in which they reported their age of first exposure to French (AFE), the number of years of instruction they had received on French (YrsInstr), the number of months they had spent living in a French-speaking environment (MthsRes), and their percentage of French use per week (%Use). The L2 learners' cloze test scores and language background information are reported in Table 1.

Table 1. L2 Learners' Cloze Test Scores and Language Background Information

	Cloze(/45)	AFE	YrsInstr	MthsRes	%Use
Mid L2	20	12.6	7.8	4	13.4
(n=16)	(3.6)	(2.9)	(2.4)	(4.2)	(12.2)
High L2	31.5	11.6	10.8	14.8	22
(n=15)	(2.9)	(3)	(3.6)	(11.8)	(13.8)

Note. Mean (Standard Deviation)

3.2 Materials

The participants completed a region-by-region self-paced reading task. The experimental stimuli were 48 unrelated sentences containing a left-dislocated noun phrase and an object clitic (*le, la,* or *les)*. Three variables were manipulated: the distance between the antecedent and the clitic (i.e., short vs. long), the number of the object clitic (i.e., singular vs. plural), and the grammaticality of the object clitic (i.e., agreeing vs. not agreeing in number with its antecedent). The number manipulation was included to determine whether L2 learners' sensitivity to agreement violations was similar for singular and plural clitics. The eight conditions resulting from crossing the three variables are illustrated in (6). Each experimental sentence contained four regions. The critical region was the second region in the short dependency condition (e.g., *Ce fruit | Marie le mangera | pour sa collation | avant l'entretien*) and the third region in the long dependency condition (e.g., *Ce fruit | avant l'entretien | Marie le mangera | pour sa collation*). All the participants completed the same version of the experiment, but each experimental item was seen in only one condition.

The antecedent of the clitic was in the first region of the experimental sentences. The lemma frequency and length of the antecedents were carefully controlled using *Lexique* (New et al., 2001) so that the conditions would be equivalent despite having different lexical words. Nouns that were likely to be known by L2 learners were selected. The antecedents were all preceded by a demonstrative adjective (*ce* 'this-masc/that-masc,' *cette* 'this-fem/that-fem,' or *ces* 'these/those') in order to avoid the definite articles *le* 'the-masc-sing,' *la* 'the-fem-sing,' and *les* 'the-plur,' which could have potentially cued the participants to the grammaticality or ungrammaticality of the object clitics. In the singular condition, half of the nouns were masculine and half were feminine so that the participants' general sensitivity to number agreement (irrespective of gender) would be assessed. The subjects were all proper nouns, and for sentences containing singular clitics, the gender of the subject was the opposite of the gender of the clitic to avoid potential confusability between the subject and the object. The subject-clitic-verb region contained verbs that were frequent and likely to be known by L2 learners, and they were controlled for length across clitic number and grammaticality conditions. The adverbial region separating the antecedent and verb regions in the dependency conditions was similarly controlled for length across clitic number and grammaticality conditions.

The experimental items were pseudo-randomized with 96 distracter items, half of which were grammatical and half of which were ungrammatical. Some of these sentences included subject clitics with right-dislocated antecedents so that the experimental sentences would not stand out.

3.3 Procedures

The participants completed the main experiment in a booth in a quiet room. The self-paced reading task was run with E-Prime (Psychology Software Tools; Schneider, Eschman, &, Zuccolotto, 2002). The sentences were presented one region at a time following a non-cumulative moving-window procedure. A region-by-region presentation was favored over a word-by-word presentation, because it results in a more natural task since object clitics do not occur in isolation in French. At the beginning of each trial, a fixation cross appeared on the left side of the screen. The participants pressed the space bar to advance to the next region. They were instructed to read at a normal pace, but to press the space bar as soon as they finished reading one region to move on to the next. After reading each sentence, the participants were asked in French whether or not the sentence they had just read made sense. It was reasoned that if the participants were sensitive to number agreement violations, they would judge that ungrammatical sentences did not make sense because the clitic in those sentences did not have an antecedent to refer to. The experiment began with 10 practice sentences. The participants received feedback on the accuracy of their responses in the practice session (in which the sentences did not contain number agreement violations involving clitics), but not in the main session. The order of test items in the main session was randomized across participants.

3.4 Data Analysis and Predictions

The dependent variable of main interest in this study is the participants' reading times. Reading times that were longer than 5,000 milliseconds were excluded from the analyses. This resulted in the exclusion of 1.8% of L2 learners' data and 1% of native speakers' data. To reduce potential inter-L2-learner variability, the reading times were then converted into standardized residual times. This was done for each subject by calculating the length of each region in the experimental items, computing a linear regression on the reading times with length as predictor, deriving residual times from the regression equation, and transforming them into z-scores (for details, see Trueswell, Tanenhaus, & Garnsey, 1994). Residual times that were larger or smaller than 2.5 standard deviations from the mean (for each group on each condition) were replaced by the value of these cut-off points. This resulted in the replacement of 1.1% of L2 learners' data and 0% of native speakers' data.

By-subject (F_1) and by-item (F_2) mixed ANOVAs were conducted on the participants' residual times in the critical regions (i.e., the second region in the short dependency conditions and the third region in the long dependency conditions) and in the post-critical regions to capture spill-over effects. Post-hoc analyses were conducted whenever the region variable (critical, post-critical) or, for L2 learners, the proficiency variable (mid, high) interacted with grammaticality. Given the different locations of the critical regions in the short and long dependency conditions, the statistical analyses were run separately for the two dependency conditions. The results of L2 learners and native speakers were also analyzed separately, because the number of participants in the two groups was not balanced. Note that the conversion of reading times into residual times eliminated the speed advantage for more proficient participants. For this reason, only the interactions between proficiency and the other within-subject factors (i.e., clitic number, grammaticality, and region) will be reported.

If participants are sensitive to number agreement violations in object clitics, they should show slower residual times at the critical region when the clitic is ungrammatical than when it is grammatical. If sensitivity to violations changes with proficiency, the effect of grammaticality should be larger for high-level L2 learners than for mid-level ones, and native speakers should be more sensitive to violations than L2 learners. If the length of the agreement dependency increases memory load, the L2 learners, and possibly the native speakers, should be more sensitive to number agreement violations in the short dependency conditions than in the long dependency ones.

4. Results

Figure 1 shows the participants' residual times and standard errors in the short dependency conditions. Note that the same example sentence is used in both conditions only for illustrative purposes. As can be seen from the residual times, all the groups slowed down at the second region when the clitic did not agree in number with its antecedent, but the high-level L2 learners did so more than the mid-level L2 learners, and they did so more when the clitic was plural than when it was singular. Furthermore, the three groups of participants appear to

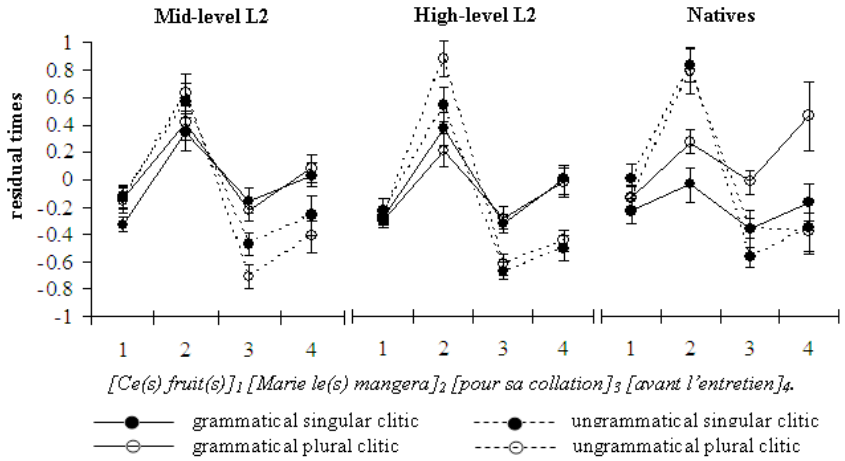

Figure 1. Residual times in the short dependency conditions

have processed the third region faster in the ungrammatical conditions than in the grammatical ones.

Mixed ANOVAs on native speakers' residual times in the second and third regions, with clitic number, grammaticality, and region as within-subject variables, revealed significant effects of clitic number in the subject analysis ($F_1(1,8)=11$, $p<.011$; $F_2(1,5)=3.88$, $p<.106$), grammaticality in the subject analysis ($F_1(1,8)=19.98$, $p<.002$; $F_2(1,5)=4.33$, $p<.092$), and region ($F_1(1,8)=48.25$, $p<.001$; $F_2(1,5)=45.34$, $p<.001$), a significant interaction between clitic number and grammaticality in the subject analysis ($F_1(1,8)=9.15$, $p<.016$; $F_2(1,5)=2.06$, $p<.211$), and a significant interaction between grammaticality and region ($F_1(1,8)=38.56$, $p< .001$; $F_2(1,5)=13.22$, $p<.015$; all other F's<1).

Given the significant interaction between grammaticality and region, subsequent mixed ANOVAs were conducted on native speakers' residual times, with clitic number and grammaticality as within-subject variables, separately for the second and third regions. This time, the alpha value was adjusted to .025 (Bonferroni correction for two comparisons). These analyses yielded a significant effect of grammaticality only in the second region (Region 2: clitic number: $F_1(1,8)=1.29$, $p<.290$; grammaticality: $F_1(1,8)=80.64$, $p<.001$; $F_2(1,5)=26.53$, $p<.001$; clitic number × grammaticality: $F_1(1,8)=2.27$, $p<.170$; $F_2(1,5)=5.44$, $p<.067$; Region 3: clitic number: $F_1(1,8)=7.11$, $p<.029$; $F_2(1,5)=1.89$, $p<.288$; grammaticality: $F_1(1,8)=7.11$, $p<.029$; $F_2(1,5)= 1.89$, $p<.288$; all other F's<1).

Mixed ANOVAs on L2 learners' residual times in the second and third regions, with clitic number, grammaticality, and region as within-subject variables and with proficiency as between-subject variable, revealed a significant effect of region ($F_1(1,29)=125.52$, $p<.001$; $F_2(1,10)=260.53$, $p<.001$), a significant interaction between grammaticality and region ($F_1(1,29)=69.1$, $p<.001$; $F_2(1,10)=40.78$, $p<.001$), a significant three-way interaction between clitic number, grammaticality, and region in the subject analyses ($F_1(1,29)=5.36$,

p<.001; $F_2(1,10)$=1.86, p<.203), and a marginally significant three-way interaction between clitic number, grammaticality, and proficiency ($F_1(1,29)$=3.56, p<.069; $F_2(1,10)$=3.55, p<.089; all other F's<1).

Since the effect of grammaticality tended to vary as a function of region and proficiency, subsequent ANOVAs with clitic number and grammaticality as within-subject variables were conducted separately on the mid- and high-level L2 learners' residual times in the second and third regions. For these analyses, the alpha value (.05) was adjusted to .0125 (Bonferroni correction for four comparisons). The results of these post-hoc analyses, summarized in Table 2, indicate that in the second region, the effect of grammaticality reached significance for the high-level L2 learners and approached significance for the mid-level L2 learners, whereas in the third region, the same effect (though in the opposite direction) reached significance for both L2 groups. Furthermore, in the third region, the high-level L2 learners showed a greater effect of grammaticality with plural clitics than with singular ones.

Table 2. Subsequent Mixed ANOVAs on the Mid- and High-Level L2 learners' Residual Times in the Second and Third Regions, Short Dependency Conditions

		Mid L2						High L2					
		Subject			Item			Subject			Item		
		df	*F*	*p*	*df*	*F*	*p*	*df*	*F*	*p*	*df*	*F*	*P*
Region 2	Clitic Number	1,15	< 1		1,5	< 1		1,14	< 1		1,5	< 1	
	Grammaticality	1,15	4.19	.059	1,5	13.10	.015	1,14	21.22	.001	1,5	4.05	.1
	Clitic Number × Grammaticality	1,15	< 1		1,5	< 1		1,14	10.06	.007	1,5	3.99	.105
Region 3	Clitic Number	1,15	7.41	.016	1,5	2.92	.148	1,14	< 1		1,5	< 1	
	Grammaticality	1,15	19.34	.001	1,5	28.27	.003	1,14	16.23	.001	1,5	9.02	.03
	Clitic Number × Grammaticality	1,15	1.08	.316	1,5	< 1		1,14	< 1		1,5		

These results indicate that the mid- and high-level L2 learners and native speakers were sensitive to number agreement violations when the object clitic was close to its antecedent, but the timing and directionality of this effect were modulated by, respectively, proficiency in French and the nature of the task. Whereas native speakers showed sensitivity to agreement violations only in the critical region, the high-level L2 learners did so in both the critical and post-critical regions and the mid-level L2 learners did so only in the post-critical region, suggesting that agreement violations were detected more rapidly with increasing proficiency. The reverse spillover effects that the L2 learners exhibited in the post-critical region were likely due to the nature of the task: the make-sense question, which was identical for each sentence, may have encouraged them to focus on the form of the sentence, and potentially on grammatical errors, rather than on its meaning; once they detected a grammatical error, they immediately proceeded to the question, resulting in faster reading times for ungrammatical sentences in the post-critical region. The high-level L2 learners also showed greater sensitivity to agreement violations in the critical region when the clitic was plural than when it was singular, a finding to which we will return.

Figure 2 shows the participants' residual times and standard errors in the long dependency conditions. Again, the same example sentence is used in both conditions only for illustrative purposes. As can be seen from the residual times, all the groups slowed down in the third region when the plural clitic was ungrammatical, but only the high-level L2 learners and native speakers slowed down in the third region when the singular clitic was ungrammatical. Again, the participants showed a reverse spill-over effect in the post-critical region.

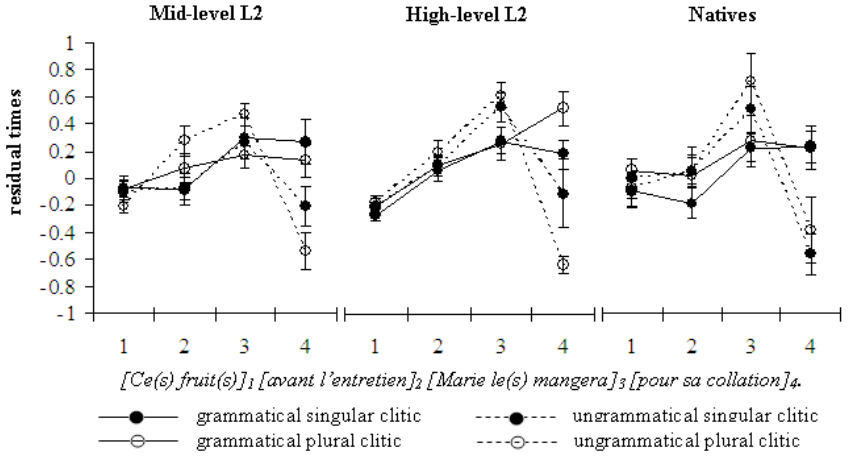

Figure 2. Residual times in long dependency conditions

Mixed ANOVAs on native speakers' residual times in the third and fourth regions, with clitic number, grammaticality, and region as within-subject variables, revealed significant effects of grammaticality in the item analysis ($F_1(1,8)=1.38$, $p<.275$; $F_2(1,5)=10.18$, $p<.024$) and region ($F_1(1,8)=13.81$, $p<.006$; $F_2(1,5)=28.37$, $p<.003$), and a significant interaction between grammaticality and region ($F_1(1,8)=16.92$, $p< .001$; $F_2(1,5)=18.31$, $p<.008$; all other F's<1).

Given the significant interaction between grammaticality and region, subsequent mixed ANOVAs were conducted on native speakers' residual times, with clitic number and grammaticality as within-subject variables, separately for the third and fourth regions. The alpha value was adjusted to .025 (Bonferroni correction for two comparisons). This time, these analyses yielded a significant (reverse) effect of grammaticality only in the post-critical region ($F_1(1,8)=18.88$, $p<.002$; $F_2(1,5)=28.8$, $p<.003$; region 3: $F_1(1,8)=3.09$, $p<.117$; $F_2(1,5)=3.63$, $p<.115$; all other F's<1).

Mixed ANOVAs on L2 learners' residual times in the third and fourth regions, with clitic number, grammaticality, and region as within-subject variables and with proficiency as between-subject variable, revealed significant effects of grammaticality ($F_1(1,29)=9.22$, $p<.005$; $F_2(1,10)=19.37$, $p<.001$) and region ($F_1(1,29)=15.88$, $p<.001$; $F_2(1,10)=68.61$, $p<.001$), a marginally significant interaction between clitic number and grammaticality in the subject analysis ($F_1(1,29)=3.08$, $p<.09$; $F_2(1,10)=1.31$, $p<.28$), a marginally significant

interaction between clitic number and region in the subject analysis ($F_1(1,29)=3.11$, $p<.088$; $F_2(1,10)=2.35$, $p<.156$), a significant interaction between grammaticality and region ($F_1(1,29)=39.05$, $p<.001$; $F_2(1,10)=45.94$, $p<.001$), a significant three-way interaction between clitic number, grammaticality, and region ($F_1(1,29)=12.62$, $p<.001$; $F_2(1,10)=14.70$, $p<.003$) and significant three-way interaction between grammaticality, region, and proficiency in the subject analysis ($F_1(1,29)=6.46$, $p<.017$; $F_2(1,10)=2.56$, $p<.141$; all other F's<1).

Again, since the effect of grammaticality tended to vary as a function of the region and proficiency, subsequent mixed ANOVAs with clitic number and grammaticality as within-subject variables were conducted separately on the mid- and high-level L2 learners' residual times in the third and fourth regions. For these analyses, the alpha value (.05) was adjusted to .0125 (Bonferroni correction for four comparisons). The results of these post-hoc analyses, summarized Table 3, indicate that in the third region, the effect of grammaticality reached significance only for the high-level L2 learners, whereas in the fourth region, it reached significance for both L2 groups. Moreover, in the fourth region, the high-level L2 learners showed a greater (reverse) effect of grammaticality with plural clitics than with singular ones. These results thus parallel those in the critical region of the short dependency conditions.

Table 3. Subsequent Mixed ANOVAs on the Mid- and High-Level L2 learners' Residual Times in the Third and Fourth Regions, Long Dependency Conditions

		Mid L2						High L2					
		Subject			Item			Subject			Item		
		df	*F*	*p*	*df*	*F*	*p*	*df*	*F*	*p*	*df*	*F*	*p*
Region 3	Clitic Number	1,15	< 1		1,5	< 1		1,14	< 1		1,5	< 1	
	Grammaticality	1,15	1.43	.25	1,5	1.24	.316	1,14	14.35	.002	1,5	6.71	.049
	Clitic Number × Grammaticality	1,15	3.02	.10	1,5	3.96	.103	1,14	< 1		1,5	< 1	
Region 4	Clitic Number	1,15	4.675	.047	1,5	7.218	.043	1,14	< 1		1,5		
	Grammaticality	1,15	16.92	.001	1,5	18.31	.008	1,14	13.07	.003	1,5	45.13	.001
	Clitic Number × Grammaticality	1,15	< 1		1,5	< 1		1,14	13.8	.002	1,5	8.76	.032

Hence, even in the long dependency conditions, all the groups were sensitive to number agreement violations in object clitics. This sensitivity was evidenced in the critical and post-critical regions for the high-level L2 learners, but only in the post-critical region for the mid-level L2 learners and native speakers. In this case, proficiency did not have a clear effect on the timing of the participants' detection of agreement violations, as native speakers did not show sensitivity to these violations in the critical region (which could be due to a lack of power, as the F values are greater than 3). In the post-critical region, the high-level L2 learners also showed a greater effect of grammaticality when the clitic was plural than when it was singular, consistent with their results in the critical region of the short dependency conditions.

5. Discussion

The results of the present study showed that when asked whether or not a sentence makes sense, both English L2 learners of French and native French speakers processed object clitics that did not agree in number with their antecedents more slowly than the corresponding grammatical clitics, in both the short and long dependency conditions. Rather than in their sensitivity to grammaticality itself, the mid- and high-level L2 learners differed in the timing of their detection of agreement violations, with more proficient L2 learners detecting them in the critical region and less proficient learners detecting them in the post-critical region. These findings are not surprising, as L2 learners are known to process sentences more rapidly as their proficiency in the target language increases (for discussion, see Frenck-Mestre, 2002).

The results also showed that L2 learners, but not native speakers, took longer to detect number agreement violations that involved singular clitics than those that involved plural clitics. This could suggest that superfluous agreement morphology (e.g., *les* in singular contexts) is easier to detect than missing agreement morphology (e.g., *le* in plural contexts), or that L2 learners' holding of plural antecedents (e.g., *Ces fruits* 'these fruits') in working memory has a detrimental effect on their detection of agreement violations in singular contexts. Contrary to predictions, the results did not provide clear evidence that L2 learners' sensitivity to number agreement violations decreased as distance between the object clitic and its antecedent increased. This might be due to nature of the task, specifically the make-sense question at the end of the sentences, which might have led the participants to focus on form rather than on meaning, resulting in sensitivity to agreement violations in both the short and long dependency conditions. Given the reverse spillover effects incurred by the task, it is uncertain from the present study whether the participants would show similar sensitivity to number agreement violations in both the short and long dependency conditions if they were asked to focus their attention on the meaning of the sentences and answer comprehension questions. Such a task would also reflect processing under more natural conditions, and it should eliminate the reverse spill-over effects observed in the present study.

Coughlin and Tremblay (to appear) used such a method to assess non-native French speakers' sensitivity to agreement violations in object clitics. Their participants were also English L2 learners of French at mid and high proficiency levels. They used a self-paced reading task with stimuli that were quite similar to those in the present study, but instead asked a comprehension question at the end of the sentence, one that was not related to the object clitic in the sentence. Such a question encouraged the participants to read for meaning rather than for form. Their results showed much weaker grammaticality effects in the short dependency conditions for both mid- and high-level L2 learners, with no interaction between grammaticality and clitic number. Furthermore, in the long dependency conditions, only the high-level L2 learners slowed down at ungrammatical clitics, suggesting that the long agreement dependencies posed difficulties for the mid-level L2 learners.

Given that the L2 learners and stimuli in the two studies were very similar, we can perhaps attribute L2 learners' sensitivity to agreement violations in the present study to the make-sense question that followed each sentence in

the self-paced reading task. Nonetheless, our results suggest that proficiency can at least influence the speed with which L2 learners detect agreement violations in online sentence comprehension, and although we did not find that long agreement dependencies reduced L2 learners' sensitivity to agreement morphology, the results of Coughlin and Tremblay (to appear) indicate that they do, suggesting that long dependencies indeed increase memory load. Such findings thus provide additional evidence in support for the computational deficit hypothesis (e.g., Foote, 2011; Hopp, 2006, 2010; McDonald, 2006; McDonald & Roussel, 2010; Sagarra & Herschensohn, 2010).

6. Conclusion

The present study investigated L2 learners' sensitivity to agreement morphology in online sentence processing. The results of a self-paced reading task with mid- and high-level English L2 learners of French and native French speakers indicated that both L2 learners and native speakers were sensitive to number agreement violations between clitics and their antecedents, irrespective of the length of the agreement dependency. However, mid-level L2 learners showed delayed sensitivity to these violations as compared to high-level L2 learners and native French speakers, suggesting that proficiency plays an important role in determining how early agreement information is used in sentence processing. These results, together with those of Coughlin and Tremblay (to appear), suggest that a computational deficit might underlie L2 learners' reduced sensitivity to agreement morphology in sentence comprehension.

References

Clahsen, H., Felser, C., Neubauer, K., Sato, M., & Silva, R. (2010). Morphological structure in native and non-native language processing. *Language Learning, 60*, 21–43.

Clahsen, H., & Felser, C. (2006). Continuity and shallow structures in language processing. *Applied Psycholinguistics, 27*, 107–126.

Coughlin, C. E., & Tremblay, A. (to appear). Proficiency- and working-memory-based explanations for non-native speakers' sensitivity to agreement in sentence processing. *Applied Psycholinguistics.*

Foote, R. (2011) Integrated knowledge of agreement in early and late English-Spanish bilinguals. *Applied Psycholinguistics, 32*, 187–220.

Frenck-Mestre, C. (2002). An on-line look at sentence processing in the second language. In R. R. Heredia & J. Altarriba (Eds.), *Bilingual sentence Processing* (pp. 217–236). Amsterdam: Elsevier.

Hopp, H. (2006). Syntactic features and reanalysis in near-native processing. *Second Language Research, 22*, 369-397.

Hopp, H. (2010). Ultimate attainment in L2 inflection: Performance similarities between non-native and native speakers. *Lingua, 120*, 901–931.

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

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

Keating, G. (2009). Sensitivity to violations of gender agreement in native and nonnative Spanish: An eye-movement investigation. *Language Learning, 59*, 503–535.

McDonald, J. L. (2006). Beyond the critical period: Processing-based explanations for poor grammaticality judgment performance by late second language learners. *Journal of Memory and Language, 55*, 381–401.

McDonald, J. L., & Roussel, C. (2010). Past tense grammaticality judgment and production in non-native and stressed native English speakers. *Bilingualism: Language and Cognition, 13*, 429–448.

Neubauer, K., & Clahsen, H. (2009). Decomposition of inflected words in a second language: an experimental study of German participles. *Studies in Second Language Acquisition, 31*, 403–435.

New B., Pallier C., Ferrand, L., & Matos R. (2001). Une base de données lexicales du français contemporain sur internet: LEXIQUE™//A lexical database for contemporary french: LEXIQUE™. *L'année psychologique, 101*, 447–462.

Sagarra, N. & Herschensohn, J. (2010). The role of proficiency and working memory in gender and number agreement processing in L1 and L2 Spanish. *Lingua*, 2022–2039.

Sato, M., & Felser, C. (2010). Sensitivity to morphosyntactic violations in English as a second language. *Second Language, 5,* 101–118.

Schneider, W., Eschman, A., & Zuccolotto, A. (2002). *E-Prime reference guide*. Pittsburgh: Psychology Software Tools Inc.

Silva, R., & Clahsen, H. (2008). Morphologically complex words in L1 and L2 processing: Evidence from masked priming experiments in English. *Bilingualism: Language and Cognition, 11*, 245–260.

Tremblay, A. (2011). Proficiency assessment standards in second language acquisition research: "Clozing" the gap. *Studies in Second Language Acquisition, 33,* 339–372.

Tremblay, A., & Garrison, M. D. (2010). Cloze-tests: A tool for proficiency assessment in research on L2 French. In M. T. Prior, Y. Watanabe, & S.-K. Lee (Eds.), *Selected proceedings of the Second Language Research Forum 2008* (pp. 73–88). Somerville, MA: Cascadilla Press.

Trueswell, J., Tanenhaus, M., & Garnsey, S. (1994). Semantic influences on parsing: Use of thematic role information in syntactic ambiguity resolution. *Journal of Memory and Language, 33*, 285–318.