Sentence Processing Among Native vs. Nonnative Speakers: Implications for Critical Period Hypothesis

Marzieh Nezakat-Alhossaini[1], Manijeh Youhanaee[2], & Ahmad Moinzadeh[3]

[1]Corresponding author, University of Isfahan, marzieh.nezakat@fgn.ui.ac.ir
[2]&[3]University of Isfahan

Received: 08/06/2014 *Accepted: 09/09/2014*

Abstract

The present study intended to investigate the processing behavior of 2 groups of L2 learners of English (high and mid in proficiency) and a group of English native speakers on English active and passive reduced relative clauses. Three sets of tasks, an offline task, and 2 online tasks were conducted. Results revealed that the high-proficiency group's performance was the same as that of the native group in all the 3 tests in terms of accuracy and processing behavior. Accordingly, it was concluded that proficient L2 learners can achieve native-like performance. Also, opposed to maturational propositions on adult L2 learning following the critical period hypothesis, the high-proficiency participants showed native-like behavior on the L2 structures. We suggest that the amount of exposure to an L2 can be a defining factor for L2 learners to perform in a native-like manner.

***Keywords*:** Sentence Processing; Reduced Relative Clauses; Native/Nonnative Speakers; Online/Offline Tasks; Critical Period Hypothesis

1. Introduction

For decades, SLA studies have been devoted to understanding the differences between L1 and L2 processing, as well as the source of these differences. Some have associated the differences to leaner attributes such as incomplete grammars (e.g., Hawkins & Chan, 1997; Hawkins & Hattori, 2006; Johnson, Shenkman, Newport, & Medin, 1996; Rah & Adone, 2010) or maturational changes during adolescence (Lenneberg, 1967), and some others to processing problems (e.g. Lardier 1998a, 1998b, 2007; Prevost & White, 2000). Therefore, there is still controversy on the different sources proposed for the observed differences between L1 and L2 performances, and on whether or not L2 learners learning an L2 after puberty are able to go beyond the L2 input.

Among different language aspects investigated in studies comparing L1 an L2, morphological features have been found to have a bearing on L2 learners' performance in the L2 (Hopp, 2010). L2 morphology is believed to be among areas prone to fossilization (Lardiere, 1998a, 1998b; White, 2003), and even advanced students exposed to an L2 for years have shown difficulty in morphological features (Larsen-Freeman, 2010). L2 learners' problems with morphological and

morphosyntactic features have been found in both offline grammaticality judgment tests (e.g., DeKeyser, 2000; Johnson & Newport, 1989) as well as online, real-time tests (e.g., Silva & Clahsen, 2008). Thus, there is still controversy on the sources of such difficulties and whether or not nonnatives can overcome these difficulties and attain native-like performances on such features.

Therefore, to find out whether the differences found in L1 and L2 processing are due to incomplete grammars, processing problems, or even maturational issues regarding the degree of automaticity calls for further research involving L2 learners in different stages of language acquisition compared with native speakers. For that reason, the present study tried to examine the processing behaviors of two groups of mid-proficiency (MP) and high-proficiency (HP) L2 learners of English (nonimmersed) with a group of English native speakers regarding an ambiguous morphosyntactic construction, that is, English active and passive reduced relative clauses. Both online and offline measures of language knowledge were applied to assess both grammatical knowledge and processing patterns of the participants.

2. Background

L2 processing of L2 structures in real time, compared with L1 processing, has not been given proper attention in L2 research literature (Felser, Roberts, Marinis, & Gross, 2003; Juffs, 1998; Papadopoulou, 2005; Rah & Adone, 2010). The way L2 learners process language structures in real time can not only guide research regarding L2 language development, but can also help understand the nature of the differences found between L1 and L2 processing.

Processing structural ambiguities has long been an area of research in understanding L1 processing, and has recently drawn L2 researchers' attention to study the differences between L1 and L2 language processing. Processing ambiguous structures has the parser deal with input compatible with more than one grammatical analysis (Clahsen & Felser, 2006a). Research concerning structural ambiguities has revolved around three main areas, namely subject vs. object ambiguities, relative clause attachment ambiguities, and main verb vs. reduced relative clause ambiguities. The main questions under study in these studies have been mostly whether or not there are differences between L1 and L2 processing strategies, and if L2 learners can develop native-like processing regarding these structures.

Regarding subject vs. object ambiguities, Juffs and Harrington (1996) conducted an online grammaticality judgment experiment on a group of advanced Chinese learners of English together with English native speakers. Investigating subject vs. object ambiguities was tested through sentences like (1):

1. *Before Mary ate the pizza arrived from the local restaurant.*

In such sentences, the verb *eat* is optionally transitive and induces a garden-path effect by making the parser comprehend *the pizza* as the direct object of the main verb. The accuracy data obtained from such sentences in this study revealed that both the native speakers and L2 learners tended to reject these sentences. The reaction times (RTs) obtained showed that both groups were surprised when reading these sentences. This finding led to the conclusion that L2 learners experienced similar processing difficulties as native speakers. Juffs (1998b) repeated the same study only this time examining L2 learners with different L1 backgrounds. The results replicated the ones in Juffs and Harrington (1996) study showing the same processing difficulties for both L1 speakers and L2 learners.

The RC attachment ambiguity is illustrated in (2) below. In this sentence, the RC *who was sitting on a chair* is preceded by a complex NP *the doctor of the patient*. The RC, then, can be attached to either of the NPs 'the doctor' or 'the patient':

2. *The man saw the doctor of the patient who was sitting on the chair.*

Felser et al. (2003) studied advanced Greek and German L2 learners of English using both offline and online tests considering RC attachments. The results revealed similar results in offline tests for both native and nonnative speakers. In the online test, however, the native and nonnative speakers showed different preferences. Papadopoulou and Clahsen (2003) found similar results concerning offline and online tests studying participants with different L1 backgrounds.

Like the other two structural disambiguities mentioned above, the garden-path effect happens by misanalysis in sentences such as *the dog tied to the tree escaped*. The problem with processing such sentences arises by the fact that the word 'tied' is considered as the main verb rather than a participle introducing a reduced relative clause, and it is not corrected far later through reanalysis (Ferreira, Christianson, & Hollingworth, 2001). For instance, sentences such as (3) and (4) below are assumed ambiguous in that the participles *raced* and *doomed* are first processed as the main verb of the sentence and need reanalysis to be found out as a participle introducing a relative clause:

3. *The horse raced past the barn fell*. (Bever, 1972)

4. *The movement doomed from the beginning came to a good ending*.

As Ferreira and Henderson (1991, 1998) propose, the sentence comprehension system assigns a thematic role to a phrase when encountering its head. In garden path sentences in which the head (here, *the horse* and *the movement*) and the error signal (here, *raced* and *doomed*) are adjacent, the thematic role is

temporarily assigned to the wrong element and if reanalysis happens it can be reversed.

Juffs (1998) conducted a study on the ambiguity resulted from the main verb and a reduced relative clause in sentences such as *the president defeated in the election resigned a day later*. The study aimed at investigating how adult learners of English as a second language process the above structures. The L2 learners participated in the study were 17 Chinese, 17 Korean or Japanese, and 17 Romance learners (Spanish, Italian, France, and Portuguese) who were all advanced in their level of proficiency in addition to a group of 17 native speakers as a comparison group. The sentences used in this study were a replication of another study by Juffs and Harrington (1995, 1996) on ESL processing of *wh*-gap constructions and Garden Path sentences. The data were collected through conducting moving window, or self-paced reading procedure. The results of the study concerned the accuracy of responses and word-by-word reading times. The native speakers and the speakers of Romance languages were equally accurate compared with the two groups of Chinese and Japanese/Korean speakers, but there was a significant difference between the two groups in favor of the first one. The results from the sentence processing tasks revealed that, first, the L2 learners were slower than the native speakers, and second, parsing was easy in ESL in relation to genetically-related languages. The overall results indicated that although the nonnative speakers were slower than the native speakers, similar patterns were found between the two in some features of the test items.

In another study by Rah and Adone (2010), garden path structures were investigated in reduced relative clauses in German and English. A group of German learners of English who had not learned English in an immersion context participated in the study. The online and offline processing of garden path sentences was examined against those of the native speakers. The results of the offline GJ (Grammaticality Judgment) task revealed no significant difference between the learners and the native speakers, indicating the existence of RC knowledge in the learners. However, a significant difference was found in terms of structure. Ambiguous structures were considered unacceptable for both groups which showed that they were less acceptable for both the learners and native speakers in comparison with unambiguous sentences. The results obtained from the online tests revealed that there were differences between the native speakers and L2 learners in processing patterns. The participants' L1, here German, could be a plausible explanation for the longer response times because RCs are not licensed in this language. Another explanation for longer response times could be that these learners were not able to construct deep syntactic representations for RCs and relied on a shallow representation on the basis of the surface information.

The studies cited above demonstrate a controversy on whether or not nonnatives can develop native-like processing patterns, in the light of similar results for both groups obtained in offline tasks (Juffs, 1998a, 1998b), and if the similar results gained in offline tasks can be indicative of the processing ability in online tasks (Rah & Adone, 2010). Therefore, whether or not the differences found are due to the L2 learners' language competence or processing strategies calls for further research. Furthermore, whether or not the amount of exposure to an L2 provides a more native-like performance compared with nonimmersed L2 learners needs to be further investigated (Dussias, 2003). Regarding the differences between the native speakers vs. nonnative speakers, the maturational aspects of language learning are also worth investigating, as it is necessary to know if more exposure to an L2 can bring automaticity which is a salient feature of native-like performance (Fernandez, 2003).

The abovementioned controversies concerning the existing differences between L1 and L2 processing strategies as well as processing speed and the possible effect of maturational factors in addition to the amount of exposure to the L2 are in need of more investigation and scrutiny. Therefore, the present study intended to have a closer look at these issues by comparing the processing of an ambiguous morphosyntactic construction, that is, English passive reduced relative clauses, between two groups of L2 learners (with mid- and high-proficiency levels) and English native speakers. The reason to have selected this structure was the literature on this structure indicating its ambiguous nature for both native and nonnative speakers which helps investigate the differences between these two groups in terms of processing strategies. Having had two different groups of L2 learners with two proficiency levels helps understanding the effect of the amount of exposure to the L2 and whether or not it can affect the processing strategies and speed among these learners.

As mentioned earlier, passive reduced relative clauses which contain the ambiguity between a main verb and a reduced participle have shown to be ambiguous in the literature (Juffs, 1998a; Macdonald, 1994; Rah & Adone, 2010). In this study, however, a compatible structure to passive reduced relative clauses, that is, active relative clauses (see example (5) below) was under investigation.

5. *The student playing in the ground is my classmate.*

The reason behind this selection was to see whether there was a difference between the processing and judging of active vs. passive voice. Passive voice is regarded to be more difficult in both comprehending and producing as there are more complex underlying structures involved (Marinis & Saddy, 2013). Consequently, this study intended to assess the participants' processing strategies as

well as their offline performance considering these two structures. This study, therefore, aimed to answer the following questions.

1. Are there differences between the offline performance of the L2 learners of English and native speakers regarding the L2 structures?
2. Are L2 learners able to acquire native-like processing strategies in terms of active vs. passive reduced relative clauses?
3. Does the amount of exposure to the L2 input help attain native-like performance on target structures?
4. Is development in L2 competence equal to L2 development in performance?

3. Method

3.1 Participants

Two groups of participants with three proficiency levels took part in this study. The first group was a group of Persian L2 learners of English studying in an English language institute. To select two separate groups of HP and MP L2 learners, the grammar part of Allen's (2004) Oxford Placement Test (OPT), containing 100 items, was conducted; scores were calculated as a proportion of 100 out of 200 based on scores provided by the OPT manual. The volunteers in score range of 71±7 out of 100 and the ones gained scores between 85.5±10 were assigned to MP (n = 15) and HP (n = 16) groups, respectively. The second group of participants were English native speakers (n = 7) recruited on the Internet through putting ads on linguistics forums, such as the Linguist List, Edling, and Cogling, as well as on academic Websites, such as Academia.edu, LinkedIn, and ResearchGate. The L2 participants were ranged in age from 25 to 35, and the native speakers were between 20 and 32 in age. All the three groups of participants took part in offline and online tests.

3.2 Procedure

Three sets of tests, two online (self-paced reading task [SPR], and stop-making-sense task [SMS[), and one offline (error-correction task) tests were conducted for the three groups of participants. All the groups did the online tests first, and they then participated in the offline test. The online tests were administered via DMDX software package (version 4.2.2.0; Forster & Forster, 2003) in two forms of local and remote testing. The local test mode was used for the participants in two groups of MP and HP, and the remote test mode was employed for the native speakers (see http://psy1.psych.arizona.edu/~jforster/dmdx/help/dmdxhremotetestingoverview.htm). A detailed explanation of each testing instrument is given below.

3.3 Instruments

3.3.1 Offline test (error correction task)

This test was conducted to evaluate the participants' competence regarding the target structures. The test was a pencil and paper one with no time pressure. The test contained 48 items consisting of 24 experimental items and 24 fillers. The fillers contained other grammatical points and were irrelevant to the L2 structures; half the items in each set were ungrammatical. The test items were checked by a native speaker for their accuracy. The participants were instructed to identify the wrong items and provide the correct forms. Examples (6) and (7) illustrate a grammatical and an ungrammatical test item:

6. **The patient who advised by his doctor to stop smoking tried to do so.*
7. *The birds noticed on the tree pecked at an insect.*

3.3.2 Self-paced reading task (moving window technique)

This test was performed to evaluate the real-time processing of the L2 structures by the participants. In a self-paced reading technique, the participants had to read sentences presented word-by-word or phrase-by-phrase as fast as possible by pressing a key on a keyboard or a button on a push box. Self-paced tasks have been widely used in studies involving online processing of different structures in SLA (e.g., Hopp, 2010; Papadopoulou & Clahsen, 2003; Roberts & Felser, 2011). These tasks are usually followed by other tasks like comprehension questions or grammaticality judgments to ensure the participants' comprehension of the sentences provided and reducing the possibility of mindless pushing of the buttons. The reaction times (RTs) obtained from pushing a key by a given participant to move from a word to the next reveal how fast a participant can process a piece of information.

The way the words or phrases are presented can take two different fashions: cumulative and noncumulative. In the former, the phrases called by the participant through pushing a button stay on the screen to the last phrase of the sentence. In the latter, on the other hand, the new phrases replace the previous ones and the participants will not be able to look back at the earlier presented phrases.

A noncumulative self-paced reading task (Just, Carpenter, & Woolley, 1982) followed by grammaticality judgments were employed in this study. The test contained 74 items including 32 items on the L2 structure and 42 fillers. The experimental items contained both grammatical and ungrammatical sentences. Example (8) illustrates a test item. The slashes represent the way the items were presented to the participants; the participants could not see the slashes. Regions (2)

and (3), which were the main target for analyses, were matched for length for all items.

8. *The young workers$_1$/supported at work place$_2$/demanded more money$_3$.*

The participants were seated in front of a 14-inch monitor and were instructed to press the Right Shift key to move through the items as quickly as possible but not too quickly to miss a segment. The grammaticality judgment provided at the end of each item was presented to the participants as two options of *grammatical* and *ungrammatical* at the right and left corner of the screen, respectively. Participants were trained to press the Right Shift key for *grammatical* and the Left shift key for *ungrammatical*. Feedback on the accuracy of answers was provided randomly and for only half the experimental and filler items. There were a number of 6 practice items at the beginning to familiarize the participants with the test procedure. The task was about 10 to 13 minutes long, depending on the participants reading speed.

3.3.3 Stop-making sense task

This test again was used to provide an examination to the participants' processing strategies and patterns. In this task, while the phrases were presented on a computer screen, the participants had to identify at which point the sentence becomes implausible by pressing a different button in which case the presentation of the sentence stops. This test was used in order for its results to be compared with that of the self-paced reading task. The logic for this compatibility was the fact that the regions in which the participants took longer to read and parse in the self-paced reading task were to be compared with the exact segment on the stop-making-sense task for which the participants had pressed the key as being plausible or implausible. As with the SPR task, there were 74 items including 32 items on the L2 structures and 42 fillers. The experimental items contained both grammatical and ungrammatical sentences. The test items were the same as the SPR task, and changes were only made to noncritical words. The same software was used for this test as well.

4. Results

The results obtained from the three tests are presented below. The results will be presented according to the accuracy of responses in both online and offline tests as well as reaction times for the online tests. Alpha was set at .05 for all statistical analyses.

4.1 Error Correction Task

As mentioned earlier, in this task, the participants were given a set of both grammatical and ungrammatical items to correct and provide the correct form as

well. The corrected items, then, received 1 and the uncorrected or wrongly corrected items received 0. A one-way ANOVA was conducted to compare the three groups' performance on this test. Table 1 demonstrates descriptive statistics for the groups.

Table 1. *Descriptive Statistics for the Error Correction Task*

	N	Mean	Std. Deviation	Lower Bound	Upper Bound	Minimum	Maximum
Mid	15	7.43	2.78	5.89	8.98	2.00	12.00
High	16	19.00	2.49	17.68	20.32	14.00	23.00
Native	7	21.14	1.54	19.71	22.57	19.50	23.00

The results of the one-way ANOVA revealed a significant difference between the groups' comparing scores in this test $F(2, 35) = 111.48$, $p < .05$. Post-hoc comparisons using the Tukey HSD test indicated that there was a significant difference between the MP group with both native and HP groups ($p < .05$); however, no significant difference was found between the HP and native speaker groups. This finding revealed similar performances for the HP L2 learners and native speakers regarding the offline performance on the target structures.

4.2 Self-Paced Reading Task

The SPR task was implemented to examine the participants' processing strategies on the target structures. The test items on this task were divided into three segments in addition to a grammaticality judgment task at the end of each test item. As stated earlier, region (3) and the GJ at the end of each item were subject to analysis. The results are presented for the accuracy of the judgments, the speed of judgments (RTs), and the RTs obtained from the third region of each test items for passive and active items, respectively.

4.2.1 Passive items

Table 2 presents the descriptive statistics for accuracy of the GJs at the end of each item for the three groups:

Table 2. *Descriptive Statistics for Accuracy Scores on Passive Items' GJs*

	N	Mean	Std. Deviation	Lower Bound	Upper Bound	Minimum	Maximum
Mid	15	4.00	2.84	3.53	4.46	2.00	5.00
High	16	5.62	2.12	4.49	6.75	2.00	8.00
Native	7	5.29	2.05	3.38	7.18	2.00	7.00

A one-way ANOVA was performed to compare the groups for the accuracy scores obtained from the SPR task. A statistically significant difference was found between the groups at the $p < .05$ level in the scores, $F(2, 35) = 3.66$, $p = .03$. The Tukey HSD revealed a significant difference between the MP group and the other two groups, $p = .03$. The HP and native groups, however, performed statistically the same ($p = .9$).

The next part of analysis for the SPR task was the speed of the judgments the participants had made for each item. Table 3 illustrates descriptive statistics for passive items:

Table 3. *Descriptive Statistics for Accuracy RTs on Passive Items' GJs*

	N	Mean	Std. Deviation	Lower Bound	Upper Bound	Minimum	Maximum
Mid	15	6869.71	2150.14	5679.00	8060.42	3545.85	11896.98
High	16	5030.94	2284.25	3813.75	6248.13	1900.46	10690.01
Native	7	3352.55	1600.51	1872.32	4832.78	1669.67	6165.09

The results of the one-way ANOVA conducted between the groups revealed a significant difference in the scores, $F(2, 35) = 7.05$, $p = .003$. The post-hoc Tukey HSD test showed a significant difference between the native speakers and the MP group, $p = .003$. No significant difference was found between the native speakers and the HP L2 learners, $p = .2$. Although the difference between the HP and MP groups was not significant, $p = .054$, the mean scores showed a considerable difference between the RTs obtained from the two groups, indicating a better performance by the former.

The last part of analysis on the SPR for passive items was the RT3 on the third regions. Table 4 below shows the descriptive statistics for the three groups:

Table 4. *Descriptive Statistics for Region 3 RTs on Passive Items*

	N	Mean	Std. Deviation	Lower Bound	Upper Bound	Minimum	Maximum
Mid	15	3276.38	636.82	2923.72	3629.04	2458.84	4556.71
High	16	2651.61	884.25	2180.43	3122.80	695.31	3973.14
Native	7	11205.78	4367.00	7166.98	15244.59	7184.80	17614.45

A statistically significant difference was found between the groups at the $p < .05$ level in RTs on this region, $F(2, 35) = 52.01$, $p < .001$. The Tukey HSD post-hoc test showed a significant difference between the native and HP L2 learners, $p < .001$. No significant difference was found between the MP and HP groups, $p = .64$.

4.2.2 Active items

As stated earlier, the same statistical analyses were performed for the active items. As with the passive items, the accuracy of the GJs for each item, the speed of the participants' judgments, and RTs on region three will be presented, respectively. Table 5 shows descriptive statistics for the accuracy of GJs on active items for the three groups.

Table 5. *Descriptive Statistics for Accuracy Scores on Active Items' GJs*

	N	Mean	Std. Deviation	Lower Bound	Upper Bound	Minimum	Maximum
Mid	15	10.80	2.78	9.26	12.34	4.00	15.00
High	16	12.56	2.09	11.44	13.67	8.00	15.00
Native	7	9.14	3.90	5.54	12.74	4.00	14.00

The one-way ANOVA conducted to compare groups on their scores revealed a significant difference regarding accuracy on GJs, $F(2, 35) = 4.07$, $p = .02$. The post-hoc Tukey HSD test showed a significant difference between the native and HP groups, $p = .02$, indicating better accuracy scores for the HP group, and no significant difference between the MP and native groups, $p = .39$, suggesting the same performance by the two groups.

The next part of the analyses concerned the speed of judgments by the participants on this test. Table 6 presents the descriptive statistics for active items:

Table 6. *Descriptive Statistics for Accuracy RTs on Active Items' GJs*

	N	Mean	Std. Deviation	Lower Bound	Upper Bound	Minimum	Maximum
Mid	15	9068.87	2889.86	7468.51	10669.22	3663.68	13883.07
High	16	8383.96	3268.60	6642.25	10125.68	3946.97	15263.76
Native	7	8592.31	3471.96	5381.29	11803.34	4840.06	14525.95

The one-way ANOVA conducted between groups did not show any significant difference in accuracy RT scores for this test, $F(2, 35) = .18$, $p = .83$.

The last statistical analysis on SPR task for active items was the RTs on regions three. Table 7 illustrates the descriptive statistics for RT scores:

Table 7. *Descriptive Statistics for Region 3 RTs on Active Items*

	N	Mean	Std. Deviation	Lower Bound	Upper Bound	Minimum	Maximum
Mid	15	35284.14	10251.87	29606.84	40961.44	22576.62	56217.58
High	16	26031.72	8384.00	21564.20	30499.24	11175.48	39276.16
Native	7	24780.34	7618.66	17734.25	31826.43	16055.80	33175.90

There was a statistically significant difference between the groups at the $p < .05$ level in RTs on region three, $F(2, 35) = 5.17$, $p = .01$. Post-hoc comparisons using the Tukey HSD test indicated that the mean scores obtained from the native and HP groups were not statistically significant ($p = .95$); however, the MP groups performed significantly different than the HP and native groups ($p = .02$ and .04, respectively).

4.3 Stop-Making-Sense Task

This online test was conducted to gain a more detailed understanding of the way participants processed test items. As stated in the Instruments subsection, the participants were trained to press a key on the keyboard to announce that the test item did not make any sense anymore. It was, therefore, hypothesized that the participants' rate of detecting the third region as ungrammatical should be low if

they mastered the target structure. Therefore, the number of third regions correctly detected was calculated for participants. Table 8 below shows descriptive statistics for the accuracy of regions three for the passive items:

Table 8. *Descriptive Statistics for Regions Three Accuracy for Passive Items*

	N	Mean	Std. Deviation	Lower Bound	Upper Bound	Minimum	Maximum
Mid	15	4.80	1.01	4.23	5.36	3.00	6.00
High	16	6.75	1.29	6.06	7.43	4.00	8.00
Native	7	7.14	1.21	6.01	8.26	5.00	8.00

A significant difference was found between groups considering RTs on the third region, $F(2, 35) = 14.37$, $p < .001$. The post-hoc test Tukey HSD test revealed a significant difference between the MP group with that of the HP and native groups ($p < .001$). No significant difference was found between the native and HP groups' performance ($p = .74$). The same procedure was implemented for active items. Table 9 below shows descriptive statistics for RTs on regions three in the SMS task:

Table 9. *Descriptive Statistics for Regions Three Accuracy for Active Items*

	N	Mean	Std. Deviation	Lower Bound	Upper Bound	Minimum	Maximum
Mid	15	12.40	1.88	11.35	13.44	9.00	15.00
High	16	13.62	2.02	12.54	14.70	10.00	16.00
Native	7	12.71	1.79	11.05	14.37	10.00	15.00

No significant difference was found between the groups in terms of RTs obtained from the third region for active items, $F(2, 35) = 1.63$, $p = .21$.

5. Discussion

This study was an attempt to investigate the processing of an ambiguous morphosyntactic structure (reduced passive relative clauses) and an unambiguous counterpart of it (active reduced relative clauses) among L2 learners of two different proficiency levels (i.e., MP and HP) and to compare them with a group of English native speakers.

The first research question put forward in this investigation was concerned with whether or not there were differences between the offline performance of the L2 learners and the native speaker group in terms of the target structures. The results obtained from the offline error correction task indicated that the MP group performed less accurately than the other two groups, whereas the HP and native speaker groups did not show any significant differences in this test.

The second research question tried to see if L2 learners were able to acquire native-like strategies in terms of the target structures involved in the study. To examine this question, two sets of online tests were conducted with both passive (ambiguous) and active (unambiguous) items, for which the statistical procedures were conducted separately. The scores obtained from the self-paced reading task were analyzed for three different aspects, namely the RTs on the third region of items, the accuracy of the judgments at the end of each item, and the speed of the judgment, for passive and active items each. The results obtained from the accuracy of the judgments for passive items revealed that the MP group could not judge the ambiguous passive items as well as the HP and native groups. The HP group, however, showed the same accuracy as the native group, indicating that the L2 participants had comprehended these items in the same way as the native speakers. This finding added to what was found in the offline test, where the HP participants' performance was similar to that of the native speaker group. The same results were obtained on active items, where the HP group, unlike the MP group, performed similarly to the native group in judging these items.

Analysis on the speed of judgments for passive and active items yielded the same results pointing to similar performances by the HP and native groups; however, comparisons between the MP and HP groups did not reveal a significant difference, despite the fact that the mean scores obtained by the HP group showed a considerably better, and in this case faster, performance (M = 6869.71 and 5030.94 for the MP and HP groups, respectively). Comparisons made regarding active items, on the other hand, showed no significant differences between groups, which surprisingly suggested the same RTs for the L2 participants and the native group on active items. As explained in the background, ambiguous structures have been shown to be problematic for L2 learners as well as native speakers, though with less difficulty. Active items, as an unambiguous counterpart to the ambiguous structure here, however, were selected to see if the same pattern is true for easier, less ambiguous structures, that is, whether the same difficulty in processing is found between the two groups. As observable from the findings on active vs. passive items here, even the MP level learners were able to read and judge these structures close to native speakers.

The last aspect of the SPR task analyzed was the speed of judgments on the third region of passive and active items. As mentioned before, RTs on this region were very important to understand the *hows* of processing these items by the participants. As it was illustrated in example (10) above, the second region of passive test items contained the reduced participle, which is believed to be taken as the main verb of the sentence. As the participants move through the third region, containing the main verb of the sentence, an increase in reading times is expected as the parser needs to reanalyze what it was comprehended on the first and second regions. Therefore, how fast they read the third region can be indicative of a faster or slower reanalysis and how much they have been garden-pathed. The same hypothesis can be true for active items only that the participle in an active RRCs (see (7) above) is less likely to be analyzed as the main verb compared with passive RRCs. This hypothesis was supported by the findings obtained from the analysis of the RTs on third regions for both target structures.

The analysis on passive items' third region RTs revealed significant differences among groups, showing that the native speakers had the lowest RTs on this region, followed by the HP and then MP groups. This finding demonstrates that the native speakers were also garden-pathed, indicated by longer RTs on the third region compared with the second region; however, this ambiguity was resolved much faster for them compared with the L2 learners. The HP learners were then faster than the MP ones, showing that ambiguity resolution can be speeded by gaining proficiency in an L2.

The same analysis for the active items showed different results. The RTs obtained from third regions in active items showed no difference between groups, suggesting that the three groups performed as fast and were not garden-pathed compared with the passive items, which was indicated by the same RTs for regions two and three for these items.

To further investigate the second research question, the third regions on the SMS task was also subject to analysis. As mentioned previously, the participants were to press a predefined key on an item to announce where it did not make sense anymore to them. To back up the results gained from the SPR task, the third regions in this test were checked for the participants' plausibility judgments. The results pointed to the same findings for the SPR task. Regarding passive items, there was a significant difference between the MP group and the other two groups; however, the native and HP groups performed the same. This was in accord with the accuracy data found for the SPR task. For the active items, however, no significant difference was found among groups, indicating the same interpretation of the structures while reading the third regions. This finding again was in line with what was found for the SPR task.

The third research question was concerned with whether the amount of exposure made an L2 learner's performance more native-like. The results obtained for the MP group, for both offline and online tasks, revealed a significant difference between this group and the HP group, which relates a better parsing with a higher proficiency level. The similar performance obtained for the HP group with that of the native group was itself evidence to the matter. The only area of difference found between the HP and native groups was the speed of processing on the third region of passive items. Having in mind the ambiguous nature of these items, the native speakers showed they were able to resolve this ambiguity faster than the HP group. However, they did not outperform the HP group in terms of accurate judgments or the way they processed these structures. This can point to the fact that the L2 learners opted for the same parsing strategies and comprehension, despite being slightly slower in processing.

The fourth research question asked whether development in L2 competence could be equal to L2 development in performance. By comparing the similar results gained from the offline error correction test, testing the L2 learners' competence, to that of online tests, examining their performance, the present study came to the idea that, regarding the target structures involved here, the L2 learners' competence was equal to their performance. That is, the MP group showed a less accurate performance in both offline and online tasks. On the other hand, the HP group showed a more accurate performance, compared with the native speakers, on both offline and online tasks.

Overall, the results of this study showed that the HP L2 participants performed similarly with the native group on the offline test implying similar metalinguistic knowledge of the target structures. This finding is in line with Juffs (1998), Rah and Adone (2010), and Hopp (2010) in which similar performances were found for advanced L2 learners with that of native speakers in terms of offline tasks. The MP L2 learners, however, were not able to perform accurately in the offline task, which implies they did not have access to the necessary grammatical information.

Regarding the results of the online tests pointing to the processing strategies employed by L2 vs. native speakers, unlike the previous studies (e.g., Juffs, 1998a; Rah & Adone, 2010) in which the L2 learners showed to be less confident in their decisions, the L2 learners in this study showed the same level of confidence as the native speakers in their decisions. This is in line with Hopp (2010, 2006) finding native-like processing strategies among proficient L2 learners. The L2 learners in this study, however, showed to be slower, or in other words, less automatic than the native speakers. Other studies in the literature (e.g., Dussias, 2003; Fernandez, 2003; Hopp, 2010, 2006; Segalowitz, 1986) have also pointed to

the fact that the L2 learners participating in their studies were slower in their processing compared with the native speakers. Some have attributed this to L1 effects on L2 processing, making L2 learners' reading slower than the native speakers.

Furthermore, the findings of this study found evidence against ideas concerning maturational issues (Lenneberg, 1967) in SLA. Some researchers have contributed the difference found between the L2 and native speakers' performances to maturational issues proposing that adult L2 learners are not able to acquire an L2 in a native-like manner (e.g. Hawkins & Chan, 1997; Silva & Clahsen, 2008; Ullman, 2004, 2005). Some others, on the other hand, have found evidence against maturational propositions and have found native-like performances by proficient L2 learners (e.g., Felser, et al., 2003; Hopp, 2006, 2010; Mueller, 2005; Papadopoulou & Clahsen, 2003; Weber-Fox & Neville, 1996; Williams, Mobiuset, & Kim, 2001). The present study also adds to the latter by finding a similar, that is, native-like, performance by the L2 learners participating in this study.

Finally, the significant differences found between the MP and HP L2 learners in this study, with a better performance for the latter, point to the fact that the amount of exposure to an L2 can lead to a native-like performance. This is in line with exposure-based processing models (MacWhinney, 1987, 1997, 2001; Mitchell, Cuetos, Corley, & Brysbaert, 1995) in which L2 processing preferences are attributed to the frequency of exposure to certain structures.

6. Conclusion

The findings of the present study revealed that the HP L2 participants were able to have a native-like performance regarding active and passive RRCs on both online and offline tests. Furthermore, the difference found between the MP and HP learners, with a better performance for the latter, suggested that more exposure to an L2 can help develop native-like proficiency. This, in turn, provided support against the maturational propositions on L2 learning suggesting that adult L2 learners cannot achieve native-like proficiency.

The differences mentioned between the MP and HP learners can have implications for language classrooms in both EFL/ESL contexts. First, this findings can be indicative of the fact that it is possible to achieve native-like proficiency, even in complex morphosyntactic constructions like RRCs. Second, the very important issue of increased exposure to an L2 becomes of outmost importance, as according to this finding, more exposure to an L2 makes a difference in linguistic performance even to a native-like level.

This study was not without its limitations. As the data were collected in a laboratory setting, caution needs to be made in generalizing the results to

naturalistic, real-life situations. Though the online tests were designed to be run as fast as possible and there was no space for distractions, the researchers did not have any control over the remote mode data, which requires cautions in exercising generalizations.

Concerning the sources of difference found between L1 and L2 processing and performance, maturational issues are considered as one of the possible explanations proposed so far. L1 transfer, specifically when it comes to complex and ambiguous structures, are arguably important to be considered as well (see Frenck-Mestre, 2002). Accordingly, more studies having L2 learners from different L1 backgrounds, different learning contexts, and proficiency levels are needed to further investigate the sources of difficulties found in L2 processing and performance.

References

Allen, D. (2004). *Oxford placement test*. Oxford: Oxford University Press.

Bever, T. G. (1972). The cognitive basis for linguistic structures. In J. R. Hayes (Ed.), *Cognition and the development of language* (pp. 279-352). New York: Wiley.

Clahsen, H., & Felser, C. (2006). Grammatical processing in language learners. *Applied Linguistics, 27*, 3-42.

DeKeyser, R. (2000). The robustness of critical period effects in second language acquisition. *Studies in Second Language Acquisition, 22*, 499-534.

Dussias, P. (2003). Syntactic ambiguity resolution in second language learners: Some effects of bilinguality on L1 and L2 processing strategies. *Studies in Second Language Acquisition, 25*, 529-57.

Felser, C., Marinis, T., & Clahsen, H. (2003). Children's processing of ambiguous sentences: A study of relative clause attachment. *Language Acquisition, 11*, 127-163.

Felser, C., Roberts, L., Marinis, T., & Gross, R. (2003). The processing of ambiguous sentences by first and second language learners of English. *Applied Psycholinguistics, 24*, 453-489.

Fernandez, E. (2003). *Bilingual sentence processing: Relative clause attachment in English and Spanish*. Amsterdam: John Benjamins.

Ferreira, F., Christianson, K., & Hollingworth, A. (2001). Misinterpretation of garden-path sentences: Implications for models of sentence processing and reanalysis. *Journal of Psycholinguistic Research, 30*(1), 3-20.

Ferreira, F., & Henderson, J. M. (1991). Recovery from misanalyses of garden-path sentences. *Journal of Memory and Language, 25*, 725-745.

Ferreira, F., & Henderson, J. M. (1998). Syntactic reanalysis, thematic processing, and sentence comprehension. In J. D. Fodor & F. Ferreira (Eds.), *Reanalysis in sentence processing* (pp. 73-100). Dordrecht: Kluwer.

Forster, K. I., & Forster, J. C. (2003). DMDX: A Windows display program with millisecond accuracy. *Behavior Research Methods, Instruments, & Computers, 35*, 116-124.

Frenck-Mestre, C. (2002). An online look at sentence processing in the second language. In R. R., Heredia & J., Altarriba (Eds.), *Bilingual sentence processing* (pp. 218-236). Amsterdam: Elsevier.

Hawkins, R., & Chan, Y. (1997). The partial availability of UG in second language acquisition: The 'failed functional features hypothesis'. *Second Language Research, 13*, 187-226.

Hawkins, R., & Hattori, H. (2006). Interpretation of English multiple *wh*-questions by Japanese speakers: A missing uninterpretable feature account. *Second Language Research, 22(3)*, 269-301.

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 nonnative and native speakers. *Lingua, 120*, 901-931.

Johnson, J., & Shenkman, K., Newport, E., & Medin, D. (1996). Indeterminacy in the grammar of adult language learners. *Journal of Memory and Language, 35*, 335–352.

Johnson, J., & Newport, E. (1989). Critical period effects in second language learning: The influence of maturational state on the acquisition of English as a second language. *Cognitive Psychology, 21*, 60-99.

Juffs, A. (1998a). Main verb vs. reduced relative clause ambiguity resolution in L2 sentence processing. *Language Learning, 48*, 107-147.

Juffs, A. (1998b). Some effects of first language argument structure and morphosyntax on second language processing. *Second Language Research, 14*, 406-24.

Juffs, A., & Harrington, M. (1995). Parsing effects in second language sentence processing: Subject and object asymmetries in *wh*-extraction. *Studies in Second Language Acquisition, 17*, 483-516.

Juffs, A., & Harrington, M. (1996). Garden-path sentences and error data in second language processing research. *Language Learning, 46*, 286-324.

Just, M.A., Carpenter, P. A., & Wolley, J. D. (1982). Paradigms and processes in reading comprehension. *Journal of Experimental Psychology: General, 111*, 228-238.

Lardiere, D. (1998a). Case and tense in the 'fossilized' steady state. *Second Language Research, 14*, 1-26.

Lardiere, D. (1998b). Dissociating syntax from morphology in a divergent L2 end-state grammar. *Second Language Research, 14*, 359-375.

Lardiere, D. (2007). *Ultimate attainment in second language acquisition: A case study*. Mahwah, NJ: Lawrence Erlbaum Associates.

Larsen-Freeman, D. (2010). Not so fast: A discussion of L2 morpheme processing and acquisition. *Language Learning, 60*(1), 221-230.

Lenneberg. E. (1967). *Biological foundations of language*. New York: John Wiley & Sons.

MacDonald, M. C. (1994). Probabilistic constraints and syntactic ambiguity resolution. *Language and Cognitive Processes, 9*, 157-201.

MacWhinney, B. (1987). Applying the competition model to bilingualism. *Applied Psycholinguistics, 8*, 315-27.

MacWhinney, B. (1997). Second language acquisition and the competition model. In J. Kroll & A. De Groot (Eds.), *Second language acquisition* (pp. 113-42). Mahwah, NJ: Erlbaum.

MacWhinney, B. (2001). The competition model: The input, the context and the brain. In P. Robinson (Ed.), *Cognition and second language instruction* (pp. 69-90). Cambridge: Cambridge University Press.

Marinis, T., & Saddy, D. (2013). Parsing the passive: Comparing children with specific language impairment to sequential bilingual children. *Language Acquisition, 20*(2), 155-179.

Mitchell, D.C., Cuetos, F., Corley, M.B., & Brysbaert, M. (1995). Exposure-based models of human parsing: Evidence for the use of coarse-grained nonlexical statistical records. *Journal of Psycholinguistic Research, 24*, 469-88.

Mueller, J. (2005). Electrophysiological correlates of second language processing. *Second Language Research, 21*, 152-174.

Papadopoulou, D. (2005). Reading-time studies of second language ambiguity resolution. *Second Language Research, 21*, 98-120.

Papadopoulou, D., & Clahsen, H. (2003). Parsing strategies in L1 and L2 sentence processing: A study of relative clause attachment in Greek. *Studies in Second Language Acquisition, 25*, 501-28.

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

Rah, A., & Adone, D. (2010). Processing of the reduced relatives clause verb ambiguity in L2 learners at different proficiency levels. *Studies in Second Language Acquisition, 32*, 79-109.

Roberts, L., & Felser, C. (2011). Plausibility and recovery from garden-paths in second language sentence processing. *Applied Psycholinguistics, 32*, 299-331.

Segalowitz, N. (1986). Skilled reading in the second language. In J. Vaid (Ed.), *Language processing in bilinguals: psycholinguistic and neurophysiological perspectives* (pp. 3-19). Hillsdale, NJ: Erlbaum.

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

Ullman, M. (2004). Contributions of memory circuits to language: The declarative/procedural model. *Cognition, 92*, 231-270.

Ullman, M. (2005). A cognitive neuroscience perspective on second language acquisition: The declarative/procedural model. In C. Sanz (Ed.), *Mind and context in adult second language acquisition: Methods, theory, and practice* (pp. 141-178). Georgetown: Georgetown University Press.

Weber-Fox, C. M., & Neville, H. J. (1996). Maturational constraints on functional specializations for language processing: ERP and behavioral evidence in bilingual speakers. *Journal of Cognitive Neuroscience, 8*, 231-256.

Williams, J. Mobiuset, P., & Kim, Ch. (2001). Native and nonnative processing of English *wh*-questions: Parsing strategies and plausibility constraints. *Applied Psycholinguistics, 22*, 509-540.

White, L. (2003). Fossilization in steady state L2 grammars: Persistent problems with inflectional morphology. *Bilingualism: Language and Cognition, 6*, 128-141.