

Relative Clause Attachment in Native and L2 Arabic¹

Abdelaadim Bidaoui
Ball State University
USA

Rebecca Foote
University of Illinois
at Urbana Champaign, USA

Mahmoud Abunasser
Oracle
USA

ملخص

تشير الدراسات السابقة إلى وجود تباين في كيفية ربط الاسم الموصول بصلته من لغة إلى أخرى ، حيث يفترض أن ذلك يرجع إلى الخصائص التركيبية بلغة معينة (Gibson et al. : 1996). للأسف معظم الأبحاث الهادفة إلى توثيق هذه الاختلافات أجريت على اللغات الهندو-أوروبية فقط. الهدف الأول من هذه الدراسة هو بحث كيفية ربط الاسم الموصول بصلته بالنسبة للناطقين باللغة العربية . كيفية ربط الاسم الموصول في هذه اللغة السامية لم يحظ بأي بحث أو دراسة أو نشر. الهدف الثاني هو بحث كيفية ربط الاسم الموصول بصلته في اللغة الثانية (L2) بالنسبة للمتعلمين للغة العربية. بعض البحوث تشير إلى أن المتعلمين يفشلون في إظهار أي تفضيل متعلق بربط الاسم الموصول بصلته في اللغة الثانية، وقد تم الافتراض أن ذلك يعود إلى عدم قدرة المتعلمين على تطبيق مبادئ تحليل العبارة (Clahsen&Felser: 2006a) مثل الناطقين باللغة العربية. ولذلك فإننا نتساءل عما إذا كان المتعلمون للغة العربية سوف يظهرون أي تفضيل في كيفية ربط الاسم الموصول بصلته ، وإذا كان الأمر كذلك، هل سيكون نفس التفضيل المذكور بالنسبة للناطقين باللغة العربية. من أجل الإجابة على هذه الأسئلة، شارك ستة عشر ناطقا باللغة العربية (L1) وستة عشر من المتعلمين (L2) للغة العربية في تجربة مع قياس زمن الإجابة- (online self-paced reading) تجربة أخرى بدون قياس زمن الإجابة (offline attachment preference). نتائج تحليل البيانات عن طريق ANOVAs و T-Tests تشير أن كلا من الناطقين والمتعلمين أظهروا نفس كيفية تفضيل ربط الاسم الموصول بصلته (high attachment) في تجربة دون قياس زمن الإجابة ، في حين أظهر فقط المتعلمون للغة العربية (L2) نفس كيفية تفضيل ربط الاسم الموصول بصلته في تجربة مع قياس زمن الإجابة. تضيف هذه الدراسة إلى معرفتنا بمعالجة الجملة بالنسبة للناطقين باللغة العربية وكذلك للمتعلمين للغة العربية التي لم تحظ بدراسة كيفية ربط الاسم الموصول بصلته بالنسبة للطرفين. نتائج هذه الدراسة تدعم التفسير التركيبي لكيفية ربط الاسم الموصول. كما أنها تعطي أدلة ضد الادعاء بأن المتعلمين لا يستخدمون مبادئ تحليل فهم الجمل شبيهة بالناطقين.

Abstract

Previous studies indicate that there is cross-linguistic variation in relative clause attachment, hypothesized to be due to language-specific structural characteristics (Gibson et al.: 1996). However, research documenting these differences has been carried out on Indo-European languages. The first goal of this study is to investigate relative clause attachment preferences in native speakers of Arabic, a Semitic language in which there are no published data on relative clause attachment preferences. The second goal is to investigate relative clause attachment preferences in second language (L2) learners of Arabic. Some research indicates that L2 learners fail to show any clear attachment preferences in their second language, which has been hypothesized to be due to their

¹ We would like to thank the anonymous reviewers for their constructive feedback and genuine evaluations. We also would like to thank Dr. Eman Saadah for her help in participant recruitment. Any remaining errors or oversights are our own.

inability to apply the phrase-structure based parsing principles (Clahsen & Felser: 2006a) that native speakers employ. We therefore ask whether learners of Arabic will show a relative clause attachment preference, and if so, whether it will be the same preference shown by native speakers. In order to answer these questions, sixteen native speakers of Arabic and sixteen late L1 English learners of L2 Arabic participated in one online task (self-paced reading) and one offline task (attachment preference).

Data were analyzed via ANOVAs and t-tests. Analyses indicated that both native speakers and L2 learners showed a preference for high attachment of relative clauses in the offline preference task, but only the L2 learners showed the same preference in the online task. This study adds to our knowledge of sentence processing in both native and L2 Arabic, an understudied language in which relatively little research exists on either population. The findings of this study support structurally-based accounts of relative clause attachment preferences. They also provide evidence against the claim that L2 learners do not use native-like parsing principles in sentence comprehension.

Keywords: L2 Arabic, sentence processing, relative clause attachment preferences, Shallow Structure Hypothesis

1. Introduction

A number of studies that have examined sentence processing in a variety of languages indicate that there are cross-linguistic differences in how speakers attach relative clauses (RCs) to their appropriate antecedents. Specifically, findings show differences in how native speakers of different languages prefer to disambiguate sentences such as the following, in which there is more than one noun phrase that can serve as the antecedent of the RC:

(1) *Someone shot the servant of the actress who was on the balcony.*

For example, in English, speakers tend to interpret the RC (“who was on the balcony”) as describing the second noun phrase (“the actress”), while in Spanish, the preference is to interpret it as describing the first (“the servant”). Researchers have hypothesized that this difference may be due to language-specific structural characteristics which, along with universal strategies, determine how speakers parse these types of structures. Most research that has been carried out in this area, however, has focused on RC disambiguation in Indo-European languages. The first goal of the current study is therefore to extend this research to Arabic, a Semitic language which has been the focus of comparatively few experimental studies, and which to our knowledge, has no published empirical data on RC attachment preferences. In the only study that we know of that

has addressed RC attachment in Arabic, Abdelghany and Fodor (1999) argue that Arabic speakers favor attaching RCs to the second noun phrase in sentences like (1). They attribute this attachment preference to the nature of Arabic prosody, claiming that it favors attachment to the second noun phrase (“low attachment”), while the prosody of other languages may favor attachment to the first noun phrase (“high attachment”). However, a different account of the mechanism behind cross-linguistic differences in RC attachment asserts that attachment preferences are based on a particular language’s structure, with languages that allow greater distances between verbs and their complements being more likely to favor high attachment (Gibson Pearlmutter, Canseco-Gonzalez, & Hickok: 1996). According to this “predicate proximity” account, Arabic should be one of the languages that favors high attachment, as it allows the verb to be distant from its complements. Our first research question is therefore: What is the RC attachment preference of native speakers of Arabic², low (as predicted by Abdelghany & Fodor, 1999) or high (as predicted by the predicate proximity account; Gibson et al.: 1996)?

The second goal of the current study is to examine the RC attachment preference of second language (L2) learners of Arabic. The nature of sentence processing mechanisms in L2 learners has been the topic of a number of recent studies, with the debate about the functioning of these mechanisms centering on whether L2 learners who have acquired their second language after puberty make use of the same types of information as native speakers do when processing sentences in the L2. According to the Shallow Structure Hypothesis (SSH; Clahsen & Felser: 2006a), the syntactic representations constructed by late L2 learners “are shallower and less detailed than those of native speakers” (p. 32). Because of this, learners are forced to rely on non-structural rather than structural cues to meaning when parsing the L2. With respect to the disambiguation of RCs, which has been examined in several L2 sentence processing studies, L2 learners do not seem to show attachment preferences in their second language that match either the L2 or their native language in the absence of disambiguating lexical cues; instead, they show no preference for either high or low attachment. Clahsen and Felser (2006a) argue that this provides evidence that late learners are not able to use the same structurally-based processing principles that native speakers do. Instead, when other types of cues to disambiguation are absent,

² The language used in the experimental sentences is Standard Arabic. Since no one currently is considered a native speaker of Standard Arabic, the L1 group consists of speakers of five dialects of Arabic from: Morocco, Jordan, Palestine, Saudi Arabia, and Yemen. These speakers are fluent in Standard Arabic and are used as a control group in this study.

they make decisions randomly. Our second research question is therefore: Do late L2 learners of Arabic show a RC attachment preference in Arabic? If so, what is their preference? Does it reflect the preference of native speakers of Arabic? In order to answer our research questions, native speakers and late L2 learners of Arabic completed two experimental tasks, an offline attachment preference task, and an online self-paced³ reading task.

2. Background

2.1. L1 Studies: Crosslinguistic Variation in RC Attachment

Studies conducted on RC attachment classify languages into two types: those that show a preference for high attachment and those that show a preference for low attachment. In general, studies investigating L1 speakers of English indicate a preference for low attachment for RC disambiguation (e.g., Carreiras & Clifton, 1999; Cuetos & Mitchell, 1988; Gibson et al.: 1996). According to Papadopoulou and Clahsen (2003), low attachment has also been observed in Norwegian, Swedish, Romanian, and Brazilian Portuguese. On the other hand, a preference for high attachment has been observed for languages such as Spanish (Cuetos & Mitchell: 1988), Dutch (Brysbaert, 1996), French (Frenck-Mestre & Pynte: 1997), and Greek (Papadopoulou & Clahsen: 2003).

In spite of the findings of distinct attachment preferences cross-linguistically, some researchers have argued that there is a uniform manner of ambiguity resolution across languages. “Right association” (Kimball: 1973), “late closure” (Frazier: 1987) and “recency” (Gibson et al.: 1996) are processing strategies that have been proposed to provide an explanation for a low RC attachment preference. The main idea behind these strategies is that, based on a need to choose whatever demands less cognitive effort, the parser will attach the relative pronoun to the second noun phrase rather than the first. This is because the second noun phrase is closer to the relative pronoun and in terms of working memory load, it consumes fewer resources. A preference for low attachment is thus argued to be universal, as it is based on cognitive needs which are not connected to language-specific rules.

In contrast to this view, other researchers have argued that language-specific characteristics play a role in relative clause attachment, in addition to the universal processing strategies outlined above. For example, Gibson et al. (1996) claim that attachment preference is

³ While the reading times are crucial in online experiments, this is not the case for offline experiments.

based on the structure of a given language in addition to the universal “late closure” or “recency” processing strategy. They describe a structurally-based account for RC disambiguation called “predicate proximity”. According to this account, ambiguous elements are attached to the constituent that is closest to the head of the predicate (i.e., the verb). This analysis gives preference to NP1 and predicts high attachment. However, this prediction depends on the strength of this principle in a particular language. These claims were based on studies of English and Spanish RC disambiguation (e.g., Cuetos & Mitchell: 1988). The results obtained for English speakers support a “late closure” analysis in which the RC is attached to the NP that is closest to the relative pronoun, but that was not the case for Spanish speakers, who favored high attachment. The difference between English and Spanish attachment preferences was attributed to the differing strengths of predicate proximity in the two languages. The strength of predicate proximity in a given language is related to the distance between the verb and its complements. Specifically, the greater this distance is in a language on average, the more strongly the predicate will be activated in order to allow attachments across longer distances. Because the predicate is more strongly activated, there is a greater preference for attachment to the predicate in ambiguous structures, and a cost when attaching to other locations (Gibson et al.: 49). English, as a language with SVO word order, does not allow long distances between the verb and its arguments. Therefore, predicate proximity in English is relatively weak, which in turn, means that the default late closure strategy is more likely to be employed, which results in low relative clause attachment. On the contrary, Spanish, a language that allows both SVO and VOS word order, permits more distance between the verb and its arguments. This gives more strength to predicate proximity and makes it more likely that high attachment of RCs will be preferred. Thus, if a language allows longer distances between a verb and its arguments, it is more likely that in this language, predicate proximity will outrank other universal factors such as late closure. In terms of how this analysis may apply to Arabic, based on the fact that Arabic allows the verb to be distant from its arguments, one can predict that predicate proximity may outrank late closure. That is, high attachment will be favored over low attachment.

2.2. RC Attachment Preferences in an L2

Building on L1 studies on RC attachment, studies on RC attachment in L2 have attempted to shed light on similarities and differences between L1 and L2 processing strategies. Dussias

(2003) examined the RC parsing strategies of L2 learners of Spanish and English in comparison to native speakers, examining whether L1 attachment preferences are transferred to the L2, or whether they are dependent on the L2 alone. She found that both L1 Spanish-L2 English and L1 English-L2 Spanish groups preferred low attachment in English and in Spanish. In contrast, high attachment was preferred by the native Spanish controls, whereas low attachment was preferred by the native English controls. Based on these results, it was argued that the learners' attachment preferences were based on a universal set of parsing strategies, regardless of their native language, due to the demands of processing an L2. Papadopoulou and Clahsen (2003) conducted a study with advanced L2 learners of Greek who had one of three L1s: German, Russian, or Spanish, all of which favor high attachment. Results indicated that, in the absence of lexical cues to disambiguation, native speakers showed a preference for high attachment, whereas the learners showed considerable variability in their attachment preferences. This lack of significant preference for a particular attachment was unexpected, given the similarity of attachment preferences in the learners' L1s to the attachment preference in Greek. Based on these results, Papadopoulou and Clahsen argued that learners are not able to apply the phrase structure-based parsing principles described in the previous section, such as late closure or predicate proximity, even when their L1 and L2 are similar in how these principles operate. Felser, Roberts, Marinis, and Gross (2003) conducted a similar study with L2 learners of English whose L1 was Greek or German, and obtained similar results; while native speakers of English typically show a preference for low attachment, L2 learners patterned differently. Specifically, in the absence of lexical cues to disambiguation, they did not show either a high or a low attachment preference. Whether their L1 was German or Greek did not play a role in the pattern of results. These results indicated that the two groups of learners failed to process genitive relative clauses the way native speakers do, providing further support for the claim that learners do not rely on structural information to the same extent that native speakers do during sentence processing.

2.3. The Shallow Structure Hypothesis

Based on the results of the studies described above along with others focused on other aspects of L2 sentence processing, Clahsen and Felser posit that in the context of L2 sentence processing, "the syntactic representations adult L2 learners compute for comprehension are shallower and less detailed than those of native speakers" (2006a: 32). They point out that,

according to proposals based on research with native speakers (e.g., Ferreira, Bailey, & Ferraro, 2002, as cited in Clahsen & Felser: 2006a), the human processing system allows for two ways of computing sentence representations during language comprehension: full parsing and shallow parsing. The latter is based on lexical and pragmatic information, whereas the former is based on grammar (Clahsen & Felser, 2006b: 117). Specifically, while full parsing involves the construction of a fully-specified syntactic representation, shallow parsing involves the identification of parts of speech, the segmentation of the input string into meaningful chunks, and the determination of what the relationships between these chunks and the main verb are. The Shallow Structure Hypothesis (SSH) argues that L2 learners are restricted to shallow rather than full parsing, either due to the fact that, though the same processing strategies are available for their L1 and their L2, their L2 knowledge source is constrained, or due to a lack or a problem in parsing mechanisms (2006b: 117). According to the SSH, this explains why processing complex syntactic representations remains a source of difficulty even for advanced learners. Because L2 learners can generally only compute grammatical representations with simple hierarchical structures, they are forced to rely on their L2 lexical and pragmatic knowledge along with their world knowledge to assign meaning to L2 sentences, instead of relying on the detailed syntactic structure that native speakers can project. In terms of RC disambiguation, this means that L2 learners are not able to use phrase structure-based parsing principles in their RC attachment choices, either because they do not project a “sufficient amount of structure” or possibly, because the structure that they do project is not in the appropriate form for their syntactic processor to operate on it (Clahsen & Felser, 2006: 33). The result of this inability, based on the studies described above, is that L2 learners do not show a RC attachment preference in their L2 when lexical cues are not present in the sentence. This is the case even when a learner’s L1 shows the same pattern of RC attachment as his or her L2, as in Papadopoulou and Clahsen (2003).

Building on this claim, the current study examines RC disambiguation in native speakers and late L2 learners of Arabic in an offline preference task, and an online self-paced reading task. The studies on which Clahsen and Felser base the SSH have all been conducted with learners of Indo-European languages. This study expands the languages under investigation to ask whether this hypothesis can be extended to apply to the L2 processing of Semitic languages. If the SSH can account for the processing of Arabic, then we expect to find that while native speakers show a clear preference for either the high or the low attachment of RCs, L2 learners will not. Before

going into more detail regarding the current study, we first describe the structure of Arabic as it relates to RCs.

2.4. RC Attachment in Arabic

Arabic is a Semitic language that allows both SVO and VSO word order (Ferhri, 2013). The default word order for subject RCs in Arabic is similar to English, with the RC introduced by a relative pronoun that agrees in gender and number with the human noun it refers to (in standard Arabic) as shown in (2):

- (2) *raʔajtu sadiiqa (NP1) ar-raʔuli (NP2) ʔal-laði jaskunu fii ʔal-bajti*
see-1m/fs friend-ms the man who-ms live-3ms in the house
'I saw the friend of the man who lives in the house.'

In sentence (2) the relative pronoun can refer to either NP1 or NP2 as the gender of both NPs is identical to the gender of the relative pronoun; this creates ambiguity in terms of which NP the RC is modifying. RC ambiguity resolution in Arabic has not been widely studied. The only research that we are aware of is Abdelghany and Fodor (1999), who argue that Arabic speakers favor low attachment. Abdelghany and Fodor's study is based on an examination of the effect of prosody on ambiguity resolution. Their claim of low attachment for Arabic stems from the Implicit Prosody Hypothesis (Quinn, Abdelghany, & Fodor: 2000). According to this hypothesis, every ambiguous sentence has a reading that the natural prosody (i.e., intonation, phrasing, and rhythm) of the sentence favors. Thus, in silent reading, a default prosodic contour can be projected onto the stimulus, where it may influence syntactic ambiguity resolution. Abdelghany and Fodor claim that the prosody of Arabic favors low attachment while the prosody of French, for example, favors high attachment. However, if Gibson et al.'s (1996) predicate proximity account applies to Arabic, then a high attachment preference would be expected, given that Arabic allows the verb to be distant from its complements. More research is needed to determine which account is correct, particularly given the lack of information readily available about the Abdelghany and Fodor's study.

3. The Current Study

The current study examines how native speakers and late L2 learners of Arabic disambiguate RCs. We ask two research questions: First, what is the RC attachment preference of

native speakers of Arabic, low or high? The only existing study on RC attachment in Arabic predicts that native speakers will favor low attachment, based on prosodic considerations. In contrast, based on the fact that Arabic allows the head verb to be distant from its arguments, the processing principles described above predict that predicate proximity may outrank late closure, so that high attachment will be favored over low attachment. The second research question is: Do late L2 learners of Arabic show a preference for relative clause attachment, and if so, what is their preference? Does it reflect the preference of native speakers? If L2 learners transfer their L1 (English) RC attachment preference, they are likely to favor low attachment. If, on the other hand, the SSH is correct, and they do not project a detailed enough representation of L2 sentences to make use of parsing principles, they may show no preference either way (in the absence of lexical cues). This is generally what has been found in previous relative clause attachment studies with L2 learners, as discussed above. Finally, they may show a preference that is consistent with the native speaker preference, which will only be distinguishable from transfer from English if the native speakers show a preference for high attachment. An offline preference task and an online self-paced reading task were employed to answer the two research questions.

4. Method

4.1. Participants

The participants in the study were sixteen native speakers of Arabic (mean age 32.9, range 28-45) and sixteen native speakers of English learning Arabic as an L2 (mean age 21, range 19-25), all students at a large Midwestern university in the United States. According to a language background questionnaire, the native speakers were from five different Arabic-speaking countries: Morocco, Jordan, Palestine, Saudi Arabia, and Yemen. Their mean length of stay in the US was 4.88 years (range 1-17), and mean years of education at the university level was 6.2 (range 4-10). Of the 16 L2 learners, 13 were studying Arabic at the time of participation, and three had taken Arabic courses in previous semesters. Their mean number of years of Arabic study was 2.75 (range 1-5). Their mean age of first exposure to Arabic was 19.66 years (range 18-22), and mean years of education at the university level was 4.33 (range 3-7). Six of the learners had studied Arabic abroad in an Arabic-speaking country.

In addition to the language background questionnaire, all participants completed a cloze test created by the first author as a rough indicator of proficiency in Arabic. The topic of the cloze

test was the current political situation in the Arab world. In the text, 23 function words and 22 content words were missing. Participants were asked to fill in the missing words. The mean native speaker score on the cloze test was 44.5 out of 45 (range 43-45); L2 learners' mean score was 36 out of 45 (range 28-45), suggesting that these learners were of intermediate to advanced proficiency in Arabic.

4.2. Materials

Participants completed two experimental tasks: an attachment preference task and a self-paced reading task. The design of each task is described below.

4.2.1. Attachment Preference Task

The attachment preference task was designed to determine participants' relative clause offline attachment preference in the absence of disambiguating cues. The task contained 24 experimental items and 48 fillers, all distinct from the items included in the self-paced reading task. In the experimental sentences, the gender of the relative pronoun and the gender of the two NPs was always feminine, and the verb did not provide any lexical cues to the disambiguation of the sentence. Participants were asked to read each sentence and answer a question to indicate their attachment preference, as in (3).

(3) *I saw the servant of the actress who usually eats a cake on the balcony.*

Who eats on the balcony?

A. The servant

B. The actress

4.2.2. Self-paced Reading Task

The self-paced reading task was designed to determine attachment preferences during online processing. This task included 24 experimental sentences, along with 120 fillers, for a total of 144 sentences. All sentences were divided into six regions⁴, which appeared on the computer screen one at a time in the center of the screen. Participants pressed the spacebar in order to move from one sentence region to another; the time in milliseconds (ms) that it took a participant to

⁴ It is common practice in self-paced tasks to divide the experimental sentences into regions with the goal of tacking the reading times of the critical regions such as the region including the verb.

move from one region to another served as the dependent variable. Two-thirds of the sentences were immediately followed by a comprehension question in order to verify that the participants were paying attention to sentence meaning while completing the experiment. Items were presented in a randomized order that was different for each participant.

All of the experimental items were sentences that contained RCs that could attach to one of two noun phrases. Two versions of each item were created: one in which the meaning of the verb disambiguated the sentence such that the RC was attached to the first NP (high attachment), and the other in which the meaning of the verb disambiguated the sentence such that the RC was attached to the second NP (low attachment). The gender of the relative pronoun and the gender of the two NPs were always feminine, so that only the verb provided disambiguating information in these sentences. The two versions (high and low attachment conditions) of an example experimental item are shown in (4a) and (4b), along with their division into the six regions. The critical region in these sentences is the one that contains the verb; this region is shown in bold below.

(4a) *High attachment*

raʔajtu sʻadiqata l-kaatibati ʔa-lati daʔiman tunaðʻifu al-bajt fi sʻ-sʻsabaah
see-1m/fs friend-fs the writer-fs who-fs usually clean-3fs the house in the morning

*I saw / the friend-fem (NP1) of the writer-fem (NP2) / who-fem / usually / **cleans***

0 1 2 3 4

***the house** / in the morning.*

5

(4b) *Low attachment*

raʔajtu sʻadiqata l-kaatibati ʔa-lati muʔaxaran naʔarat kitaaban ʔan tʻ-tʻulaab
see-1m/fs friend-fs the writer-fs who-fs recently publish-3fs a book about students

*I saw / the friend-fem (NP1) of the writer-fem (NP2) / who-fem / recently / **published a***

0 1 2 3 4

***book** / about students.*

5

Two counterbalanced experimental lists, as shown in the two examples above, were created so that a particular participant saw only one version of each experimental item, the high or the low attachment version.

The logic of this task is as follows: when a disambiguating cue is present in the sentence, if a participant makes use of that cue, he or she should slow down when the cue leads to a disambiguation that does not match his or her attachment preference in comparison to when the cue leads to a disambiguation that does match his or her attachment preference. This slowdown in reading time may take place in the region where the cue is located, or the region following the one where the cue is located (the spillover region).

4.3. Procedure

Participants were tested individually in a quiet room. First, they were asked to fill out a questionnaire in which they were asked to provide personal information such as age, education, languages spoken, and years of exposure to Arabic. The goal behind having a questionnaire is to make sure that the participants have the expertise required to take part in the experiment. Participants were then asked to complete the cloze test. After that, they completed the self-paced reading task, followed by the attachment preference task⁵.

5. Results

5.1. Attachment Preference Task

For the purposes of analyzing participants' RC preferences, responses were assigned a "0" if they indicated a preference for the second NP as the referent of the relative clause, and "1" if they indicated a preference for the first NP as the referent. Hence, a mean preference score from 0.0 to 0.5 indicates low attachment and from 0.5 to 1.0 indicates high attachment. According to this score, the native speakers and L2 learners both showed a preference for high attachment, with a mean of 0.64 for each group (standard error of 0.05). Two different statistical analyses were conducted on the data, one to determine whether there were statistical differences between groups in their attachment preferences, and one to determine whether participants' mean scores indicated a clear attachment preference, high or low. First, an independent samples t-test was conducted on preference scores with group as the between-participants variable. The analysis confirmed that there were no differences between native speakers and L2 learners in terms of attachment preference [$t(30)=0.03$, $p=.973$]⁶. The second analysis was performed in order to

⁵ Exampmls of experimental sentences are provided in (4a) and (4b).

⁶ Since the dependent variable is actually nominal rather than continuous, and thus these data violate the assumptions of the independent samples t-test, we also employed a mixed-effect logistic regression model in R to analyze the data

determine whether participants' mean preference scores were significantly different from 0.5 (which would indicate no preference). A one-sample t-test with 0.5 as the test value was conducted, and confirmed that the mean preference scores did differ from 0.5 [$t(31)=3.78$, $p<.01$]. Based on this analysis, it seems that native Arabic speakers as well as L2 learners show a preference for high RC attachment. This contradicts the claims of Abdelghany and Fodor (1999), but coincides with the predictions of Gibson et al.'s (1996) predicate proximity account. Now we turn to the results of the self-paced reading task in order to examine online RC attachment preferences.

5.2. Self-paced Reading Task

5.2.1. Comprehension Questions

Both groups were relatively accurate in their responses to the comprehension questions⁷, although as expected, the L2 learners were less so. Table 1 displays mean accuracy rates by group.

Table 1. Mean percentage accuracy on self-paced reading task comprehension questions by group (standard deviation in parentheses)

	Comprehension question response accuracy
Native speakers	90.1 (6.2)
L2 learners	77.4 (10.8)

Two different statistical analyses were conducted on the accuracy rates. The first was to determine whether there were significant differences in accuracy between the two groups. The second was to determine whether each participant group's accuracy rate was significantly above chance. This analysis is important to conduct, as a rate that did not significantly differ from chance would indicate that participants were not reading the sentences for meaning. For the first

(R Development Core Team, 2015), using the *afex* library in order to obtain p values for significance testing (Singmann, Bolker, & Westfall, 2015). The results were the same as the t-test ($p = .97$), confirming no differences between groups in their attachment preference.

⁷ It is important to note that the comprehension questions are related to the content of the experimental sentences and have nothing to do with the attachment preferences. The goal of the comprehension questions is just to make sure that the participants are not randomly doing the experiment.

analysis, an independent samples t-test was conducted, with group as the between-participants variable. This analysis confirmed that the L2 learners were significantly less accurate than the native speakers in their responses to the comprehension questions [$t(30)=4.34, p<.001$]⁸. One-way t-tests similar to the one conducted on the preference task data confirmed that both participant groups' accuracy rates were significantly above chance (50%) [native speakers: $t(15)=26.34, p<.001$; L2 learners: $t(15)=10.18, p<.001$], indicating that participants were reading sentences for meaning. We now turn to analyses of the reading times, which are examined at the sentence regions before, at, and after the region with the disambiguating verb. The region before is included as a baseline at which there should be no difference between the sentences that differ in their verbs. The region after is included in order to capture any spillover effect.

5.2.2. Reading Times

Residual reading times at each region by group are listed in Table 2 and depicted in Figures 1 and 2⁹.

Table 2. Residual reading times in ms at each sentence region by group

	Before		At		After	
	Low	High	Low	High	Low	High
Native speakers	-5	1	66	75	6	-9
L2 learners	84	-45	563	442	-99	-341

Figure 1. Residual reading time in ms at each sentence region: native speakers

⁸ As with the analysis of the preference task data, the accuracy rates for the comprehension questions were analyzed using a mixed-effect logistic regression model in R. Again, results matched those of the t-test ($p<.001$), confirming that the L2 learners were significantly less accurate than the native speakers in their responses.

⁹ The residual reading time is the mean difference between the raw reading time and the reading time expected on the basis of region length. It is calculated by determining a linear regression equation between the word length and reading time for each participant, and then subtracting the predicted reading time from the observed reading time for each data point (see, for example, Ferreira & Clifton, 1986). This means that positive residual reading times are slower than predicted based on the regression equation, while negative residual reading times are faster than predicted.

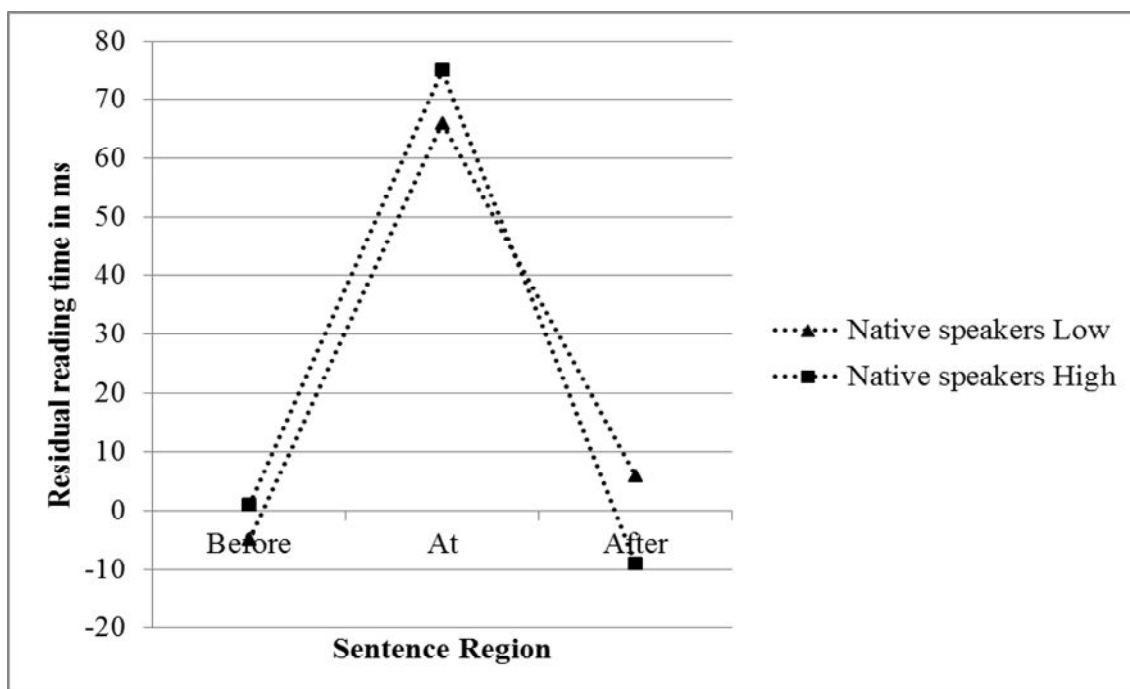
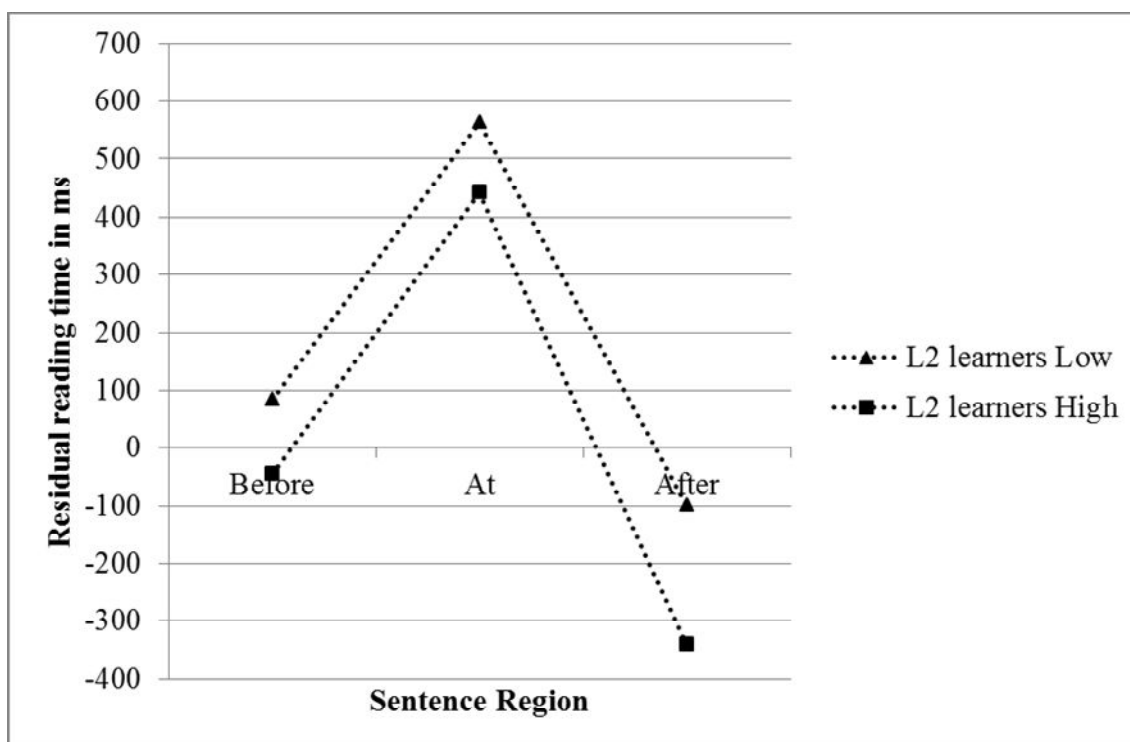


Figure 2. Residual reading time in ms at each sentence region: L2 learners



If participants prefer high RC attachment in the absence of disambiguating cues, they should show a slowdown when reading the portion of the sentence that contains the verb when it indicates that low attachment is correct. Based on Figures 1 and 2, it does seem that both groups' reading times are slower in the region immediately after the verb, when that verb indicates low attachment. L2 learners' reading times also seem to be slower in the other regions of sentences with verb cues that indicate low RC attachment. Mixed ANOVAs were conducted on residual reading times at each sentence region. In Region 1 (before the occurrence of the disambiguating verb), there was no effect of attachment [$F(1, 30)=2.21, p=.147$], no effect of group [$F(1, 30)=0.22, p=.646$], and no interaction between the two [$F(1, 30)=2.72, p=.110$]. In Region 2 (the region where the disambiguating verb occurred), again there was no effect of attachment [$F(1, 30)=0.29, p=.596$], although in this region there was an effect of group [$F(1, 30)=17.71, p<.001$], because the L2 learners' reading times were longer than predicted. There was no interaction between attachment and group, however [$F(1, 30)=0.39, p=.538$]. In contrast to the other regions, in Region 3 (the region immediately after the disambiguating verb), there was a main effect of attachment [$F(1, 30)=6.65, p<.05$], with faster reading times when the verb indicated high attachment than when it indicated low attachment. There was an effect of group [$F(1, 30)=20.02, p<.001$], this time because the L2 learners' reading times were shorter than predicted. There was also an interaction between attachment and group [$F(1, 30)=5.19, p<.05$]; while both groups' reading times were faster when the verb indicated high attachment, this effect was larger in the L2 group (a difference in residual reading times of 242 ms versus only 15 ms in the native speaker group). Paired-samples t-tests conducted separately on residual reading times in this region revealed that the effect was significant in the L2 learner group [$t(15)=2.52, p<.05$], but not in the native speaker group [$t(15)=0.56, p=.585$]. This pattern of results indicates a clear online preference for high attachment in the L2 learners that reflects their offline preference. However, the native speakers do not show a clear online preference, although numerically, reading times suggested a preference for high attachment.

6. DISCUSSION

While the only existing study on RC attachment in Arabic predicts that native speakers will favor low attachment based on prosodic considerations (Abdelghany & Fodor: 1999), Gibson

et al.'s (1996) predicate proximity account predicts a high attachment preference based on allowed word orders in Arabic. In the present study, the results of the offline attachment preference task indicated that native speakers show a statistically significant preference for high attachment of RCs, in line with the predicate proximity account. This suggests that the fact that Arabic allows the head verb to be distant from its arguments is more influential in how syntactic ambiguity is resolved during silent reading of Arabic than the effects of an implicit default prosodic contour¹⁰. It also implies that the universal parsing principle of late closure is overridden by the strength of predicate proximity in Arabic, at least in this type of task. In the online reading task, native speakers showed a numeric preference for the high attachment of relative clauses, but this difference did not reach statistical significance.

As for the L2 learners in the current study, although their L1 (English) tends toward a low attachment preference, they showed the same high attachment preference in Arabic as the native speakers did. This suggests that their attachment preference was not influenced by their L1 (although we did not test them in English, so we cannot state with certainty that this is the case). Instead, they seem to have adopted a nativelike attachment preference in Arabic. The lack of an L1 transfer effect is in line with what other researchers have found in L2 relative clause processing (e.g., Dussias, 2003; Felser et al., 2003; Papadopoulou & Clahsen: 2003). However, the significant preference for high over low attachment differs from previous research that seemed to support the SSH. While the L2 learners of Felser et al. (2003) and Papadopoulou and Clahsen (2003) did not show a clear preference for either high or low attachment in the absence of lexical cues, the learners of the current study did, suggesting that L2 learners are able to make use of parsing principles like native speakers do, both in offline and online tasks.

One question that arises with respect to the native speaker data in this study is why the native speaker participants did not show a statistically significant attachment preference in the self-paced reading task. One possibility is that, although the offline attachment preference task indicated a significant preference for high attachment, this preference may not have been enough to withstand the time pressure and burden on working memory that a phrase-by-phrase self-paced reading task imposes¹¹. Recall that low attachment is proposed to be universal due to the parser's

¹⁰ It is important to note here that any study based on the "prosodic contour" should focus on a particular variety of Arabic. Though this study is based on "predicate proximity", all the experimental items in this study are in SA.

¹¹ This is due to the fact that in the offline task the participants read the whole experimental sentence at the same time, while in the self-paced reading task they see region after region without being able to go back.

preference to choose whatever demands the least cognitive effort. Predicate proximity is hypothesized to be able to override this tendency in languages that allow greater distances between a verb and its predicates. If, however, the burden on working memory becomes too great, as it could in this type of task, it is possible that native speakers that prefer high attachment under normal circumstances could resort to low attachment in some cases in an effort to reduce the load on working memory. This would result in a weakened high attachment preference in the task, or possibly even a low attachment preference. In the case of the present study, it seems that the high attachment effect was simply weaker in the self-paced reading task, such that it did not reach statistical significance in the native speaker group. Clahsen and Felser (2006b) referenced a study that shows a similar pattern of results. In this study, they note that offline tasks and online tasks may be tapping different types of preferences. According to them, offline tasks tend to indicate which interpretation of the sentence is preferred ultimately, while online tasks give an indication of initial interpretation preferences. They reference a study in which native speakers of Italian showed a preference for low attachment in terms of their reading times in a self-paced reading task, but a preference for high attachment in their responses to comprehension questions about the sentences that they read immediately before the questions were presented (DeVincenzi & Job: 1993). According to DeVincenzi and Job, this pattern of results can be explained by the notion that initial sentence parsing is carried out based on universal syntactic preferences (i.e., low attachment or late closure), while final interpretation can be affected by other factors that may override this initial interpretation (p. 204). What this means in the context of the current study is that our native speakers of Arabic may have varied in whether they initially assigned an interpretation to the experimental sentences based on universal parsing preferences or predicate proximity, potentially due to the time pressure they experienced on an individual level, or due to their working memory limitations. Since we did not measure either factor, it is difficult to draw a clear conclusion on this point.

Turning to the results for the L2 learners on the self-paced reading task, unlike the native speakers, they did show a significant effect of attachment in their reading times. This result contradicts the predictions of the SSH, that L2 learners are unable to make use of structurally-based parsing principles in the same way as native speakers, such that they show no RC attachment preference in the absence of lexical cues. One question that arises in the interpretation of the L2 results, however, is why the L2 learners would maintain their preference for high

attachment in the self-paced reading task when the native speakers did not. The explanation of the difference between the native results in the two tasks offered above is related to an increased burden on memory¹², and it seems logical that the L2 learners would have experienced this increased burden to a much greater degree than the native speakers, given their lower proficiency level. If we examine mean raw response times on the self-paced reading task, we see that the native speakers responded much more quickly on average than the L2 learners (691 vs. 2061 ms), suggesting that the native speakers felt more pressure to respond quickly than the L2 learners did. What is more important, however, in determining whether this is the case, is the relative difference in response times between the self-paced reading task and the attachment preference task. Although we did not report response times for the preference task, given its offline nature, we did collect these data. The average response time on this task was 3598 ms for the native speakers, and 6845 ms for the L2 learners. This means that while the L2 learners responded only 3.3 times as quickly on the self-paced reading task as they did on the preference task, the native speakers responded 5.2 times as quickly on the reading task as on the preference task, implying that they were more subject to increased time pressure and therefore an increased burden on working memory than the L2 learners were. However, this interpretation of the response times for the two tasks must be taken with caution, given that the response times for the preference task are for reading a whole sentence, while response times for the reading task are for reading parts of sentences. Nevertheless, this difference in the proportional speed-up in the two groups may be able to explain the difference between the groups in terms of the presence of an attachment effect in the self-paced reading task.

7. CONCLUSION

With respect to RC attachment preferences in Arabic, the results of this study indicate that both native speakers and L2 learners prefer high attachment over low attachment, in line with the predictions of the “predicate proximity account” of RC attachment (Gibson et al.: 1996). According to this account, the strength of “predicate proximity” in a given language is related to the distance between the verb and its complements. Thus, the greater this distance is in a given language the more strongly the predicate will be activated in order to allow attachments across longer distances. Due to the fact that Arabic allows the verb to be distant from its arguments,

¹² For more details on working memory load refer to Reichle, Tremblay, and Coughlin (2016).

“predicate proximity” outranks universal accounts such as “late closure”. It is also important to note that the results of this study contradict the prosody account of RC attachment in Arabic (Abdelghany and Fodor, 1999). According to the prosody account, Arabic speakers favor low attachment. Abdelghany and Fodor attribute this attachment preference to the nature of Arabic prosody. The results of the offline task indicate that both native speakers and L2 learners prefer high attachment in Arabic. In terms of the SSH, the results of the two tasks contradict its predictions for L2 learners, given that our learners showed a clear preference for high RC attachment in Arabic. However, since the native speakers did not show a statistically significant preference for high attachment, further research should investigate the circumstances under which these two groups show clear RC attachment preferences in self-paced reading tasks.

References:

- Abdelghany, H., & Fodor, J. D. (1999). Low attachment of relative clauses in Arabic. Poster presented at AMLaP (Architectures and Mechanisms of Language Processing), Edinburgh, UK.
- Brysbaert, M. (1996). Modifier attachment in sentence parsing: Evidence from Dutch. *The Quarterly Journal of Experimental Psychology: Section A*, 49(3), 664-695.
- Carreiras, M., & Clifton, C. (1999). Another word on parsing relative clauses: Eyetracking evidence from Spanish and English. *Memory & cognition*, 27(5), 826-833.
- Clahsen, H. & Felser, C. (2006a). Grammatical processing in language learners. *Applied Psycholinguistics*, 27 (01), 3-42.
- Clahsen, H. & Felser, C. (2006b) . Continuity and shallow structures in language processing: *Applied Psycholinguistics*, 27(1), 107-126.
- Cuetos, F., & Mitchell, D. C. (1988). Cross-linguistic differences in parsing: Restrictions on the use of the Late Closure strategy in Spanish. *Cognition*, 30(1), 73-105.
- De Vincenzi, M., & Job, R. (1993). Some observations on the universality of the late-closure strategy. *Journal of Psycholinguistic Research*, 22(2), 189-206.
- Dussias, P. E. (2003). Syntactic ambiguity resolution in L2 learners. *Studies in Second Language Acquisition*, 25(04), 529-557.
- Fehri, A. F. (2013). *Issues in the structure of Arabic clauses and words*. Springer Science & Business Media.

- 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*(03), 453-489.
- Ferreira, F., Bailey, K. G., & Ferraro, V. (2002). Good-enough representations in language comprehension. *Current directions in psychological science*, *11*(1), 11-15.
- Ferreira, Fernanda & Charles Clifton. 1986. The independence of syntactic processing. *Journal of memory and language*. *25* (3). 348-368.
- Frazier, L. (1987). Sentence processing: A tutorial review. In M. Coltheart (ed.), *Attention and performance XII*, 559-586. Hillsdale, NJ: Erlbaum.
- Frenck-Mestre, C., & Pynte, J. (1997). Syntactic ambiguity resolution while reading in second and native languages. *The Quarterly Journal of Experimental Psychology A*, *50*(1), 119-148.
- Gibson, E., Pearlmutter, N., Canseco-Gonzalez, E., & Hickok, G. (1996). Recency preference in the human sentence processing mechanism. *Cognition*, *59*(1), 23-59.
- Kimball, J. (1973). Seven principles of surface structure parsing in natural language. *Cognition*, *2*(1), 15-47.
- Papadopoulou, D., & Clahsen, H. (2003). Parsing strategies in L1 and L2 sentence processing. *Studies in Second Language Acquisition*, *25*(04), 501-528.
- Quinn, D., Abdelghany, H., & Fodor, J. D. (2000, March). More evidence of implicit prosody in reading: French and Arabic relative clauses. In Poster presented at the 13th Annual CUNY Conference on Human Sentence Processing, La Jolla, CA, March.
- R Development Core Team (2015). R: A language and environment for statistical computing (Version 3.2.2) [Computer software]. Retrieved from <http://www.R-project.org>.
- Reichle, R. V., Tremblay, A., & Coughlin, C. (2016). Working memory capacity in L2 processing. *Probus*, *28*(1), 29-55.
- Singmann, H., Bolker, B., & Westfall, J. (2015). afex: Analysis of factorial experiments (Version 0.14-2) [Computer software]. Retrieved from <https://github.com/singmann/afex>.