DOI: 10.26346/1120-2726-199

Transfer of L1 strategies in L2 processing of long-distance dependencies

Rania Al-Aqarbeh

Department of English Language and Literature, Mutah University, Jordan
<rania11@mutah.edu.jo>

This study investigated the online processing of the relative clause dependency. The participants were native speakers of southern Jordanian Arabic, a grammatical resumption language, and were advanced learners of English as a second language, an intrusive resumption language. Another relevant difference between these two languages is that resumption in southern Jordanian Arabic ameliorates the relative clause island effect, while it does not in English. Two offline acceptability judgment tasks and two online eyetracking reading tasks were conducted in Jordanian Arabic (L1) and English (L2). The results revealed that the L2 learners had the pre-requisite grammatical knowledge to process the relative clause dependency in English. They seemed to posit a resumptive pronoun to resolve the dependency inside a relative clause island in their L1 and L2 alike. This result demonstrated that the L2 learners in the current study exhibited detailed syntactic processing. By doing so, they diverged from the processing behavior native speakers of English exhibited in processing similar stimuli in previous studies (Traxler & Pickering 1996; Omaki & Schulz 2011). In conclusion, the different processing behavior the L2 learners in this study manifested can be attributed to L1 transfer rather than to shallower, less detailed syntactic processing as proposed by the Shallow Structure Hypothesis.

KEYWORDS: resumption, dependency, islands, reading, grammatical processing strategies.

1. Introduction

The processing of online formation of long-distance dependencies has been investigated in psycholinguistic studies from different perspectives by native speakers as well as L2 learners. Long-distance dependencies, also known as filler-gap dependencies, refer to constructions like relative clauses in which a displaced constituent (the filler) *whom* is associated with a gap. This gap is represented by the underscore, in its extraction site, irrespective of the theory of movement adopted, as illustrated in (1) below. Gaps are represented by an underscore in all the examples throughout the paper.

(1) *The woman*$_i$ whom$_i$ *I met* ___$_i$ *yesterday is beautiful.*

In intrusive resumption languages, like English, resumption does not constitute a part of the grammar. Instead, the extraction site has a gap, a phonetically null variable.

Conversely, in grammatical resumption languages, like Jordanian Arabic (JA), resumption is a part of the grammar. This canonical extraction site can be filled by a resumptive pronoun (RP), a phonetically overt pronominal variable, as demonstrated in example (2) below (Chao & Sells 1983; Sells 1984; McCloskey 2006). RPs are deitalicized in the relevant examples throughout the paper.

(2) *ʃifit* *is-sajjarah*$_i$ *illi* *iftareiti*-ha$_i$/ *___$_i$ *imba:riħ.*
see.1SG the-car that buy.2SG.F-it$_i$/ ____$_i$ yesterday.
'I saw the car that you bought (it) yesterday.'

The mainstream assumption in previous syntax literature is that gaps render constructions with islands ungrammatical, and RPs ameliorate island effects (Ross 1967; Kroch 1981; Sells 1984; Engdahl 1985; Choueiri 2018; McCloskey 2017). Islands are constructions out of which A'-movement is prohibited, and are therefore considered diagnostics of A'-movement. Islands, in effect, are grammatical constraints on A'-movement, making gaps inside these constructions not allowed (Ross 1967).

Many psycholinguistic studies have targeted the strategies the parser employs in order to resolve long-distance dependencies in the absence of islands. In essence, the parser incrementally processes the dependency. Upon identifying a filler, the parser starts the search and posits a gap as soon as it encounters the first subcategorizer, a transitive verb that subcategorizes for an object argument (Frazier 1987; Garnsey *et al.* 1989; Pickering & Traxler 2003; Wagers & Phillips 2009) or even before the first subcategorizing category (Omaki & Schulz 2011). These claims have been supported by several types of evidence in previous psycholinguistic research such as the filled-gap effect. For the sake of illustration, consider the following example. In an online self-paced reading study, Stow (1986: 239) employed experimental pair items as presented in Examples (3-4) below; (3a) and (4a) involved target items with a dependency while (3b) and (4b) were baseline control items that lacked a dependency and had an embedded *if*-clause instead. Stowe used an overt lexical NP, *us* and *Greg's*, respectively, after the first subcategorizer in both items. An increase in reading times (RTs) was only found on the lexical NP in the experimental item (3a). This increase in RTs has been known as the filled-gap effect in previous psycholinguistic literature on intrusive languages. This effect is assumed to result from a reanalysis by

virtue that the parser predicts a gap and posits it as early as possible, but when it finds a lexical NP in the position where it predicts a gap to resolve the dependency, it does a reanalysis and starts the search again for another site to posit the gap. The filled-gap effect was not found in control items like (3b), that lacked a dependency because the parser did not predict a gap. Therefore, no reanalysis was in effect. Interestingly, when the parser of intrusive resumption languages encountered an island, like the bracketed subject island which disallows gaps in a subject phrase in (4a), it refrained from positing a gap inside the island. Rather, it preferred a late site for a gap over positing an early gap inside an island. The explanation advanced was that the parser is sensitive to grammatical constraints, e.g. islands, and therefore does not predict a gap inside islands that disallow gaps. Thus, no reanalysis is involved in parsing such constructions. All the regions of interest in all the relevant examples are written in bold throughout the paper.

(3) a. *My brother wanted to know who Ruth will bring* ***us*** *home to ___ at Christmas.*
b. *My brother wanted to know if Ruth will bring* ***us*** *home to Mom at Christmas.*

(4) a. *The teacher asked what [the silly story about* ***Greg's*** *older brother] was supposed to mean ___.*
b. *The teacher asked if [the silly story about* ***Greg's*** *older brother] was supposed to mean anything.*

In a nutshell, the parser of intrusive resumption languages posits a gap as early as possible on the condition that no grammatical constraints banning gaps intervene.

Comparatively, previous psycholinguistic research on the processing of similar constructions in grammatical resumption languages like Hebrew, despite the fact that there have been many fewer relevant studies conducted on intrusive resumption languages, shows that the parser of grammatical resumption languages "prefers an early RP over a later gap" in islands that allow gaps (Keshev & Meltzer-Asscher 2017: 549). They employed a self-paced reading experiment using Stowe's (1986) filled-gap design, as demonstrated in (5-6) below, from Keshev & Meltzer-Asscher (2017: 559). They found a filled-gap effect, or in their terms a filled-RP effect, on the lexical NP inside a Complex NP (CNP), an island in which resumption is grammatical. However, there was no such effect inside a Coordinate Structure (CS), an island where is ungrammatical.

CNP-RC condition:

(5) *Ha-šotrim mekirim et-ha-iša še-ha-xašudim [še-takaf et-ha-melcar]*
the-cops know ACC-the-woman that-the-suspects that-attacked ACC-the-waiter]
kilel ***et-ha-tabax*** *axrey še-hu daxaf ota be-mis'ada*
cursed ACC-the-cook after that-he pushed her in-restaurant

yokratit *be-Tel Aviv*
upscale in-Tel Aviv
'The cops know the woman who the suspect [who attacked ***the waiter***] cursed the cook after he had pushed her in an upscale restaurant in Tel Aviv.'

CS-RC condition:

(6) *Ha-šotrim* *mekirim* *et-ha-iša* *še-ha-xašudim* *takaf* ***et-ha-melcar***
the-cops know ACC-the-woman that-the-suspects attacked ACC-the-waiter
[ve-kilel ***et-ha-tabax****]* *axrey* *še-hu* *daxaf* *ota*
[and-cursed ACC-the-cook] after that-he pushed her
be-mis'ada *yokratit* *be-Tel Aviv*
in-restaurant upscale in-Tel Aviv
'The cops know the woman who the suspect who attacked ***the waiter*** and [cursed ***the cook***] after he had pushed her in an upscale restaurant in Tel Aviv.'

The difference in strategies adopted by the parser of grammatical resumption languages in contrast with the parser of intrusive resumption languages in processing long-distance dependency formation inside islands provides a suitable and suggestive testing ground for the Shallow Structure Hypothesis (SSH) (Clashen & Felser 2006a; 2006b; 2006c; 2018). This study aims at testing the main claims of the SSH by studying the processing of L2 English, an intrusive resumption language, by highly proficient L2 learners of English who are native speakers of southern JA, a grammatical resumption language.

The paper is organized as follows. Section 2 reviews the methodologies and results of two relevant previous psycholinguistic studies whose methodologies are partially adopted in the current study. Section 3 provides a brief description of southern JA properties pertinent to the issues under investigation. Section 4 summarizes the main claims of the SSH. Section 5 portrays a detailed presentation of the offline and online experiments conducted in this study. Section 6 concludes with general discussion of the findings of these experiments and their potential theoretical implications.

2. Previous studies

This section reports on the theoretical claims, methodologies, and findings of two studies that constitute the foundation of the methodologies adopted in designing the experiments of the current study: (a) Traxler & Pickering's (1996) Experiment 2, which examined the processing of long-distance dependencies inside islands by native speakers of

English, and (b) Omaki & Schulz's (2011) research, which adopted and modified Traxler & Pickering's (1996) experiment, and applied it to the processing of similar constructions by native speakers of English as well as L2 learners.

Traxler & Pickering (1996) conducted an eyetracking reading study to investigate the processing of long-distance dependency formation by manipulating two independent variables: Plausibility (Plausible *vs* Implausible) and Islandhood (Relative Clause Island *vs* Non-Island) as demonstrated in the following item set. They manipulated the Plausibility in these examples through the use of the nouns *book* vs *city* on the grounds that a plausible interpretation was triggered by matching *the book* with the verb *wrote*, as in (7a) and (7c), whereas matching the noun *the city* instead rendered the interpretation implausible as in (7b) and (7d). The transitive verb whose direct object position that represented the first potential gap site appeared in non-island conditions in (7a-b), yet inside the bracketed island conditions in (7c-d). They used the relative clause island (RC-island) as a strong island in all their island conditions. The regions where a misanalysis could have been reconsidered, and thus ambiguity was resolved were *about while*, in non-island conditions, and *saw while*, in island conditions. Three critical regions were identified: Region 3 *wrote unceasingly*, and Region 4 *with great dedication*, Region 5 *about while / saw while*.

(7) a. *We like **the book** that the author **wrote** __unceasingly and with great dedication about ___ while waiting for a contract.*

b. *We like **the city** that the author **wrote** unceasingly and with great dedication about _____ while waiting for a contract.*

c. *We like **the book** that the author [who **wrote** ___ unceasingly and with great dedication] saw while waiting for a contract.*

d. *We like **the city** that the author [who **wrote** unceasingly and with great dedication] saw ____while waiting for a contract.*

They found a plausibility effect in non-island conditions, namely, longer first-fixation, first-pass fixation, and total reading times on the first subcategorizer the parser encountered. They considered this effect as an indication that the parser posited a gap as a complement to the verb. Nonetheless, no similar plausibility effect was depicted when the verb was inside a strong island, e.g. RC-island.

They concluded that the parser adopted a filler-verb immediate association to resolve the long-distance dependency once it encountered a verb unless this association was ruled out by island constraints. The

plausibility mismatch effect on the verb *wrote* increased reading times when the filler was implausible, *the city*, contra to when it was plausible, *the book*. The strategy which the parser adopts once it encounters a filler and begins to actively search for a potential position for a gap in order to resolve the dependency is referred to in the literature as the active-gap creation strategy. It has been attested that the active-gap creation strategy is either driven by processing principles by which the parser resolves the dependency as soon as possible (De Vincenzi 1991; Pritchett 1992) or by reducing costly processing due to maintaining the filler in the memory (Gibson 1998) using several psycholinguistic measures in L1 processing (Crain & Fodor 1985; Frazier 1987; Lee 2004; Stowe 1986; Pickering & Traxler 2003; Phillips & Wagers 2007) and L2 processing (Jackson & Bobb 2009; Jackson & Dussias 2009; Williams, Mobius & Kim 2001; Williams 2006; Dallas & Kaan 2008). On the other hand, several previous studies have supported the lack of active-gap creation inside islands in L1 processing of the long-distance dependency (Stowe 1986; Yoshida 2006), yet less has been investigated in L2 processing in corresponding contexts.

Omaki & Schulz (2011) investigated L2 processing of the long-distance dependency in the presence of strong islands and compared it to L1 processing of the same constructions. Their main goal was to examine whether advanced L2 learners actually construct shallow representations that are impoverished of structural details as claimed by Clahsen & Felser (2006a; 2006b; 2006c). To this end, they combined two methodologies: an offline acceptability judgment task and online self-paced reading experiments, in which they adapted Traxler & Pickering's (1996) stimuli. They argue that combining offline and online methodologies constitutes an ideal ground to test the SSH for two reasons. First, the offline task is essential to investigate whether L2 learners have the essential target-like grammatical knowledge of rules and constraints without being under time pressure. Second, the online task is inevitable to explore whether these grammatical rules and constraints are employed in real-time grammatical processing. This standing has inspired adopting and combining both offline and online measures in the current study. Omaki & Schulz (2011) applied their tasks to advanced native speakers of Spanish, which allows an overt *wh*-movement that, essentially, is sensitive to island constraints, who were learning English as an L2. They applied both tasks to two groups of participants: an experimental group which involved L2 learners of English who were native speakers of Spanish and a control group that consisted of native speakers of English.

Omaki & Schulz (2011: 572) ran an offline grammaticality judgment that required the participants to judge the grammaticality of English sentences with and without RC-islands as demonstrated below on a 5-point Likert Scale.

(8) a. Grammatical sentence
The murder case$_i$ that the law students [$_{RC}$ who learned about the constitution] discussed ___$_i$ was going to be on the exam.

b. Ungrammatical sentence*:*
The murder case$_i$ that the law students [$_{RC}$ who learned about ___$_i$] discussed ___$_i$ was going to be on the exam.

They found that L2 learners treated grammatical and ungrammatical conditions similar to how native speakers of English judged them. Accordingly, they concluded that L2 learners had the pre-requisite grammatical knowledge.

Then, they conducted an online self-paced reading task using Traxler & Pickering's (1996) stimuli with slight modification as represented below from (Omaki & Schulz 2011: 575). The plausible conditions involved a noun *the book* that matched the first subcategorizer *wrote*, whereas the implausible conditions involved a noun *the city* that did not match the first subcategorizer. The non-island conditions contained one relative clause; the island conditions encompassed two relative clauses with an RC-island embedded inside the other relative clause. They considered two critical regions: the first subcategorizer *wrote* and the immediately-following adverb *regularly*. The first subcategorizers were all transitive verbs requiring an object while the adverb was considered critical for any potential spill-over effect.

(9) a. Non-Island.Implausible
The city that the author wrote regularly about __ was named for an explorer.

b. Non-Island.Plausible
The book that the author wrote regularly about __ was named for an explorer.

c. Island.Implausible
The city that the author who wrote regularly saw __was named for an explorer.

d. Island.Plausible
The book that the author who wrote __ regularly saw was named for an explorer.

They found that both native speakers of English and L2 learners demonstrated evidence of a plausibility mismatch effect on the first transitive verb in the non-island conditions while no such effects were

found in the island conditions. Based on these findings, they argue that advanced L2 learners build abstract detailed structural representations and that the active-gap search is constrained by grammatical constraints like strong islands. More precisely, they argue that L2 learners do not exhibit shallower structure contrary to the assumptions of the SSH. Rather, they pattern with native speakers in these regards.

Omaki & Schulz's (2011: 575) findings replicated the findings of previous studies concerning L2 processing and lending further support that "L2 learners whose L1 has overt wh-movement are sensitive to island constraints" (Martohardjono & Gair 1993; Schachter 1990; White 1988; Belikova & White 2009). Nonetheless, both languages investigated in their study are intrusive resumption languages and allow overt *wh*-movement that constrain positing gaps inside islands. Therefore, the results drawn cannot be clearly attributed either to the learners' L1 transfer of processing strategies or exhibiting the processing strategies typical of the L2. Therefore, I contend that investigating languages that are typologically different could result in a stronger challenge to the SSH's claims.

3. Grammatical resumption in southern JA

The current study investigates the processing of relative clauses by native speakers of a southern variety of JA as spoken at Mutah University in the south of Jordan. JA is a Levantine variety of Arabic. It is a grammatical resumption language in which resumption is part of its grammar. Southern JA allows gap and resumption strategies to alternate in the *wh*-interrogative dependency (10a). The filler *weiʃ* 'what' is related to a gap, represented by an underscore, or an RP written in deitalicized script. Both are grammatical. However, I contend that only the resumption strategy is allowed in the relative clause dependency (10b) in southern JA. Here, the gap strategy is ungrammatical. Nonetheless, some scholars may argue that both strategies are available for forming relative clauses in Arabic varieties in general including Levantine Arabic varieties (Aoun *et al.* 2010, and references therein). Therefore, it is imperative to conduct a formally designed task to explore which strategies are actually available to resolve the relative clause dependency in southern JA.

(10) a. *weiʃ* *ga:lat* *lajla* *innu* *sami:r* *iʃtra:*-ah/ __?
what say.3SG.F Laila that Sameer buy.3SG.M-it/___?
'What did Laila say that Sameer bought (it)?'

b. *gareit il-ikta:b illi iʃtreit*-uh/ *__ *min il-maktabih*
read.1SG the-book that buy.1SG-it/ __ from the-bookshop
'I read the book that I bought it from the book shop.'

The mainstream theoretical accounts in the literature on the syntax of Arabic is that the long-distance dependency in which a fronted constituent is related to a gap in its extraction site is formed via *wh*-movement; however, the presence of an RP reveals that the fronted element is base-generated in its surface position, and is related to the RP via binding or other non-movement mechanisms (Shlonsky 1992; Malkawi & Guilliot 2007; Aoun *et al.* 2010; Choueiri 2018). To the best of my knowledge, there has been no previous study that applies formally designed tasks to explore the acceptability of gaps in relative clauses or whether resumption actually ameliorates island effects like the RC-island in southern JA. Therefore, running an offline acceptability judgment task ahead of the other tasks involved in this study is warranted.

4. Shallow Structure Hypothesis (SSH)

The Shallow Structure Hypothesis (SSH) was originally put forth to explain the observed L1-L2 differences in processing by proposing two main arguments (Clahsen & Felser 2006a, 2006b, 2006c, 2018). The first argument assumes that the "representations adult L2 learners compute for comprehension are *shallower* and *less* detailed than those of native speakers …. and rely more on non-structural information in parsing" (Clahsen & Felser 2006a: 3f). Clahsen & Felser (2018) clarify that this claim does not imply that L2 speakers do not employ syntactic representations in processing sentences in their L2. Instead, they emphasize that "L2 learners underuse syntactic information during real-time processing" (Clahsen & Felser 2018: 3) and rely more on non-grammatical (i.e. semantic, pragmatic, etc.) sources of information, giving them more priority than native speakers during online processing. Clahsen & Felser (2018: 6) assert that L2 learners exhibit a consistently robust asymmetrical processing pattern across different L1-L2 settings. More precisely, L2 learners manifest "nativelike online sensitivity to lexical-semantic disambiguation cues"; however, they exhibit a non-nativelike online processing behavior when these lexical-semantic cues are not present.

The second argument that Clahsen & Felser (2018) address is learners' L1 transfer. They argue that assumptions reported in the literature regarding the lack of L1 transfer that have often been attributed to the SSH (Jegerski, VanPatten & Keating 2011) are unwarranted

and misunderstood. They claim that this is not what the SSH actually implies. Rather, they assume that the SSH implies that "L1 transfer [should] influence processing only indirectly" (Clahsen & Felser 2018: 5). However, they do not show what indirect L1 influence means.

Most of the evidence reported with respect to sentence processing in the literature on the SSH comes from filler-gap dependencies and resolving ambiguities. Clahsen & Felser (2006c: 31-32) argue that L2 sentence processing is dominated by shallow processing and this is evident, particularly, in "parsing complex hierarchical structures" that involve abstract elements like gaps. They clarify that sentences can be successfully interpreted by segmenting the incoming elements from the sentence online into meaningful parts based on lexical, semantic, pragmatic, and other non-grammatical sources of information. However, applying structure-driven parsing requires computing a syntactic representation of the sentence being processed and this "includes hierarchical phrase structures as well as abstract elements such as empty categories." They resort to Marinis, Roberts, Felser, and Clahsen's (2005) self-paced reading study in which they investigated the processing of a long-distance *wh*-dependency by native speakers of English (a *wh*-movement language) and advanced, adult L2 learners with different L1s including languages with *wh*-movement, like German and Greek, and languages with *wh*-in-situ, like Chinese and Japanese. Their experimental stimuli are illustrated in Example (11) below. In (11a), the filler *who* is extracted from the direct object position in the complement clause and there is an intermediate gap that breaks up the long dependency into shorter ones. However, there is no such intermediate gap in the corresponding example in (11b).

(11) a. *The nurse **who** the doctor argues __ that the rude patient had angered __ is refusing to work late.*
b. *The nurse **who** the doctor argument about the rude patient had angered __ is refusing to work late.*

They found that native speakers' processing was characterized by shorter RTs at the subcategorizing verb in the conditions that involved an intermediate gap (11a), but not in conditions that lacked an intermediate gap (11b). On the other hand, no such difference in RTs was depicted for L2 learners' processing of both conditions with or without intermediate gaps. Marinis *et al.* (2005) found that the lack of intermediate gap effect was consistent regardless of similarity or difference between the structure of the learners' L1 and L2. Marinis *et al.* (2005) assume that while native speakers postulate an intermediate gap at the

clause boundary, L2 learners did not demonstrate evidence of postulating an intermediate gap like the native speakers.

Clahsen & Felser (2006a) explain Marinis *et al.*'s (2005) findings as follows. They argue that the parser of native speakers posits a gap that is bound by the filler *who* as it encounters the complementizer *that* which marks the beginning of a subordinate clause obeying the Subjacency Principle. They add that the cyclic integration of the filler at intermediate gap sites makes the semantic integration of the filler with its subcategorizer easier than sentences that lack intermediate gap sites as syntactically represented in (12). On the other hand, L2 processing does not involve consulting a mental representation of the filler into the intermediate gap position which makes the syntactic processing of such a sentence shallower and less detailed. Rather, the parser of non-native L2 learners relies more on semantic and lexical segmentation of the sentence that involves computing predicate-argument relations and thematic roles. This is sketched in (13) below from Clahsen & Felser (2006a: 32) which shows that L2 sentence comprehension proceeds incrementally as the parser immediately integrates each new element into the semantic structure of the sentence by assigning thematic roles. They add that the lack of an intermediate gap in L2 processing does not mean that L2 learners do not have the grammatical knowledge of the Subjacency Principle, but it "follows from the learners' failure to project the syntactic structure necessary for accommodating intermediate gaps."

(12) *[DP The nurse [CP [who$_i$] the doctor argued [CP [e$_2$] that the rude patient had angered [e$_1$]]]]... is refusing to work late*

(13) a. *[The nurse] who [the doctor]* argued *[that ……*
AGENT THEME

b. *[the rude patient] had* angered ……
THEME

c. *[The nurse] who [the doctor] argued [that …*
[the rude patient] had angered] is refusing to work late.
EXPERIENCER

Regarding the L1 transfer of processing strategies, the phenomenon that has usually been reported is the relative clause ambiguity resolution. Clahsen & Felser (2006a, 2006b, 2006c, 2018) reported relevant findings from previous studies on relative clause attachment ambiguities that involve genitive (NP_1-of-NP_2) antecedents as in *Someone saw the servant of the actress who was on the balcony* (Felser, Roberts, Marinis & Gross 2003; Papadopoulou & Clahsen 2003). These studies found that L1 native

speakers adopt a phrase-structure-based attachment principle preferring NP_2 disambiguation (Carreiras & Clifton 1999; Fernandez 2003; Roberts 2003), whereas advanced adult L2 learners from different L1 backgrounds exhibit no preference to either NP_1 or NP_2 disambiguation strategies and so fail to show native-like processing strategies. Even though the findings they often reported from relevant previous studies show that L2 processing of sentences is not influenced by learners' L1, Clahsen & Felser (2018) comment that L1 influence on L2 grammar is limited and that there are no claims made within the SSH that it is lacking. In fact, their claims regarding L1 transfer are not clear and their terms of indirect and limited influence do not imply the absence of L1 transfer. Furthermore, Clahsen & Felser (2018) argue that some claims made in the literature such as the one made by Omaki & Schulz (2011: 576) that "L2 learners …. should not respect the RC island constraint because there is no RC representation in their analysis" are overstated and not made in the SSH.

5. The current study

This study aims to investigate to what extent the SSH may account for the L2 processing of relative clauses in English by learners whose L1 is typologically different. The following sub-goals are targeted first in order to situate the findings within the main tenets of the SSH.

1) To identify the dependency-resolving strategy in the relative clause dependency in southern JA.
2) To figure out whether resumption ameliorates islands like the RC-island.
3) To establish whether the advanced L2 learners have the pre-requisite grammatical knowledge of the dependency-resolving strategy in the relative clause dependency in English.
4) To determine whether the advanced adult L2 learners have the pre-requisite grammatical knowledge of the RC-island in English.
5) To identify what characterizes the online processing of the relative clause dependency in southern JA.
6) To pinpoint whether L2 processing of relative clauses in English is similar to or different from L1 processing of this construction in southern JA.

The study hinges on conducting both offline and online measures of island constraints on long-distance filler-gap formation. The offline method is necessary to explore whether the learners have the pre-

requisite target-like grammatical knowledge of the constraints at issue, whereas the online task is necessary to provide real-time evidence of processing strategies. Overall, the order of the tasks undertaken is: (a) an offline acceptability judgment task of southern JA, (b) an offline acceptability judgment task in English, (c) an eyetracking reading study in southern JA, and (d) an eyetracking reading study in English.

5.1. Experiment 1: Acceptability judgment task in southern JA

At the outset of the current study, an offline acceptability judgment task was conducted in order to account for the following research questions.

I. Does southern JA allow both gaps and resumption in a relative clause dependency?
II. Does resumption exhibit sensitivity to the RC-island in southern JA?

Materials, design, and procedure

In order to account for the aforementioned research questions, I constructed the experimental sentences in accordance with a 2*2 factorial design that involved two independent variables with two levels each: ISLAND (RC_Island *vs* RC_NoIsland) and TAIL, a dependency-resolving strategy, (RP *vs* Gap). The design resulted in four conditions as demonstrated in (14) with a full item set. The RC_Island conditions were the target items and their corresponding RC_NoIsland conditions were the baseline control items created for the sake of comparison.

(14) a. RC_NoIsland.RP
ʃifit illawħah illi salma rasamat-ha ʕalg-uː-ha ʕ-al-ħeiṭ
see.1SG the.portrait that Salma draw.3SG.F-it hang-3PL.M-it on-the-wall
'I saw the portrait that Salma drew it they hang it on the wall.'

b. RC_Island.RP
ʃifit illawħah illi salma [illi rasamat-ha*]* ʕalg-uː-ha ʕ-al-ħeiṭ
see.1SG the.portrait that Salma [who draw.3SG.F-it] hang-3PL.M-it on-the-wall
'I saw the portrait that Salma [who drew it] they hang it on the wall.'

c. RC_NoIsland.Gap
**ʃifit illawħah illi salma rasamat_* ʕalg-uː-ha ʕ-al-ħeiṭ
see.1SG the.portrait that Salam draw.3SG.F_hang-3PL.M-it on-the-wall
'I saw the portrait that Salma drew _ they hang it on the wall.'

d. RC_Island.Gap
**ʃifit illawħah illi salma [illi rasamat _]* ʕalg-uː-ha ʕ-al-ħeiṭ
see.1SG the.portrait that Salam [that draw.3SG.F _] hang-3PL.M-it on-the-wall
'I saw the portrait that Salma [who drew _] they hang it on the wall.'

Each participant saw 4 items on each condition; thus, 16 experimental items were created. Each of the four items within an item set were lexically matched and differed only in the manipulation of the values of the independent variables. Twenty-four filler items were also included with ratio of 1.5:1 of filler items to experimental sentences. The filler items involved 12 grammatical sentences and 12 ungrammatical sentences. The ungrammatical sentences contained incorrect manipulation of word order, incorrect agreement between subjects and verbs and modifiers with nouns, incorrect usage of pronouns and demonstratives, and incorrect derived forms of words. Each list also contained three practice items. In total, the task involved 43 items: 3 practice items, 16 experimental items, and 24 filler items. Sixteen item sets were constructed and distributed into four lists. The order of presentation was randomized so that participants would not see two consecutive items on the same condition in the same list. I created all of the items. Then all of the items were judged by 5 scholars of modern linguistics who are also native speakers of southern JA. The task was emailed to the participants in the format of a questionnaire. They were instructed to judge the acceptability of the sentences on a 7-point Likert Scale (*1 = absolutely unacceptable*; *7 = perfectly acceptable*).

Participants

Since the same participants were required to take part in all the experiments, they were required to complete a survey with detailed, relevant demographic information before I could confirm their participation. I personally contacted 55 students in the Department of English Language and Literature at Mutah University whose GPA is (76+), namely, excellent students. Their overall proficiency in English was measured by means of a C-test. Since there was no control group of native speakers of English, I used the results of the C-test administered to native speakers of English and L2 learners of English by other scholars as follows for comparison's sake. Schulz (2006) applied the C-test to 30 native speakers of English, and reported that the average score was 50-60 (*SD* = 7.7; range: 26-59). Omaki & Schulz reported that the average score of the C-test as administered to advanced L2 learners of English was 42.5 out of 60 (*SD* = 10.1; range 24-58). In my study, twelve participants from those recruited for the current study scored very low (18-), and were therefore excluded from the very beginning. Furthermore, 5 participants reported that they have some vision problems and 2 refused to visit the lab to take part in the experiments. The average score of the remaining 36 participants of the current study was 43 out of 60 points (*SD* = 7.8; range: 30-52). The average score of the

L2 learners in this experiment was close to the average score of the advanced L2 learners in Omaki & Schulz's (2011) study. Consequently, the participants recruited for this experiment could be considered advanced adult L2 leaners of English.

All 36 participants gave consent to participate in all of the experiments, and were all given course credits for their participation. They all reported the following information. They were all native speakers of the southern JA variety at issue and it was the only variety used at their homes and by their parents and families. They all have average vision with no vision problems. They all started learning English at the age of 5 and only in a classroom setting. Their first exposure to English was at KG2 when they were five years old, and they all had received between 15 and 16 years of instruction in English at school and university. None of them lived or even traveled to any English-speaking countries, but they were used to English movies, songs, and games. They all had access to English as used by native speakers only through media and had this access on a regular basis (around 5-6 hours per week). All of them used online social media platforms like Facebook, Instagram, and WhatsApp on a daily basis. They all reported that they chatted using southern JA. The confirmed sample included 25 females and 11 males. The age range was 19-21 with the mean age of 19.3. First, they visited the Psycholinguistics Lab at Mutah University where they completed the demographic survey and proficiency test. Then, they were emailed the acceptability judgment task in southern JA and were required to return it via email in less than 24 hours. A week later, they did the acceptability judgment task in English following the same procedure. Two weeks later, they participated in an eyetracking reading experiment in southern JA. They were tested individually at the lab using the SR EyeLink 1000 Plus eye-trackers (SR Research Toronto, Ontario, Canada). A month later, they took part in the fourth experiment which was an eyetracking reading study in English in the same lab. The temporal intervals that were maintained among tasks were to control for the participants getting familiar with the items.

Results

The average rating of grammatical fillers was 6.31 (SD = 1.12); the average rating of ungrammatical fillers was 1.23 (SD = 1.07). The null hypotheses to be tested in the statistical analysis of this experiment were:

(i) H_0: there is no difference in acceptability ratings of resumption in relative clause dependency in southern JA between Island conditions and their corresponding NoIsland conditions.

(ii) H_0: there is no difference in acceptability ratings between RPs and gaps in the relative clause dependency in southern JA.

To compensate for the relatively small number of participants, four tokens on each condition were rated. Then, the ratings of the tokens per condition per participant were averaged ahead of the analysis. This was applied to all the experiments. To test the above hypotheses, a paired-sample *t* test was conducted and the results are reported in Table 1. The results showed that the difference in rating the acceptability of RPs in Island conditions in contrast with NoIsland conditions was statistically non-significant as shown also by the slight effect size (dividing the mean by the SD as a standardizer) and the CI that went through zero. According to these results, the first null hypothesis in (i) is accepted. The lack of significant difference in using the resumptive strategy in presence or absence of islands revealed that the resumptive strategy does ameliorate the RC-island effect in southern JA. The difference in using the gap strategy between Island versus NoIsland conditions was statistically non-significant and this finding was supported by the slight effect size, Cohen's *d*, and the CI that went through zero. This finding might look surprising, but a closer inspection of Figure 1 below showed that the gap strategy was rated unacceptable in both Island and NoIsland conditions. This result can be interpreted as showing that the gap strategy is not a valid strategy to resolve a relative clause dependency in southern JA altogether. These results were, further, boosted by the third and fourth pairwise comparisons between resumptive and gap strategies in Island and NoIsland conditions, respectively. Here, the difference was statistical and the effect size was relatively high as well. Accordingly, the second null hypothesis in (ii) is rejected.

PAIRED-COMPARISONS	*M (SD)*	*t*	*p*	*df*	*d*	*95% CI*
RC_Island.RP *vs* RC_NoIsland.RP	.111 (.523)	1.276	.210	35	.212	[-.066, .288]
RC_Island.Gap *vs* RC_NoIsland.Gap	.167 (.697)	1.435	.160	35	.239	[-.069, .402]
RC_Island.RP *vs* RC_Island.Gap	5.500 (.561)	58.864	.000	35	9.803	[5.310, 5.690]
RC_NoIsland.RP *vs* RC_NoIsland.Gap	5.444 (.695)	47.026	.000	35	7.833	[5.209, 5.679]

Table 1. The results of the statistical analysis of acceptability judgment of L1 data.

The average rating of each condition is presented in Figure 1. As shown in this figure, the participants rated resumption in Island as well as NoIsland conditions as perfectly acceptable (above 6). However, they rated the gap strategy in Island and NoIsland conditions as completely unacceptable (below 2).

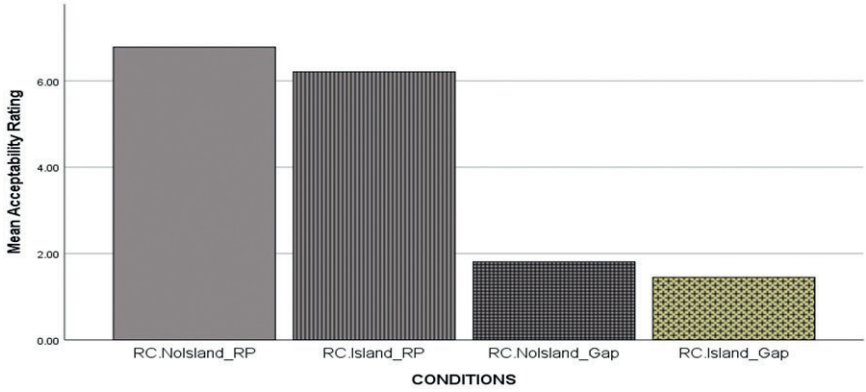

Figure 1. Acceptability judgment ratings of the conditions in L1 (Southern JA).

In a nutshell, the following conclusions can be drawn from the results of Experiment 1:

a. Resumption is the only dependency-resolving strategy of the relative clause dependency in southern JA.
b. Resumption ameliorates the RC-island effect. Resumption does not exhibit sensitivity to the RC-island in southern JA.
c. Gaps are considered unacceptable in Island and NoIsland conditions in relative clauses in southern JA.

5.2. Experiment 2: Acceptability judgment task in L2 English

The second offline acceptability judgment task was administered to account for the following research questions.

I. Do L2 learners of English whose L1 is southern JA have the pre-requisite grammatical knowledge of the availability of the gap strategy only in the English relative clause dependency?
II. Do these L2 learners have the required grammatical knowledge of the RC-island constraint on long-distance dependency formation in the English relative clause dependency?

Materials, design, and procedure

I constructed the experimental sentences in light of Omaki & Schulz's (2011) sentences like those reported earlier in example (8) in this study. However, because both Spanish and English, the languages at issue in their study, are intrusive resumption languages, their examples included only sentences with gaps. I added examples with RPs because the learners' L1 in my study is a grammatical resumption language and I aimed at identifying whether those L2 learners have the pre-requisite grammatical knowledge that English only allows the gap strategy in these constructions. Consequently, I created the experimental items according to a 2*2 factorial design that consisted of two independent variables with two values each akin to the design of Experiment 1: ISLAND (RC_Island *vs* RC_NoIsland) and TAIL, a dependency-resolving strategy, (RP *vs* Gap). The design yielded four conditions as detailed in the item set in (15). The gap in (15a) and the RP in (15c) occurred inside an embedded declarative *that*-clause. However, their first occurrence was inside an RC-island in (15b, d), respectively. However, their second occurrence was outside the island to resolve the dependency.

(15) a. RC_NoIsland.Gap
The novel$_i$ that the students [$_{RC}$ whom we met yesterday] read ___ $_i$ is interesting.

b. RC_Island.Gap
**The novel$_i$ that the students [$_{RC}$ who learned about ___ $_i$] read ___ $_i$ is interesting.*

c. RC_NoIsland.RP
The novel$_i$ that the students [$_{RC}$ whom we met yesterday] read **it$_i$ is interesting.*

d. RC_Island.RP
The novel$_i$ that the students [$_{RC}$ who learned about **it$_i$] read **it**$_i$ is interesting.*

Each participant saw 4 tokens on each condition; thus, they judged 16 experimental items in total. I created 16 experimental items besides 3 practice items at the outset of the task. Therefore, the task included 43 items: 3 practice items, 16 experimental items, and 24 filler items (with 1.5:1 ratio of filler sentences to experimental sentences). Sixteen experimental items were created for each of the 4 items within an item set. They were lexically matched to control for any confounding factors that may result from lexical consideration. I created all of the items myself in a similar manner to Omaki & Schulz's (2011) examples. Then, all of the experimental, practice, and filler items were judged by 3 scholars of linguistics who are native speakers of American English. The 36 participants who participated in Experiment 1 took part in Experiment 2 and returned it via email within 24 hours. They were instructed to judge the

grammaticality of the sentences on a 7-point Likert Scale (*1 = absolutely unacceptable*; *7 = perfectly acceptable*).

Results

The average rating of grammatical fillers was 5.76 (*SD* = 1.81); the average rating of ungrammatical fillers was 2.11 (*SD* = 1.69). The null hypotheses to be tested in the statistical analysis of this experiment were:

(i) H_0: There is no difference in the acceptability ratings of resumption in Island versus NoIsland conditions.
(ii) H_0: There is no difference in the acceptability ratings of gaps in Island versus NoIsland conditions.
(iii) H_0: There is no difference in the acceptability ratings of gaps versus resumption in relative clause dependency in English.

The results obviously showed that the first null hypothesis is accepted and actually the difference in rating the acceptability of the resumptive strategy in Island versus NoIsland conditions was statistically non-significant. This result was, further, backed by the slight effect size, Cohen's *d*, (dividing the mean by the SD as a standardizer) and the CI going through zero. This reveals that the L2 learners in this study did not accept resumption as a valid strategy to resolve relative clause dependency in English in the presence or absence of an RC-island. On the other hand, the second null hypothesis is rejected on the basis of the statistically significant difference in using the gap strategy in NoIsland versus Island conditions. This finding was, moreover, supported by the large effect size and the CI. The third null hypothesis is rejected with respect to NoIsland conditions because the difference between the gap and resumptive strategies was statistically significant, the effect size was large, and the CI did not go through zero. This indicated that the L2 learners in the current study had the grammatical knowledge necessary to identify the gap strategy as the only available resolving strategy in the dependency at issue in English. On the other hand, the difference between these two strategies was statistically non-significant in Island conditions. This result could not be interpreted by virtue that L2 learners treated the resumptive strategy as ameliorating island effects, whereas the gap strategy did not. Rather, it can be mainly concluded that the L2 learners rated the resumptive strategy as completely unacceptable in English.

PAIRED-COMPARISONS	*M (SD)*	*t*	*p*	*df*	*d*	*95% CI*
RC_Island.RP *vs* RC_NoIsland.RP	.139 (.543)	1.536	.134	35	.256	[-.045, .322]
RC_Island.Gap *vs* RC_NoIsland.Gap	5.361 (.593)	54.249	.000	35	9.040	[5.160, 5.562]
RC_NoIsland.RP *vs* RC_NoIsland.Gap	-5.306 (.525)	-60.660	.000	35	-10.11	[-5.483, -5.128]
RC_Island.RP *vs* RC_Island.Gap	-.083 (.554)	-.902	.373	35	.149	[-.271, .104]

Table 2. The results of the statistical analysis of acceptability judgment of L2 data.

The average rating of each condition is visualized in Figure 2 below. As shown in Figure 2, the participants rated resumption in Island as well as NoIsland conditions as unacceptable (below 2 on the Likert scale used). Additionally, the L2 learners rated the NoIsland condition with the gap strategy as perfectly acceptable (above 6), yet as completely unacceptable in the corresponding island conditions (below 2).

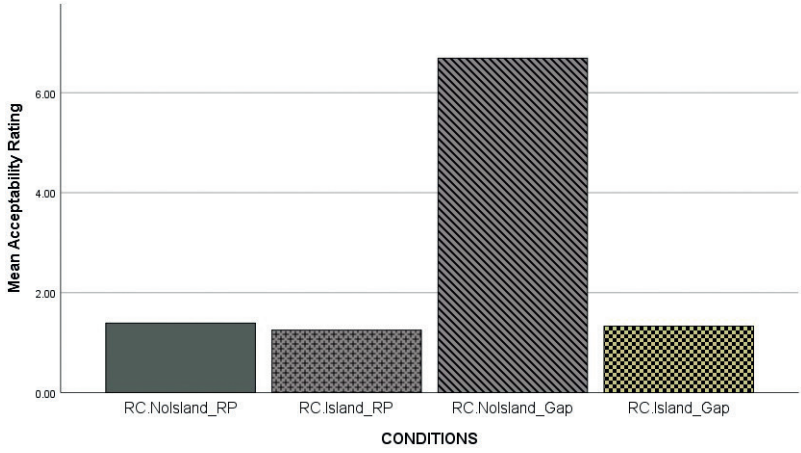

Figure 2. Acceptability judgment ratings of the conditions in L2 (English).

In a nutshell, the results of the offline task concerning English relative clauses reveal that the advanced adult L2 learners whose L1 is southern JA have the pre-requisite grammatical knowledge of the following:

a. Only the gap strategy is allowed to resolve the relative clause dependency in English in NoIsland conditions.
b. Gaps are rated unacceptable inside the RC-island. These L2 learners know that RC-island is a strong island in English that bans gaps.
c. The resumptive strategy is not allowed in Island or NoIsland conditions.

5.3. Experiment 3: Eyetracking reading study in L1 Jordanian Arabic

There are no previous online eyetracking reading studies on southern JA; thus, the characteristics of the processing of long-distance dependencies in this variety have not been investigated so far. To meet this end, an eyetracking reading experiment targeting online processing of southern JA was conducted to answer the following research questions:

1) What characterizes the online processing of resolving the relative clause dependency in southern JA?
2) Does the southern JA parser exhibit sensitivity to the RC-island when resolving the dependency at issue?

As stated in Section 3, southern JA speakers are getting more and more familiar with reading and writing in this variety due to the extensive use of their L1 in social media platforms like Facebook, WhatsApp, Instagram, Twitter, etc. This makes conducting an online reading task in JA possible and legitimate. The characters used in spelling are exactly identical to the Standard Arabic alphabet. No special spelling characters are peculiar to JA.

Materials, design, and procedure

I adopted a 2*2 factorial design with two independent variables (factors) with two levels each as parenthesized: ISLAND (RC_Island *vs* RC_NoIsland) and PLAUSIBILITY (Plausible *vs* Implausible). This design resulted in 4 conditions. I constructed the material similarly to Traxler & Pickering's (1996) stimuli. In all the examples I created, the first subcategorizers for the dependency-resolving strategy were transitive verbs. There was the region of interest (ROI) in the item set given in (16). Plausibility was manipulated through the use of the nouns *il-gi*ṣ*ah* 'the story' and *il-madi:nih* 'the city' by virtue that the noun, *il-gi*ṣ*ah* 'the story', and the first subcategorizer, i.e. transitive verb, *katab* 'wrote', plausibly matched (16a,c). Conversely, the noun, *il-madi:nih* 'the city', and the subcategorizer, *katab* 'wrote', plausibly mismatched (16b,d). As a consequence, the use of the noun, *il-gi*ṣ*ah* 'the story', rendered

the interpretation semantically plausible, whereas the use of the noun, *il-madi:nih* 'the city', resulted in a semantically implausible interpretation. Additionally, the first subcategorizer, *katab* 'wrote', occurred inside an RC-island (16c-d), but it occurred inside the counterpart NoIsland conditions (16a-b). Ambiguity and misanalysis in the implausible conditions were resolved on the preposition along with the RP it attached to. Since resumption is the only dependency-resolving strategy, the relative clause dependency in all of the examples was resolved by RPs. Both nouns at issue in every item set matched in number and gender features in the sense that either both were masculine and singular or both were feminine and singular to control for any confounding factors regarding agreement issues.

(16) a. RC_NoIsland.Plausible

aʕjab-at-ni:	*il-giṣah*	*illi:*	*il-ka:tib*	*biku:n*	katab-ha	*il-ʕa:m*
admire-3SG.F-me	the-story	that	the-author	be.3SG.M	write.3SG.M-it	the-year
wa-ħaka	*ʕan*-ha	*li-il-ʤami:ʕ*				
and-talk.3SG.M	about-it	to-the-all				

'I admired the story that the author wrote it last year and talked about it to everyone.'

b. RC_NoIsland.Implausible

#aʕjab-at-ni:	*il-madi:nih*	*illi:*	*il-ka:tib*	*biku:n*	katab-ha	*il-ʕa:m*
admire-3SG.F-me	the-city	that	the-author	be.3SG.M	write.3SG.M-it	the-year
wa-ħaka	*ʕan*-ha	*li-il-ʤami:ʕ*				
and-talk.3SG.M	about-it	to-the-all				

'I admired the city that the author wrote it last year and talked about it to everyone.'

c. RC_Island.Plausible

aʕjab-at-ni:	*il-giṣah*	*illi:*	*il-ka:tib*	*[illi*	katab-ha	*il-ʕa:m]*
admire-3SG.F-me	the-story	that	the-author	[that	write.3SG.M-it	the-year]
wa-ħaka	*ʕan*-ha	*li-il-ʤami:ʕ*				
and-talk.3SG.M	about-it	to-the-all				

'I admired the story that the author [who wrote it last year] and talked about it to everyone.'

d. RC_Island.Implausible

#aʕjab-at-ni:	*il-madi:nih*	*illi:*	*il-ka:tib*	*[illi*	katab-ha	*il-ʕa:m]*
admire-3SG.F-me	the-city	that	the-author	[that	write.3SG.M-it	the-year]
wa-ħaka	*ʕan*-ha	*li-il-ʤami:ʕ*				
and-talk.3SG.M	about-it	to-the-all				

'I admired the city that the author [who wrote it last year] and talked about it to everyone.'

All of the NoIsland experimental items involved one relative clause, whereas all their Island counterparts involved two relative clauses. To compensate for the difference in the number of words across conditions, in the NoIsland conditions I included a four-letter auxiliary *biku:n* 'has been' that was identical in number of characters and position to the

additional four-letter relative pronoun in the Island conditions. Both occurred as the fifth word in all conditions. Therefore, the ROI was linearly identical across the experimental items. Each experimental item included 10 words with 50 characters in total. The same applied to all practice and filler items. There were 26 characters between the beginning of the sentence and the ROI. All the items matched in length. Island and NoIsland experimental items matched in several relevant respects like the number of characters, the position of the ROI, length of words used, etc. Each participant saw seven experimental items on each condition. Each list had 3 practice items to familiarize the participants with the task. I also included 42 filler items in a 1.5:1 ratio of filler items to the experimental items. The task included 73 items as follows: 3 practice items, 28 experimental items, and 42 filler items. Twenty-eight experimental item sets were created and distributed into 7 lists using a Latin-Square design so that each list involved one version of each item. All the items were judged by the same 5 scholars of modern linguistics who are native speakers of southern JA who judged the items in Experiment 1. The lists were distributed across participants. The order of presenting items was randomized across participants. Each item (practice, filler, and experimental) was followed by a *yes/no* comprehension question to assure that the participants attended to the task.

The eyetracking reading task was conducted in the Psycholinguistic Lab at Mutah University. The participants were individually tested and the task was presented on an LCD monitor placed 70 cm from the participants. All the items were displayed on the screen as single-line sentences. The data were recorded at a sampling rate of 1000 Hz. The calibration and validation were performed at the beginning of the task and after twenty items were presented. The participants were instructed that these items were in southern JA taken from real chats among Jordanians on social media channels like Facebook. They were also told that they should read the sentences at their own pace and answer the comprehension questions by pressing certain keys on the keyboard. An average session lasted for 45 minutes. The same 36 participants took part in this task.

Results

The mean comprehension accuracy for the participants was 91.2%. Before running the statistical analysis, the trials that involved blinks, no fixated data or fixation, less than 90 ms in some regions were all excluded resulting in 14.1% of trials excluded. Filler and practice items were also excluded from the analysis. The collected eye-movement data were analyzed using SPSS IBM version 26. A paired-sample *t* test was employed to test the following null hypotheses with respect to each

reading measure, but they are put together here to save space. The reading measures are discussed below in this order: first-fixation duration (FFD), regression path duration, and total reading time (total RT).

(i) H_0: There are no differences in reading measures between plausible and implausible conditions in the online processing of relative clauses in southern JA.
(ii) H_0: There are no differences in reading measures with respect to plausibility mismatch between Island conditions and their NoIsland counterparts in the online processing of relative clauses in southern JA.

As far as FFD is concerned, the results of the statistical analysis are presented in Table 3 below. The results demonstrated a statistical difference between plausible and implausible conditions in Island as well as NoIsland conditions. This showed that a plausibility mismatch effect was found regardless of the presence or absence of islands. This finding was also supported by the line graph in Figure 3 with a considerable peak in implausible conditions in Region 6, the first subcategorizer, the transitive verb *katabha* 'wrote it'. In accordance with this finding, the first null hypothesis is rejected. On the other hand, the presence of islands did not prove an influential factor on the FFD and the second null hypothesis is accepted, as there was no statistical difference that can be attributed to the presence of the RC-island in southern JA.

PAIRED-COMPARISONS	*M (SD)*	*t*	*p*	*df*	*d*	*95% CI*
RC_NoIsland.Plausible *vs* RC_NoIsland.Implausible	-133.417 (34.967)	-22.893	.000	35	-3.816	[-145.248, -121.585]
RC_Island.Plausible *vs* RC_ Island.Implausible	-133.194 (34.208)	-23.362	.000	35	-3.894	[-144.769, -121.620]
RC_NoIsland.Plausible *vs* RC_Island.Plausible	-1.083 (2.771)	-2.346	.025	35	-.391	[-2.021, -.146]
RC_NoIsland.Implausible *vs* RC_Island.Implausible	-.861 (3.235)	-1.597	.119	35	-.266	[-1.956, .233]

Table 3. The results of the statistical analysis of FFD of the L1 (southern JA) reading data.

Figure 3 below portrays the average FFD per region per condition. A closer inspection of the line graph brings to light that there is a sharp increase in FFD on the ROI, Region 6, in implausible conditions compared to plausible conditions. On the other hand, the presence of the

RC-island does not have an effect, as obviously visualized in the figure where both lines representing RC.Island_Plausible and RC.NoIsland_ Plausible perfectly align together, as the implausible ones. The figure also portrays an increase in Region 9 at which ambiguity is resolved in implausible conditions.

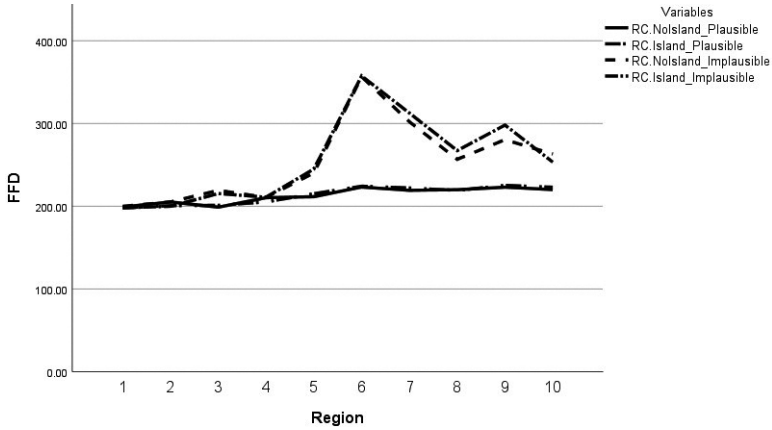

Figure 3. Mean FFD by region and condition of the L1 (southern JA) reading data. Region 6 corresponds to the first subcategorizer *katab* 'write' in the example sentences. Region 9 corresponds to the disambiguating preposition with the RP *ʕanha* 'about it'.

Turning to the regression path duration, the results are displayed in Table 4. The results showed that a plausibility mismatch effect was also found since there was a statistically significant difference in regression duration between plausible and implausible conditions in Island as well as NoIsland conditions, as shown by the statistical analysis of the first and second pairwise comparisons presented. Consequently, the first null hypothesis is rejected in accordance with the statistical difference between plausible conditions compared to implausible conditions regardless of the presence or absence of the RC-island. On the other hand, there was no effect for the presence of islands as the difference between Island versus NoIsland conditions in plausible conditions (third comparison) as well as implausible conditions (fourth comparison) was statistically non-significant. This finding was supported by the notably slight effect size. Therefore, the second null hypothesis is accepted, and I concluded that there was no effect of the RC-island on the regression path duration.

PAIRED-COMPARISONS	*M (SD)*	*t*	*p*	*df*	*d*	*95% CI*
RC_NoIsland.Plausible *vs* RC_NoIsland.Implausible	-139.750 (34.665)	-24.188	.000	35	-4.031	[-151.479, -128.021]
RC_Island.Plausible *vs* RC_Island.Implausible	-138.250 (36.887)	-22.488	.000	35	-3.748	[-150.731, -125.769]
RC_NoIsland.Plausible *vs* RC_Island.Plausible	-1.111 (5.492)	-1.214	.233	35	-.202	[-2.969, .747]
RC_NoIsland.Implausible *vs* RC_Island.Implausible	.389 (18.377)	.127	.900	35	.021	[-5.829, 6.607]

Table 4. The results of the statistical analysis of regression path duration of the L1 reading data.

Figure 4 below provides a visual representation of the average regression path duration per region per condition. It unmasks the lack of significance of the island effect because plausible conditions (both Island and NoIsland) tightly match, as do implausible conditions. The clear increase in the reading measure is in implausible conditions in the ROI region in contrast with the plausible conditions because of the reanalysis required. Another increase is in the 9th region due to disambiguation.

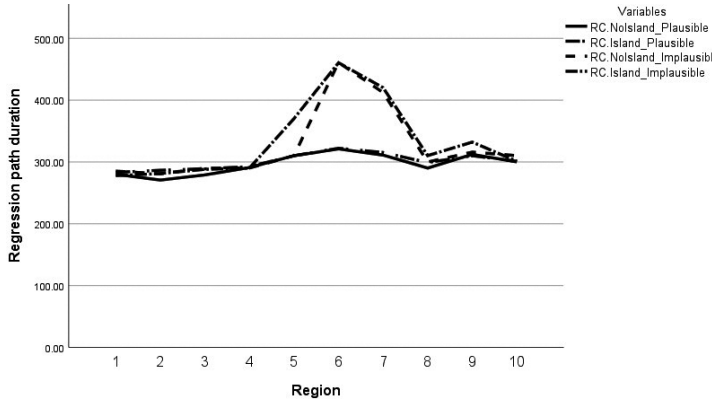

Figure 4. Mean regression path duration by region and condition of the L1 reading data. Region 6 corresponds to the first subcategorizer *katab* 'write' in the example sentences. Region 9 corresponds to the disambiguating preposition with the RP *ʕanha* 'about it'.

The last reading measure to consider with respect to the learners' L1 data is the total RTs. The relevant results are displayed in Table 5. Similar to the previous reading measures, a statistically significant difference in total RTs between plausible versus implausible conditions was found. This was not affected by the presence or absence of the RC-island since the plausibility mismatch effect was only found when plausible conditions

were compared to implausible conditions (first and second pairwise comparisons) in Island and NoIsland conditions. This finding led to the rejection of the first null hypothesis. The lack of significance of the island was further illuminated by the results of the third and fourth comparisons by which plausible conditions were compared in Island versus NoIsland conditions. This yields no statistical difference, nor implausible conditions. Accordingly, the second null hypothesis is accepted.

PAIRED-COMPARISONS	*M (SD)*	*t*	*p*	*df*	*d*	*95% CI*
RC_NoIsland.Plausible *vs* RC_NoIsland.Implausible	-148.556 (48.411)	-18.412	.000	35	-3.069	[-164.935, -132.176]
RC_Island.Plausible *vs* RC_Island.Implausible	-146.306 (47.728)	-18.393	.000	35	-3.066	[-162.454, -130.157]
RC_NoIsland.Plausible *vs* RC_Island.Plausible	-2.222 (6.655)	-2.003	.053	35	-.334	[-4.474, .030]
RC_NoIsland.Implausible *vs* RC_Island.Implausible	.028 (11.360)	.015	.988	35	.002	[-3.816, 3.872]

Table 5. The results of the statistical analysis of total RTs of the L1 reading data.

Like the results of the previous reading measures of this experiment, Figure 5 plainly exposes the increase in total RTs in the ROI, the 6th region, in implausible conditions in contrast with plausible conditions. Lines representing Island and NoIsland conditions align altogether in plausible as well as implausible conditions. This promotes the lack of significance of the island effect in southern JA data.

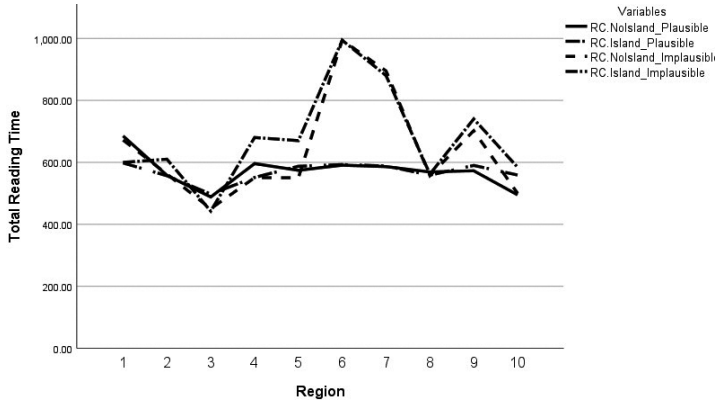

Figure 5. Mean total RTs by region and condition of the L1 reading data. Region 6 corresponds to the first subcategorizer *katab* 'write' in the example sentences. Region 9 corresponds to the disambiguating preposition with the RP *ʕanha* 'about it'.

All in all, the following conclusions can be derived from the results of Experiment 3.

1. There is a plausibility mismatch effect represented by the evident increase in reading measures (FFD, regression path duration, and total RTs) in implausible conditions compared to their corresponding plausible conditions.
2. L2 learners demonstrated evidence of postulating RPs as the parser encounters a potential subcategorizer, even if inside an RC-island.
3. L2 learners build abstract detailed structural representations as they project syntactic structure necessary for accommodating an RP.
4. An RC-island does not ban the postulation of an RP in southern JA. RPs are grammatical in this construction in southern JA.

5.4. Experiment 4: Eyetracking reading study in L2 English

The current study was undertaken during the COVID-19 pandemic, and it was not possible for me to recruit native speakers of English to be the control group. Therefore, I used the findings concerning native speakers' processing of relative clauses in English from Traxler & Pickering's (1996) and Omaki & Schulz's (2011) studies as the basis of comparison with native speakers' online processing. This experiment aimed at accounting for the following research questions:

a. Does the parser of the L2 learners resolve the relative clause dependency in English in the same way as the parser of native speakers of English in accordance with the reported behavior in previous studies?
b. Is the parser of the L2 learners constrained by the RC-island in processing the English relative clause dependency?

Material, design, and procedure

The design and stimuli of Traxler & Pickering's (1996: 465) were adopted in this experiment. A 2*2 factorial design was employed that contained two independent variables with two values each as follows: PLAUSIBILITY (Plausible *vs* Implausible) and ISLAND (RC_Island *vs* RC_NoIsland). Below is a typical sample of an item set used in Traxler & Pickering's (1996) study which I also employed, and constructed all of the experimental items accordingly. Here the noun *the book* and the first subcategorizer *wrote* plausibly matched resulting in a semantically plausible interpretation (17a,c). However, the noun *the city* and the first subcategorizer did not match, and so they rendered the interpretation semantically implausible (17b, d). The ROI of concern in this study was the first potential gap site *wrote*. The disambiguating regions were *about*

while in NoIsland conditions (17 a-b) and *saw while* in Island conditions (17 c-d). Similar to Experiment 3, the task contained 73 items distributed as follows: 3 practice items, 28 experimental items, and 42 filler items. All the items were judged by the same 3 scholars of linguistics who were native speakers of American English and judged the items of Experiment 2. The vertical slashes separating the words in example (17a) represent into which regions the experimental sentences were divided in a way to match the ROI and disambiguating region in Traxler & Pickering's (1996) stimuli while maintaining 10 regions like the southern JA experimental items.

(17) a. RC_NoIsland.Plausible
*We like | the book | that | the author| **wrote** unceasingly | and | with great dedication |*
1 2 3 4 5 6 7
about while | waiting for | a contract.
8 9 10
b. RC_NoIsland.Implauisble
*We like | the city | that | the author | **wrote** unceasingly | and | with great dedication |*
1 2 3 4 5 6 7
about while | waiting for | a contract.
8 9 10
c. RC_Island.Plausible
*We like | the book | that | the author | [who **wrote** unceasingly | and | with great*
1 2 3 4 5 6 7
dedication] | saw while | waiting for | a contract.
8 9 10
d. RC_Island.Implausible
*We like | the city | that | the author | [who **wrote** unceasingly | and | with great*
1 2 3 4 5 6 7
dedication] | saw while | waiting for | a contract.
8 9 10

Twenty-eight sets of experimental sentences were utilized; they were distributed into 7 lists of items using a Latin-Square design so that each list presented one version of each item. The same practice and filler items were used in all lists. The order in which the items were presented was randomized across participants. Each item was followed by a *yes/no* comprehension question. Overall, this experiment was mainly an eye-tracking reading task, and so the same procedure of Experiment 3 was adopted here; the same 36 participants did this task.

Results

The mean comprehension accuracy of the participants was 89.1%. Before conducting the statistical analysis, the trials that involved blinks, no fixated data, or fixation less than 90 ms in some regions were all excluded resulting in 17.3% of the trials were excluded. Filler and

Practice items were not included in the analysis. The collected eye-movement data were statistically analyzed using SPSS IBM version 26. A paired-sample *t* test was utilized to test the following null hypotheses with respect to each individual reading measure; however, they are grouped together here to save space.

(i) H_0: There are no differences in reading measures between plausible and implausible conditions.
(ii) H_0: There are no differences in reading measures with respect to plausibility between Island conditions and their NoIsland counterparts.

Table 4 below presents the results of the statistical analysis concerning the FFD on the 5th region, the ROI in this experiment. The results delineated that there was a statistically significant difference between plausible versus implausible conditions in NoIsland conditions (first pairwise comparison) and Island conditions (second comparison). Consequently, the first null hypothesis (i) should be rejected. This result provided evidence of a plausibility mismatch effect. Nonetheless, the participants who showed that they had the grammatical knowledge in the offline task that the RC-island is a strong island in English which bans postulating gaps inside, they did not manifest this grammatical knowledge in their L2 online processing. Instead, the results showed no effect of the island. The first and second pairwise comparisons unmasked a statistical difference between plausible and implausible conditions irrespective of the presence of the RC-island. Furthermore, the third and fourth comparisons revealed that the difference within plausible conditions or implausible conditions between Island versus NoIsland conditions was statistically non-significant undermining the role of the RC-island in L2 processing. As a result, the second null hypothesis is accepted.

PAIRED-COMPARISONS	*M (SD)*	*t*	*p*	*df*	*d*	*95% CI*
RC_NoIsland.Plausible *vs* RC_NoIsland.Implausible	-129.972 (32.526)	-23.975	.000	35	-3.996	[-140.978, -118.967]
RC_Island.Plausible *vs* RC_Island.Implausible	-132.556 (39.560)	-20.104	.000	35	-3.351	[-145.941, -119.170]
RC_NoIsland.Plausible *vs* RC_Island.Plausible	4.250 (15.315)	1.665	.105	35	.278	[-.932, 9.432]
RC_NoIsland.Implausible *vs* RC_Island.Implausible	1.667 (22.101)	.452	.654	35	.075	[-5.811, 9.145]

Table 6. The results of the statistical analysis of the FFD of the L2 (English) reading data.

The line graph in Figure 6 below shows that the lines representing plausible conditions in Island and NoIsland contexts align together, as do their implausible counterparts. The clear difference is in Region 5 in which an increase in FFD is depicted in implausible conditions in contrast with plausible ones regardless of the presence of the RC-island. Another clear increase in the FFD measured is in Region 8 where disambiguation is supposed to occur.

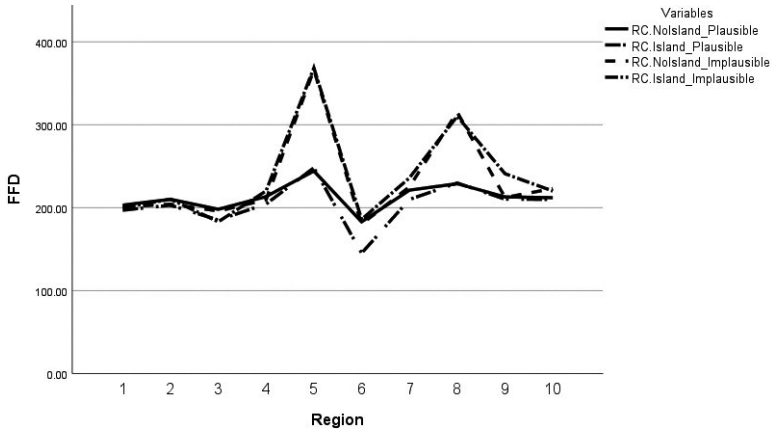

Figure 6. Mean FFD by region and condition of the L2 reading data. Region 5 corresponds to the first subcategorizer *wrote unceasingly* in the example sentences. Region 9 corresponds to the disambiguating regions *about while* in NoIsland conditions and *saw while* in island conditions.

The results of the statistical analysis of the regression path duration in the 5th region, the ROI, illuminated a statistically significant difference in implausible conditions compared to plausible conditions regardless of the presence of the RC-island. Thus, the first null hypothesis in (i) should be rejected, as a plausibility mismatch effect was found. The difference between Island conditions versus NoIsland conditions within plausible (third pairwise comparison) and implausible conditions (fourth comparison) was statistically non-significant. Accordingly, the second null hypothesis in (ii) is accepted.

PAIRED-COMPARISONS	*M (SD)*	*t*	*p*	*df*	*d*	*95% CI*
RC_NoIsland.Plausible *vs* RC_NoIsland.Implausible	-177.667 (23.652)	-45.070	.000	35	-7.512	[-185.669, -45.070]
RC_Island.Plausible *vs* RC_Island.Implausible	-184.250 (31.294)	-35.326	.000	35	-5.888	[-194.838, -173.662]
RC_NoIsland.Plausible *vs* RC_Island.Plausible	.750 (12.290)	.366	.716	35	.061	[-3.408, 4.908]
RC_NoIsland.Implausible *vs* RC_Island.Implausible	-5.833 (30.850)	-1.135	.264	35	-.189	[-16.272, 4.605]

Table 7. The results of the statistical analysis of the regression path duration of the L2 reading data.

Figure 7 below provides a visual representation of the average regression path duration per region per condition. The lines representing implausible conditions exhibit evident increase in the 5th and the 8th regions. The lines representing island conditions for plausible versus implausible conditions align together indicating that the effect of the RC-island is slight.

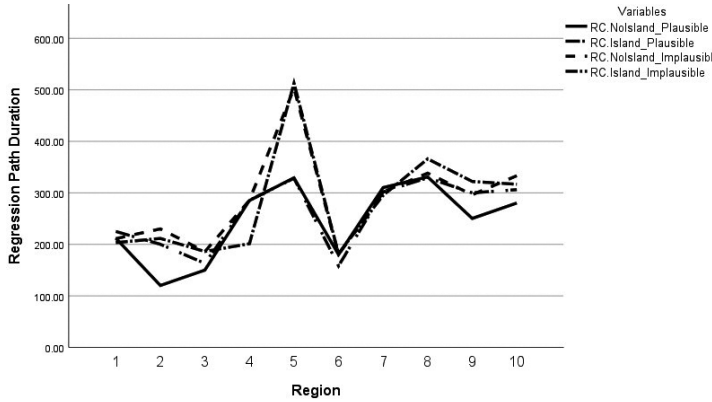

Figure 7. Mean regression path duration by region and condition of the L2 reading data. Region 5 corresponds to the first subcategorizer *wrote unceasingly* in the example sentences. Region 9 corresponds to the disambiguating regions *about while* in NoIsland conditions and *saw while* in island conditions.

The statistical analysis of the final reading measure, total RTs, is displayed in Table 8 below. The results were in line with all of the results concerning the other reading measures in Experiment 3 and 4. There was a statistical difference between plausible and implausible conditions.

Therefore, the first null hypothesis should be rejected since a plausibility mismatch effect, represented by the increase in the total RTs, was found. Similar to the other reading measures, no effect of the presence of the RC-island was found; thus, the second null hypothesis should be accepted.

PAIRED-COMPARISONS	*M (SD)*	*t*	*p*	*df*	*d*	*95% CI*
RC_NoIsland.Plausible *vs* RC_NoIsland.Implausible	-177.361 (60.231)	-17.668	.000	35	-2.945	[-197.740, -156.982]
RC_Island.Plausible *vs* RC_Island.Implausible	-176.861 (58.516)	-18.135	.000	35	-3.022	[-196.660, -157.062]
RC_NoIsland.Plausible *vs* RC_Island.Plausible	-8.806 (39.622)	-1.333	.191	35	-.222	[-22.212, 4.601]
RC_NoIsland.Implausible *vs* RC_Island.Implausible	-8.306 (42.706)	-1.167	.251	35	.194	[-22.755, 6.144]

Table 8. The results of the statistical analysis of the total RTs of the L2 reading data.

The line graph in Figure 8 below showcases the average total RTs per region per condition. Similar to the results of the other reading measures of Experiments 3 and 4, the lines representing the implausible conditions highlight a crystal-clear rise in the 5th and 8th regions. There was no increase with respect to the presence of the RC-island. The lines representing Island and NoIsland conditions tightly align together for plausible as well as implausible conditions.

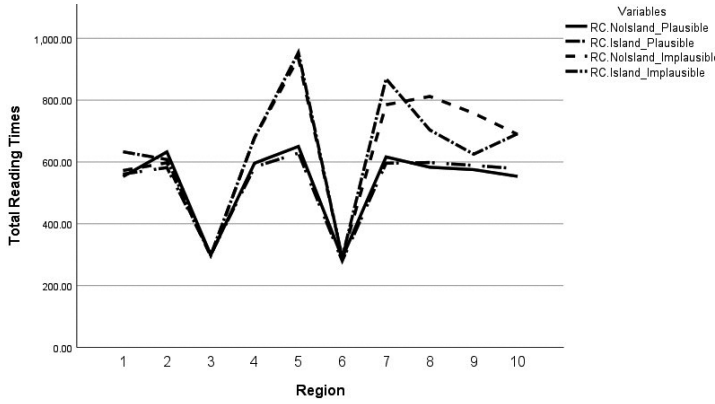

Figure 8. Mean total RTs by region and condition of the L2 reading data. Region 5 corresponds to the first subcategorizer *wrote unceasingly* in the example sentences. Region 9 corresponds to the disambiguating regions *about while* in NoIsland conditions and *saw while* in island conditions.

To sum up, the following conclusions can be drawn from the results of Experiment 4.

1) A plausibility mismatch effect was depicted in implausible conditions in contrast with their corresponding plausible conditions. This effect was indicated by the statistically significant increase in reading measures (FFD, regression path duration, and total RTs).
2) L2 learners did not demonstrate evidence of reflecting their offline grammatical knowledge that the RC-island is a strong island in English when processing the English relative clause dependency in an online task.
3) The processing behavior the L2 learners manifested in their online processing of the English relative clause dependency was similar to a great extent to their processing behavior in their online processing of the same dependency in their L1s. Therefore, it seems likely that these learners postulated RPs when they encountered the first potential subcategorizer even if it was inside an RC-island.
4) Similar to what they did in the processing of this dependency in their L1, L2 learners demonstrated evidence that they built abstract detailed structural representations as the parser projected the syntactic structure necessary for accommodating an RP. In short, what characterized their L2 processing was likely the L1 transfer of processing strategies rather than a shallow, less detailed structural processing.

5. Conclusions and general discussion

The current research was set up to test the assumptions of the SSH postulated by Clahsen & Felser (2006a, 2006b, 2006c, 2018) about L2 processing during real-time processing of long-distance relative clause dependency formation with an RC-island. The language-pair tested was L1 (southern JA) and L2 (English) because they are typologically different and employ different parsing strategies. Southern JA employs only the resumptive strategy, whereas English utilizes the active-gap creation strategy. In accordance to the methodologies of Traxler & Pickering's (1996) and that of Omaki & Schulz's (2011) studies, I conducted 4 related experiments: 2 offline acceptability judgment tasks and 2 online eyetracking reading experiments. The underlying goal of the offline tasks was to measure whether the participants had the pre-requisite grammatical knowledge of the dependency-resolving strategy of the relative clause dependency in their L1 and L2, and whether the depend-

ency is sensitive to an RC-island or not in each language when under no time pressure. The findings reveal that the target participants judge the resumptive strategy as the only dependency-resolving strategy of the relative clause dependency in their L1, yet they exhibit the required grammatical knowledge that only the gap strategy is the acceptable strategy to resolve this dependency in their L2. Furthermore, they had sufficient knowledge that RPs are allowed inside an RC-island in southern JA but not in English. On the other hand, the online reading tasks were employed to measure the participants' real-time processing of these constructions. The results delineate that a plausibility mismatch effect, measured by an increase in all the reading measures considered (first-fixation duration, regression-path duration, and total reading time), was found in Island and NoIsland conditions during online processing of these constructions in southern JA. This can be explained by assuming that the southern JA parser predicts an RP immediately upon encountering the first subcategorizer. However, the semantic mismatch between the displaced NP *ilmadi:nih* 'the city' and the first subcategorizer *katab* 'wrote' entails that the parser does a reanalysis and starts a new search while holding the displaced elements. This processing behavior was found in Island and NoIsland conditions entailing that the parser predicts and posits an RP upon encountering the first subcategorizer regardless of the presence of the RC-island because in southern JA RPs are allowed and ameliorate island violations in relative clauses. Likewise, a similar plausibility mismatch effect was found in L2 online processing of the corresponding construction in English inside both Island and NoIsland conditions. This finding is different from the processing of native speakers of English in the aforementioned studies in which a plausibility mismatch effect was found only in NoIsland conditions. The account proposed accordingly is that the parser of English predicts and posits a gap in NoIsland conditions when encountering the first subcategorizer and the implausible mismatch interpretation leads to a reanalysis which results in an increase in the reading measures. However, the parser does not posit a gap inside islands, and that is why no plausibility mismatch effect is found. As far as the results drawn from the eyetracking experiments in this study, the potential account is that L2 processing involves the same processing strategies of the learners' L1. In both cases, the parser predicts an RP and since RPs are not sensitive to the RC-island in southern JA, it seems that the parser posits the RP regardless of the presence of islands.

Putting all the findings together, the target participants have the grammatical knowledge necessary to resolve the relative clause dependency differently under no time pressure. Nonetheless, they seem to

transfer their L1 processing strategies during their L2 processing of the long-distance dependency. The conclusions drawn from the current study lend further support and evidence to the SSH with respect to the L1 transfer as an account of L1-L2 processing differences. However, the current study raises some questions and concerns regarding both arguments and claims postulated within the SSH. First, Clahsen & Felser (2018: 5) claim that "L1 transfer [should] influence processing only indirectly". The findings of the current study unmask a direct transfer of online processing strategies. L2 learners in this study apply the same L1 parsing strategies during their L2 online processing of the relative clause dependency. The second concern that the conclusions of the current study highlights is the SSH's main argument of shallow, less-detailed structural processing of L2 processing. Clahsen & Felser (2006c) themselves argue that full parsing involves detailed syntactic computation of the incoming elements of the sentence including abstract elements like gaps, whereas shallow parsing involves more reliance on semantic and pragmatic sources of information. The L2 processing in the current study evidently involves positing an RP when the first subcategorizer is encountered. This RP is akin to the abstract category like gaps. There is no evidence that parsing is only complete towards the end of the sentences when the final correct subcategorizer in the implausible conditions is reached. This unmasks that L2 learners do not mainly rely on semantic and pragmatic sources of information rather than incremental syntactic computation. Therefore, their parsing should be described as full rather than shallow in light of Clahsen & Felser's (2006c) clarification of this distinction. Furthermore, it cannot be the case that the L2 learners do not have the grammatical knowledge of the relevant peculiarities of English because the offline task clearly shows that they have the pre-requisite grammatical knowledge they need to process these constructions similarly to native speakers. All in all, the current study provides further evidence to the claim that the L1-L2 processing differences can be in part attributed to the L1 transfer of processing strategies when under time pressure.

Abbreviations

CNP = Complex Noun Phrase; CS = Coordinate Structure; FFD = first-fixation duration; JA = Jordanian Arabic; RTs = reading times; ROI = region of interest; RC = relative clause; RP = resumptive pronoun; SSH = Shallow Structure Hypothesis.

Bibliographical References

Aoun, Joseph; Benmamoun, Elabbas & Choueiri, Lina 2010. *The syntax of Arabic*. Cambridge: Cambridge University Press.

Belikova, Alyona & White, Lydia 2009. Evidence for the fundamental difference hypothesis or not? Island constraints revisited. *Studies in Second Language Acquisition* 31. 199-223.

Carreiras, Manuel & Clifton, Charles 1999. Another word on parsing relative clauses: Eyetracking evidence from Spanish and English. *Memory and Cognition* 27. 826-833.

Chao, Wynn & Sells, Peter 1983. On the Interpretation of Resumptive Pronouns. *North East Linguistics Society* 13. Issue 1, article 6. <scholarworks.umass.edu/nels/vol13/iss1/6>.

Choueiri, Lina 2018. Resumption in varieties of Arabic. In Benmamoun, Elabbas & Bassiouney, Reem (eds.), *The Routledge Handbook of Arabic Linguistics*. Abingdon: Routledge. 131-154.

Clahsen, Harald & Felser, Claudia 2018. Some notes on the Shallow Structure Hypothesis. *Studies in Second Language Acquisition* 40,3. 693-706.

Clahsen, Harald & Felser, Claudia 2006a. Grammatical processing in language learners. *Applied Psycholinguistics* 27. 3-42.

Clahsen, Harald & Felser, Claudia 2006b. Continuity and shallow structures in language processing. *Applied Psycholinguistics* 27. 107-126.

Clahsen, Harald & Felser, Claudia 2006c. How native-like is non-native language processing? *Trends in Cognitive Sciences* 10. 564-570.

Crain, Stephen & Fodor, Janet Dean 1985. How can grammars help parsers? In Dowty, David; Kartunnen, Lauri & Zwicky, Arnold (eds.), *Natural language parsing: Psychological, computational and theoretical perspectives*. 94-28. New York: Cambridge University Press.

Dallas, Andrea & Kaan, Edith 2008. Second language processing of filler-gap dependencies by late learners. *Language and Linguistics Compass* 2. 372-388.

De Vincenzi, Marica 1991. *Syntactic parsing strategies in Italian*. Dordrecht: Kluwer.

Engdahl, Elisabeth 1985. Parasitic gaps, resumptive pronouns, and subject extractions. *Linguistics* 23. 3-44.

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

Fernandez, Eva 2003. *Bilingual sentence processing*. Amsterdam: John Benjamins.

Frazier, Lyn 1987. Syntactic processing: Evidence from Dutch. *Natural Language & Linguistic Theory* 5. 51-59.

Garnsey, Susan; Tanenhaus, Michael & Chapman, Robert 1989. Evoked potentials and the study of sentence comprehension. *Journal of Psycholinguistic Research* 18. 51-60.

Gibson, Edward & Pearlmutter, Neal 1998. Constraints on sentence comprehension. *Trends in Cognitive Sciences* 2. 262-268.

Jackson, Carrie & Bobb, Susan 2009. The processing and comprehension of *wh*-questions among second language speakers of German. *Applied*

Psycholinguistics 30. 603-636.
Jackson, Carrie & Dussias, Paola 2009. Cross-linguistic differences and their impact on L2 sentence processing. *Bilingualism: Language and Cognition* 12. 65-82.
Jegerski, Jill; VanPatten, Bill & Keating, Gregory 2011. Cross-linguistic variation and the acquisition of pronominal reference in L2 Spanish. *Second Language Research* 27. 481-507.
Keshev, Maayan & Meltzer-Asscher, Aya 2017. Active dependency formation in islands: How grammatical resumption affects sentence processing. *Language* 93. 549-68.
Kroch, Anthony 1981. On the role of resumptive pronouns in amnestying island violations. In CLS 17: Papers from the 17th Regional Meeting of the Chicago Linguistic Society. 125-135. Chicago, Illinois: Chicago Linguistic Society.
Lee, Ming-Wei 2004. Another look at the role of empty categories in sentence processing (and grammar). *Journal of Psycholinguistic Research* 33. 51-73.
Malkawi, Nouman & Nicolas Guilliot 2007. Reconstruction and islandhood in Jordanian Arabic. In Mughazy, Mustafa (ed.), *Perspectives on Arabic linguistics XX: Papers from the 20th Annual Symposium on Arabic linguistics*. Amsterdam: John Benjamins. 87-104.
Martohardjono, Gita & Gair, James 1993. Apparent UG accessibility in second language acquisition: Misapplied principles or principled misapplications. In Eckman, Fred (ed.), *Confluence: Linguistics, L2 acquisition and speech pathology*. Amsterdam: Benjamins. 79-103.
Marinis, Theodore; Roberts, Leah; Felser, Claudia & Clahsen, Harald 2005. Gaps in second language sentence processing. *Studies in Second Language Acquisition* 27. 53-78.
McCloskey, James 2006. Resumption. In Everaert, Martin & van Riemsdijk, Henk (eds.), *The Blackwell Companion to Syntax*. Blackwell Publishing. 94-117.
McCloskey, James 2017. Resumption. In Everaert, Martin & van Riemsdijk, Henk (eds.), *The Blackwell Companion to Syntax*. Blackwell Publishing. 84-117.
Omaki, Akira & Schulz, Barbara 2011. Filler-gap dependencies and island constraints in second language sentence processing. *Studies in Second Language Acquisition* 33. 563-588.
Papadopoulou, Despina & Clahsen, Harald 2003. Parsing strategies in L1 and L2 sentence processing. *Studies in Second Language Acquisition* 25. 501-528.
Phillips, Colin & Wagers, Matt 2007. Relating structure and time in linguistics and psycholinguistics. In Gaskell, Gareth (ed.), *Oxford Handbook of Psycholinguistics*. Oxford: Oxford University Press. 739-756.
Pickering, Martin & Traxler, Matthew 2003. Evidence against the use of subcategorization frequency in the processing of unbounded dependencies. *Language and Cognitive Processes* 18. 469-503.
Pritchett, Bradley 1992. Parsing with grammar: Islands, heads, and garden paths. In Goodluck, Helen & Rochemont, Michael (eds.), *Island constraints: Theory, acquisition, and processing*. Dordrecht: Kluwer Academic. 321-349.
Roberts, Leah 2003. *Syntactic processing in learners of English*. PhD dissertation. University of Essex, Colchester.
Ross, John Robert 1967. *Constraints on variables in syntax*. PhD dissertation.

Massachusetts Institute of Technology, Cambridge.
Traxler, Matthew & Pickering, Martin 1996. Plausibility and the processing of unbounded dependencies: An eye-tracking study. *Journal of Memory and Language* 35. 454-475.
Schachter, Jacquelyn 1990. On the issue of completeness in second language acquisition. *Second Language Research* 6. 93-124.
Schulz, Barbara 2006. *Wh-scope marking in English interlanguage grammars: Transfer and processing effects on the second language acquisition of complex wh-questions*. PhD dissertation. University of Hawaii, Honolulu.
Shlonsky, Ur 1992. Resumptive pronouns as last resort. *Linguistic Inquiry* 23. 443-468.
Sells, Peter 1984. Syntax and semantics of resumptive pronouns. PhD dissertation. Massachusetts Institute of Technology, Cambridge.
Stowe, Laurie 1986. Evidence for on-line gap location. *Language and Cognitive Processes* 1. 227-245.
Wagers, Matt & Phillips, Colin 2009. Multiple dependencies and the role of the grammar in real-time comprehension. *Journal of Linguistics* 45. 395. <doi:10.1017/ S0022226709005726>.
White, Lydia 1988. Island effects in second language acquisition. In Flynn, Suzanne & O'Neil, Wayne (eds.), *Linguistic theory in second language acquisition*. Dordrecht: Kluwer Academic. 144-172.
Williams, John 2006. Incremental interpretation in second language sentence processing. *Bilingualism: Language and Cognition* 9. 71-88.
Williams, John; Möbius, Peter & Kim, Choonkyong 2001. Native and non-native processing of English wh- questions: Parsing strategies and plausibility constraints. *Applied Psycholinguistics* 22. 509-540.
Yoshida, Masaya 2006. *Constraints and mechanisms in long-distance dependency formation*. PhD dissertation. University of Maryland, College Park.