UNIVERSITY *of* York

This is a repository copy of *Minding the Gap:Similarity-based Interference in L1 and L2 Processing of Long-distance Wh-dependencies*.

White Rose Research Online URL for this paper:
https://eprints.whiterose.ac.uk/221000/

Version: Published Version

Article:

White Rose
university consortium
Universities of Leeds, Sheffield & York

eprints@whiterose.ac.uk
https://eprints.whiterose.ac.uk/

CAMBRIDGE
UNIVERSITY PRESS

ORIGINAL ARTICLE

Minding the gap: Similarity-based interference in L1 and L2 processing of long-distance Wh-dependencies

Ehsan Solaimani

Department of Language and Linguistic Science, University of York, York, UK
Email: ehsan.solaimani@york.ac.uk

Abstract

A central question in the second language (L2) processing literature has been whether and under what conditions readers reactivate copies of syntactic movement operations at structurally defined gap sites. The present study contributes to the debate by examining the role of similarity-based interference in processing intermediate copies in long-distance dependencies with either three similar description noun phrases (NPs) (*the nurse*, *the doctor*, *the patient*) or two similar (*the nurse*, *the doctor*) and one dissimilar NP (*John*). Sixty-nine advanced L2 readers of English with either French (+ wh-movement) or Persian (− wh-movement) as their L1 and 33 native English readers (+ wh-movement) participated in a self-paced reading task involving long-distance dependencies. The results indicate that L2 readers process wh-dependencies in the same way as native readers, both in structures with similar and dissimilar NPs. This suggests that highly advanced L2 readers reactivate moved elements at inter-causal boundaries and process long-distance wh-dependencies in the same way as native readers, especially when the NPs involved in a dependency relation are sufficiently distinguishable.

Keywords: Second language processing; similarity-based interference; wh-dependency; working memory

Introduction

Many studies have investigated the processing of filler-gap dependencies, i.e., structures where a verbal argument (wh-filler) is moved from its canonical position (gap) to an earlier position. Evidence relying on a variety of experimental paradigms suggests that syntactic representations computed during the processing of these structures include abstract elements, such as copies of moved wh-fillers (priming: Dekydtspotter & Miller, 2013; Felser & Roberts, 2007; Roberts et al., 2007; pupillometry: Fernandez et al., 2018; functional magnetic resonance imaging: Pliatsikas et al., 2017; self-paced reading (SPR): Keine, 2020; Marinis et al., 2005; Pliatsikas & Marinis, 2013). For illustration, consider

an example of a long-distance filler-gap dependency (Gibson & Warren, 2004, Appendix A, p. 76: Experimental Items):

(1) *The general [$_{CP}$ who [$_{TP}$ the adviser thought [$_{CP}$ <who> that [$_{TP}$ the sergeant's message had angered <who>]]]] was attempting to appear calm.*[1]

The movement of the displaced filler *<who>* in wh-movement languages such as English undergoes successive cyclicity, that is, before moving to the clause initial position prior to *the adviser*, it lands on the specifier position of the embedded clause phrase (CP). This is because complement clauses such as the one in *(1)* (i.e., *that the sergeant's message had angered*) create what are known as island configurations (Ross, 1967). In these structures, the filler *who* can only cross one of either CP or tense phrase (TP) at a time, resulting in two smaller movements in *(1)*. Initially, the filler *<who>* crosses the CP (*that the sergeant's message had angered*) and then the TP (*the adviser thought that the sergeant's message had angered*). In other words, the fronting of the direct object of *angered* leaves a copy behind prior to *that*, which breaks the long-distance dependency into two smaller movements (Chomsky, 1995).[2]

For native (L1) processing, empirical evidence in reaction time (RT) experiments for the reactivation of the intermediate copy in wh-movement languages is observed as longer RTs on regions that include the intermediate structure, suggesting additional processing of encoding *<who>*, as well as faster RTs at the subcategorizing verb (*angered*) as a result of the intermediate gap mediating the long-distance dependency (Gibson & Warren, 2004). For example, Gibson and Warren (2004) examined RTs in an SPR task in sentences such as *(1)* and *(2)* (Gibson & Warren, 2004, p. 61).

(2) *The general who the adviser's thoughts about the sergeant's message had angered <who> was attempting to appear calm.*

Because there is no intervening clause between the wh-filler and its canonical positions in *(2)*, its integration with the subcategorizing verb *angered* crosses a longer dependency distance than in *(1)*. Gibson and Warren (2004) observed that native English readers had longer RTs on the complementizer *that* in *(1)* compared to the preposition *about* in *(2)*, while the subcategorizer *angered* in *(1)* was read significantly faster than the same word in *(2)*, thus suggesting RT facilitation due to the presence of the intermediate structure (see Keine, 2020, for a recent extension of Gibson & Warren's study). As for L2 processing, some have suggested that unlike native readers who reactivate movement copies only at structurally defined gap sites (specifier position of the embedded CP), the displaced filler remains active in L2 readers' memory throughout the entire sentence regardless of L1 (+/− wh-movement) (e.g., Felser & Roberts, 2007; Marinis et al., 2005), although this has been disputed in other studies (e.g., Berghoff, 2023; Pliatsikas & Marinis, 2013).

Crucially, forming long-distance dependencies such as *(1)* and *(2)* poses a unique challenge to the sentence processing mechanism, not least because of the need to retain a filler in memory across referentially complex intervening materials, which can cause processing overload. One potential source of processing difficulty in

(1) and *(2)* may be the similarity of the intervening noun phrases (NPs) *the general, the adviser,* and *the sergeant,* all of which provide descriptive information about the referent's occupational title. This similarity can create interference effects by disrupting the integration of <*who*> with its subcategorizer *angered* (Gordon et al., 2002, 2006).

The present study builds on and extends previous research to investigate L1 and L2 processing of long-distance wh-dependencies by examining the impact of interference caused by the similarity of intervening NPs.

Background

The present study is a partial replication of Marinis et al. (2005). As such, we take a close look at the details of the reported experiment. Marinis et al. (2005) adapted Gibson and Warren's (2004) materials to investigate in an SPR task L2 processing of English long-distance dependencies by adult L1 readers of Chinese (− wh-movement), Japanese (− wh-movement), German (+ wh-movement), and Greek (+ wh-movement), as well as by a group of L1-English readers (Marinis et al., 2005, p. 61):

(3a) *The nurse who the doctor argued <who> that the rude patient had angered <who> is refusing to work late.* Extraction-VP
(3b) *The nurse who the doctor's argument about the rude patient had angered <who> is refusing to work late.* Extraction-NP
(3c) *The nurse thought the doctor argued that the rude patient had angered the staff at the hospital.* Non-extraction-VP
(3d) *The nurse thought the doctor's argument about the rude patient had angered the staff at the hospital.* Non-extraction-NP

The extraction sentences involved the movement of the wh-filler <*who*> across either a VP *argued (3a)* or an NP *argument (3b),* whereas the corresponding non-extraction structures *(3c)* and *(3d)* did not involve a similar movement. Marinis et al. (2005) hypothesized that if readers access intermediate copies, extraction-VP structures *(3a)* should elicit longer RTs on the third region *(that, about)* compared to non-extraction-VP structures *(3c)* due to the intermediate copy, whereas no such RT difference should be observed between the two NP structures *(3b)* and *(3d).* Additionally, following Gibson and Warren (2004), Marinis et al. (2005) predicted that the subcategorizer region 5 *angered* should be read faster in extraction-VP than in extraction-NP structures, since the intermediate copy in *(3a)* reduces the dependency distance between *who* and *angered.* By contrast, no such difference should be observed between *(3c)* and *(3d).*

Marinis et al. (2005) observed longer RTs for the L1-English readers on the third region *(that, about)* in extraction compared to non-extraction conditions. In addition, RTs at the subcategorizer region 5 for L1-English participants were faster in extraction-VP than extraction-NP conditions. On the other hand, Marinis et al. (2005) reported that none of the L2 groups showed longer RTs at region 3 *(that, about)* in extraction compared to non-extraction conditions. Furthermore, RTs

were not faster in extraction-VP compared to extraction-NP structures for any of the L2 groups. Marinis et al. (2005) interpreted their results as suggesting that L2 dependency formation in long-distance wh-structures is not facilitated by the intermediate copy, unlike the case of L1 processing of these structures which involves reactivation of the intermediate copy at the spec (CP) of the embedded clause. These results are consistent with the shallow structure hypothesis (Clahsen & Felser, 2006, 2018), according to which L2 learners are ultimately capable of acquiring certain grammatical rules (predominantly morphological), but they appear to rely more heavily on non-syntactic (e.g., subcategorization information at the verb *angered*) rather than abstract syntactic information (e.g., intermediate copy) to comprehend syntactically complex structures.

However, Marinis et al.'s (2005) results are open to interpretation. First, Martinis et al. directly compared L1 and L2 RTs on the same regions 3 (*that, about*) and 5 (*angered*), assuming that the two reflect the same computational moments in L1 and L2 processing. However, L2 processing tends to be slower than L1 processing, which makes direct comparisons between L1 and L2 processing potentially unrevealing (c.f. "comparative fallacy in L2 processing research," Dekydtspotter et al., 2006). To illustrate this, Dekydtspotter et al. (2006) reanalyzed Marinis et al.'s (2005) data on the region immediately following the complementizer, i.e., region 4 (*the rude patient*), and found that the L1-Japanese and L1-German readers had longer RTs in extraction-VP compared to non-extraction-VP sentences, suggesting that at least some L2 readers in Marinis et al.'s (2005) study processed the intermediate copy, albeit with delay.

Second, the participants in Marinis et al. (2005) were reportedly at an upper intermediate proficiency level, thus raising the possibility that they might not have been sufficiently advanced (Solaimani & Marefat, 2024) and immersed in English to display native-like processing (Pliatsikas & Marinis, 2013). Pliatsikas and Marinis (2013) replicated Marinis et al.'s (2005) study on two groups of advanced L2-English readers of L1-Greek: one with only classroom exposure to L2 English and another with a mean of 9 years of exposure who were more proficient in English than the first group. They reported that the group with limited exposure to English showed similar processing behavior as Marinis et al.'s participants (2005). However, the group with more exposure behaved like L1-English readers, such that their RTs were facilitated at the subcategorizing verb due to the intermediate copy.

Finally, a closer look at the materials in Marinis et al. (2005) reveals that all of their experimental sentences involved three [+human] NPs (the nurse, the doctor, and the patient), two preceding region 3 (*that, about*) and one between region 3 and 5 (*angered*), which might have created similarity-based interference effects. According to cue-based parsing models (Lewis et al., 2006), memory retrieval of the sentence subject *the nurse* in (3) as the object of the verb *had angered* causes difficulty during processing. That is, the verb *angered* triggers the retrieval of the sentence subject *the nurse*, but the linearly closer NPs *the doctor* and *the rude patient* might interfere in this retrieval, also known as similarity-based interference. While similarity-based interference has been studied extensively in L1 processing, much less is known about how it influences L2 processing of long-distance wh-structures. It is possible that due to their less automatic lexical access (Hopp, 2018), L2 readers' lack of RT facilitation at the subcategorizer *angered* in Marinis et al. (2005) reflects

additional difficulty associated with the processing of three similar NPs (Cunnings, 2017), rather than access or lack thereof to the intermediate copy.

To investigate the impact of NP similarity, Cunnings and Fujita (2021) manipulated whether the local subject and object in relative clauses were proper names (e.g., *Rebecca*) or definite descriptions (e.g., *the girl*), as below (ibid, p. 8). The participants were native English readers and intermediate-to-advanced L2 English readers of various L1 backgrounds.

> **(4a) Subject-extraction**
> *The boy that <who> saw the girl/Rebecca the other day, walked through the park.*
> **(4b) Object-extraction**
> *The boy that the girl/Rebecca saw* <who> *the other day, walked through the park.*

Cunnings and Fujita (2021) reported that while RTs in both L1 and L2 were unaffected in subject extractions due to NP similarity, longer RTs were observed at the relative clause region (*that the girl/Rebecca saw the other day, that saw the girl/ Rebecca the other day*) in object extractions with matched NPs (two description NPs), compared to object extractions with mismatched NPs (one description, one proper NP). Cunnings and Fujita (2021) concluded that the reduction of similarity between different NPs in a long-distance dependency facilitates L1 and L2 processing of these structures.

The current study

Against this background, the present study investigates L2 processing of intermediate copies of movement. Specifically, we examine three different factors as the source of potential L1-L2 differences in processing intermediate copies in English. The first is the influence of L1, as also investigated by Marinis et al. (2005). To this aim, we compare L1-French and L1-Persian readers of L2-English in how they process sentences such as *(3)*. French is similar to English in that it is an SVO language that derives RCs by wh-movement operations in its standard form (Rowlett, 2007), although colloquial French tends to leave wh-words in situ by avoiding the "Qu'est-ce que" structure (c.f. "Qu'est-ce que tu manges?" for "What are you eating?" vs. "Tu manges quoi?" for "you are eating what?"). By contrast, Persian is an SOV language (Karimi, 2005) that does not allow wh-movements in RCs (Karimi & Taleghani, 2007).[3]

Following Juffs (2005), we consider two possibilities for potential L1 transfer effects. One is that the underlying syntactic representations of RCs may transfer from L1 to L2. If there is influence of L1 on L2 grammar, it is expected that L2 readers of English with no wh-movement L1s such as Persian would be less likely to compute intermediate copies, compared to L2 readers of English with wh-movement L1s such as French. This is because long-distance dependencies in Persian do not involve movement, and therefore, no intermediate copy is posited at inter-clausal boundaries, as in the Persian example (5)[4].

(5) *pæræstɒr-i*ᵢ [CP *ke* [TP *doʊktoʊr goʊft-Ø* [CP *ke* [TP *bimɒr-e bi-ædæb (u-ra)*ᵢ
    æsæbɒni kærd]]]]
nurse-RES that doctor argued-3SG that patient-EZAF without-politeness
(him/her-ACC) angry did
"the nurse who the doctor argued that the rude patient angered"

This is in contrast to French long-distance RCs as in *(6)*, where a null wh-operator <Op> moves from the position following *avait fâchée* (anger) to the intermediate position prior to *que* (that) and then to the frontal position at spec (CP). Although in the French example, it is a null wh-operator that undergoes movement, unlike English where an overt wh-pronoun moves, both French and English derive long-distance dependencies by means of wh-movement operations (Rowlett, 2007). Therefore, under an L1-transfer account, we expect L1-French readers to perform more similarly to L1-English readers, compared to L1-Persian readers.

(6) *L'infirmière* [CP <Op> *que* [TP *le docteur prétendait* [CP <Op> *que* [TP *le
    malade malpoli avait fâchée* <Op>]]]].
the nurse <Op> that the doctor argued <Op> that the patient rude
angered <Op>
"the nurse who the doctor argued that the rude patient angered"

The other possibility is that L2 readers may transfer their L1 processing preferences while reading sentences in their L2, despite potentially target-like syntactic representations (Solaimani et al., 2023). Juffs (2005) argues that processing RCs in head-final languages pose more WM load to the parser than processing RCs in head-initial languages. Specifically, Persian syntactically allows optional resumptives (e.g., "u-ra," equivalent to him/her in English) in object RCs such as *(5)* (Solaimani et al., 2023; Taghvaipour, 2005). It is possible that L1-Persian readers process English RCs according to the demands of their L1, that is, they may expect a case-marked resumptive pronoun prior to the verb. The fact that resumptive pronouns are not grammatical in standard English in sentences such as *(3)* may lead to additional processing difficulty by L1-Persian readers (compared to L1-French readers), whose L1 does not allow resumptive pronouns (Rowlett, 2007).

A second potential factor, which has not been previously examined in studies on intermediate copies, is the effect of memory overload due to similarity-based interference. Upon encountering the wh-filler *who* in *(3)*, the parser needs to encode and store it in working memory while processing the intervening materials, until it can be successfully retrieved with the subcategorizing verb *angered*. We investigate the extent to which computation of the intermediate copy at the clause boundary—prior to *that/about* in *(3)*—is affected by the additional memory load incurred by processing three similar NPs (*the nurse, the doctor, and the patient*) vs. two similar (*the nurse, the doctor*) and one dissimilar NP (*John*). We examine the effect of NP similarity by manipulating the descriptive vs. proper name status of the NP subject of the most embedded clause and explore potential RT facilitation at the subcategorizing verb *angered* due to this manipulation.

Finally, as the third potential factor, we also investigate how individual differences in proficiency and length of immersion experience affect L2 processing of intermediate copies. This will allow us to provide a more comprehensive understanding of L2 processing of intermediate copies. Overall, the research questions (RQs) are as follows:

(RQ1) Is there any difference between L1-English and L2-English readers (L1-French and L1-Persian) in processing intermediate copies in English long-distance wh-dependencies?

(RQ2) How does L1 (French, Persian) impact the processing of intermediate copies in L2 English?

(RQ3) How does the reduction of similarity-based interference between the three NPs in long-distance dependencies such as *(3)* impact the processing of intermediate copies in L1 and L2 English?

(RQ4) What are the impacts of proficiency in English and length of immersion experience in an English-speaking environment on L2 processing of intermediate copies?

Method

We first attempted to replicate Marinis et al. ([2005](#)) by including four experimental conditions, repeated below, crossing extraction (extraction, non-extraction), and phrase (VP, NP):

(3a) *The nurse who the doctor argued <who> that the rude patient had angered <who> is refusing to work late.* Extraction-VP (matched)

(3b) *The nurse who the doctor's argument about the rude patient had angered <who> is refusing to work late.* Extraction-NP (matched)

(3c) *The nurse thought the doctor argued that the rude patient had angered the staff at the hospital.* Non-extraction-VP (matched)

(3d) *The nurse thought the doctor's argument about the rude patient had angered the staff at the hospital.* Non-extraction-NP (matched)

Notice that the above sentences involve three matched NPs (*the nurse, the doctor, and the patient*) and therefore are expected to create similar levels of similarity-based interference.

We also included two additional extraction (mismatched) conditions *(7a)* and *(7b)* to investigate the impact of similarity-based interference. These conditions had similar NPs and VPs as the extraction-VP and extraction-NP structures *(3a)* and *(3b)*, but the subject of the most embedded clause always involved a proper name (*John had fascinated, John's argument about the journalist had fascinated*) (for a full list of materials, see the OSF link[5]). Therefore, these conditions involved mismatched NPs, i.e., two similar NPs (i.e., *the politician, the journalist*) and one dissimilar NP prior to the subcategorizing verb (i.e., *John*).

Table 1. Participants' biographical information, c-test scores, and comprehension accuracy on SPR task

	L1-English (*n* = 32)		L1-French (*n* = 33)		L1-Persian (*n* = 36)	
	Mean (SD)	Range	Mean (SD)	Range	Mean (SD)	Range
Age (years)	35.00 (3.42)	20–51	33.00 (3.67)	19–39	33.00 (5.47)	18–63
Age of immigration to English-speaking countries	—	—	22.90 (2.77)	17–29	28.80 (7.13)	17–41
Months of residence in English-speaking countries (immersion)	—	—	88.80 (50.20)	19–221	56.20 (68.10)	2–277
Proficiency (c-test score, max 10)	7.68 (.54)	6.09–8.38	7.43 (.63)	5.96–8.44	7.03 (.62)	5.46–8.31
SPR accuracy	.87 (.06)	.78–.97	.85 (.06)	.70–.98	.85 (.06)	.73–.98

(7a) *The politician who the journalist stated that **John** had fascinated is calling a press conference.* Extraction (VP), mismatched

(7b) *The politician who **John's** statement about the journalist had fascinated is calling a press conference.* Extraction (NP), mismatched

Participants

There were three groups of participants: 34 L1-French, 41 L1-Persian, and 33 L1-English readers. All the L1-French, L1-English, and 5 of L1-Persian participants were recruited through Prolific (www.prolific.co). The remaining L1-Persian readers were recruited through social media. All participants were naive with respect to the purpose of the experiment. The L1-English group reported that they were fluent only in their L1, and the L2 participants reported that they were not fluent in any additional language other than their L1s (French or Persian) and L2-English.

Data were included from those participants who scored above 70% on the SPR task comprehension questions, leading to the removal of data from 1 participant in the English group (3.03%), 1 in the French group (2.94%), and 5 in the Persian group (12.20%). The remaining data reported below came from 32 English natives and 69 L2 readers of English (L1-French: 33; L1-Persian: 36). Table 1 provides a summary of the participants' biographical information as well as proficiency and comprehension accuracy scores on the SPR task.

Pre-tasks

  a. ***Background information:* Questionnaire**

All participants completed a language history questionnaire to provide information about their experience learning English. The questionnaire was composed of items on participants' L1, number of years living in an English-speaking country, and other

languages known. A t-test showed that the L1-French participants had significantly longer immersion experience than the L1-Persian participants (t(64) = 2.31, *p* = .02).

b. *Proficiency: C-test*

Additionally, a c-test was administered to all participants (Keijzer, 2007), which required the completion of five truncated passages (Cronbach's alpha = .91). A c-test is a variant of a cloze task where the second half of every second word is omitted, and the participants are expected to fill in the gaps to reconstruct the passage (Klein-Braley & Raatz, 1984). Many studies have suggested that c-tests tap into both lower and higher order processing skills (for review, see Trace, 2020), although this is not completely uncontroversial (e.g., see Park, 1998). Overall, there were 100 items (20 per passage). The responses were assessed on a 9-point scale, ranging from 0 for a left-blank item to 9 for an item where the elicited and expected response matched completely (see the OSF page for full details). The overall proficiency score was derived by calculating the average score across all items.

An analysis of variance showed that there was a significant difference in c-test scores among the groups (F(2, 98) = 10.46, *p* < .001): Tukey comparisons indicated that L1-English readers scored higher than L1-French and L1-Persian readers, even though this difference was statistically significant only for the Persian group (estimate = .67, SE = .15, t(98) = 4.53, *p* < .001, *d* = 1.10). The difference between the English and French readers was not significant (t(98) = 1.78, *p* = .18), whereas the French readers scored significantly higher than the Persian readers (estimate = .40, SE = .15, t(98) = 2.73, *p* = .02, *d* = .66).

Materials

The participants read a total of 60 sentences in the SPR task, of which 2 were practice items, 36 experimental items, and 22 fillers. Of the total of experimental items, 24 had definite description NPs as in *(3)* (matched), adapted from Marinis et al. (2005), and 12 had a mixture of proper names and definite descriptions as in *(5)* (mismatched). The experimental items and fillers were distributed across 6 lists in a Latin square design. Each list contained 6 sentences per condition (6 * 6 = 36), and the items in each list were pseudorandomized to disguise the purpose of study. All experimental items and fillers were followed by *yes-no* comprehension questions to ensure that the participants were attentive to the task. Each sentence was segmented into 6 presentation regions in a noncumulative technique as in *(6)*:

Region:	1	2	3	4	5	6
(6)	*The nurse who*	*the doctor argued*	*that*	*the rude patient*	*had angered*	*is refusing to work late*
Did the doctor's argument anger the nurse?			Yes/ No			

Procedure

All data were collected online using Qualtrics, version (2020), and Ibex Farm (Zehr & Schwarz, 2018). The different tasks were administered in two separate sessions,

with approximately 5 days in between. Initially, all participants completed the language history questionnaire on Qualtrics and the c-test on Ibex. Subsequently, the participants completed the SPR task on Ibex.

Predictions

Regarding RQ1, we expected L1-English readers to encode the movement copy at region 3 and reactivate it at region 5 when reading the critical verb (*had angered*), consistent with previous studies supporting the psychological reality of intermediate copies (see "2. Background" for review). Recall that the intermediate copy occurs at the spec (CP) of the embedded clause (prior to *that*) in extraction-VP conditions. Therefore, longer RTs were expected at region 3 in extraction-VP than in extraction-NP conditions, as compared with the corresponding non-extraction conditions. Given the results of Pliatsikas and Marinis (2013) and Dekydtspotter et al. (2006), we remained open to the possibility that this effect may be detected at or carried over to region 4. As for the possibility of the reactivation of the intermediate copy, it was expected that extraction-VP conditions at region 5 should be read faster than extraction-NP conditions at this region, compared to the corresponding non-extraction conditions. This should be evidenced by an extraction by phrase interaction, which we will refer to as the copy reactivation effect. If L2 readers process intermediate copies differently than L1-English readers, we should find interactions between group (L1, L2) at either region 3, where the intermediate copy is encoded in memory, or region 5, where the intermediate copy is integrated with the critical verb.

For RQ2, we hypothesized that the L1-Persian group would exhibit different RT patterns compared to the L1-French group, if processing movement copies is influenced by either L1 grammar or processing preferences. Specifically, while we expect the L1-French readers to have RT profiles similar to L1-English readers, L1-Persian readers should not show evidence of encoding the intermediate copy at region 3 or reactivating the copy at region 5, if they transfer the [-wh] feature from their L1. Alternatively, if L1-Persian readers' process object RCs in the same way as they would do in Persian, by assuming optional resumptive pronouns prior to the critical verb at region 5, they were expected to show no RT difference between extraction and non-extraction structures. This is because by inserting a resumptive pronoun to indicate the direct object in extraction conditions, verbal arguments will be equally distanced to the critical verb (*had angered*) in non-extraction and extraction conditions, and therefore, no additional difficulty is expected in extraction conditions.

As for RQ3, if copy reactivation is mediated by the reduction of similarity-based interference, mismatched extraction-VP structures should elicit faster RTs at region 5 than mismatched extraction-NP structures, compared to the corresponding matched conditions. That is, if the copy reactivation effect is modulated by similarity-based interference, the difference between *(5a)* and *(5b)* at region 5 should be larger than the difference between *(3a)* and *(3b)* at this region. Finally, for RQ4, we expected proficiency in English and length of immersion experience to contribute to native-like processing of intermediate copies. Specifically, it was

expected that more proficient L2 readers who had lived in an English-speaking country for a longer period should be more likely to encode the intermediate copy at region 3 and integrate it with the verb at region 5.

Analysis

Data from participants with substandard performance were excluded (as in Marinis et al., 2005; Pliatsikas & Marinis, 2013). The analyses were carried out only on those participants whose mean accuracy on comprehension questions were above 70%, which resulted in the deletion of 6.54% of the total data. Additionally, outliers were excluded based on individual data points beyond 2 standard deviations of the mean values for each condition per subject and item, leading to the reduction of 2.53% of the remaining data. Finally, it was found that one of the items with mismatched NPs had a coding mistake and was deleted before further analysis, leading to the exclusion of 2.46% of the remaining data. The analysis was performed only on items whose comprehension questions were answered correctly.

Initially, raw RTs were log transformed to reduce skew. Nested linear mixed-effects models were constructed on log RTs of regions 3, 4, and 5, using the *lmerTest* package (Bates et al., 2015) in R (R Core Team, 2020), which adds p-values to the output of the *lme4* package (Bates et al., 2015) using Satterthwaite's approximation of degrees of freedom. We created two sets of models. First, we examined whether our results replicate Marinis et al. (2005), and thus focused only on matched conditions as in *(3)* (extraction-VP, extraction-NP, non-extraction-VP, and non-extraction-NP). All models included sum-coded (−0.5, +0.5) fixed effects of extraction (non-extraction, extraction), phrase (NP, VP), and group (L1, L2), as well as their interactions. By contrast, the second set of models examined the impact of interference in extraction conditions at region 5 (*had angered*), and as such, included extraction matched *(3a)* and *(3b)* and extraction mismatched conditions *(5a)* and *(5b)*. These models included sum-coded (−0.5, +0.5) fixed effects of match (matched, mismatched), phrase (NP, VP), group (L1, L2), and their interactions.[6] To provide a more comprehensive understanding of individual differences in L2 processing of intermediate copies, we also carried out a separate set of analysis similar to the above but only on the L2 data, and included scaled covariates of length of immersion experience and proficiency scores.[7] In case significant interactions were found, we conducted follow-up analysis by constructing additional models to locate the source of these interactions. All models were assessed with the maximal random effects structure that converged (Barr et al., 2013), including by-subject and by-item adjustments to the intercept and slopes.

Results

We first explain our results for regions 3, 4, and 5, focusing on the matched conditions: matched non-extraction NP, matched non-extraction-VP, matched extraction NP, and matched extraction VP. These conditions correspond to those used in Marinis et al. (2005), allowing us to examine whether intermediate copies were accessed at region 3 (*that/about*), and whether they facilitated dependency
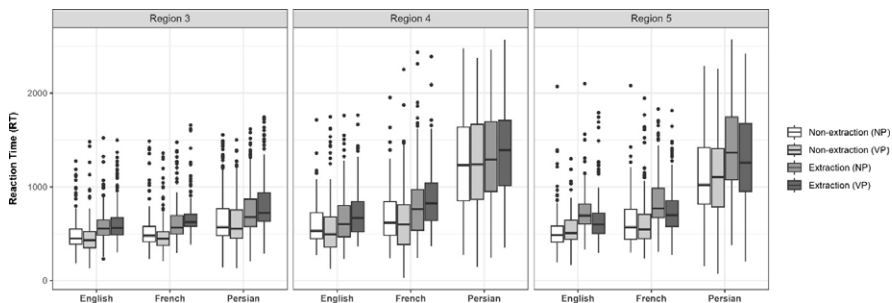
Figure 1. Box plots showing raw RTs per extraction (non-extraction, extraction), phrase (NP, VP), group (L1-English, L1-French, and L1-Persian), and regions (3, 4, 5).

formation at region 5 (*had angered*). We also analyzed region 4 to examine a potential spill-over from region 3 (Figure 1; for full results, see Table A1 in Appendix).

Regions 3 (that/about) and 4 (the rude patient). For these regions, there was a main effect of L1, with significantly faster log RTs for L1-English readers than L2 groups (region 3: estimate = .167, SE = .051, t(97) = 3.28, p = .001, d = .43; region 4: estimate = .429, SE = .084, t(98) = 5.12, $p < .001$, d = .76). The main effect of extraction was also significant, with slower log RTs for extraction than non-extraction conditions (region 3: estimate = .517, SE = .055, t(88) = 9.41, $p < .001$, d = 1.32; region 4: estimate = .431, SE = .080, t(89) = 5.41, $p < .001$, d = .77). Furthermore, there was a significant interaction between extraction and phrase, and extraction-VP conditions were read significantly slower than extraction-NP conditions, relative to the corresponding non-extraction conditions (region 3: estimate = .156, SE = .045, t(1740) = 3.44, $p < .001$, d = .40; region 4: estimate = .263, SE = .054, t(1719) = 4.85, $p < .001$, d = .47). Follow-up analysis revealed that in non-extraction conditions, VP structures were read significantly faster than NP structures (region 3: estimate = .092, SE = .035, t(867) = 2.65, p = .008, d = .23; region 4: estimate = .133, SE = .043, t(846) = 3.08, p = .002, d = .22), but in extraction conditions, VP structures were read significantly slower (region 3: estimate = .059, SE = .029, t(840) = 2.06, p = .040, d = .18; region 4: estimate = .124, SE = .035, t(863) = 3.54, $p < .001$, d = .25), thus suggesting that the intermediate copy was processed at regions 3 and 4. No other main or interaction effects were significant (ps > .31), indicating that there was no difference between L1 and L2 readers in how they processed the intermediate copy at these regions.

Regions 5 (had angered). Similar to the previous regions, the main effect of L1 was significant, and L2 readers had slower log RTs than L1-English readers (estimate = .399, SE = .078, t(98) = 5.14, $p < .001$, d = .78). The main effects of extraction and phrase were also significant, with slower log RTs for extraction than non-extraction conditions (estimate = .533, SE = .066, t(65) = 8.10, $p < .001$, d = 1.04) and faster log RTs for VP than NP conditions (estimate = .138, SE = .048, t(1727) = 2.88, p = .004, d = .27). Crucially, there was a significant interaction between extraction and phrase, and extraction-VP structures were read significantly faster than extraction-NP structures, compared to the corresponding non-extraction conditions (estimate = .141, SE = .048, t(1725) = 2.95, p = .003, d = .27).
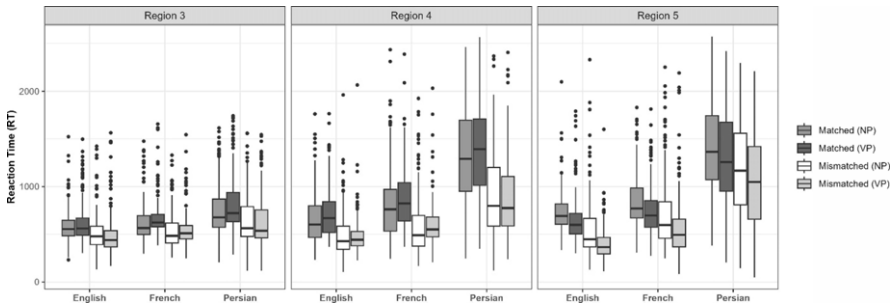
Figure 2. Box plots showing raw RTs per match (matched, mismatched), phrase (NP, VP), group (L1-English, L1-French, and L1-Persian), and regions (3, 4, 5).

Follow-up analysis showed that unlike the previous regions, in non-extraction conditions, the difference in log RTs between NP and VP structures was not significant (estimate = .005, SE = .043, t(81) = .11, p = .91, d = .01). By contrast, in extraction conditions, VP structures were read significantly faster than NP structures (estimate = .136, SE = .041, t(44) = 3.28, p = .002, d = .30), thus suggesting that the intermediate copy in extraction-VP structures was reactivated to break the long-distance dependency into two smaller dependencies.

To assess the impact of similarity-based interference, we then focused only on extraction conditions and compared NP and VP conditions in extraction-matched and extraction-mismatched items (Figure 2; for full results, see Table A2 in Appendix).

There were significant main effects of L1, phrase, and match at region 5. Specifically, L2 readers had slower log RTs than L1 readers (estimate = .469, SE = .082, t(97) = 5.73, $p < .001$, d = .89). Similarly, VP conditions were read significantly faster than NP conditions (estimate = .466, SE = .061, t(91) = 7.65, $p < .001$, d = .80), and mismatched conditions had faster log RTs than matched conditions (estimate = .823, SE = .101, t(50) = 8.14, $p < .001$, d = 1.41). Additionally, the interaction between L1 and match was significant, with faster log RTs for L1 readers in mismatched than matched conditions, compared to the L2 groups (estimate = .248, SE = .069, t(1616) = 3.62, $p < .001$, d = .43). Importantly, the interaction between match and phrase was also significant, and extraction-VP structures were read significantly faster than extraction-NP structures in mismatched conditions compared to the corresponding matched conditions (estimate = .191, SE = .056, t(1609) = 3.43, $p < .001$, d = .33). This suggests that integrating the intermediate copy with its lexical subcategorizer posed less of a challenge in mismatched conditions. No other main or interaction effects were observed, therefore suggesting that L1 and L2 groups processed the intermediate copy in a similar way: both groups reactivated the copy at region 5, which was facilitated by the reduction of similarity-based interference in mismatched items.

Finally, regarding individual differences, we present the results from the models for each group separately, focusing on the effects of proficiency and immersion experience. The only significant effect was that of an interaction at region 5 between c-test proficiency and phrase for the L1-French group in extraction structures with

matched vs. mismatched NPs (estimate = .137, SE = .050, t(529) = 2.71, p = .006, d = .30; all other ps > .051). This indicates that highly proficient L1-French readers had significantly faster log RTs in VP than NP structures. However, the 3-way interaction between proficiency, phrase, and match was not significant (estimate = .062, SE = .050, t(529) =1.24, p = .22, d = .14), demonstrating that VP structures were read faster than NP structures, overall, and high proficiency did not affect matched and mismatched extraction conditions differently. In addition, the 3-way interaction between proficiency, extraction, and phrase was not significant in matched conditions, confirming that higher proficiency among the L1-French readers did not significantly contribute to the processing of the intermediate copy.

In summary, the results showed that both L1 and L2 readers processed the filler at regions 3 and 4 and integrated the intermediate copy with its lexical subcategorizer at region 5, which was facilitated by the reduction of similarity between the intervening NPs.

Discussion

This study investigated L1 and L2 processing of long-distance filler-gap dependencies and examined whether readers access the intermediate structure in sentences with matched (three description NPs) and mismatched NPs (one proper name and two description NPs). The results showed that as far as dependencies with matched NPs are concerned, all groups processed the intermediate structure prior to the complementizer *that* in extraction-VP conditions, since log RTs were slower in extraction-VP structures (containing an additional intermediate copy) compared to extraction-NP structures (with no intermediate copy). Therefore, our results are compatible with previous studies that highlighted the psychological reality of intermediate copies. The same pattern of results was also found at region 4, suggesting that this effect carried over to the subsequent region, which is compatible with other studies suggesting that the effect of the intermediate copy is not bound only to its canonical location (Dekydtspotter et al., 2006; Pliatsikas & Marinis, 2013). Importantly, however, this pattern was reversed in the subsequent region 5, when the subcategorizing verb (*had angered*) was encountered. Recall that region 5 was the point where verbal arguments, including the filler *who*, need to be associated with the verb, and that extraction-VP conditions included an intermediate copy breaking the long-distance dependency into two smaller dependencies. At region 5, extraction-VP conditions were read faster than extraction-NP conditions, thus suggesting the facilitatory effect of the intermediate copy.

Notice that in VP conditions, the subject *the patient* was always adjacent to the verb *had angered*. This is in contrast to NP conditions, where an additional NP *the rude patient* intervened between the two (*the doctor's argument about the rude patient had angered*). Therefore, it may be argued that faster log RTs in extraction-VP conditions may reflect easier processing of the subject-verb agreement relationship in VP structures, rather than access to the intermediate copy. However, we argue that while the subject-verb adjacency might have contributed to relative ease of processing extraction-VP conditions, this does not explain the lack of

a significant difference in log RTs between non-extraction-NP and non-extraction-VP conditions. The two non-extraction conditions mirrored the extraction conditions in terms of subject-verb distance in the embedded clause. Therefore, a different effect, other than that of subject-verb adjacency in VP conditions, must underlie the RT facilitation effect found at extraction-VP conditions. Given that extraction-VP structures involved an intermediate copy which was absent in the other conditions, we argue that the above effect is due to the presence of the intermediate copy, resulting in a shorter filler-gap dependency at region 5 between *who* and *had angered.* This is compatible with previous research investigating the role of intermediate structures in long-distance dependencies, which found a facilitatory effect of the intermediate copy at the subcategorizing verb (Dekydtspotter & Miller, 2013; Felser & Roberts, 2007; Fernandez et al., 2018; Keine, 2020; Marinis et al., 2005; Pliatsikas et al., 2017; Pliatsikas & Marinis, 2013; Roberts et al., 2007).

Importantly, we did not find any L1-L2 differences, neither in the initial processing of the intermediate copy at *that* nor in its ultimate integration with the verb at *had angered.* This suggests that similar to the L1 group, L2 readers managed to process the intermediate copy, both at regions 3 and 5. Therefore, as far as RQ1 is concerned (i.e., whether there is a difference between L1 and L2 readers), we argue that there is no L1-L2 difference in the processing of the intermediate copy in long-distance dependencies such as *(1)* since both groups encoded the copy at region 3, held it active in memory in region 4, and reactivated it in region 5 to facilitate filler-gap integration. This is not compatible with previous studies that found that L2 readers directly associate the displaced filler with the verb by relying on the verb's subcategorization information (Felser & Roberts, 2007; Marinis et al., 2005). To explain the results, we note that the L2 participants in this study were highly advanced and had a relatively high L2 experience, compared to the L2 participants in previous studies that did not find a facilitatory effect of the intermediate copy. The lack of L1-L2 difference is compatible with previous studies that argued L2 readers process long-distance wh-dependencies in the same way as L1 readers, given sufficient linguistic experience and proficiency (Pliatsikas & Marinis, 2013).

Regarding RQ2 (i.e., whether there is a difference between L1-French and L1-Persian readers), we found that both L1-French and L1-Persian readers patterned similarly to L1-English readers. Both groups encoded the movement copy at region 3 and reactivated it at region 5 for filler-gap integration. Therefore, the results suggest that even if there are L1 effects, they are superable at an advanced proficiency. Specifically, we did not find a difference between L1-French and L1-Persian readers in encoding of the intermediate copy at region 3, nor in integrating the copy with the critical verb at region 5. Thus, the absence/presence of wh-movement in L1 does not explain the native-like processing of intermediate copies for the L2 readers. In addition, the L1-Persian readers had longer RTs in extraction than non-extraction conditions, which further casts doubt on the hypothesis that they process object RCs by assuming an L1-like resumptive pronoun prior to the critical verb. Overall, both L2 groups behaved similarly to the L1-English readers regarding the initial processing of the filler at region 3 and its ultimate integration at region 5.

As for RQ3 (i.e., whether copy reactivation is easier in dependencies with mismatched NPs than in dependencies with matched NPs), we found that extraction-VP structures in mismatched conditions were read faster at region 5 than extraction-NP structures, compared to the corresponding structures in matched conditions. This suggests that accessing the intermediate copy was facilitated by the reduction of similarity-based interference in mismatched conditions. Importantly, this effect did not interact with L1, and therefore, the results suggest that L1 and L2 processing of long-distance dependencies were affected to the same extent by similarity-based interference. Recall that mismatched conditions involved two descriptive NPs and one proper name, as opposed to matched conditions which had three descriptive NPs. Therefore, we argue that the discourse status (proper name and description) of NPs impacts the relative ease with which they are processed in long-distance dependencies. Specifically, the less similar the NPs are in a long-distance dependency, the more distinct their representations will be in memory (Gordon et al., 2002, 2006), which creates less difficulty during both L1 and L2 processing of dependency structures.

Notice that the interaction between match and phrase was observed at the verb region 5 and crucially not at the preceding region 4 when the proper name is first encountered in mismatched conditions. In fact, the RTs at regions 3 and 4 (proper name/descriptive NP) were very similar in mismatched and matched conditions. Therefore, it might appear that the similarity-based interference effect observed at region 5 reflects *retrieval* interference at the verb as opposed to *encoding* interference following the complementizer *that*. However, there is no reason why verbs (e.g., *anger*) may cue a particular type of noun as their object (e.g., *John*) more strongly than another type of noun (*e.g., the rude patient*). Thus, we remain speculative as to the nature of the interference effect observed at region 5. It is possible that the encoding of NPs and their later retrieval rely on two distinct sets of cues. For example, it might well be the case that the discourse status of NPs is not critical to their successful encoding in memory (at region 3), but this feature is accessed during retrieval to distinguish between different NPs for dependency completion (at region 5), especially when processing becomes difficult. More research is required to tease apart the underlying processes for encoding and retrieval interference.

These findings are also consistent with the view that movement out of islands are regulated by computational principles, and thus question the assumed (un)grammaticality of island configurations (Belikova & White, 2009; Perpiñán, 2020). Notice that the extraction-VP conditions such as *(3a)* involved movement out of complement clauses, which act as barriers to wh-movement operations. If islands were purely syntactic phenomena, we would not have observed any processing effects related to access to the intermediate copy, since the match/mismatch manipulations did not target the presence/absence of the hypothesized intermediate copy. Instead, we found that both L1 and L2 groups (regardless of [+/−wh-movement]) showed processing profiles indicating easier access to the intermediate copy in mismatched than matched conditions. This supports the view that island configurations are at least partially a processing matter, rather than solely a grammatical one.

Regarding RQ4 (i.e., whether c-test proficiency and length of immersion experience predicted the processing of intermediate copies among the L2 group), none of the effects reported above interacted with proficiency or length of immersion experience. This suggests that regardless of proficiency and length of immersion, the L2 readers in this study had access to the intermediate copy as they encoded it at region 3 and reactivated it at region 5 to facilitate filler-gap integration. We note that while a significant majority of L2 readers had c-test scores within the same range of L1-English readers (97.0% among L1-French and 97.2% among L1-Persian readers), average immersion experience in this study was quite low (7 years among L1-French and 4 years among L1-Persian readers). This stands in contrast with the average immersion experience on previous studies that reported a reliable interaction between immersion and copy reactivation (e.g., Pliatsikas, & Marinis, 2013 reported a minimum of 13 years immersion). Therefore, it is likely that the L2 readers in this study were highly advanced, to the point that they processed intermediate copies in the same way as L1-English readers, whereas the range of immersion experience was too low to display a reliable interaction with the processing of intermediate copies. This also highlights that L2 readers of English may achieve highly advanced proficiency levels in processing complex syntactic structures independently of length of immersion in an English-speaking environment. However, we do not draw a strong conclusion regarding the relationship between proficiency, immersion, and processing of intermediate copies, as the range of c-test scores and immersion experience was relatively limited in this study. Rather, we argue that both groups were at a near-native proficiency level and displayed the same processing pattern as L1-English readers while reading long-distance wh-dependencies.

Conclusion

This study employed an SPR task to investigate long-distance filler-gap dependencies in L1- and L2-English. The results suggested that readers encode the intermediate copy during processing these structures and later access it to facilitate dependency completion. We provided evidence that copy reactivation is stronger in structures that include more distinct NPs (*the doctor, John, and the patient*), compared to structures with more similar NPs (*the doctor, the nurse, and the patient*). Overall, the findings suggest that L2 readers access hierarchically complex syntactic information by reactivating the phonologically null intermediate copy, especially when retrieval operations are more manageable because of lower similarity-based interference. The findings challenge the argument that L2 grammars do not provide the type of syntactic information required to process wh-dependencies in a native-like fashion (Clahsen & Felser, 2006, 2018).

Notes

1 Adopting a generative syntactic framework, this study assumes that a sentence is a tense phrase (TP), which is dominated by a complementizer phrase (CP). The findings do not depend on this assumption.

2 Although Chomskyan generative approaches have an explicit expression of a "punctuated" path of extraction (i.e., phases), other approaches, such as nontransformational ones, also employ some sort of metric with respect to incremental, cyclic parsing (see Chaves & Putnam, 2020: Chapter 5 for an overview of these approaches). In these approaches, the presence of a null category does not derive from movement operations but is encoded within phrase structure rules themselves. We do not aim to distinguish between these different theoretical syntactic accounts, but highlight that the findings of this study are relevant to all formal approaches to successive cyclicity.

3 Persian RCs are obligatorily introduced by the invariant complementizer *ke* ("*that*"), which, unlike the case of relative pronouns, does not agree with the head RC in number, gender, and case (Aghaei, 2006). Furthermore, Persian does not respect locality conditions of movement (Raghibdoust, 1993, pp. 55-69). Therefore, following earlier studies (Karimi & Taleghani, 2007; Solaimani et al., 2023), it is assumed in this study that Persian is not a wh-movement language.

4 The suffix "-" is represented by "-RES" as it serves to distinguish between restrictive and nonrestrictive RCs in Persian (Taghvaipour, 2005). We follow Karimi, 2005) to indicate "−e" as EZAF (addition) in the gloss as it is added to the head noun in Persian adjective-noun phrases.

5 https://osf.io/uejyb/

6 For the second set of models comparing long-distance dependencies with matched and mismatched NPs, we only report the results of models on region 5, even though full results are available in Table A2 of the Appendix for regions 3, 4, and 5. This is because it is only after region 5 that all the three NPs involved in the dependency are processed across the four conditions of matched-NP, matched-VP, mismatched-NP, and mismatched-VP structures.

7 We created the models for L1-French and L1-Persian group separately by adding immersion and proficiency scores. This was done to avoid multicollinearity issues and overparameterization which might mask the effects of interest.

References

Aghaei, B. (2006). *The syntax of ke-clause and clausal extraposition in Modern Persian*. (MA thesis, University of Texas at Austin).

Barr, D. J., Levy, R., Scheepers, C., & Tily, H. J. (2013). Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language*, **68**(3), 255–278. https://doi.org/10.1016/j.jml.2012.11.001

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, **67**(1), 1–48. https://doi.org/10.18637/jss.v067.i01

Belikova, A., & White, L. (2009). Evidence for the fundamental difference hypothesis or not? Island constraints revisited. *Studies in Second Language Acquisition*, **31**(2), 199–223. https://doi.org/10.1017/S0272263109090287

Berghoff, R. (2023). Wh-dependency processing in a naturalistic exposure context: Sensitivity to abstract syntactic structure in high working-memory L2 speakers. *Studies in Second Language Acquisition*, **45**(2), 586–598. https://doi.org/10.1017/S0272263122000304

Chaves, R. P., & Putnam, M. T. (2020). Non-movement-based approaches. In R. D. Van Valin (Ed.), *Unbounded dependency constructions: Theoretical and experimental perspectives* (pp. 156–202). Oxford University Press. https://doi.org/10.1093/oso/9780198784999.001.0001

Chomsky, N. (1995). *The minimalist program*. MIT Press.

Clahsen, H., & Felser, C. (2006). Grammatical processing in language learners. *Applied Psycholinguistics*, **27**(1), 3–42. https://doi.org/10.1017/S0142716406060024

Clahsen, H., & Felser, C. (2018). Some notes on the shallow structure hypothesis. *Studies in Second Language Acquisition*, **40**(3), 693–706. https://doi.org/10.1017/S0272263117000250

Cunnings, I. (2017). Processing and working memory in bilingual sentence processing. *Bilingualism: Language and Cognition*, **20**(4), 659–678. https://doi.org/10.1017/S1366728916000675

Cunnings, I., & Fujita, H. (2021). Similarity-based interference and relative clauses in second language processing. *Second Language Research*, **39**(2), 539–563. https://doi.org/10.1177/02676583211063534

Dekydtspotter, L., & Miller, A. K. (2013). Inhibitive and facilitative priming induced by traces in the processing of wh-dependencies in a second language. *Second Language Research*, **29**(3), 345–372. https://doi.org/10.1177/0267658312467030

Dekydtspotter, L., Schwartz, B. D., & Sprouse, R. A. (2006). The comparative fallacy in L2 processing research. In M. G. O'Brien, C. Shea, & J. Archibald (Eds.), *8th Generative Approaches to Second Language Acquisition Conference (GASLA 2006)* (pp. 33–40). Cascadilla Proceedings Project.

Felser, C., & Roberts, L. (2007). Processing wh-dependencies in a second language: A cross-modal priming study. *Second Language Research*, **23**(1), 9–36. https://doi.org/10.1177/0267658307071600

Fernandez, L., Höhle, B., Brock, J., & Nickels, L. (2018). Investigating auditory processing of syntactic gaps with L2 speakers using pupillometry. *Second Language Research*, **34**(2), 201–227. https://doi.org/10.1177/0267658317722386

Gibson, E., & Warren, T. (2004). Reading-time evidence for intermediate linguistic structure in long-distance dependencies. *Syntax*, **7**(1), 55–78. https://doi.org/10.1111/j.1368-0005.2004.00065.x

Gordon, P. C., Hendrick, R., Johnson, M., & Lee, Y. (2006). Similarity-based interference during language comprehension: Evidence from eye tracking during reading. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, **32**(6), 1304–1321. https://psycnet.apa.org/doi/10.1037/0278-7393.32.6.1304

Gordon, P. C., Hendrick, R., & Levine, W. H. (2002). Memory-load interference in syntactic processing. *Psychological Science*, **13**(5), 425–430. https://doi.org/10.1111/1467-9280.00475

Hopp, H. (2018). The bilingual mental lexicon in L2 sentence processing. *Second Language*, **17**, 5–27. https://doi.org/10.11431/secondlanguage.17.0_5

Juffs, A. (2005). The influence of first language on the processing of wh-movement in English as a second language. *Second Language Research*, **21**(2), 121–151. https://doi.org/10.1191/0267658305sr255oa

Karimi, S. (2005). *A minimalist approach to scrambling: Evidence from Persian*. Walter de Gruyter.

Karimi, S., & Taleghani, A. (2007). Wh-movement, interpretation, and optionality in Persian. In: Karimi, S., Samiian, V., and Wilkins, W.K. (eds). *Clausal and phrasal architecture: syntactic derivation and interpretation* (pp. 166–187). Benjamins.

Keijzer, M. (2007). Last in first out?: an investigation of the regression hypothesis in Dutch emigrants in Anglophone Canada. *Toegepaste Taalwetenschap in Artikelen*, **78**(1), 131–139. https://doi.org/10.1075/ttwia.78.12kei

Keine, S. (2020). Locality domains in syntax: Evidence from sentence processing. *Syntax*, **23**(2), 105–151. https://doi.org/10.1111/synt.12195

Klein-Braley, C., & Raatz, U. (1984). A survey of research on the C-Test1. *Language Testing*, **1**(2), 134–146. https://doi.org/10.1177/026553228400100202

Lewis, R. L., Vasishth, S., & Van Dyke, J. A. (2006). Computational principles of working memory in sentence comprehension. *Trends in Cognitive Sciences*, **10**(10), 447–454. https://doi.org/10.1016/j.tics.2006.08.007

Marinis, T., Roberts, L., Felser, C., & Clahsen, H. (2005). Gaps in second language sentence processing. *Studies in Second Language Acquisition*, **27**(1), 53–78. https://doi.org/10.1017/S0272263105050035

Park, J. (1998). *The C-test: usefulness for measuring written language ability of non-native speakers of English in high school* (Doctoral Dissertation, Iowa State University).

Perpiñán, S. (2020). Wh-movement, islands, and resumption in L1 and L2 Spanish: is (un) grammaticality the relevant property?. *Frontiers in Psychology*, **11**, 501618. https://doi.org/10.3389/fpsyg.2020.00395

Pliatsikas, C., Johnstone, T., & Marinis, T. (2017). An fMRI study on the processing of long-distance wh-movement in a second language. *Glossa: A Journal of General Linguistics*, **2**(1): 101. https://doi.org/10.5334/gjgl.95

Pliatsikas, C., & Marinis, T. (2013). Processing empty categories in a second language: When naturalistic exposure fills the (intermediate) gap. *Bilingualism: Language and Cognition*, **16**(1), 167–182. https://doi.org/10.1017/S136672891200017X

R Core Team. (2020). *R: A language and environment for statistical computing [software]*. R Foundation for Statistical Computing. Available at: http://www.R-project.org (accessed March 2023).

Raghibdoust, S. (1993). *Interrogative constructions in Persian*. (MA thesis, University of Ottawa).

Roberts, L., Marinis, T., Felser, C., & Clahsen, H. (2007). Antecedent priming at trace positions in children's sentence processing. *Journal of Psycholinguistic Research*, **36**, 175–188. https://doi.org/10.1007/s10936-006-9038-3

Ross, J. R. (1967). *Constraints on variables in syntax*. Doctoral dissertation, MIT Press, Cambridge, MA.

Rowlett, P. (2007). *The syntax of French*. Cambridge University Press.

Solaimani, E., & Marefat, H. (2024). Definiteness matters as a discourse cue in L1 and L2 processing of relative clauses. *Pragmatics & Cognition*, **31**(1), 185–204. https://doi.org/10.1075/pc.00044.sol

Solaimani, E., Myles, F., & Lawyer, L. (2023). Testing the Interpretability Hypothesis: Evidence from acceptability judgments of relative clauses by Persian and French learners of L2 English. *Second Language Research*. https://doi.org/10.1177/02676583231162783

Taghvaipour, M.A. (2005). *Persian relative clauses in head-driven phrase structure grammar*. Doctoral Dissertation, University of Essex, Colchester, UK.

Trace, J. (2020). Clozing the gap: How far do cloze items measure?. *Language Testing*, **37**(2), 235–253. https://doi.org/10.1177/0265532219888617

Zehr, J., & Schwarz, F. (2018). *Penncontroller for internet-based experiments (ibex)*. https://doi.org/10.17605.

Appendix

Results of statistical analysis comparing matched extraction and matched non-extraction conditions (extraction-NP, extraction-VP, non-extraction-NP, and non-extraction-VP) (Table A1) and analysis comparing matched and mismatched NP and VP conditions (matched-NP, matched-VP, mismatched-NP, and mismatched-VP) (Table A2).

Table A1. Results of statistical analysis on matched conditions at regions 3, 4, and 5

	Region											
	3				4				5			
	Estimate	SE	t	p	Estimate	SE	t	p	Estimate	SE	t	p
Intercept	6.239	0.042	147.332	**<0.001**	6.372	0.073	87.641	**<0.001**	6.362	0.066	96.889	**<0.001**
Group	0.167	0.051	3.281	**0.001**	0.429	0.084	5.120	**<0.001**	0.399	0.078	5.140	**<0.001**
Extraction	0.517	0.055	9.408	**<0.001**	0.431	0.080	5.409	**<0.001**	0.533	0.066	8.098	**<0.001**
Phrase	0.039	0.045	0.863	0.388	0.017	0.054	0.322	0.747	0.138	0.048	2.878	**0.004**
Extraction*Phrase	0.156	0.045	3.440	**0.001**	0.263	0.054	4.853	**<0.001**	−0.141	0.048	−2.948	**0.003**
Group*Extraction	−0.018	0.067	−0.275	0.783	−0.058	0.097	−0.595	0.552	−0.000	0.072	−0.004	0.997
Group*Phrase	−0.056	0.055	−1.015	0.310	0.042	0.067	0.624	0.533	−0.003	0.058	−0.048	0.961
Group*Extraction*Phrase	−0.013	0.055	−0.242	0.809	−0.108	0.067	−1.630	0.103	0.072	0.058	1.224	0.221

Table A2. Results of statistical analysis on extraction conditions with matched and mismatched NPs at regions 3, 4, and 5

	Region											
	3				4				5			
	Estimate	SE	t	p	Estimate	SE	t	p	Estimate	SE	t	p
Intercept	6.264	0.043	146.533	**<0.001**	6.300	0.064	98.851	**<0.001**	6.289	0.071	88.920	**<0.001**
Group	0.157	0.050	3.128	**0.002**	0.390	0.074	5.264	**<0.001**	0.469	0.082	5.734	**<0.001**
Match	0.410	0.063	6.543	**<0.001**	0.714	0.088	8.147	**<0.001**	0.824	0.101	8.144	**<0.001**
Phrase	−0.025	0.050	−0.508	0.612	−0.169	0.079	−2.125	**0.034**	0.466	0.061	7.648	**<0.001**
Match*Phrase	−0.096	0.047	−2.060	**0.040**	−0.078	0.079	−0.978	0.328	−0.191	0.056	−3.429	**0.001**
Group*Match	0.022	0.057	0.385	0.700	0.071	0.063	1.121	0.263	−0.248	0.069	−3.615	**<0.001**
Group*Phrase	−0.040	0.061	−0.650	0.516	0.080	0.063	1.263	0.207	−0.134	0.075	−1.792	0.073
Group*Match*Phrase	0.017	0.057	0.296	0.767	0.075	0.063	1.190	0.234	0.040	0.069	0.589	0.556

Cite this article: Solaimani, E. (2025). Minding the gap: Similarity-based interference in L1 and L2 processing of long-distance Wh-dependencies. *Applied Psycholinguistics*. https://doi.org/10.1017/S0142716424000456