Canadian Journal of Applied Linguistics
Revue canadienne de linguistique appliquée

Canadian Journal of Applied Linguistics (CJAL)
Revue canadienne de linguistique appliquée (RCLA)

Impact of the Intrinsic Complexity and Prior Linguistic Knowledge on the Acquisition of Relative Clauses

Golpar Bahar et Gero Kunter

Volume 27, numéro 3, hiver 2024

URI : https://id.erudit.org/iderudit/1118852ar
DOI : https://doi.org/10.37213/cjal.2024.33406

Aller au sommaire du numéro

Éditeur(s)

University of New Brunswick

ISSN

1920-1818 (numérique)

Découvrir la revue

Citer cet article

Bahar, G. & Kunter, G. (2024). Impact of the Intrinsic Complexity and Prior Linguistic Knowledge on the Acquisition of Relative Clauses. *Canadian Journal of Applied Linguistics / Revue canadienne de linguistique appliquée*, *27*(3), 80–101. https://doi.org/10.37213/cjal.2024.33406

Résumé de l'article

L'étude explore dans quelle mesure la complexité intrinsèque des clauses relatives (RC) et les connaissances linguistiques préalables influencent l'acquisition des RC par les apprenants de L2. L'étude examine les principales sources des types erronés et évités de RC anglais produits par des apprenants d'anglais persan à trois niveaux de compétence. La tâche d'obtention de données était un test de traduction composé de six types de RC modélisés sur les types de RC dans la hiérarchie d'accessibilité des expressions nominales. Pour analyser les données, les fréquences d'occurrence des RC formés correctement et par erreur ont été comptées et les RC évités ont été identifiés dans chaque type de RC. Ensuite, une analyse précise des erreurs a été réalisée. L'analyse statistique de 3840 RC a montré que les types d'erreurs les plus courants étaient (i) la formation de RC anglais avec des pronoms résomptifs et (ii) la modification de RC plus marqués avec un ordre non canonique des mots en RC moins marqués avec un ordre canonique des mots. Les erreurs sont interprétées comme une preuve de l'impact à la fois du transfert L1 et des contraintes intrinsèques universelles des RC. L'analyse des types de RC évités, principalement les plus marqués, indique que l'évitement est principalement lié aux contraintes intrinsèques universelles des RC.

Impact of the Intrinsic Complexity and Prior Linguistic Knowledge on the Acquisition of Relative Clauses

Golpar Bahar
University of Siegen

Gero Kunter
University of Siegen

Abstract

The study explores the extent to which the intrinsic complexity of relative clauses (RCs) and prior linguistic knowledge impact the acquisition of RCs by L2 learners. The study investigates the main sources of the erroneous and avoided types of English RCs produced by Persian-speaking learners of English at three proficiency levels. The data elicitation task was a translation test comprised of six types of RCs modeled on the RC types in the Noun Phrase Accessibility Hierarchy. To analyze the data, the occurrence frequencies of the correctly and erroneously formed RCs were counted and the avoided RCs were identified in each RC type. Then, a precise error analysis was done. The statistical analysis of 3840 RCs showed that the most common error types were (i) forming English RCs with resumptive pronouns and (ii) altering more-marked RCs with non-canonical word order to less-marked RCs with canonical word order. The errors are interpreted as evidence for the impact of both L1 transfer and the universal intrinsic constraints of RCs. The analysis of the avoided RC types, mostly the more marked RCs, indicates that avoidance is mainly linked to the universal intrinsic constraints of RCs.

Résumé

L'étude explore dans quelle mesure la complexité intrinsèque des clauses relatives (RC) et les connaissances linguistiques préalables influencent l'acquisition des RC par les apprenants de L2. L'étude examine les principales sources des types erronés et évités de RC anglais produits par des apprenants d'anglais persan à trois niveaux de compétence. La tâche d'obtention de données était un test de traduction composé de six types de RC modélisés sur les types de RC dans la hiérarchie d'accessibilité des expressions nominales. Pour analyser les données, les fréquences d'occurrence des RC formés correctement et par erreur ont été comptées et les RC évités ont été identifiés dans chaque type de RC. Ensuite, une analyse précise des erreurs a été réalisée. L'analyse statistique de 3840 RC a montré que les types d'erreurs les plus courants étaient (i) la formation de RC anglais avec des pronoms résomptifs et (ii) la modification de RC plus marqués avec un ordre non canonique des mots en RC moins marqués avec un ordre canonique des mots. Les erreurs sont interprétées comme une preuve de l'impact à la fois du transfert L1 et des contraintes intrinsèques universelles des RC. L'analyse des types de RC évités, principalement les plus marqués, indique que l'évitement est principalement lié aux contraintes intrinsèques universelles des RC.

Introduction

Several hypotheses have been proposed regarding the order of acquisition of relative clauses, encompassing theories such as the Noun Phrase Accessibility Hypothesis (NPAH), Absolutive Hypothesis (AH), Perceptual Difficulty Hypothesis (PDH), Subject Object Hierarchy Hypothesis (SOHH), Linear Distance Hypothesis (LDH), Structural Distance Hypothesis (SDH), and Word Order Difference Hypothesis (WDH). While these hypotheses primarily address theories of first language acquisition of RCs, efforts have been made to assess their applicability in second language acquisition contexts. However, due to the limited scope of cross-linguistic studies incorporating non-L1 data, scholars such as Gass and Lee (2007) argue that investigations utilizing L2 data could significantly advance our understanding of second language acquisition. The present study specifically focuses on the NPAH and WDH due to their potential insights into the acquisition processes of L1 Persian and L2 English learners, particularly regarding resumption differences. Persian is a representative language permitting resumption, contrasting with English, which lacks this feature.

According to the NPAH, the relativizability of a noun phrase (NP) is linked to its syntactic position, and some syntactic positions are more accessible to relativization than others. Based on the universal accessibility hierarchy, which Keenan and Comrie (1977) propose, the subject position is the most accessible position to relativization followed by direct object (DO), indirect object (IO), oblique (OBL), genitive (GEN), and object of comparison (OCOMP) in order. The NPAH "reflects the psychological ease of comprehension" (Keenan & Comrie, 1977, p. 88); thus, RCs formed on subject positions (henceforth, SU relatives) are the easiest RC types and OCOMP relatives are the most difficult RC types to learn (Izumi, 2003). Although the NPAH was originally conceived as a typological universal, not intended to predict the acquisition order of RCs, it was extended to reflect the natural acquisition order of RCs in both L1 and L2 contexts based on the notion that marked items are acquired later than unmarked ones. This extension was supported by several studies in L1 (Brown, 1971; Prideaux & Baker, 1986; Romaine, 1984; Tavakolian, 1981) and L2 (Bahar, in press, 2024; Bahar, 2023; Doughty, 1991; Doughty & Long, 2003; Eckman, 1977; Eckman et al., 1988; Gass, 1979, 1980, 1982; Hamilton, 1994; Hawkins, 2007; Ioup & Kruse 1977; Pavesi, 1986). Consequently, the NPAH has been advocated as a robust framework for understanding linguistic phenomena concerning RCs and elucidating their acquisition in both L1 and L2 contexts (Izumi, 2003).

Regarding the comprehension of RCs, MacDonald and Christiansen (2002) propose a perspective rooted in connectionist theory. They introduce the Word Order Difference Hypothesis (WDH), which postulates that a combination of linguistic exposure and innate factors shapes the processing abilities of individual language learners. Emphasizing the distinction between canonical and non-canonical word order, they argue that non-canonical structures present greater processing challenges. Specifically, they highlight the relative ease of processing subject RCs, attributed to their conformity with the word order found in common English sentences. This familiarity, derived from the frequent occurrence of simple active sentences, facilitates the acquisition of subject RCs. Conversely, object RCs, characterized by less typical word order patterns, pose greater difficulties due to learners' lack of exposure. It is important to acknowledge that preceding the proposal of the WDH, previous studies such as those conducted by Bever, 1970, MacWhinney et al., 1984, and Slobin & Bever, 1982, have delved into similar areas, primarily within the realm of first language

acquisition. Rahmany et al. (2011) contribute to this discourse by investigating RC acquisition among Persian-speaking children. Their findings indicate that Persian children encounter greater challenges with object and genitive RCs featuring non-canonical word order than subject RCs.

In addition to the hypotheses proposed on the acquisition order of RCs, some theories assume that language transfer from previously learned languages affects the acquisition of RCs in the language being learned. Based on these theories, L2 learners draw on their prior linguistic knowledge when engaging with the target language. Language transfer is well documented in second language research, with much of the literature focused on identifying errors and examining how transfer interacts with other factors (Leung & Williams, 2013). Ellis (1994, p. 300) claims that no theory of L2 acquisition "can be considered complete" if it overlooks the impact of learners' prior linguistic knowledge on acquiring the target language.

Relativization in Persian

In most world languages, a relative clause can modify an NP appearing in different syntactic positions. In Standard English, an NP modified by an RC can occur in one of the following syntactic positions: subject (SU) (see example 1), direct object (DO) (see example 2), indirect object (IO) (see example 3), oblique (OBL) (see example 4), genitive (GEN) (see example 5), or object of comparison (OCOMP) (see example 6). Henceforth, SU relatives, DO relatives, etc. denote NPs functioning as subjects, direct objects, etc. In the examples below, _ indicates a gap, created by *wh*-movement.

(1) The person who _ sent the key…
(2) The person who(m) I met _…
(3) The person who(m) I gave the key _…
(4) The person from whom we got the key _…
(5) The person whose _ friend sent the key…
(6) The person who I am taller than _…

In Persian, RCs follow the NPs they modify. They are introduced by the invariant relative marker *ke*, which functions similarly to "that" in English. The marker *ke* is used universally, regardless of the animacy, gender, grammatical role, or number of the noun modified by the RC (Taghvaipour, 2004, p. 267). In Persian, the presence of *ke* is mandatory, and RCs without *ke* are not permitted.

In contrast to English, where resumptive pronouns are not permitted, Persian allows them in RC constructions. They appear when one constituent of the RC is missing due to the *wh*-movement (Andrews, 2007). Resumptive pronouns appear in all RC types mentioned above, except in subject relatives. In subject relatives, only a gap is left in the clause (shown as _ in example 7). However, gap and resumptive pronouns are permitted in DO relatives and can be used interchangeably (example 8). In IO relatives (example 9), OBL relatives (example 10), GEN relatives (example 11), and OCOMP relatives (example 12), no gap can occur. In the following examples, the suffix *-i* that attaches to the end of the NPs modified by a restrictive RC is called demonstrative *-i* (Lazard, 1957, p. 66). Demonstrative *-i* can attach to the NP of restrictive RCs regardless of whether the noun is definite or indefinite. In the examples below, bolded elements indicate resumption.

(7) Persar-i ke _ email rā[1] ferestād…
boy-DEM REL email OM sent. 3SG
'The boy who sent the email…'
(8) Pesar-i ke (man) mišenās-am-(**aš**)…
boy-DEM REL (I) know-1SG-RES
'The boy whom I know…'
(9) Pesar-i ke (man) email rā barā-y**aš** ferestād-am…
boy-DEM REL (I) email OM for- RES sent-1SG
'The boy whom I sent the email…'
(10) Pesar-i ke (man) ketab rā az **u** gereft-am…
boy-DEM REL (I) book OM from RES got-1SG
'The boy from whom I got the book…'
(11) Pesar-i ke mādar-**aš** ketab rā ferestad…
boy-DEM REL mother-his book OM sent.3SG
'The boy whose mother sent the book…'
(12) Pesar-i ke (man) az **u** kootah-tar hast-am…
boy-DEM REL (I) from RES short-er be-1SG
'The boy who I am shorter than…'

Table 1 comprehensively compares the gap and resumptive pronoun strategies employed within relative clauses in Standard English and Persian.

Table 1
Comparison of the Distribution of Gap Strategy and Resumptive Pronoun Strategy in RCs between English and Persian

		Relative Clauses					
		SU	DO	IO	OBL	Gen	OCOMP
Gap	English	√	√	√	√	√	√
Resumption		X	X	X	X	X	X
Gap	Persian	√	√	X	X	X	X
Resumption		X	√	√	√	√	√

Note. The X and √, respectively, indicate the presence or absence of gap and resumptive pronoun availability for each RC type in both languages. SU: Subject; DO: Direct Object; IO: Indirect Object; OBL: Oblique; GEN: Genitive; OCOMP: Object of Comparison

Objectives of the Study

Many studies have examined the predictions proposed by the NPAH and the WDH (Birney et al., 2006; Brandt et al., 2008; Chan et al., 2011; Diessel & Tomasello, 2005; Hawkins, 2007; Izumi, 2003; Kidd & Bavin, 2002; Marefat & Rahmany, 2009; Ozeki & Shirai, 2007; Yas, 2016). Moreover, several studies have explored how the characteristics of RCs in familiar languages influence the acquisition of RCs in a second language. However, previous research has primarily focused on assessing the accuracy of relative clause structures produced by L2 learners. The identification and examination of erroneous types of RCs, particularly those that speakers actively avoid, have received limited attention. In general, avoidance remains a relatively understudied aspect of language acquisition research and merits further investigation. Additionally, from a typological standpoint, Persian stands out as a language worthy of investigation. Similar to English, Persian employs relative clauses to post-modify noun phrases.

However, a notable distinction arises in their treatment of resumption: while English prohibits resumption, Persian allows resumptive pronouns. This study explores the formation of English RCs by Persian-speaking learners of English, aiming to specifically examine the erroneous and avoided types of relative clauses in their performance. Importantly, previous studies on both L1 Persian and L2 English acquisition have largely overlooked the analysis of avoided and erroneous types of relative clauses, making this exploration particularly novel and valuable. The study addresses important theoretical questions regarding the impact of (i) general learnability, based on the assumption of the universal intrinsic constraints of RCs, and (ii) prior linguistic knowledge on RC acquisition. The study explores whether the errors Persian-speaking learners of English make in the formation of English RCs and the types of RCs they avoid forming can reflect the impact of the universal intrinsic constraints of RCs and/or L1 transfer. The study addresses the following research questions:

(i) What are the most common error types made by Persian-speaking learners when constructing English RCs?
(ii) What RC types do Persian-speaking learners of English mostly avoid forming?

To address the research questions, the following prediction is formulated: The frequency of erroneous RCs and the avoidance of specific types of RCs by Persian-speaking learners of English reflect the combined effects of L1 transfer and the intrinsic complexity of the RC types during the acquisition process. While our focus is on L2 acquisition, it is imperative to acknowledge the interplay between L1 transfer, the NPAH, and the WDH in shaping the acquisition patterns observed in our study. If the most common errors in the participants' performance resemble properties of Persian RCs, this may indicate the influence of L1 transfer. However, we recognize that the NPAH and WDH would have also influenced the original acquisition of RCs in Persian. Therefore, the interaction between L1 transfer and these inherent linguistic constraints could contribute to the acquisition patterns exhibited by our participants. Additionally, if the most prevalent erroneous types of RCs are associated with more complex structures, this implies that such structures present greater intrinsic difficulties for learners. This finding aligns with the predictions of the NPAH and WDH. Likewise, if learners tend to avoid certain types of relative clauses that are more marked, this avoidance may stem from their increased complexity. Overall, our study aims to elucidate the subtle relationship between L1 transfer, the NPAH, and the WDH in the acquisition of RCs by Persian-speaking learners of English, shedding light on the underlying mechanisms shaping their acquisition patterns.

Method

Participants

The study included 147 Persian school graduates, comprising 61 males and 86 females, aged between 17 and 19, from various schools in Iran. They were all native speakers of Persian who had finished their studies at school and were preparing for university entrance exams. They volunteered to participate in the experiment. They were not paid for their participation and did not receive any benefits. They had all studied the same English school textbooks with the same syllabi, had not spent any time

in English-speaking countries, had not had any in-person contact with native English speakers, and had not taken any extra English language courses. The input they had received on English RCs was largely identical. Their textbooks included only subject and direct object relatives, and their teachers strictly adhered to the same topics presented in the textbooks. A preliminary grammar placement test developed by Cambridge University Press was conducted before the experiment. The placement test, provided by the language institute where the study was conducted, was administered to assess participants' general English proficiency and classify them into three distinct proficiency levels. The placement test consisted of 120 multiple-choice items and the allocated time was 40 minutes following the instructions for the test. 19 out of 147 participants were excluded from the study due to their placement test scores falling below 15.8% of the total score, which was established as the minimum requirement for inclusion (scores under 20). This decision was made before the experiment to ensure a certain level of proficiency required for the study's objectives. This led to a final participant count of 128, comprising 72 females and 56 males. Participants were grouped into three proficiency levels - high, intermediate, and low – based on their scores from the placement test. Specifically, 21 students scoring 81 and above were classified as high proficiency, 72 students scoring between 41 and 80 were labeled intermediate proficiency, and 35 students scoring between 20 and 40 were categorized as low proficiency. To ensure participant privacy, individuals were informed that they need not provide their names on answer sheets. Additionally, even if names were provided, answer sheets were pseudonymized to prevent any identification in analysis sheets.

Elicitation Tasks

The data elicitation task in this study was a translation test (see Appendix, Table A1). The test was designed based on six types of RCs in the NPAH and comprised 30 test sentences, with five sentences representing each type of RC. That is 5 × SU, 5 × DO, 5 × IO, 5 × OBL, 5 × GEN, and 5 × OCOM (see examples 13-18 below for each RC type, respectively). The syntactic function of the noun phrase in the matrix clause (NPmat role) is not the focus of the NPAH; however, to ensure the homogeneity of the test items, all the NPs in the test items had the same NPmat role of a predicate nominal. The test sentences were in Persian and the participants were asked to translate the sentences into English. The order of the test sentences was randomized. Distractor items were excluded from this experiment due to the time-consuming nature of their inclusion. On average, the participants spent 40 minutes on the test; therefore, incorporating many distractor items would have resulted in an excessively lengthy experiment. The researcher recognizes that omitting distractor items may have made participants aware of the experiment's purpose, which could have implications for the results. For practical reasons related to time and space, it was necessary to either reduce the number of test sentences- potentially compromising the integrity of the results- or exclude distractor items. The latter option was deemed the more reasonable choice. Using the translation test in this study allowed for a focused examination of the production of rare types of relative clauses. While employing various elicitation methods can lead to more precise conclusions, the present study was limited to a single elicitation measurement due to challenges in recruiting participants willing to engage in different testing formats.

(13) In mard-i ast ke Sārā rā mi-shenās-ad.
this man- DEM is REL Sara OM PRS-know-3SG
'This is the man who knows Sara.'
(14) In mard-i ast ke Sārā oo rā mi-shenās-ad.
this man- DEM is REL Sara him.RES OM PRS-know-3SG
'This is the man whom Sara knows well.'
(15) In mard-i ast ke Sārā be oo ketāb-i dād.
this man- DEM is REL Sara to him.RES book-INDEF gave.3SG
'This is the man to whom Sara gave a book.'
(16) In mard-i ast ke shomā dishab darbārāreye oo sohbat=kard-id.
this man- DEM is REL you last night about him.RES talked-2PL
'This is the man about whom you spoke last night.'
(17) In pesar-i ast ke khāhar-ash dar kelās-e mā bud.
this boy- DEM is REL sister-his in class-EZ our was
'This is the boy whose sister was in our class.'
(18) In dokhtar-i ast ke Mary az oo bāhush-tar ast.
this girl- DEM is REL Mary than her clever-more is
'This is the girl who Mary is smarter than.'

Procedure

First, the translation test was administered to all 128 participants, with each participant assigned 30 test sentences, resulting in a total of 3840 RCs. Before the test, participants were familiarized with the content and informed that the test was for research purposes. During the translation test, participants read each of the 30 Persian sentences and translated them into English within a 40-minute time limit. After the test, the researcher counted and verified the number of correctly translated RCs across the six types. Each correct response was assigned a score of 1. Since there were five test sentences for each type of RC, participants received a score ranging from 0 to 5 for their correct formation of each RC type. A score of 0 was given when no sentences were correctly formed, while a score of 5 was assigned when all five sentences were correctly translated. Minor errors, such as changes in tense, number, or definiteness, as well as grammatical or lexical mistakes that did not affect the overall structure and content of the RCs, were not counted as incorrect. For example, sentences that altered the non-canonical word order of an object relative to the canonical order of a subject relative, used a resumptive pronoun, or omitted a relative pronoun (such as "whose") were considered incorrect (examples 19, 20, 21).

(19) * This is the man who knows Sara.
(instead of: This is the man who Sara knows)
(20) * This is the man who Sara knows him well.
(instead of: This is the man who Sara knows well)
(21) * This is the man's son that had an accident.
(instead of: This is the man whose son had an accident)

In addition to analyzing correctly formed RCs, an avoidance and error analysis was conducted. The analysis aimed to determine whether the errors made by the Persian-speaking school graduates and the types of RC they tended to avoid could provide further insights into the acquisition of English RCs and the strategies employed by learners at different proficiency levels. To perform the error and avoidance analysis,

the number of erroneously formed RCs and the number of avoided RCs (i.e., unanswered test sentences) were counted separately for each type of RC. The errors identified in the data were then categorized based on their specific types for further investigation.

Results

Analysis of the Correctly Formed RCs

To assess the Persian school graduates' competence in the formation of the six types of English RCs in the NPAH (SU, DO, IO, OBL, GEN, and OCOMP), the participants completed a translation test in which they translated 30 test sentences from Persian into English. Table 2 presents the translation test's absolute frequency of correctly formed RCs across the three proficiency levels. The table also displays the proportions of correctly formed RCs compared to the total possible number for each RC type at different proficiency levels. To calculate these proportions, the total occurrences of each RC type were divided by the maximum possible number of that RC type at each proficiency level. The total possible number of each RC type at each level of proficiency was calculated by multiplying the number of participants at each level by five, which is the number of test sentences for each RC type. For example, as there were 21 participants at the high proficiency level and every participant produced 5 sentences, there were 105 sentences for the high proficiency level. Therefore, the 103 correctly formed SU relatives represent 98.09% of that.

Table 2
Frequency and Proportion of the Correctly Formed English RC Types at the Three Proficiency Levels

Proficiency levels (Number of participants)	SU	DO	IO	OBL	GEN	OCOMP
High level ($N = 21$)	103 (98.09%)	91 (86.67%)	66 (62.86%)	69 (65.71%)	50 (47.62%)	42 (40%)
Intermediate level ($N = 72$)	258 (71.67%)	191 (53.06%)	105 (29.17%)	75 (20.83%)	41 (11.39%)	14 (3.89%)
Low level ($N = 35$)	69 (39.43%)	35 (20%)	4 (2.28%)	0 -	9 (5.14%)	0 -

Note. SU: Subject, DO: Direct object, IO: Indirect object, OBL: Oblique, GEN: Genitive, OCOMP: Object of comparison; *N*: Number of participants at each proficiency level

IO and OBL relatives are structurally different but their NPs have identical semantic roles. However, in the analysis of the data in the current study, it was observed that nearly all the participants formed English IO relatives using a prepositional phrase (only 2 participants formed 4 instances of IO relatives without using prepositions). Therefore, the researchers decided to follow the previous studies (Diessel & Tomasello, 2005; Izumi, 2003; Keenan & Comrie, 1977; Kim & O'Grady, 2016; Ozeki & Shirai, 2007; Yas, 2016), in which RCs with prepositional phrases that corresponded semantically to IOs were counted as correctly formed IO relatives. The findings presented in Table 2 suggest the operational presence of the NPAH, particularly evident

among participants at the high and intermediate proficiency levels. Within these levels, the observed frequencies of correctly used RCs closely adhere to the hierarchy postulated by the NPAH. At the lower proficiency level, a slight discrepancy is evident following the IO relatives. Nonetheless, this marginal deviation, attributed to the limited sample size and inherent constraints associated with the low proficiency group, does not seem to undermine the validity of the NPAH.

To statistically test whether the participants' performance at the three proficiency levels in Table 2 agrees with the order predicted by the NPAH, we analyzed the frequencies of correctly formed RCs. This analysis was conducted using an Aligned Rank Transform (ART) ANOVA with the ARTool package (Kay et al., 2021) in the statistical software R (R Core Team, 2021). The model included the main effects of proficiency level and RC types, an interaction term between the two variables, and a random effect for participants. ART ANOVA is a non-parametric equivalent to a conventional ANOVA. The reason for choosing this test was that both Shapiro-Wilk's normality test (Razali & Wah, 2011; Shapiro & Wilk, 1965) and a visual inspection of the histograms revealed that the frequencies were not normally distributed for the three proficiency levels across the RC types (see Table 3 for skewness, Kurtosis, and their standard errors) (Cramer, 1998; Cramer & Howitt, 2004; Doane & Seward, 2011).

Table 3

Skewness and Kurtosis Measures Obtained from Shapiro-Wilk's Normality Test per Proficiency Level and Relative Clause Type

		Skewness	Standard error	Kurtosis	Standard error
Subject	High	-2.97	0.50	7.56	0.97
	Intermediate	-1.02	0.28	-0.90	0.55
	Low	0.42	0.39	-1.67	0.77
Direct Object	High	-2.14	0.50	4.42	0.97
	Intermediate	-0.15	0.28	-1.85	0.55
	Low	1.56	0.39	1.23	0.77
Indirect Object	High	-0.59	0.50	-1.68	0.97
	Intermediate	0.94	0.28	-0.95	0.55
	Low	3.98	0.39	14.75	0.77
Oblique	High	-0.66	0.50	-1.39	0.97
	Intermediate	1.45	0.28	0.33	0.55
	Low				
Genitive	High	-0.02	0.50	-2.10	0.97
	Intermediate	2.03	0.28	2.52	0.55
	Low	2.75	0.39	6.52	0.77
Object of Comparison	High	0.30	0.50	-1.40	0.972
	Intermediate	4.16	0.28	17.34	0.55
	Low				

Note. SU: Subject; DO: Direct Object; IO: Indirect Object; OBL: Oblique; GEN: Genitive; OCOMP: Object of Comparison

The ART ANOVA showed a main effect for RC type ($F = 54.423$, $df = 5$, $p < 0.0000$), a main effect for the proficiency level of the participants ($F = 30.458$, $df = 2$, $p < 0.0000$), and a significant interaction between the two ($F = 10.779$, $df = 10$, $p < 0.0000$). A Tukey post hoc test using the *art.con()* function was performed to determine which pairwise comparisons of RC types differed significantly within each proficiency

level. Table 4 summarizes the significance of each Tukey comparison (significant differences are printed in bold). It shows that SU relatives had significantly higher mean ranks than all the other RC types. A higher mean rank of SU relatives indicates more correctly formed SU relatives.

Table 4
Post Hoc Pairwise Comparisons of RC types at each Proficiency Level

	RC type	SU	DO	IO	OBL	GEN	OCOMP
High level	SU	-	0.99	0.00	0.01	0.00	0.00
	DO	0.99	-	0.08	0.42	0.00	0.00
	IO	0.00	0.08	-	0.99	0.69	0.47
	OBL	0.01	0.42	0.99	-	0.20	0.09
	GEN	0.00	0.00	0.69	0.20	-	1.00
	OCOMP	0.00	0.00	0.47	0.09	1.00	-
Intermediate level	SU	-	0.00	0.00	0.00	0.00	0.00
	DO	0.00	-	0.00	0.00	0.00	0.00
	IO	0.00	0.00	-	0.70	0.00	0.00
	OBL	0.00	0.00	0.70	-	0.64	0.00
	GEN	0.00	0.00	0.00	0.64	-	0.89
	OCOMP	0.00	0.00	0.00	0.00	0.89	-
Low level	SU	-	0.15	0.00	0.00	0.00	0.00
	DO	0.15	-	0.053	0.00	0.28	0.00
	IO	0.00	0.053	-	1.00	0.99	1.00
	OBL	0.00	0.00	1.00	-	0.99	1.00
	GEN	0.00	0.28	0.99	0.99	-	0.99
	OCOMP	0.00	0.00	1.00	1.00	0.99	-

Note. The given values are the results of the Tukey comparisons. SU: Subject; DO: Direct Object; IO: Indirect Object; OBL: Oblique; GEN: Genitive; OCOMP: Object of Comparison

Regarding the interaction between the two variables, relative clause type and proficiency level, the following results were obtained (see Table 5). As the table presents, post hoc comparisons demonstrated statistically significant differences in the mean ranks of all the RC types formed by the participants at the three proficiency levels, except for OBL, GEN, and OCOMP relatives between low and intermediate levels, and SU relatives between high and intermediate levels of proficiency.

Table 5
Post Hoc Pairwise Comparisons of each RC Type at the Three Proficiency Levels

		SU	DO	IO	OBL	GEN	OCOMP
High level	Intermediate level	0.19	0.01	0.01	0.00	0.01	0.00
	Low level	0.00	0.00	0.00	0.00	0.00	0.00
Low level	Intermediate level	0.00	0.00	0.02	0.17	1.00	1.00

Note. The given values are the results of the Tukey comparisons. SU: Subject; DO: Direct Object; IO: Indirect Object; OBL: Oblique; GEN: Genitive; OCOMP: Object of Comparison

Error and Avoidance Analysis

In addition to the correctly formed RCs formed by the participants, the erroneously formed RCs and the avoided RCs were counted and analyzed. Table 6 presents the frequency of the erroneously formed RCs and the avoided RCs at the three proficiency levels. It also shows the proportions of the erroneously formed RCs concerning all possible numbers of each RC type at each level of proficiency. To calculate the proportion of the erroneously formed RCs, the number of occurrences of each RC type was divided by the total possible number of that RC type at each proficiency level. The total possible number of each type of RC at each proficiency level was calculated by multiplying the number of participants at that level by five, corresponding to the number of test sentences for each RC type.

Table 6
Frequency and Proportion of the Erroneously Formed RCs and the Avoided RCs at the Three Proficiency Levels

Proficiency Levels (Number of Participants)		SU	DO	IO	OBL	GEN	OCOMP
High level (N = 21)	Avoided RCs	0	1 (0.95%)	10 (9.53%)	15 (14.28%)	29 (27.62%)	22 (20.96%)
	Erroneous RCs	2 (1.90%)	13 (12.38%)	29 (27.62%)	21 (20%)	26 (24.77%)	41 (39.04%)
Intermediate level (N = 72)	Avoided RCs	31 (8.61%)	60 (16.67%)	129 (35.83%)	201 (55.83%)	266 (73.89%)	279 (77.5%)
	Erroneous RCs	71 (19.72%)	109 (30.28%)	126 (35%)	84 (23.34%)	53 (14.73%)	67 (18.62%)
Low level (N = 35)	Avoided RCs	26 (14.86%)	63 (36%)	94 (53.71%)	112 (64%)	125 (71.43%)	116 (66.29%)
	Erroneous RCs	80 (45.72%)	77 (44%)	77 (39.42%)	63 (36%)	41 (23.43%)	59 (33.71%)

Note. The number of test sentences for each RC type was 5. SU: Subject; DO: Direct Object; IO: Indirect Object; OBL: Oblique; GEN: Genitive; OCOMP: Object of Comparison

Errors within each type of RC were categorized by error type (see Table 7). Table 7 shows the proportion of each erroneously formed RC type to all the other RC types in the data. Examples of the error types listed in Table 7 are presented in Table C1 in the Appendix. The analysis of the entire error patterns revealed that the three most common error types were related to incorrect usage of resumptive pronouns, changing relative clauses with non-canonical word order to the ones with canonical word order, and omitting relative markers. Furthermore, the distribution of error types within each RC type reveals noticeable patterns. In SU relatives, the most prevalent error was the absence of the predicate within the matrix clause. The predominant error for DO, IO, and OBL relatives involved the incorrect use of resumptive pronouns. The primary error type was a significant tendency to omit relative markers for GEN relatives. In OCOMP relatives, changes in word order within the RC emerged as the main error type.

Table 7

Frequency and Percentage Distribution of Error Types in each RC Type

Error types	SU	DO	IO	OBL	GEN	OCOMP	Total
Using resumptive pronouns	8 (5.22%)	71 (35.68%)	149 (64.22%)	86 (51.19%)	27 (22.5%)	31 (18.56%)	372
Changing word order in the RC	1 (0.65%)	78 (39.19%)	39 (16.81%)	24 (14.28%)	--	76 (45.50%)	218
Missing relative markers	36 (23.53%)	3 (1.51%)	1 (0.43%)	2 (1.19%)	53 (44.16%)	1 (0.60%)	96
Missing relative clause	29 (18.95%)	19 (9.55%)	10 (4.31%)	2 (1.19%)	14 (11.67%)	19 (11.38%)	93
Missing the predicate of the matrix clause	62 (40.52%)	11 (5.53%)	7 (3.01%)	--	--	12 (7.18%)	92
Wrong constituent order/ wrong preposition place	10 (6.53%)	7 (3.52%)	16 (6.90%)	13 (7.74%)	3 (2.5%)	5 (3.00%)	54
Missing preposition	--	--	--	41 (24.40%)	--	--	41
Using a wrong relative pronoun	3 (1.96%)	5 (2.51%)	9 (3.88%)	--	20 (16.67%)	--	37
Missing *than*	--	--	--	--	--	17 (10.18%)	17
Missing constituents in the relative clause	4 (2.61%)	5 (2.51%)	--	--	3 (2.5%)	5 (3.00%)	17
Using extra prepositions/verbs	--	--	1 (0.45%)	--	--	1 (0.60%)	2

Note. The numbers indicating the frequency of error types in each RC type are the sum of the erroneous RCs in each RC type at the three proficiency levels. SU: Subject; DO: Direct Object; IO: Indirect Object; OBL: Oblique; GEN: Genitive; OCOMP: Object of Comparison

One might argue that another factor that needs to be investigated is whether some test sentences show a particularly increased error rate. Therefore, the researcher checked whether all the errors related to each RC type have occurred in the same test

sentence of each RC type or were distributed among the 5 test sentences of each type. An investigation of the distribution of errors shows that errors can occur for all 5 sentences regardless of RC types (see Appendix for full details). Tables B1, B2, and B3 in the Appendix show the number and the percentage of errors at each test sentence of each particular RC type at the three proficiency levels. The results demonstrate that the errors in each RC type were distributed across all the test sentences of the same RC type.

Discussion

As the frequencies and proportions of correctly formed RC types in this study indicate (see Table 2), the occurrence rates of the correctly formed SU and DO relatives were the highest across the three proficiency levels. Many studies have explored the natural acquisition of RCs, examining various languages and genres among children and adult learners. These investigations have scrutinized the comprehension and production of RCs by both L1 and L2 learners. Noteworthy contributions include works by Birney et al. (2006), Clancy et al. (1986), Doughty (1991), Eckman et al. (1988), Gass (1979), Hawkins (2007), Hyltenstam (1984), Izumi (2003), Marefat & Rahmany (2009), Ozeki & Shirai (2007), Pavesi (1986), and Sadighi (1994). Within the realm of L2 acquisition, many studies have demonstrated alignment between the acquisition order of RCs and the natural order of acquisition.

In this study, participants at the higher proficiency levels correctly formed some test sentences with marked RCs (IO, OBL, GEN, and OCOMP), but their occurrence was lower than that of the less marked RCs (SU and DO). This is in line with the NPAH. The substantial differences between the occurrence of the less marked RCs and more marked RCs at the three proficiency levels could be due to a variety of reasons such as the absence of less marked RCs in the input and the content of textbooks of English, the learners' avoidance of using more marked RCs, or the intrinsic higher difficulty level of these RC types. This will be expanded on below.

To make a more informed assessment of the participant's performance in the production of RCs, the type of errors the learners made and the type of RCs they avoided were closely analyzed. The results showed a noticeable growth in erroneously formed RCs from less marked RCs to more marked ones. A closer inspection of the error patterns revealed that the three most common error types were related to:

(i) using resumptive pronouns in the construction of English RCs, for example, (**This is the man that Sara gave him a book*),
(ii) changing RCs with non-canonical word order to RCs with canonical word order, for example, (**This is the boy who met my sister*) instead of (*This is the boy who(m) my sister met*),
(iii) missing relative markers, for example, (**This is the student got a good mark*) instead of (*This is the student who got a good mark*).

Using resumptive pronouns may reveal an L1 transfer. Unlike in English, resumption is allowed in the structure of object relatives in Persian. Thus, the main reason for this type of error is the difference in the construction of RCs between the learners' first language (Persian) and the target language (English). Concerning the limited incidence of errors associated with the incorrect resumptive pronouns in subject-RCs, it is noteworthy to highlight that such errors were attributable solely to two individuals classified within the low proficiency group. Specifically, one participant

exhibited erroneous constructions across all five test sentences, and another committed three errors. Importantly, no other participants across any proficiency group replicated such mistakes. These isolated errors can be understood within a broader context and appear to have minimal effect on interpreting our results due to their infrequent occurrence.

Altering non-canonical word order to canonical word order supports the *Word Order Difference Hypothesis*, according to which non-canonical word order causes more difficulty than canonical word order (MacDonald & Christiansen, 2002). Changing non-canonical word order to canonical indicates a type of avoidance strategy employed by learners of English: avoidance of more marked types of RCs, which are intrinsically more difficult to learn and use than the less marked ones. Learners alter the more marked RC type, DO relative, to the less marked one, SU relative.

The omission of relative markers, the third most common type of error, does not seem to exhibit L1 transfer. The relative marker *ke* in Persian cannot be omitted from the sentence. The predominant occurrence of omitted relative markers primarily arises from the absence of the genitive relative marker *whose* within GEN relatives, with a secondary occurrence observed in the omission of relative pronouns within SU relatives. Elucidating the rationale behind the exclusion of *whose* in genitive relatives presents a challenge. One possible explanation relates to learners' tendency to use possessive pronouns as substitutes for *whose* due to a perceived equivalence in expressing possession. This tendency might originate from not fully grasping the differences between possessive and relative pronouns in English grammar. The Word Order Difference Hypothesis could explain the omission of relative markers in SU relatives, the second most common error of this type. In this error type, learners add "this is" to declarative clauses to form SU relatives, taking advantage of the similar word order between SU relatives and basic declarative sentences. The aggregate results obtained from the analysis of errors in all RC types in the present study suggest that both L1 transfer and the universal accessibility hierarchy of RCs account for the erroneous English RCs formed by Persian-speaking learners of English.

The analysis of error patterns as a whole, along with the distribution of error types within each RC type, revealed intriguing findings. Principally, the most prevalent error, the incorrect use of resumptive pronouns, stood out prominently in DO, IO, and OBL relatives. Subsequently, the second most common error, altering word order, was notably frequent in OCOMP relatives. Lastly, the third most common error, missing relative markers, was predominantly observed in GEN relatives.

In addition, the results obtained from the analysis of the avoided types of RCs demonstrated that from more marked RCs to less marked ones, the number of test sentences that remained unanswered substantially increased, especially at the intermediate and low levels of proficiency. The learners seem to have avoided translating the test sentences because they might have found these RC types more difficult than SU and DO relatives. This highlights the learners' avoidance of forming more marked RCs. Furthermore, less-proficient learners used the avoidance strategy significantly more than more-proficient learners. The question that needs to be raised here is why Persian-speaking learners of English have such a high tendency to avoid more marked RCs. This avoidance might be determined by either the higher frequency distributions of less marked RC types in input or the inherent difficulty of more marked RCs. Based on the aggregate results attained, the first prediction, according to which the frequency of the erroneous RCs and that of the avoided types of RCs by Persian-speaking learners of English display the effect of L1 transfer and the intrinsic complexity of the RC types, is confirmed.

Another significant finding of the study was that some participants at the high level and a few at the intermediate and low levels could produce RCs that were not addressed in the grammar sections of their textbooks. This observation raises the question of how these constructions were acquired. The formation of these untaught constructions may be attributed to several factors. One possibility is that participants at higher proficiency levels might have picked up these constructions unintentionally. For instance, these constructions may have appeared in the texts of the books they have read, been utilized by their English teachers, or been featured in the movies, TV programs, or games they have watched. Another possibility is that the participants at higher proficiency levels might have been able to extend their knowledge of the formation of the less marked RCs (SU and DO relatives) to the formation of more marked RCs.

Conclusion

This study provides a detailed examination of how the inherent complexity of RCs, coupled with prior linguistic knowledge, shapes the acquisition process of RCs in L2 contexts. By focusing on Persian-speaking learners of English, the study offers unique insights into the factors influencing error patterns and avoidance strategies in L2 acquisition. In contrast to conventional discussions that primarily focus on L1 transfer and universal constraints, this study emphasizes the critical role of differences in RC construction between learners' native and target languages. These disparities significantly contribute to the production of erroneous RCs in L2 contexts. Additionally, the study highlights avoidance as a prominent strategy employed by learners when faced with the intricacies of more complex RC structures in the target language. An important finding of this study is the intricate relationship between avoidance and the universal intrinsic constraints of RCs, as outlined in the NPAH and WDH. This sheds light on the underlying mechanisms behind avoidance behaviors in RC acquisition and emphasizes the significance of linguistic complexity in instructional planning and pedagogical strategies. In summary, this study offers a fresh perspective on the challenges encountered by L2 learners in mastering complex syntactic structures. By clarifying the specific effects of linguistic disparities and avoidance strategies on acquiring RCs, this study offers valuable insights that can help inform more targeted language teaching methods and curriculum design in L2 contexts.

Correspondence should be addressed to Golpar Bahar.
Email: Golparbahar.et@gmail.com

Notes

[1] OM: Object modifier

References

Andrews, A. (2007). Relative clauses. In T. Shopen (Ed.), *Language typology and syntactic description* (pp. 206–236). Cambridge University Press.

Bahar, G. (2023). Acquisition of English relative clauses by native speakers of Kurdish Sorani. *Canadian Journal of Linguistics/Revue canadienne de linguistique, 68*(3), 414–434. https://doi.org/10.1017/cnj.2023.21

Bahar, G. (2024). *Examining the noun phrase accessibility hierarchy and the perceptual difficulty hypothesis in English relative clauses formed by native speakers of Central Kurdish*. In A. Grond, & S. Gündoğdu (Eds.), *Current issues in Kurdish linguistics II* (pp. 84–94). Praesens. [In press].

Bever, T. (1970). The cognitive basis for linguistic structures. In J. R. Hayes (Ed.), *Cognition and development of language* (pp. 279–362). Wiley.

Birney, D., Halford, G., & Andrews, G. (2006). Measuring the influence of complexity on relational reasoning: The development of the Latin square task. *Educational and Psychological Measurement, 66*(1), 146–171. https://doi.org/10.1177/0013164405278570

Brandt, S., Diessel, H., & Tomasello, M. (2008). The acquisition of German relative clauses: A case study. *Journal of Child Language, 35*(2), 325–348. https://doi.org/10.1017/S0305000907008379

Brown, D. (1971). Children's comprehension of relativized English sentences. *Child Development, 42*(6), 1923–1936. https://doi.org/10.2307/1127595

Chan, A., Matthews, S., & Yip, V. (2011). The acquisition of relative clauses in Cantonese and Mandarin. In E. Kidd (Ed.), *The acquisition of relative clauses: Processing, typology, and function* (pp. 197–226). John Benjamins Publishing Company. https://doi.org/10.1075/tilar.8.10cha

Clancy, P., Lee, H., & Zoh, M.-H. (1986). Processing strategies in the acquisition of relative clauses: Universal principles and language-specific realizations. *Cognition, 24*(3), 225–262. https://doi.org/10.1016/S0010-0277(86)80003-8

Cramer, D. (1998). *Fundamental statistics for social research*. Routledge.

Cramer, D., & Howitt, D. (2004). *The SAGE dictionary of statistics*. SAGE.

Diessel, H., & Tomasello, M. (2005). A new look at the acquisition of relative clauses. *Language, 81*(4), 882–906. https://doi.org/10.1353/lan.2005.0169

Doane, D., & Seward, L. (2011). Measuring skewness: A forgotten statistic?. *Journal of Statistics Education, 19*(2), 1–18. https://doi.org/10.1080/10691898.2011.11889611

Doughty, C. (1991). Second language instruction does make a difference: Evidence from an empirical study of SL relativization. *Studies in Second Language Acquisition, 13*(4), 431–469. https://doi.org/10.1017/S0272263100010287

Doughty, C., & Long, M. (Eds.) (2003). *The handbook of second language acquisition*. Blackwell. https://doi.org/10.1002/9780470756492

Eckman, F. (1977). Markedness and the contrastive analysis hypothesis. *Language Learning, 27*(2), 315–330. https://doi.org/10.1111/j.1467-1770.1977.tb00124.x

Eckman, F., Bell, L., & Nelson, D. (1988). On the generalization of relative clause instruction in the acquisition of English as a second language. *Applied Linguistics, 9*(1), 1–20. https://doi.org/10.1093/applin/9.1.1

Ellis, R. (1994). *The study of second language acquisition*. Oxford University Press.

Gass, S. (1979). Language transfer and universal grammatical relations. *Language Learning, 29*(2), 327–344. https://doi.org/10.1111/j.1467-1770.1979.tb01073.x

Gass, S. (1982). From theory to practice. In M. Hynes & W. Rutherford (Eds.), *On TESOL 81: Selected papers from the fifteenth annual conference of teachers of English to speakers of the languages* (pp. 1–15). Teachers of English to Speakers of Other Languages.

Gass, S. (1980). An investigation of syntactic transfer in adult second language learners. In R. Scarcella, & S. Krashen (Eds.), *Research in second language acquisition: Selected papers from the Los Angeles Second Language Acquisition Research Forum* (pp. 132–141). Newbury House.

Gass, S., & Lee, J. (2007). Second language acquisition of relative clauses. *Studies in Second Language Acquisition, 29*(2), 329–335. http://www.jstor.org/stable/44487155.

Hamilton, R. (1994). Is implicational generalization unidirected and maximal? Evidence from relativization instruction in a second language. *Language Learning, 44*(1), 123–157. http://doi.org/10.1111/j.1467-1770.1994.tb01451.x

Hawkins, R. (2007). The noun phrase accessibility hierarchy: Lame duck or dead duck in theories of SLA? *Studies in Second Language Acquisition, 29*(2), 345–349. http://www.jstor.org/stable/44487157.

Hyltenstam, K. (1984). The use of typological markedness conditions as predictors in second language acquisition: The case of pronominal copies in relative clauses. In R. Andersen (Ed.), *Second languages: A crosslinguistic perspective* (pp. 39–58). Newbury House.

Ioup, G., & Kruse, A. (1977). Interference vs. structural complexity in second language acquisition: Language universals as a basis for natural sequencing. In H. Brown, C. Yorio, & R. Crymes (Eds.), *On TESOL 77: Teaching and learning English as a second language: Trends in research and practice* (pp. 159–171). TESOL.

Izumi, S. (2003). Processing difficulty in comprehension and production of relative clauses by learners of English as a second language. *Language Learning, 53*(2), 285–323. https://doi.org/10.1111/1467-9922.00218

Kay, M., Elkin, L., Higgins, J., & Wobbrock, J. (2021). ARTool: Aligned rank transform for nonparametric factorial ANOVAs (R package version 0.11.1). https://doi.org/10.5281/zenodo.594511

Keenan, E., & Comrie, B. (1977). Noun phrase accessibility and universal grammar. *Linguistic Inquiry, 8*(1), 63–99. http://www.jstor.org/stable/4177973

Kidd, E., & Bavin, E. L. (2002). English-speaking children's comprehension of relative clauses: Evidence for general-cognitive and language-specific constraints on development. *Journal of Psycholinguistic Research, 31*, 599–617. https://doi.org/10.1023/A:1021265021141

Kim, C.-E., & O'Grady, W. (2016). Asymmetries in children's production of relative clauses: Data from English and Korean. *Journal of Child Language, 43*(5), 1038–1071. http://doi.org/10.1017/S0305000915000422

Lazard, G. (1957). *Grammaire du persan contemporain*. Klincksieck.

Leung, J., & Williams, J. (2013). Prior linguistic knowledge influences implicit language learning. *Proceedings of the Annual Meeting of the Cognitive Science Society, 35*. https://escholarship.org/uc/item/6kh2b1dw

Marefat, H., & Rahmany, R. (2009). Acquisition of English relative clauses by Persian EFL learners. *Journal of Language and Linguistic Studies, 5*(2), 21–48. https://www.jlls.org/index.php/jlls/article/view/80

MacWhinney, B., Bates, E., & Kliegl, R. (1984). Cue validity and sentence interpretation in English, German, and Italian. *Journal of Verbal Learning and*

Verbal Behavior, 23(2), 127–150. https://doi.org/10.1016/S0022-5371(84)90093-8

MacDonald, M., & Christiansen, M. (2002). Reassessing working memory: Comment on Just and Carpenter (1992) and Waters and Caplan (1996). *Psychological Review, 109*(1), 35–54. https://doi.org/10.1037/0033-295X.109.1.35

Ozeki, H., & Shirai, Y. (2007). Does the noun phrase accessibility hierarchy predict the difficulty order in the acquisition of Japanese relative clauses? *Studies in Second Language Acquisition, 29*(2), 169–196. https://doi.org/10.1017/S0272263107070106

Pavesi, M. (1986). Markedness, discoursal modes, and relative clause formation in a formal and an informal context. *Studies in Second Language Acquisition, 8*(1), 38–55. https://www.jstor.org/stable/44486850

Prideaux, G., & Baker, W. (1986). *Strategies and structures: The processing of relative clauses*. John Benjamins.

R Core Team. (2021). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing. https://www.R-project.org/

Razali, N. M., & Yap, W. B. (2011). Power comparisons of Shapiro-Wilk, Kolmogorov-Smirnov, Liliefors, & Anderson-Darling tests. *Journal of Statistical Modeling and Analytics, 2*(1), 21–33.

Romaine, S. (1984). Relative clauses in child language, pidgins, and creoles. *Australian Journal of Linguistics, 4*(2), 257–281.

Rahmany, R., Marefat, H., & Kidd, E. (2011). Persian-speaking children's acquisition of relative clauses. *European Journal of Developmental Psychology, 8*(3), 367–388. https://doi.org/10.1080/17405629.2010.509056

Sadighi, F. (1994). The acquisition of English restrictive relative clauses by Chinese, Japanese, and Korean adult native speakers. *IRAL-International Review of Applied Linguistics in Language Teaching, 32*(2), 141–153.

Shapiro, S. S., & Wilk, M. B. (1965). An analysis of variance test for normality (complete samples). *Biometrika, 52*(3/4), 591–611. https://doi.org/10.2307/2333709

Slobin, D., & Bever, T. (1982). Children use canonical sentence schemas: A crosslinguistic study of word order and inflections. *Cognition, 12*(3), 229–265. https://doi.org/10.1016/0010-0277(82)90033-6

Taghvaipour, M. A. (2004). An HPSG analysis of Persian relative clauses. *Proceedings of the 11th International Conference on Head-Driven Phrase Structure Grammar*, 274–293. https://doi.org/10.21248/hpsg.2004.16

Tavakolian, S. (1981). The conjoined clause analysis of relative clauses. In T. S. (Ed.), *Language acquisition and linguistic theory* (pp. 167–187). MIT Press.

Yas, E. (2016). *Acquisition of English relative clauses by German L1 and Turkish L1 speakers* [Unpublished doctoral dissertation]. Free University of Berlin, Berlin, Germany. http://dx.doi.org/10.17169/refubium-8862.

Appendix

Table A1
Test Sentences in the Translation Test Employed in the Study

Relative Clause Type	Number	Test Sentence
Subject	1	This is the man who knows Sara.
	2	This is the boy who met Sara yesterday.
	3	This is the cat that played in the park yesterday.
	4	This is the student who got a good mark.
	5	This is the woman who came to the library.
Direct Object	1	This is the man whom Sara knows well.
	2	This is the boy whom my sister met last week.
	3	This is the girl whom Tom loved so much.
	4	This is the book that I bought yesterday.
	5	This is the school subject which I liked a lot.
Indirect Object	1	This is the man to whom Sara gave a book.
	2	This is the girl to whom I gave my doll.
	3	This is the woman to whom Sara sent a letter.
	4	This is the library to which I gave the books.
	5	This is the person to whom I showed the house.
Object of Preposition	1	This is the place from which we bought the books.
	2	This is the man about whom you spoke last night.
	3	This is the table on which he put his bag yesterday.
	4	This is the teacher from whom we learned a lesson.
	5	This is the task on which the students worked.
Genitive	1	This is the boy whose sister was in our class.
	2	This is the singer whose song was the best.
	3	This is the man whose son had an accident.
	4	This is the girl whose mom came to our school.
	5	This is the dog whose picture was in the newspaper.
Object of Comparison	1	This is the girl who Mary is smarter than.
	2	This is the rival who I am better than.
	3	This is the flat which my house is smaller than.
	4	This is the student who Mary is more intelligent than.
	5	This is the boy who Perter is younger than.

Note: The test sentences were in Persian, and the English translations are listed here.

Table B1

Number and Percentage of Errors at Each Test Sentence of Each RC Type at the High Proficiency Level

Number of test sentences	SU		DO		IO		OBL		GEN		OCOMP	
	No	%	No	%	No	%	No	%	No	%	No	%
1	1	50	1	7.7	5	17.24	4	19.05	3	11.54	9	21.95
2	0	-	2	15.39	6	20.69	3	14.29	6	23.07	7	17.07
3	1	50	4	30.77	7	24.14	6	28.58	6	23.07	7	17.07
4	0	-	5	38.46	5	17.24	3	14.29	8	30.77	10	24.40
5	0	-	1	7.7	6	20.69	5	23.80	3	11.54	8	19.51
Total number of errors	2		13		29		21		26		41	

Note. Number of errors at each test sentence of each particular RC type

Table B2

Number and Percentage of Errors at Each Test Sentence of Each RC Type at the Intermediate Proficiency Level

Number of test sentences	SU		DO		IO		OBL		GEN		OCOMP	
	No	%	No	%	No	%	No	%	No	%	No	%
1	19	26.76	31	28.45	34	26.99	14	16.67	10	18.87	16	23.89
2	17	23.94	25	22.94	28	22.22	21	25	10	18.87	12	17.91
3	11	15.50	22	20.18	28	22.22	17	20.24	12	22.64	13	19.40
4	10	14.08	17	15.60	14	11.11	18	21.43	8	15.09	14	20.90
5	14	19.72	14	12.85	22	17.47	14	16.67	13	24.53	12	17.91
Total number of errors	71		109		126		84		53		67	

Table B3

Number and Percentage of Errors at Each Test Sentence of Each RC Type at the Low Proficiency Level

Number of test sentences	SU		DO		IO		OBL		GEN		OCOMP	
	No	%	No	%	No	%	No	%	No	%	No	%
1	18	22.5	19	24.68	16	20.78	15	23.80	8	17.78	12	20.34
2	15	18.75	20	25.98	13	16.88	16	25.40	8	17.78	12	20.34
3	14	17.5	15	19.49	19	24.67	10	15.87	8	17.78	12	20.34
4	14	17.5	13	16.89	12	15.59	14	22.23	9	20	11	18.65
5	19	23.75	10	12.98	17	22.08	8	12.70	8	17.78	12	20.34
Total number of errors	80		77		77		63		41		59	

Table C1

Examples of the Error Types Listed in Table 7

Error types	SU	DO	IO	OBL	GEN	OCOMP
Using resumptive pronouns	This is the man who he knows Sara	This is the man whom Sara knows him well.	This is the woman to whom Sara sent her a letter.	This is the man you spoke about him last night.	This is the boy whose her sister was in our class.	This is the girl who Mary is smarter than her.
Changing word order in the RC	This is the man who Sara knows.	This is the man who knows Sara well.	This is the man who gave a book to Sara.	This is the man who spoke about you last night.	--	This is the girl who is smarter than Mary.
Missing relative markers	This is the man knows Sara.	This is the boy's sister met last week.	This is the man to Sara gave a book	This is the man about you spoke last night.	This is the boy his sister was in our class.	This is the student is more intelligent than.
Missing relative clause	This man knows Sara.	This man Sara knows well.	This man to Sara gave a book	This man you spoke about last night.	The boy's sister was in our class.	This girl is Mary smarter than.
Missing the predicate of the matrix clause	This man who knows Sara	This man whom Sara knows well.	The man to whom Sara gave a book.	--	--	This girl who Mary is smarter than
Wrong constituent order/ wrong preposition place	The boy is who met Sara yesterday.	The man is whom Sara knows well.	This is man whom to Sara gave a book.	This is the man whom you about spoke.	The boy was in whose sister in our class.	This is Mary who the girl is smarter than.
Missing preposition	--	--	--	This is the task that the students worked.	--	--
Using a wrong relative pronoun	This is the man which knows Sara.	This is the man which Sara knows well.	This is the man which Sara gave a book to.	--	This is the boy who his sister was in our class.	--
Missing *than*	--	--	--	--	--	This is the girl who Mary is smarter.

Missing constituents in the relative clause	This is the man who knows.	This is the man whom knows well.	--	--	This is the boy whose was in our class.	This is the girl who is smarter than.
Using extra prepositions /verbs	--	--	This is the man to whom Sara gave a book to.	--	--	This is the girl who is Mary is smarter than.