Choi, Eunji & Lee, Sunyoung. 2025. Development of the relative clause among Chinese-speaking learners of Korean. *Glossa: a journal of general linguistics* 10(1). pp. 1–36. DOI: https://doi.org/10.16995/glossa.9598

Open Library of Humanities

Development of the relative clause among Chinese-speaking learners of Korean

Eunji Choi, Department of Korean Language and Literature, Korea University, KR, postino78@korea.ac.kr

Sunyoung Lee, Division of Korean Language and Multiculture, Korea Cyber University, KR, sylee41@koreacu.ac.kr

The development of relative clauses (RCs) has been examined through various theoretical frameworks, including the Markedness Hypothesis, the Surface Configuration Hypothesis, the Dependency Locality Theory, and the Absolutive Hypothesis. This study tested these hypotheses using 18 months of spontaneous spoken data collected from Chinese-speaking learners of Korean. A total of 1,178 sentences containing RCs were extracted and analyzed. The study investigated the frequency of occurrence, the timing of initial emergence, and the syntactic error rates of different RC patterns over time. Predicate types were classified based on their obligatory arguments, and the frequency and error rates for each type were examined. Additionally, the study examined the distribution of predicate types, as well as the relationship between predicate types and the grammatical roles played by head nouns within the relative clause. The results revealed that among the 1,178 RCs, subject RCs occurred most frequently, followed by direct object and oblique RCs. However, no clear differences were observed between direct object RCs and oblique RCs in terms of their timing of appearance and error rates. These results suggest that the Markedness Hypothesis, the Surface Configuration Hypothesis, and the Dependency Locality Theory may have limitations in explaining the observed patterns. Notably, the study found that RCs involving predicates with simpler argument structures were acquired more easily than those with more complex ones. Specifically, the developmental sequence of RCs followed this order: intransitive verbs without obligatory oblique arguments > transitive verbs > intransitive verbs with obligatory oblique arguments > ditransitive verbs. Overall, the results support the Absolutive Hypothesis in explaining the development of RCs across different predicate types.

1 Introduction

This study examines several hypotheses concerning the development of relative clauses (RCs), including the Markedness Hypothesis, the Surface Configuration Hypothesis, the Dependency Locality Theory, and the Absolutive Hypothesis. To identify factors influencing RC development, the study analyzes spontaneous spoken data collected over 18 months from three Chinese-speaking learners of Korean.

RC development has been approached from diverse perspectives. The Markedness Hypothesis is the most well-known hypothesis, grounded in the Noun Phrase Accessibility Hierarchy (NPAH) (Keenan & Comrie 1977). The NPAH, a linguistic typological framework, posits that certain types of RCs are more accessible than others, leading to predictions about their relative difficulty in second language acquisition. The markedness hierarchy suggested by the NPAH predicts the difficulty of acquiring RCs (Gass 1979a; 1979b; Pavesi 1986; Eckman et al. 1988; Izumi 2003; Xia et al. 2020).

Scholars have examined RC development from a psycholinguistic perspective (Tarallo & Myhill 1983; Hawkins 1989; Gibson 1998; 2000). The Surface Configuration Hypothesis and Dependency Locality Theory are among the various possibilities discussed in the psycholinguistic literature. Tarallo & Myhill (1983) and Hawkins (1989) developed the Surface Configuration Hypothesis, which states that the difficulty of acquiring RCs is determined by the distance between the head noun and the extraction site within the RC. Gibson (1998; 2000) proposed the Dependency Locality Theory, asserting that the greater the psycholinguistic burden, consisting of memory and integration costs, the more challenging it is to acquire RCs.

Furthermore, scholars have attempted to explain RC development in relation to the RC predicate argument structure (Fox 1987; Diessel & Tomasello 2005). Under Fox's (1987) Absolutive Hypothesis, head nouns were analyzed in conjunction with the RC predicate argument to explore whether they were intransitive subjects, agents, or patients. According to this hypothesis, intransitive subjects and patients precede agents in the hierarchy. Diessel & Tomasello (2005) analyzed head nouns based on the RC predicate argument, whereby the RC difficulty was attributed to their similarity to simple non-embedded sentences. Fox (1987) and Diessel & Tomasello (2005) demonstrated the need to analyze RC development based on the RC predicate argument.

The current study, which is based on spontaneous spoken conversation data collected longitudinally from Chinese-speaking learners of Korean, aims to evaluate previous claims regarding RC development. As this study analyzed longitudinal data, it was possible to closely examine learners' developmental patterns over time. Furthermore, Chinese and Korean share common features. Both are prenominal languages in forming RCs; however, their basic syntactic structures differ. The study will explore RC development in Chinese-speaking learners of Korean

by determining whether the differences between Korean and Chinese affect development. In addition, the results of this study, which targets second language learners, will be compared to those of previous studies that have examined RC development in native speakers.

2 Theoretical background

2.1 Research on the sentence component of RC head nouns and development

Keenan & Comrie (1977; 1979) developed the NPAH, a universal linguistic hierarchy for the RC, after analyzing more than 50 languages. Their study stated that the RCs can be categorized as subject, direct object, indirect object, oblique, genitive, and object of comparison RCs, depending on the grammatical role of the head noun within the clause. The accessibility hierarchy, which follows the order of subject > direct object > indirect object > oblique > genitive > object of comparison, can be derived from this classification.

(1) a. The man who met her (subject RC)
b. The man that she met (direct object RC)
c. The man who she gave books to (indirect object RC)
d. The man who she played with (oblique RC)
e. The man whose books are numerous (genitive RC)
f. The man who she is taller than (object of comparison RC)

According to the NPAH, if any RC type is possible in a specific language, the preceding types are also possible within this hierarchy. For instance, in languages such as Korean, it is possible to relativize oblique RCs. Therefore, in Korean, RCs of the subject, direct object, and indirect object types, which precede oblique types, can be created. Thus, the NPAH is a linguistic typological hypothesis that categorizes languages based on their ability to relativize according to this hierarchy. The NPAH originated in discussions of linguistic typology; however, it has since spurred claims that the hierarchy is linked to frequency and cognitive load. RC types that appear earlier in the hierarchy are assumed to be more frequent or less cognitively demanding during human sentence processing. In this regard, the NPAH, which was initially a linguistic typological hypothesis, has been applied to the study of foreign languages. The Markedness Hypothesis, which builds on the NPAH, posits that the hierarchy predicts the degree of difficulty and the order of RC acquisition. In this hierarchy, the preceding types are considered more universal and less marked than the latter ones. More precisely, learners can more easily acquire the less marked types that appear earlier in the hierarchy (Eckman et al. 1988; Doughty 1991). Accordingly, several studies (Gass 1979a; 1979b; Pavesi 1986; Izumi 2003; Xia et al. 2020) have found that typological markedness in the NPAH provides a useful predictor of RC learning difficulties.

However, most studies on linguistic typological markedness have focused on languages with postnominal RC structures, such as English. Only a few studies have examined languages

with prenominal RC structures, such as Korean, Japanese, and Chinese, and have produced inconsistent results. Some studies on prenominal RC languages have supported the predictive validity of the NPAH (second language learners of Chinese: Xu 2014; native speakers of Chinese: Hu et al. 2016a; 2016b). Nevertheless, other studies have shown that RC difficulty or acquisition order does not necessarily follow the predictions of the Markedness Hypothesis (second language learners of Japanese: Ozeki & Shirai 2007; second language learners of Japanese and Chinese: Tarallo & Myhill 1983; native speakers of Chinese: Hsiao & Gibson 2003; Chen et al. 2008; Gibson & Wu 2013).

These findings have given rise to psycholinguistic hypotheses and theories about RC complexity. Using grammatical judgment tasks, Tarallo & Myhill (1983) examined the RC acquisition of native English speakers learning German and Portuguese (postnominal languages) and Chinese and Japanese (prenominal languages). The results showed that German and Portuguese learners achieved the highest level of accuracy for the subject RC type, whereas Chinese and Japanese learners achieved the highest level of accuracy for the direct object RC type. To explain the difference in the results between prenominal and postnominal RC languages, Tarallo & Myhill (1983) used the degree of proximity between the head noun and the extraction site. The closer the extraction site and the head noun are in the surface configuration, the easier it is to acquire the RC type. As found in (2)–(4), in contrast to English, in Korean, Japanese, and Chinese, the distance between the extraction site and the head noun is shorter for the direct object RC type than for the subject RC type, which results in higher accuracy for the direct object RC type.

(2) English
   a. HN [___ V O]$_{REL}$ (subject RC)
   b. HN [S V ___]$_{REL}$ (direct object RC)

(3) Korean and Japanese
   a. [___ O V]$_{REL}$ HN (subject RC)
   b. [S ___ V]$_{REL}$ HN (direct object RC)

(4) Chinese
   a. [___ V O]$_{REL}$ HN (subject RC)
   b. [S V ___]$_{REL}$ HN (direct object RC)

Hawkins (1989) supported this view and presented data from genitive RC acquisition by English speakers learning French. In French, the distance between the head noun and the extraction site of the genitive RC is shorter when the genitive RC functions as an embedded subject noun phrase. Accordingly, Hawkins (1989) hypothesized and verified that learners achieved higher accuracy when the genitive RC was used as an embedded subject noun phrase.

Gibson (1998; 2000) explained this phenomenon through the Dependency Locality Theory from a psycholinguistic perspective. The Dependency Locality Theory is centered on memory and integration costs. The memory cost becomes greater as the distance over which the predicted syntactic category must be retained increases until the prediction becomes correct. The integration cost increases as the distance between the elements to be integrated increases. Gibson (1998) explained that the subject RC is less complex in English because both integration and memory costs are lower in subject RCs than in direct object RCs.

However, the predictions made within the Dependency Locality Theory about the difficulty of the subject RC and the direct object RC in Korean and Japanese are somewhat unclear. Regarding integration cost, Ishizuka et al. (2006) found that in Korean and Japanese, the subject RC, with a greater distance between the empty noun phrase and the head noun, incurs higher integration cost than the direct object RC. In contrast, Miyamoto & Nakamura (2003) argued that integration occurs when the verb and the noun phrase are encountered, and that the empty noun phrase functions as a placeholder for the upcoming head noun. Therefore, they claimed that there is little difference in integration cost between subject RCs and direct object RCs. Predictions concerning the relative difficulty of the subject RC and the direct object RC with respect to memory cost are also mixed. Gibson & Wu (2013) argued that in languages such as Korean and Japanese, the direct object RC is more difficult than the subject RC. This is due to ambiguity that may arise at the beginning of a direct object RC, which might initially be interpreted as a main clause. The direct object RC, which begins with a subject, is likely to be interpreted as the subject of the main clause. This ambiguity persists until a verb or head noun is encountered. However, the subject RC, which begins with an object, is interpreted as a RC from the start, incurring little memory cost. However, in Korean, where main clauses can begin with either a subject or an object, the claim by Gibson & Wu (2013) that direct object RCs involve higher memory costs than subject RCs is still debatable.

2.2 Study on predicate arguments in RCs and their development

As discussed above, the Markedness Hypothesis, the Surface Configuration Hypothesis, and the Dependency Locality Theory focus on the grammatical roles assigned to head nouns. Fox (1987), Diessel & Tomasello (2005), and Hogbin & Song (2007) have examined RCs with a focus on predicate arguments.

Fox (1987) showed that, in natural English discourse, intransitive subjects and patients were more likely to be preferred as the head noun of a RC than agents. According to Fox (1987), this phenomenon is explained by the Absolutive Hypothesis, which is based on Du Bois' (1981a; 1981b; 1985; 1987; 2003) Preferred Argument Structure Hypothesis and claims that absolutive arguments, such as intransitive subjects and patients, are preferred as RC head nouns. Du Bois (2003) examined core arguments in spoken discourse across multiple languages. Lexical arguments were more frequent in intransitive subjects and patients than in agents. By contrast,

agents were more likely to appear as pronouns or to be omitted, rather than to be realized as lexical arguments. Based on these findings, the Absolutive Hypothesis posits that "a language must be able to relativize on S and P if it has a strategy for relativization at all"[1] (Fox 1987: 869). The hypothesis states that absolutive arguments, such as intransitive subjects and patients, occupy the leftmost position in the hierarchy, which is distinct from the subject position defined by the NPAH.

Similarly, Hogbin & Song (2007) analyzed written English narrative data from the 18th and 20th centuries. Their study confirmed Fox's (1987) findings, revealing that the frequency of head nouns in RCs followed the order: intransitive subject/patient > oblique > agent > genitive > indirect object/object of comparison (Hogbin & Song 2007: 229).

While Fox (1987) and Hogbin & Song (2007) provided meaningful insights into the hierarchy of RCs based on predicate arguments rather than grammatical roles, they did not directly examine the development of RCs. In contrast, the study by Diessel & Tomasello (2005) is more relevant to RC development, as it explored how children acquire RCs based on predicate arguments. A sentence repetition task involving RCs was conducted with four-year-old children who were native speakers of English and German. The error rates followed the order: intransitive subject < agent < patient, indirect object, oblique < genitive. Diessel & Tomasello (2005) accounted for the lower error rates in intransitive subjects and agents from a syntactic perspective. As shown in (5), agent RCs were structurally similar to simple, non-embedded sentences, which made them easier to process.

(5) a. [The man]$_{NP1}$ saw [the boy]$_{NP2}$. (simple sentence)
b. [The man]$_{NP1}$ who saw [the boy]$_{NP2}$ (agent RC)
c. [The man]$_{NP1}$ who [the boy]$_{NP2}$ saw (patient RC)

(Diessel & Tomasello 2005: 901)

Among intransitive subjects and patients, intransitive subjects showed fewer errors. Because intransitive subjects require only one argument, whereas patients require at least two, the latter are conceptually more complex.

RC development has been widely discussed and explained from a variety of perspectives. In particular, drawing on the work of Fox (1987), Hogbin & Song (2007), and Diessel & Tomasello (2005), it is important to examine whether a detailed analysis of predicate argument structure can provide a new explanation for RC development. Fox (1987), Hogbin & Song (2007), and Diessel & Tomasello (2005) studied RC development and use by native speakers. It would be valuable to explore how Chinese-speaking learners of Korean as a second language develop RCs through an analysis of predicate argument structure.

[1] S refers to intransitive subjects, and P to patients of transitive verbs

2.3 Research on RC development in Korean language learners

Similar to other studies on prenominal languages, studies on RC development in Korean language learners have yielded inconsistent results about whether RC acquisition follows the Markedness Hypothesis. O'Grady et al. (2003), Lee (2005), Jeon & Kim (2007), and Kweon & Lee (2008) reported that RC development follows the Markedness Hypothesis. O'Grady et al. (2003) investigated whether learners preferred subject RC to direct object RC. In total, 53 learners were asked to listen to sentences and select pictures. The results indicated that they preferred the subject RC. Furthermore, Lee (2005) asserted that more subject RCs than direct object RCs or oblique RCs were found in the writing data of 17 learners, supporting the Markedness Hypothesis. Jeon & Kim (2007) examined oral production data from 40 learners, revealing that subject RC was preferred over direct object RC and used more precisely. In addition, Kweon & Lee (2008) argued that their online processing experiment conducted on learners demonstrated that subject RC was easier than direct and indirect object RCs, supporting the Markedness Hypothesis.

Conversely, several studies have reported that the order of learners' RC development is inconsistent with the Markedness Hypothesis (Kim 2010; Cho 2012; Han 2015; Kim & Kim 2016; Lee & Choi 2020). Kim (2010) reported that the frequency of learners' spoken and written language data often appeared in the order of subject RC > direct object RC > oblique RC, which seemingly matched the NPAH. However, an experiment assessing the difficulty through sentence combination production tasks did not yield statistically significant results indicating that subject RC was easier than direct object RC. Moreover, in picture comprehension tasks, learners understood oblique RC more accurately and found subject RC and direct object RC difficult. Thus, the order was subject RC > direct object RC > oblique RC in the corpus; direct object RC = subject RC > oblique RC in the sentence combination production tasks; and oblique RC > direct object RC = subject RC in the picture comprehension tasks. Furthermore, Cho (2012) conducted a grammatical judgment experiment with Chinese-speaking intermediate and advanced-level learners. The study reported that because the differences in subject RC, direct object RC, and oblique RC types were not statistically significant, the results did not support the Markedness Hypothesis. In addition, Han (2015) analyzed a corpus of beginner to advanced-level learners and found that subject RC dominated direct object RC. Nonetheless, direct object RC was more dominant than subject RC for post-first-grade learners at the beginner's level, contrary to the predictions of the Markedness Hypothesis. Kim & Kim (2016) examined colloquial and written language output data of native English, Japanese, and Chinese speakers. They found that oblique RC occurred more frequently than or at a similar rate to direct object RC, which differed partially from the NPAH. In an analysis of a free speech corpus of beginner learners, Lee & Choi (2020) found that subject RC appeared first and was associated with the highest frequency and a significantly lower error rate compared to the other types, whereas direct object RC and oblique RC were similar in frequency and error rate, concluding that the order was subject RC > direct object RC = oblique RC.

These findings suggest that the Markedness Hypothesis is insufficient for fully explaining the direct object RC and oblique RC hierarchy in Korean. Studies supporting the Markedness Hypothesis (O'Grady et al. 2003; Lee 2005; Jeon & Kim 2007; Kweon & Lee 2008) have found that subject RC is more easily learned than direct object RC; however, they could not determine that direct object RC is easier than oblique RC. Furthermore, discussions from the psycholinguistic perspective have failed to explain RC development in Korean. No clear evidence shows that subject RC is acquired earlier or with greater accuracy than direct object RC, as suggested by Gibson & Wu (2013). Moreover, there is no clear evidence that direct object RC develops more easily than subject RC in Korean, as argued by Tarallo & Myhill (1983), Hawkins (1989), and Ishizuka et al. (2006). Hence, it is necessary to consider and analyze the RC predicates and their arguments to identify whether a predicate argument-based explanation can account for RC development in Korean.

To date, a variety of research methods have been employed to study the RC acquisition of Korean language learners. Experimental studies have assessed learners' comprehension and production through sentence comprehension tasks (O'Grady et al. 2003; Kweon & Lee 2008; Kim 2010), sentence combination production tasks (Kim 2010), picture-based production tasks (Jeon & Kim 2007), and grammaticality judgment tasks (Cho 2012). In addition, several studies have analyzed learner-produced corpus data (Lee 2005; Kim 2010; Han 2015; Kim & Kim 2016; Lee & Choi 2020). Methodologically, learner corpus data are valuable because they reflect learners' actual use of RCs in communicative contexts. However, previous corpus-based studies on Korean have rarely addressed how learners' developmental patterns evolve over time. Therefore, this study investigates RC development by analyzing a longitudinal oral corpus.

2.4 Arguments of the predicate and the RC

Fox (1987), Hogbin & Song (2007), and Diessel & Tomasello (2005) classified RC predicates according to whether they were intransitive or transitive verbs and raised concerns about the predicate argument structure in RC research. The argument of a predicate is the type of sentence constituent required by the predicate. Depending on the required arguments, Korean predicates can be divided into four types: intransitive verbs without an obligatory oblique argument (INTR1), transitive verbs (TRAN), intransitive verbs with an obligatory oblique argument (INTR2), and ditransitive verbs (DTRAN). For example, the intransitive verb *wuntonghata* (to exercise) requires only one subject as an obligatory argument. Some intransitive verbs require two arguments: a subject and an oblique noun phrase. For example, the intransitive verb *tanita* (to attend) requires a subject and an oblique noun phrase to indicate the destination. A transitive verb, such as *sata* (to buy), requires two arguments: a subject and an object. A ditransitive verb, such as *pillyecwuta* (to lend), requires three arguments: a subject, an object, and an oblique.[2]

[2] Among the arguments of a ditransitive, the sentence constituent that corresponds to the indirect object in English is realized as an oblique in Korean.

(6) a. *Na-nun kongwen-eyse wuntongha-nun Chelswu-lul manna-ss-ta.* (subject RC)
I-TOP park-LOC exercise-AE Chelswu (HN)-ACC meet-PST-DECL
'I met Chulsoo who was exercising in the park.'

b. *Na-nun Chelswu-ka wuntongha-nun kongwen-ey ka-ss-ta.* (oblique RC)
I-TOP Chelswu-NOM exercise-AE park (HN)-LOC go-PST-DECL
'I went to the park where Chul-soo exercises.'

(7) a. *Sicang-eyse kwail-ul sa-n wuli-nun cip-ulo tolawa-ss-ta.* (subject RC)
Market-LOC fruit-ACC buy-AE we (HN)-TOP home-LOC come back-PST-DECL
'We who bought fruit at the market came back home.'

b. *Wuli-ka sicang-eyse san kwail-un maywu singsinghay-ss-ta.* (direct object RC)
We-NOM market-LOC buy-AE fruit (HN)-TOP very fresh-PST-DECL
'The fruits we bought at the market were very fresh.'

c. *Wuli-ka kwail-ul san sicang-un maywu pwumpy-ess-ta.* (oblique RC)
We-NOM fruit-ACC buy-AE market (HN)-TOP very crowded-PST-DECL
'The market where we bought fruits was very crowded.'

(8) a. *Hakkyo-ey tani-nun tongsayng-un kathi ka-ci*
School-DAT attend-AE younger brother (HN)-TOP together go-CONN
moshay-ss-ta. (subject RC)
cannot-PST-DECL
'My younger brother who attended school could not go together.'

b. *tongsayng-i tani-nun hakkyo-nun cokum mel-ta.* (oblique RC)
younger brother-NOM attend-AE school (HN)-TOP a little far-DECL'
'The school my brother attends is a little far away.'

(9) a. *Cinkwu-eykey wusan-ul pillyecwu-n kunun kenmwul-lo*
Friend-DAT umbrella-ACC lend-AE he (HN)-TOP building-LOC
tuleka-ss-ta. (subject RC)
enter-PST-DECL
'He who lent an umbrella to his friend entered the building.'

b. *Ku-ka chinkwu-eykey pillyecwu-n wusan-ul*
HE-NOM friend-DAT lend-AE umbrella (HN)-ACC
tollyepata-ss-ta. (direct object RC)
get back-PST-DECL
'He got back the umbrella that he lent his friend.'

c. *Ku-ka wusan-ul pillyecwu-n chinkwu-ka nwukwu-inci*
HE-NOM umbrella-ACC lend-AE friend (HN)-NOM who-CONN
molun-ta. (oblique RC)
do not know-DECL
'I don't know his friend to whom he lent his umbrella.'

In (6), *wuntonghata* (to exercise), an INTR1 predicate, requires only a subject as an obligatory argument. An RC with *wuntonghata* (to exercise) as its predicate can thus be a subject RC. As shown in (6b), *wuntonghata* (to exercise) can take an oblique as an optional argument and thus form a direct object RC. Conversely, *sata* (to buy), a TRAN predicate in (7), requires both a subject and an object as obligatory arguments. An RC involving *sata* (to buy) as its predicate can be of the subject, direct object, or oblique RC. The predicate *tanita* (to attend) in (8), an INTR2 predicate, requires both a subject and an oblique argument; hence, an RC containing it can be either a subject or an oblique RC. In (9), *pillyecwuta* (to lend), a DTRAN predicate, requires a subject, an object, and an oblique as obligatory arguments. As a result, it can yield subject, direct object, and oblique RCs.

The RC type realized in this manner may vary depending on the arguments required by the RC predicates (**Table 1**). Thus, the arguments of the RC predicate must be considered when analyzing clause types in corpus data.

Predicate type		**INTR1**	**TRAN**	**INTR2**	**DTRAN**
Example		*wuntonghata* (to exercise)	*sata* (to buy)	*tanita* (to attend)	*pillyecwuta* (to lend)
Obligatory argument		subject	subject, object	subject, oblique	subject, object, oblique
Possible RC types	subject RC	✓	✓	✓	✓
	direct object RC	–	✓	–	✓
	oblique RC	✓	✓	✓	✓

Table 1: Predicate Types and their possible RC realizations.

3 Research method

3.1 Research questions

This study investigated the following research questions by examining the RC development patterns of Chinese-speaking learners of Korean:

RQ1) Can the development of RCs based on head noun types in learners' spontaneous speech data be explained by linguistic typological or psycholinguistic hypotheses?
RQ2) Can the development of RCs based on predicate types in learners' spontaneous speech data be explained by the Absolutive Hypothesis or syntactic complexity?

As discussed in Section 2.3, unlike studies on postnominal RCs, the literature on Korean, a language with prenominal RCs, has not yielded similar findings. Thus, it is meaningful to examine whether the development of RCs in Korean language learners can be adequately explained by

the Markedness Hypothesis from a linguistic typological perspective or the Surface Configuration Hypothesis and the Dependency Locality Theory from a psycholinguistic perspective. Therefore, to answer RQ1, the frequency of occurrence, timing of initial emergence, and syntactic error rate over time were examined to understand learners' RC development.

Section 2.4 discussed how the RC type may vary depending on the RC predicate. Therefore, RC analysis, based on predicate type, can confirm whether RC development can be explained by the Absolutive Hypothesis, which is grounded in the Preferred Argument Structure Hypothesis, or by syntactic complexity. Thus, to answer RQ2, RC predicate types were classified according to their argument structure, and RC development was examined based on these classifications.

3.2 Research data

In this study, three Chinese-speaking learners of Korean were surveyed, and spontaneous conversational speech was collected across 19 sessions over an 18-month period, between October 2018 and March 2020. The participants were women in their 20s and native Chinese speakers. They began studying Korean at a Korean language education institution affiliated with a university in Seoul in September 2018. All three participants were assigned to the lowest level of Korean (Level 1). They participated in this research two months after they started learning the language. They used the *Fun Fun Korean* series as their regular course textbook, and their experience with other L2 languages was not controlled. In the curriculum, RC was taught four months after they began learning the language.

Researchers collected participants' speech through one-on-one meetings held once or twice a month, during which they engaged in spontaneous conversation. Each conversation lasted approximately 60 minutes, and the topics ranged from daily life and school-related experiences to past experiences, hobbies, and jobs.[3] Transcripts, automatic morpheme analysis, and post-processing were used to compile the collected speech data into a longitudinal corpus containing 93,471 sentences produced by the participants (**Table 2**).

	2–3 months	**4–6 months**	**7–9 months**	**10–12 months**	**13–15 months**	**16+ months**	**Total**
Participant 1	5,393	5,877	6,357	6,757	7,148	6,824	38,356
Participant 2	2,315	2,317	4,013	3,897	2,766	3,271	18,579
Participant 3	4,421	5,425	7,258	7,240	7,116	5,076	36,536
Total	12,129	13,619	17,628	17,894	17,030	15,171	93,471

Table 2: Number of sentences produced by each learner across periods.

[3] These sessions were conducted solely for research purposes, independent of the coursework or academic credit offered by the educational institution. The researchers and participants interacted on an equal footing, rather than in a hierarchical instructor-student relationship. Informed consent was obtained regarding participation and recording, and participants were compensated approximately $24 per session.

The researchers extracted RCs from the learner longitudinal corpus using AntConc,[4] identifying 3,376 clauses marked with the adnominal endings *-nun* and *-un*. However, these clauses were extracted mechanically using AntConc, without checking the context of the utterances, as in some cases it was unclear whether the participants intended to produce RCs. Therefore, the researchers manually reviewed the data and selected participants' RC utterances by excluding 2,235 cases from the initially extracted 3,376 clauses, some of which are exemplified in **Table 3**.

Criteria for exclusion	**Cases**
Mimicking a teacher's utterance	T: *Kal swu issn-un kos-i manhi iss-eyo.* Can go-AE place (HN)-NOM many be-DECL 'There are many places to go.' Participant 1: *Kal swu issn-un kos-i?* Can go-AE places (HN)-NOM 'Places to go?' [Participant 1, 3 months]
Repeating and modifying self-utterance	Participant 1: *Ce-nun cwungkwuk-eyse ama coh-un* I-TOP China-LOC maybe good-AE *tayhakkyo-ey ama coh-un tayhakkyo-ey* university-DAT maybe good-AE university (HN)-DAT *coh-un tayhakkyo-lul colephay-se animyen…* good-AE university-ACC graduate-CONN or 'I graduated from a good university in China maybe, maybe a good university, good university or…' [Participant 1, 11 months]
Appositive clause	Participant 1: *myengcel-ey cokum ponay-ko cip-eyse* holidays-TEMP some spend-CONN home-LOC *kongpwuha-l kyeyhoyk-i iss-nuntey cikum-un…* study-AE plan (HN)-NOM have-CONN now-TOP 'I have plans to spend some time during the holidays and study at home, but now…' [Participant 1, 17 months]
Noun clause	Participant 3: *Kuliko e pokose cal calha-ciman palum* And uh report well good-CONN pronunciation *e malha-nun ke te yensuphay-ya ha-ko…* uh talk-AE thing (HN) more practice-CONN must-CONJ… 'And I'm good at writing reports, but I have to practice pronunciation and speaking more…' [Participant 3, 12 months]

(Contd.)

[4] AntConc is a freeware corpus analysis toolkit for text analysis, created by Laurence Anthony. It is useful for lexical and grammatical studies, as researchers can obtain frequency data and contextual information on relevant words by inputting corpus data and entering keywords or parts of speech.

Criteria for exclusion	**Cases**
Chunk expression	Participant 1: *Ku chinkwu-nun hankwuk hankwuk-ey on-ci* The friend-TOP Korea Korea-LOC come-CONN *sa nyen sa nyenli sa nyen-i toye-ss-eyo.* 4 year 4 year-? 4 year-NOM have been-PST-DECL 'It's been 4 years since the friend came to Korea.' [Participant 1, 7 months]
Noun phrase	Participant 2: *Um ci cina-n tal e cina-n tal* Um ? pass-AE month (HN) uh pass-AE month (HN) *um cina-n pen-ey po-nun TOPIK-to…* um pass-AE time (HN) take-AE TOPIK-also 'Last month, last month, last TOPIK test, also…' [Participant 2, 17 months]
Impossible to interpret	Participant 2: *E a sahoypat-un ney ttalumyen salamtul-uy* Uh ah socialize-AE yes according people-GEN *sungsan sungsan kwani-lul pakkwe-ss-eyo.* ? ? ? -ACC change-PST-DECL 'Uh, socialized, yes, according to that, people's ( ), ( ), the ( ) have changed.' [Participant 2, 16 months]
Incorrect analysis	Participant 3: *Kunyang kath-un e hal haswuksa haswuksa-eyse* Just same-AE uh ? dorm dorm-LOC *salamtul iyakihay-ss-eyo.* people talk-PST-DECL 'I just talked to people in the same dorm.' [Participant 3, 8 months]
Other	T: *Hankwuke kongpwuhay-ss-eyo?* Korean study-PST-Q 'Did you study Korean?' Participant 3: *Wuli chayk wuli chayk 'caymiiss-nun hankwuke'.* Our book our book Fun-AE Korean (HN) 'Our book Our book 'Fun Korean'' [Participant 3, 3 months]

Table 3: Criteria and cases for excluding the relative clause utterance.

Cases in which the participants imitated another person's speech or repeated or modified their speech during the conversation were excluded. Appositive clauses were excluded because they express the same meaning as the head noun. They were excluded because their structure differs from that of RCs, in which the head noun appears outside the clause.[5] In addition, noun clauses, such as *malha-nun ke* (speaking), were excluded because they are action noun constructions formed by combining the adnominalized verb form *malha-nun* with the dependent noun *ke* (thing).[6] Chunk expressions are combinations of adnominal endings and dependent nouns, such as *-on ci* and *-ul ttay*. Such cases were excluded because they differ from RCs in both structure and meaning. Other excluded cases included utterances that were uninterpretable due to learner errors, misanalyses by AntConc's parser, and instances in which an RC formed part of a proper noun, such as the book title *Fun Korean*.

In addition, the researchers manually examined the transcribed data and identified 37 additional clauses that appeared to be intended as RCs but were not detected by AntConc due to the absence of adnominal endings. A total of 1,178 RCs were selected for analysis.

---

[5] Scholars—for example, Matsumoto (1990: 117–118)—regarded RCs as adnominal clauses that semantically define the head noun and considered appositive clauses to be included in the RC. As shown in (iib), appositive clauses in Korean (e.g., "a plan to study at home") are structurally and semantically distinct from typical RCs, such as those in (ib) (e.g., "a plan that I made"). In the former, the head noun does not correspond to any syntactic element within the clause but rather represents the entire proposition or action expressed by the clause. Therefore, such appositive clauses cannot be analyzed using frameworks from previous RC studies, which rely on identifying the syntactic role (e.g., subject, object) of the head noun within the clause. Accordingly, these clauses were treated as a separate type and excluded from the analysis in this study.

(i) a. *Nay-ka kyeyhoyk-ul seywe-ss-ta.* (simple sentence)
I-NOM plan-ACC make-PST-DECL
'I made a plan.'

b. Nay-ka seywu-n kyeyhoyk (direct object RC)
I-NOM make-AE plan (HN)
'A plan I made.'

(ii) a. *Kyeyhoyk-un (nay-ka) cip-eyse kongpwuha-nun kes-ita.* (simple sentence)
Plan-TOP (I-NOM) home-LOC study-AE thing-PRED
'The plan is to study at home.'

b. (*nay-ka) cip-eyse kongpwuha-nun kyeyhoyk*
(I-NOM) home-LOC study-AE plan (HN)
'A plan to study at home.'

[6] The noun clause corresponds to the internal RC of Jeon & Kim (2007). Jeon & Kim (2007) divided Korean relative clauses into head-external RCs and head-internal RCs. Among these, the head-external type corresponds to the RCs examined in this study, while the head-internal type corresponds to what is referred to here as noun clauses. According to Jeon & Kim (2007: 253), the head-internal type "has its lexical head in the RC and is marked by the complementizer 'kes'." However, it is difficult to identify the lexical head, and since these clauses function as nouns, they are referred to as noun clauses in this study.

3.3 Analysis methodology

For RQ1, the authors classified the head noun into three categories—subject, direct object, and oblique—based on its syntactic role within the RC.[7] The following are examples of learner utterances classified as a subject RC, a direct object RC, and an oblique RC.[8]

(10) a. *Ku um Hankwuke calha-nun Cwungkwuk salam-un chepen ttay hankwuk-ey*
That um Korean be good at-AE Chinese (HN)-TOP last time Korea-LOC
wase. (subject RC)
come-CONJ
'A Chinese person who is good at Korean came to Korea last time and…'
[Participant 3, 7 months]

b. *Cey ceyil cohaha-nun aitol nwu, nwukwu-ya?* (direct object RC)
? most like-AE idol (HN) ? who-Q
'Who's your favorite idol?'

c. *Kuntey yocum yensup yensupha-nun sikan ep-sese um…* (oblique RC)
But these days practice practice-AE time (HN) not have-CONJ um
'But I don't have time to practice these days, um…'
[Participant 3, 7 months]

---

[7] This study classified the syntactic roles of head nouns within RCs into subject, direct object, and oblique. In Korean, as shown in example (9), there is no sentence component corresponding to the English indirect object, and such elements are therefore classified as oblique. In Korean, genitive expressions may or may not undergo adnominalization, as shown in (i) and (ii). Adnominalization is limited to possessions that are closely associated with the owner or the owner's body. Therefore, genitive RCs cannot be generalized in Korean. Moreover, objects of comparison cannot be adnominalized in Korean, as shown in (iii). For these reasons, Jeon & Kim (2007), Kim (2010), Cho (2012), and Lee & Choi (2020) classified Korean RCs into subject, direct object, and oblique.

(i) a. *nay-ka Jiho-uy son-ul cap-ass-ta.*
I-NOM Jiho-GEN hand-ACC hold-PST-DECL
'I held Jiho's hand.'

b. *nay-ka son-ul cap-un Jiho* (genitive RC)
I-NOM hand-ACC hold-AE Jiho (HN)
'Jiho whose hand I held.'

(ii) a. *nay-ka Jiho-uy cha-lul wuncenhay-ss-ta.*
I-NOM Jiho-GEN car-ACC drive-PST-DECL
'I drove Jiho's car.'

b. **nay-ka cha-lul wuncenha-n Jiho* (genitive RC)
I-NOM car-ACC drive-AE Jiho (HN)

(iii) a. *nay-ka Jiho-pota khu-ta.*
I-NOM Jiho-OCOMP tall-DECL
'I'm taller than Jiho.'

b. **nay-ka khu-n jiho* (object of comparison RC)
I-NOM tall-AE Jiho (HN)

[8] *Calhata* (to be good at) in (10a), *cohahata* (to like) in (10b), and *yensuphata* (to practice) in (10c) all belong to the predicate type TRAN in Korean.

After classifying the RCs based on the syntactic roles of the head nouns, the frequency of occurrence and the timing of initial emergence of the RCs across the learning period was analyzed. Moreover, RCs containing syntactic errors were examined to calculate the error rate. Examples of syntactic errors in RCs include the following:

(11) a. *Um wuli yumyeng siktang-ey ka ka-yo.*
Um we famousness restaurant (HN)-LOC ? go-REQ
'Well, let's go to a *famousness* restaurant.'
[Participant 1, 3 months]

b. *Chenchenhi tulama silh-eyo.*
Slowly dramas (HN) not like-DECL
'I don't like *slowly* dramas.'
[Participant 1, 6 months]

c. *Pisushay-yo tosi iss-eyo.*
Similar-DECL city (HN) be-DECL
'*Be similar* there's a city.'
[Participant 1, 4 months]

Syntactic errors in RCs occur when the root form of a word or an oblique is placed in front of the head noun instead of the RC predicate, as shown in examples (11a) and (11b). Another common error involves using a sentence-final ending rather than an adnominal ending before the head noun, as illustrated in example (11c).[9] These errors violate the syntactic rules that govern RCs in Korean, where the RC must precede the head noun and the adnominal ending must be attached to the RC predicate. Syntactic errors of these types were identified, and both the overall error rate and the error rate by period were examined.

To address RQ2, the predicates in RCs were classified into four types, as shown in **Table 4**.

In addition, the frequency of occurrence, the timing of initial emergence, and the period-specific frequency for subject, direct object, and oblique RC types based on the predicates were analyzed.

---

[9] Appropriate target structures for the errors in (11a), (11b), and (11c) are given in (i).

(i) a. *yumyengha-n siktang*
famous-AE restaurant (HN)
'a famous restaurant'

b. *chenchenhi cenkaytoy-nun tulama*
slowly unfold-AE drama (HN)
'a slow-paced drama'

c. *pisusha-n tosi*
similar-AE city (HN)
'a similar city'

To ensure validity, the authors conducted both an independent analysis and a cross-examination. Of the 1,178 sentences, 139 (11.8%) were under discussion. When disagreements arose during cross-examination, the researchers resolved them through discussion to unify the analysis criteria.[10]

Predicate Types	Examples of Predicate
INTR1	*kakkapta* (to close), *kanunghata* (to be possible), *kantanhata* (to be simple), *kathta* (to be same), *kilta* (to be long), *napputa* (to be bad), *nophta* (to be high), *tayanghata* (to be diverse) etc.
TRAN	*kacita* (to have), *kekcenghata* (to worry), *kongpwuhata* (to study), *tamtanghata* (to take charge), *taysinhata* (to replace), *mannata* (to meet), *mantulta* (to make), *malhata* (to say) etc.
INTR2	*kata* (to go), *kyelhonhata* (to marry), *naota* (to come out), *tanita* (to go around), *tolakata* (to go back), *tulekata* (to go in), *tuleota* (to come in), etc.
DTRAN	*kaluchita* (to teach), *nehta* (to insert), *tamta* (to put), *patta* (to receive), *mwulepota* (to ask), *ponayta* (to send), *iyakihata* (to talk), *cwuta* (to give) etc.

Table 4: Classification of predicate types.

4 Results

4.1 RC development based on head noun types

Table 5 illustrates how head nouns are classified according to their syntactic roles in RCs produced by the participants in colloquial speech. Of the 1,178 clauses, 860 were subjects, 224 were direct objects, and 94 were obliques. Subjects accounted for the largest proportion (73.0%) of all head nouns, followed by direct objects (19.0%) and obliques (8.0%).

[10] A representative example of a predicate that required cross-checking is *kathta* (to be same). There has been a discussion about whether this predicate belongs to the INTR1 type, which requires only a subject, or the INTR2 type, which requires a subject and an oblique. In the simple sentence in (ia), *kathta* requires two arguments to be compared, one of which is realized as the subject and the other as the oblique. However, as shown in (ic), the RC in which the oblique argument serves as the head noun is ungrammatical in Korean, and an RC with this predicate can appear only as the subject type. Therefore, we classified it as an INTR1 predicate.

(i) a. simple sentence: *Na-nun ku-wa khi-ka kath-ta.*
I-TOP he-COM height-NOM same-DECL
'I am the same height as him.'

b. subject RC: *ku-wa khi-ka kath-un na.*
He-COM height-NOM same-AE me (HN)
'He who is the same height as me.'

c. oblique RC: **na-nun khi-ka kath-un ku*
me-TOP height-NOM same-AE he (HN)
'Me who he is the same height as.'

d. subject RC: *na-wa khi-ka kath-un ku*
me-COM height-NOM same-AE him (HN)
'I who is the same height as him.'

Head noun	subject	direct object	oblique	Total
Participant 1	379 (80.5%)	64 (13.6%)	28 (13.3%)	471 (100%)
Participant 2	174 (80.9%)	27 (12.6%)	14 (6.5%)	215 (100%)
Participant 3	307 (62.4%)	133 (27.0%)	52 (10.6%)	492 (100%)
Total	**860 (73.0%)**	**224 (19.0%)**	**94 (8.0%)**	**1178 (100%)**

Table 5: Distribution of RC head noun types.

To test whether the frequency distribution of RC types by the syntactic roles of head nouns was uniform, a chi-square goodness-of-fit test was conducted using SPSS (version 21.0). Expected frequencies were set equal across the three categories (1:1:1). The results revealed a statistically significant difference in distribution, and the null hypothesis of equal proportions was rejected, $\chi^2(2) = 822.8, p < .001$. The significance level was set at $\alpha = .05$.

Table 6 shows the frequency of head noun types across the learning period. The analysis was conducted by dividing the 18-month learning period into 3-month intervals.

Head noun	2–3 months	4–6 months	7–9 months	10–12 months	13–15 months	16+ months	Total
subject	30 (100.0%)	60 (78.9%)	114 (68.2%)	270 (75.6%)	223 (70.1%)	163 (70.9%)	860 (73.0%)
direct object	0 (0.0%)	10 (13.2%)	32 (19.2%)	59 (16.5%)	74 (23.3%)	49 (21.3%)	224 (19.0%)
oblique	0 (0.0%)	6 (7.9%)	21 (12.6%)	28 (7.8%)	21 (6.6%)	18 (7.8%)	94 (8.0%)
Total	30 (100.0%)	76 (100.0%)	167 (100.0%)	357 (100.0%)	318 (100.0%)	230 (100.0%)	1178 (100.0%)

Table 6: Frequency and percentage of head noun types by period.

Within the first 2–3 months of learning, 30 RCs were produced, all of which involved subject types.[11] Subject RCs first appeared at 2 months, as observed in all three participants. By 4–6 months, other RC types began to emerge, including direct object RCs (13.2%) and oblique RCs (7.9%). Direct object RCs first appeared at 5 months for Participants 1 and 3, and at 8 months for Participant 2. Oblique RCs also appeared at 5 months for Participants 1 and 3, and at 9 months for Participant 2.

[11] As noted in Section 3.2, participants were taught how to produce RCs during the fourth month. However, RCs had already been used 30 times during the 2–3 month period, indicating that learners attempted to use RCs prior to receiving explicit instruction in the classroom.

For all participants, subject RCs emerged first, followed by direct object and oblique RCs. The latter two appeared at similar times for Participants 1 and 3, but with a slight delay for Participant 2.

Figure 1 presents the change in the proportion of subject, direct object, and oblique RCs over time, corresponding to **Table 6**. As the number of sessions increases, the proportion of subject RCs decreases, that of direct object RCs gradually increases, and that of oblique RCs increases and then decreases.

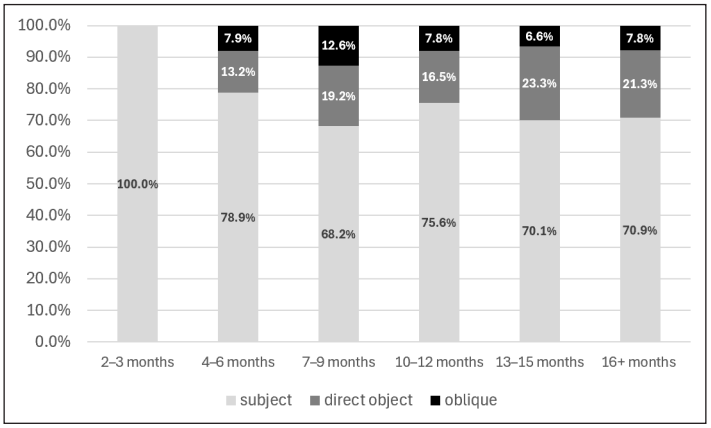

Figure 1: Changes in the percentage of head noun types by period.

Table 7 presents the number and percentage of syntactic errors in participants' RC production. A total of 35 syntactic errors (3.0%) were found among the 1,178 RCs.[12] When categorized by head noun type, 32 errors occurred in subject RCs, accounting for 3.7% of all subject RCs. Three errors were found in direct object RCs, and no errors occurred in oblique RCs.[13]

Head noun	subject	direct object	oblique	Total
All	860	224	94	1178
Errors	32	3	0	35
Error rate	3.7%	1.3%	0%	3.0%

Table 7: Syntactic errors and error rates by head noun types.

[12] A total of 146 errors (12.4%) were found among the 1,178 RCs. Of these, vocabulary errors accounted for the largest proportion (54 cases, 4.6%). There were 43 form errors (3.7%), 35 syntactic errors (3.0%), and 14 pronunciation errors (1.2%). Among these, only syntactic errors, which are directly relevant to the purpose of this study, are reported.

[13] Examples of syntactic errors in subject and direct object RCs are provided below, with the target forms shown in parentheses.

(i) a. Subject RC: *mayw-e umsik (maywu-n umsik)*
be spicy-DECL food (HN) be spicy-AE food (HN)
'a spicy food'

b. Direct object RC: *hwuhoy il (hwuhoyha-nun il)*
regret thing (HN) regret-AE thing (HN)
'something to regret'

Table 8 displays the number and proportion of syntactic errors by learning period. Most of the 35 errors occurred in the early stages of learning, specifically in the 2–3 and 4–6 month periods, with 18 and 13 errors, respectively. Only two errors were found in both the 7–9 and the 10–12 month periods, and the number of errors declined substantially after 7 months. No subject RC errors were observed from 13–15 months onward, and no direct object RC errors occurred from 10–12 months onward.

		2–3 months	**4–6 months**	**7–9 months**	**10–12 months**	**13–15 months**	**16+ months**	**Total**
subject	All	30	60	114	270	223	163	860
	Errors	18	12	0	2	0	0	32
	Error rate	60.0%	20.0%	0.0%	0.8%	**0.0%**	**0.0%**	3.7%
direct object	All	–	10	32	59	74	49	224
	Errors		1	2	0	0	0	3
	Error rate		10.0%	4.6%	**0.0%**	**0.0%**	**0.0%**	1.3%
oblique	All	–	6	21	28	21	18	94
	Errors		0	0	0	0	0	0
	Error rate		**0.0%**	**0.0%**	**0.0%**	**0.0%**	**0.0%**	0.0%
Total	All	30	76	167	357	318	230	1178
	Errors	18	13	2	2	0	0	35
	Error rate	60.0%	17.1%	1.2%	0.6%	**0.0%**	**0.0%**	3.0%

Table 8: Syntactic error rates of head noun types by period.
Note. Dashes (–) indicate that no RCs of that type were produced during the corresponding period.

4.2 RC development based on predicate types

4.2.1 Development of the RC according to the predicate type

Table 9 shows that INTR1 predicates accounted for 60.2% of all RCs, followed by TRAN (26.9%), INTR2 (9.4%), and DTRAN (3.5%). A chi-square goodness-of-fit test was conducted using SPSS (version 21.0) to examine whether the frequency distribution of RCs—INTR1, TRAN, INTR2, and DTRAN—was uniform. The total number of RCs was 1,178, with observed frequencies of 709 for INTR1, 317 for TRAN, 111 for INTR2, and 41 for DTRAN. Expected frequencies were set equally across the four categories (1:1:1:1). The analysis revealed a statistically significant deviation from uniformity, and the null hypothesis was rejected, $\chi^2(3) = 917.660, p < .001$. The significance level was set at $\alpha = .05$.

Predicate	INTR1	TRAN	INTR2	DTRAN	Total
Participant 1	301 (63.9%)	113 (24.0%)	47 (10.0%)	10 (2.1%)	471 (100%)
Participant 2	146 (67.9%)	42 (19.5%)	24 (11.2%)	3 (1.4%)	215 (100%)
Participant 3	262 (53.3%)	162 (32.9%)	40 (8.1%)	28 (5.7%)	492 (100%)
Total	709 (60.2%)	317 (26.9%)	111 (9.4%)	41 (3.5%)	1178 (100%)

Table 9: Distribution of RC predicate types.

The next step was to examine how frequently each predicate type appeared across different learning periods (**Table 10** and **Figure 2**). During the 2–3 month period, all 30 RCs were of the INTR1 type. INTR1 predicates first emerged at 2 months for all participants. TRAN predicates appeared during the 4–6 month period (Participant 1: 4 months; Participant 2: 8 months; Participant 3: 5 months), as did INTR2 predicates (Participant 1: 5 months; Participant 2: 8 months; Participant 3: 4 months). DTRAN predicates emerged during the 7–9 month period (Participant 1: 8 months; Participant 2: 16 months; Participant 3: 7 months).

These results indicate that the participants tended to prefer predicates with fewer obligatory arguments during the early stages of RC acquisition. In the initial phase of learning Korean (2–3 months), they predominantly used INTR1 predicates. RC predicates with more complex argument structures, such as TRAN and INTR2, began to appear at 4–6 months, followed by DTRAN at 7–9 months.[14] This developmental progression is reflected in the timing of each predicate's first appearance for each participant. For Participant 1, INTR1 appeared at 2 months, TRAN at 4 months, INTR2 at 5 months, and DTRAN at 8 months. For Participant 2, INTR1 appeared at 2 months, TRAN and INTR2 at 8 months, and DTRAN at 16 months.[15] For Participant 3, INTR1 appeared at 2 months, INTR2 at 4 months, TRAN at 5 months, and DTRAN at 7 months. In all three cases, INTR1—the predicate type with the fewest obligatory arguments—emerged first, and DTRAN—the one with the most—appeared last. TRAN and INTR2 emerged around the same time, albeit in slightly different orders.

Table 10 and **Figure 2** present the frequency distribution of predicate types over time. INTR1 predicates, which accounted for 100% of RCs during the initial 2–3 months, declined

[14] Participants were introduced to all four predicate types (i.e., INTR1, TRAN, INTR2, and DTRAN) through Level 1 textbooks used in the first three months of Korean learning. When RCs were introduced in the fourth month, instruction likewise incorporated all four predicate types.

[15] Participant 2 exhibited a distinct overall pattern compared to the other two participants, with RCs occurring less frequently and being acquired later than in Participants 1 and 3.

to 72.4% at 4–6 months and further to 58.1% at 7–9 months, as the frequencies of TRAN and INTR2 predicates increased. DTRAN predicates, which first appeared at 7–9 months, gradually increased from the 10–12 month period onward.

Predicate	2–3 months	4–6 months	7–9 months	10–12 months	13–15 months	16+ months	Total
INTR1	30 (100.0%)	55 (72.4%)	97 (58.1%)	233 (65.3%)	161 (50.6%)	133 (57.8%)	709 (60.2%)
TRAN	0 (0.0%)	15 (19.7%)	49 (29.3%)	87 (24.4%)	111 (34.9%)	55 (23.9%)	317 (26.9%)
INTR2	0 (0.0%)	6 (7.9%)	11 (6.6%)	31 (8.7%)	35 (11.0%)	28 (12.2%)	111 (9.4%)
DTRAN	0 (0.0%)	0 (0.0%)	10 (6.0%)	6 (1.7%)	11 (3.5%)	14 (6.1%)	41 (3.5%)
Total	30 (100.0%)	76 (100.0%)	167 (100.0%)	357 (100.0%)	318 (100.0%)	230 (100.0%)	1178 (100.0%)

Table 10: Frequency and percentage of predicate types by period.

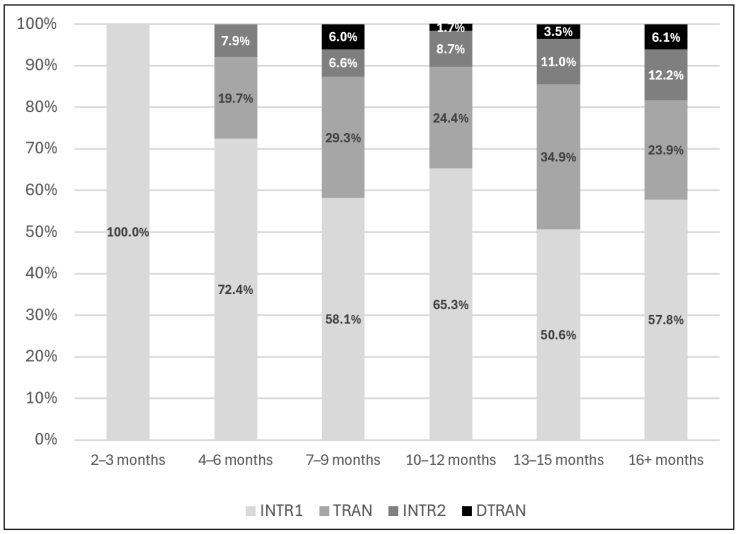

Figure 2: Changes in the percentage of predicate types by period.

INTR1 was predominantly used during the early stages of learning, although its accuracy was extremely low. Its error rate was 60.0% at 2–3 months, decreasing to 23.6% at 4–6 months. By contrast, TRAN emerged later, during the 4–6 month period, and exhibited a relatively low error rate from the outset. The error rates for INTR2 and DTRAN remained at 0.0% (**Table 11**).

	2–3 months		4–6 months	7–9 months	10–12 months	13–15 months	16+ months	Overall
INTR1	All	30	55	97	233	161	133	709
	Errors	18	13	0	0	0	0	31
	Error rate	60.0%	23.6%	**0.0%**	**0.0%**	**0.0%**	**0.0%**	4.4%
TRAN	All	–	15	49	87	111	55	317
	Errors		0	2	2	0	0	4
	Error rate		**0.0%**	4.1%	2.3%	**0.0%**	**0.0%**	1.3%
INTR2	All	–	6	11	31	35	28	111
	Errors		0	0	0	0	0	0
	Error rate		**0.0%**	**0.0%**	**0.0%**	**0.0%**	**0.0%**	**0.0%**
DTRAN	All	–	–	10	6	11	14	41
	Errors			0	0	0	0	0
	Error rate			**0.0%**	**0.0%**	**0.0%**	**0.0%**	**0.0%**
Total	All	30	76	167	357	318	230	1178
	Errors	18	13	2	2	0	0	35
	Error rate	60.0%	17.1%	1.2%	0.6%	**0.0%**	**0.0%**	3.0%

Table 11: Error rate of predicate types by period.
Note. Dashes (–) indicate that no RCs of that type were produced during the corresponding period.

4.2.2 Predicate types and head noun types

Table 12 shows how the distribution of head noun types varies across predicate types.

	Subject	direct object	oblique	Total
INTR1	699 (98.6%)	–	10 (1.4%)	709 (100.0%)
TRAN	81 (25.6%)	198 (62.5%)	38 (12.0%)	317 (100.0%)
INTR2	70 (63.1%)	–	41 (36.9%)	111 (100.0%)
DTRAN	10 (24.4%)	25 (61.0%)	6 (14.6%)	41 (100.0%)

Table 12: Frequency and percentage of head noun types by predicate type.
Note. Dashes (–) indicate unavailable structures for that predicate type.

Of the 709 total INTR1 cases, 699 (98.6%) involved subject head nouns, which constitute the obligatory argument for INTR1 predicates. A chi-square goodness-of-fit test comparing the frequencies of subject and oblique types revealed a statistically significant difference in

distribution ($\chi^2(1) = 669.564, p < .001$), with significance assessed at the $\alpha = .05$ level. Subject RCs with INTR1 predicates first appeared at 2 months (Participants 1, 2, and 3: 2 months) while oblique RCs first appeared at 8 months (Participant 1: 10 months; Participant 2: 12 months; Participant 3: 8 months).

In the TRAN types, the direct object accounted for a significantly higher proportion, with 198 cases (62.5%), followed by the subject with 81 cases (25.6%) and the oblique with 38 cases (12.0%). A chi-square goodness-of-fit test revealed that the observed frequencies for subject, direct object, and oblique types deviated significantly from an equal expected distribution ($\chi^2(2) = 129.773, p < .001$), with significance assessed at the $\alpha = .05$ level. In TRAN RCs, direct object types first appeared at 4 months (Participant 1: 4 months; Participant 2: 8 months; Participant 3: 5 months), followed by oblique types at 5 months (Participant 1: 5 months; Participant 2: 10 months; Participant 3: 5 months) and subject types at 7 months (Participant 1: 12 months; Participant 2: 9 months; Participant 3: 7 months). Overall, the order of first appearance was direct object, oblique, and then subject RCs.

In INTR2, subject RCs accounted for 63.1% and oblique RCs for 36.9% of cases. A chi-square test assuming equal expected frequencies revealed a statistically significant difference in their distribution ($\chi^2(1) = 7.577, p = .006$), with significance evaluated at the $\alpha = .05$ level. By comparison, the proportion of oblique RCs in INTR1—where the oblique is an optional element—was much lower, at 1.41%, while it was relatively high in INTR2. In INTR2 RCs, subject types first appeared at 4 months (Participant 1: 5 months; Participant 2: 8 months; Participant 3: 4 months), and oblique types at 6 months (Participant 1: 6 months; Participant 2: 10 months; Participant 3: 6 months).

Finally, 41 instances of DTRAN RCs were observed, representing a relatively small proportion of the total. Direct object RCs comprised 61.0% of these cases, followed by subject RCs (24.4%) and oblique RCs (14.6%). As with TRAN RCs, direct object was the most frequent head noun type in DTRAN. A chi-square analysis assuming an equal distribution across subject, direct object, and oblique types revealed a statistically significant difference ($\chi^2(2) = 14.683, p = .001$), with the alpha level set at .05. In DTRAN RCs, direct object types first appeared at 11 months (Participant 1: 14 months; Participant 2: no occurrence; Participant 3: 11 months), subject types at 9 months (Participant 1: 9 months; Participant 2: 17 months; Participant 3: 13 months), and oblique types also at 9 months (Participant 1: 9 months; Participant 2: 17 months; Participant 3: 14 months). Due to the variability in the timing of initial emergence across participants, no consistent order of appearance was observed.

The predicate type analysis showed that the subject was the most frequent grammatical role of the head noun in INTR1 and INTR2 RCs. In contrast, in TRAN and DTRAN RCs, the direct object appeared more frequently than the subject (**Figure 3**).

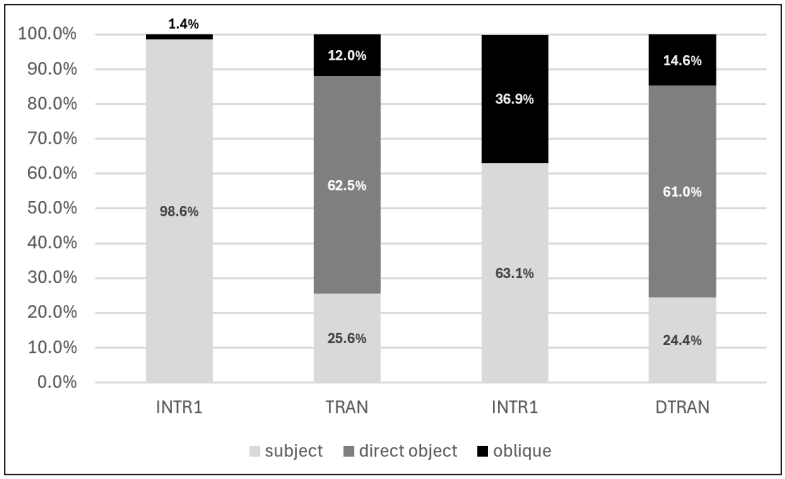

Figure 3: Proportion of head noun types by predicate type.

5 Discussion

5.1 RC development based on head noun types

This study examined the development of RCs in Chinese-speaking learners of Korean, focusing on how different types of head nouns emerged during RC development. Based on an analysis of 1,178 RCs, the RC types occurred in the following order of frequency: subject > direct object > oblique. In terms of timing, the subject type was the first to emerge, within 2–3 months of learning. Subsequently, within 4–6 months, the direct object and oblique types emerged. For the subject type, the syntactic error rate was high; however, after four months, the rate stabilized, and few errors were found across all RC types. The high error rate was due to a phenomenon observed in the early stages of RC learning, rather than to any inherent difficulty with the subject type. The emergence of direct object and oblique types around 4–6 months, followed by a sharp decline in errors, indicates the impact of instruction. Learners were introduced to various RC types around the fourth month; subsequently, types other than the subject emerged, and error rates decreased.

These results do not align with the Markedness Hypothesis, which predicts RC development from a typological perspective. While the overall frequency followed the order predicted by the NPAH (subject > direct object > oblique), the direct object and oblique RC types emerged at similar stages of development, and no difference was observed in their error rates. This finding is consistent with previous studies. Kim (2010), Cho (2012), Han (2015), and Lee & Choi (2020) also found the frequency order of subject > direct object > oblique; however, their findings were inconsistent regarding relative difficulty, based on sentence comprehension, sentence combination, picture-based production, and grammaticality judgment tasks. Like many studies that contradict the predictions of the Markedness Hypothesis, the present study also faced the

challenge of identifying the developmental sequence between direct object and oblique RCs, as well as any difference in their relative difficulty.

According to the Surface Configuration Hypothesis proposed by Tarallo & Myhill (1983) and Hawkins (1989), the shorter the distance between the NP extraction site and the head noun, the easier it is to acquire the RC. Therefore, in Korean, direct object RCs and oblique RCs should be easier to learn than subject RCs. The overall findings of the present study show that subject RCs develop first, which is contrary to the prediction of the Surface Configuration Hypothesis. In TRAN RCs, where subject and direct object development can be directly compared, direct object RCs appear earlier and more frequently than subject RCs, seemingly supporting the prediction put forward by the Surface Configuration Hypothesis. However, this hypothesis does not account for the developmental relationship between subject and oblique RCs. In the subject–oblique comparison, the distance between the empty NP and the head noun is greater for the subject than for the oblique, suggesting that subject RCs should be more difficult than oblique RCs. Nevertheless, the results of the present study do not support this prediction — either in terms of the overall developmental order among subject, direct object, and oblique RCs, or in the case of the INTR2 RCs in particular.

According to Gibson's (2000) Dependency Locality Theory, which adopts a psycholinguistic perspective, there are various predictions about the difficulty of subject RCs and direct object RCs in Korean. For example, Ishizuka et al. (2006) predicted that subject RCs with a large distance between the empty NP and the head noun would be more difficult than direct object RCs. The results of this study, which show that subject RCs emerged before direct object RCs, contradict the prediction of Ishizuka et al. (2006). The results of Miyamoto & Nakamura (2003), who assumed that there would be no difference in processing difficulty between subject and direct object RCs, and Gibson & Wu (2013), who also predicted that subject RCs would be easier than direct object RCs based on the Dependency Locality Theory, are not in line with the findings of this study. This suggests that Gibson & Wu's (2013) account does not adequately explain the findings of the present study, in which direct object RCs emerged more frequently and appeared earlier than subject RCs in transitive constructions.

5.2 Development of RCs by predicate type

By analyzing the RCs in learners' spontaneous conversational speech according to the predicate type, the order INTR1 > TRAN = INTR2 > DTRAN was identified for both frequency of occurrence and timing of initial emergence. INTR1 RCs appeared between 2 and 3 months of learning Korean, whereas TRAN and INTR2 RCs appeared between 4 and 6 months. DTRAN RCs, which have the most complex argument structure, appeared last, after 7 months. Regarding the syntactic error rate, the error rate of INTR1 RCs was high; however, after four months, learners showed stable

syntactic usage patterns in the INTR1 RCs and other RC types. Hence, RC development among Chinese-speaking learners of Korean followed the order: INTR1 > TRAN = INTR2 > DTRAN.

RC development with respect to these predicate types appears to be closely related to the development of the head nouns of subjects, direct objects, and obliques. INTR1 is similar to the subject type, whereas TRAN and INTR2 are similar to the direct object and oblique types, in terms of frequency, timing of initial emergence, and error rates, respectively. The timing of the subject's emergence was the same as that of INTR1, and the emergence of the direct object and the oblique coincided with TRAN and INTR2, respectively.

INTR1 had the simplest argument structure with only the subject as the obligatory argument, while DTRAN had the most complex argument structure, requiring the subject, direct object, and oblique as obligatory arguments. Therefore, learners' spontaneous conversational speech data revealed a tendency to acquire INTR1 RCs first and DTRAN RCs last.[16]

It is difficult to find evidence that the sequential development of predicate types (e.g., INTR1, TRAN, INTR2, DTRAN) occurs in sentence types other than RCs. The research participants were using all four predicate types within the corpus after approximately 2–3 months of learning Korean. Textbooks included not only INTR and TRAN but also DTRAN predicates (e.g., *mwutta* (to ask), *pakkwuta* (to change), *patta* (to receive), *ponayta* (to send), and *billida* (to borrow)) as target vocabulary to be learned within three months. No studies could be found that examine the frequency and development of predicates by Korean language learners with respect to argument structure. Furthermore, it was difficult to find clear evidence that INTR1, TRAN, INTR2, and DTRAN develop in a sequential order, even in studies of early vocabulary development in children. Fukuda & Choi (2009) investigated early language development in Japanese and Korean children and found that they used more INTR1 than TRAN, regardless of parental input. Lee et al. (2008) found that Korean children aged 18 and 24 months predominantly used INTR1 in their early verb development. Nonetheless, the authors reported that children aged 18 months began to use TRAN predicates (e.g., *an-a* (hug), *salang-hay* (I love you), *cha* (kick), *kke* (turn off), *mek-e* (eat), and *sin-e* (wear)). Studies on early language development in children have found that children use INTR1 more than TRAN; however, they do not explain why the development of TRAN predicates is slower than that of INTR1 predicates.

Despite the difficulty of finding evidence of sequential development of INTR1, TRAN, INTR2, and DTRAN in non-RC usage, the RCs of these predicate types appeared sequentially in this study. The findings showed that the complexity of argument structure had a greater impact on

[16] Meanwhile, whether the later appearance of the DTRAN type can be attributed to a limited number of DTRAN predicates remains unclear. DTRAN predicates are found among the basic verbs that learners typically learn and use; however, the number of predicate types is smaller than that of INTR or TRAN. This study could not confirm such points based solely on learners' spontaneous conversational speech data. This issue should be addressed in future research.

RC development than on non-RC development. In non-RCs, sentences can be composed using predicates with omitted arguments and without an explicit understanding of argument structure. However, when constructing RCs, knowledge of the arguments associated with RC predicates seems to have a greater influence.

Table 13 presents the order of frequency and timing of initial emergence of head noun types according to RC predicate types. Certain patterns of development can be observed.

Predicate	Frequency	Timing of initial emergence
INTR1	subject > oblique	subject (2 months), oblique (8 months)
TRAN	direct object > subject > oblique	direct object (4 months), oblique (5 months), subject (7 months)
INTR2	subject > oblique	subject (4 months), oblique (6 months)
DTRAN	direct object > subject > oblique	subject (9 months), oblique (9 months), direct object (11 months)

Table 13: Order of frequency and first appearance of head noun type according to predicate type.

In INTR1 and INTR2 RCs, the subject type appeared more frequently and considerable earlier than the oblique type. In TRAN and DTRAN RCs, the direct object type appeared more frequently than the subject type. Regarding the order of appearance, the direct object type appeared earliest in the TRAN RCs, whereas the subject type appeared after the oblique type. In the DTRAN RCs, the subject type appeared earliest; however, there was considerable variation among learners.[17]

Several explanations can be proposed for this phenomenon. One possible explanation is the influence of the learners' first language. As illustrated in (12) and (13), the subject RC in TRAN differs in word order between Chinese and Korean, whereas the direct object RC shares the same word order in both languages. Hence, these syntactic similarities and differences between Chinese and Korean may influence learners' preference for direct object RCs over subject RCs in TRAN and DTRAN types.

(12) Subject type in TRAN RC
  a. $[\ \text{V O}\ ]_{\text{REL}}$ HN (Chinese)
  b. $[\ \text{O V}\ ]_{\text{REL}}$ HN (Korean)

(13) Direct object type in TRAN RC
  a. $[\ \text{S V}\ ]_{\text{REL}}$ HN (Chinese)
  b. $[\ \text{S V}\ ]_{\text{REL}}$ HN (Korean)

[17] The head noun types in DTRAN RCs varied across learners. By the time the DTRAN type first emerged at 9 months, learners had already become familiar with producing various types of RCs. Therefore, the order of appearance of head noun types in DTRAN RCs may not fully reflect the developmental trajectories of the learners.

However, this explanation is insufficient to account for the phenomenon observed in INTR2 RCs. In INTR2 RCs, the subject RC differs in word order between Korean and Chinese (14), whereas the oblique RC exhibits the same word order in both languages (15). Nevertheless, in INTR2 RCs, the oblique RC is not preferred to the subject RC. Therefore, it is difficult to attribute the findings of the study solely to first language influence.

(14) Subject type in INTR2 RC
  a. [V A]$_{REL}$ HN (Chinese)
  b. [A V]$_{REL}$ HN (Korean)

(15) Oblique type in INTR2 RC
  a. [S V]$_{REL}$ HN (Chinese)
  b. [S V]$_{REL}$ HN (Korean)

Similar to Diessel & Tomasello (2005), the results of the present study can be explained from the perspective of syntactic similarity. While the outcome is similar to that of Diessel & Tomasello (2005), there are also notable differences. Diessel & Tomasello (2005) reported that the error rates followed the order: intransitive subject < agent < patient, indirect object, oblique < genitive. However, the findings of the present study differed —the agent developed less frequently and appeared later than the patient[18] (**Table 14**).

Class	Frequency	Timing of initial emergence	Note
intransitive subject	769	2 months	INTR1 subject INTR2 subject
patient	223	4 months	TRAN direct object DTRAN direct object
oblique	95	5 months	oblique
agent	91	7 months	TRAN subject DTRAN subject

Table 14: Frequency and timing of initial emergence of RCs by argument type.

[18] In **Table 14**, the timing of initial emergence reflects when each participant first produced the relevant utterance. The overall order of emergence was intransitive subject > patient > oblique > agent, although the relative timing of the oblique and agent types varied among participants. For Participant 1, the order was intransitive subject (2 months) > patient (4 months) > oblique (5 months) > agent (9 months); for Participant 3, it was intransitive subject (2 months) > patient (5 months) > oblique (5 months) > agent (7 months). However, Participant 2 showed a different order between the oblique and agent types: intransitive subject (2 months) > patient (8 months) > agent (9 months) > oblique (10 months). Because the difference in frequency between the oblique and agent types was not substantial, and their order of emergence varied among learners, it is difficult to determine which emerged earlier.

Diessel & Tomasello (2005) found that intransitive subject/agent RCs had lower error rates compared to the other types due to their structural similarity to simple non-embedded sentences. However, such logic does not straightforwardly apply to the structure of the Korean language. In Korean, the head noun in a simple non-embedded sentence must be artificially displaced to a later position when forming any type of RC (16). Therefore, applying the logic of Diessel & Tomasello (2005) to Korean is problematic.

(16) a. [*Ku salam-i*]$_{NP1}$ [*sonyen-ul*]$_{NP2}$ *po-n-ta*. (simple sentence)
The man-NOM boy-ACC see-PRS-DECL
'The man sees the boy.'

b. [*Sonyen-ul*]$_{NP1}$ *po-nun* [*ku salam*]$_{NP2}$ (agent RC)
Boy-ACC see-AE the man (HN)
'The man who sees the boy.'

c. [Ku salam-*i*]$_{NP1}$ *po-nun* [sonyen]$_{NP2}$ (patient RC)
The man-NOM see-AE boy (HN)
'The boy whom the man sees.'

Finally, it is worth considering whether Fox's (1987) Absolutive Hypothesis can accounts for the observed patterns. The results of this study align with Fox's Absolutive Hypothesis, which posits that intransitive subjects and patients are preferred as head nouns in RCs over agents. In addition, Hogbin & Song (2007) argued that frequency followed the order: intransitive subject/direct object > oblique > agent > genitive > indirect object/object of comparison. Similar to Hogbin & Song (2007), the overall findings of this study demonstrate that frequency and first appearance follow the order: intransitive subject > direct object > oblique, agent[19] (see **Table 14**).

Fox's (1987) Absolutive Hypothesis explains this phenomenon in terms of the preferred argument structure. More precisely, as the agent of transitive verbs often conveys old information in discourse, it appears less frequently than either the subject of INTR1 or the patient. In the corpus of this study, there were 135 cases (60.5%) in which the subject of the RC was omitted out of the total 223 patient-type RCs, and 56 cases (25.1%) in which it was realized as a pronoun, accounting for most of the cases. In only 32 cases (14.3%), the lexical argument was used. Thus, the fact that the subjects in INTR RCs and objects in TRAN or DTRAN RCs appeared more frequently and developed earlier can be explained by the Absolutive Hypothesis.

Fox (1987) and Hogbin & Song (2007) explored the use of RCs in corpora compiled from native English speakers. Given that the results of the present corpus study involving second

[19] As discussed in note 19, the agent and oblique types do not differ significantly in frequency, and the order of their initial appearance varies across learner. While it is difficult to determine whether the oblique or the agent type develops earlier, both emerge later than the intransitive subject or the direct object types.

language learners are similar to these previous findings, it is likely that RC development in a second language follows a pattern similar to that observed in first language acquisition. However, the present results differ from those of Diessel & Tomasello's (2005) study of RC development in a native language, especially with respect to the agent type. This difference could be attributed to differences in research methodologies. Diessel & Tomasello (2005), an experimental study in which participants were asked to listen to and repeat RC sentences, provided evidence on the relative difficulty of RC types. However, it is worth considering that the difficulty of RCs and the order of their development do not necessarily correlated. It is important to consider whether discursive factors, such as preferred argument structure, influence RC development in actual use, regardless of the difficulty level identified in the experiments.

6 Conclusion

The objectives of this study were to explore the development patterns of RCs in Chinese-speaking learners of Korean by: i) verifying whether RC development according to head noun types in spontaneous spoken data aligns with a linguistic typological hypothesis or a psycholinguistic hypothesis, and ii) examining whether RC development according to predicate types can be explained by the Absolutive Hypothesis based on Preferred Argument Hypothesis or syntactic complexity. This study utilized a learner corpus consisting of 19 recordings of spontaneous speech collected over 18 months from three Chinese speakers in their twenties. Based on this data, the study examined the frequency of occurrence, timing of initial emergence, and syntactic error rates of RC patterns over time. In addition, predicate types were classified according to the obligatory arguments of each predicate, and the frequency and error rate of each type were investigated. The study further examined the frequency of the predicate types according to learning time and the ratio of sentence components for head nouns according to the predicate types.

In terms of frequency, the 1,178 observed RCs appeared in the order: subject > direct object > oblique. The subject type RCs were the most frequent, and they developed earlier than other types. However, no significant differences were observed in the timing of initial emergence or error rates between direct object and oblique. As a result, the findings could not be fully explained by the Markedness Hypothesis. Similarly, psycholinguistic explanations such as the Surface Configuration Hypothesis and Dependency Locality Theory did not align with the study's outcomes. This study showed that RCs headed by predicates with simpler argument structures were acquired more easily than those with more complex structures. Accordingly, RC development appeared to follow the order: INTR1 > TRAN, INTR2 > DTRAN. Moreover, the overall results and the results of each predicate type supported Fox's (1987) Absolutive Hypothesis, which asserts that the intransitive subject or patient precedes the agent in the hierarchy.

Follow-up experimental studies should further investigate the complexity of argument structure and preferred arguments in RC predicates within the context of RC development. This is particularly important because it is difficult to control for the frequency of specific predicate types in natural conversations when using corpus data. Future research should also examine whether the findings are generalizable across languages and how argument-related factors influence non-RC development.

Abbreviations

AE = adnominal ending, COM = comparative, CONN = connective ending, CONJ = conjunction, DTRAN = ditransitive verb, HN = head noun, INTR1 = intransitive verb without obligatory oblique argument, INTR2 = intransitive verb with obligatory oblique argument, NP = noun phrase, NPAH = Noun phrase accessibility hierarchy, O = object, PST = past tense, REL = relative clause boundary marker, REQ = request, RC = relative clause, S = subject, TEMP = temporal marker, TRAN = transitive verb, V = verb, For the other abbreviations, I followed the conventions outlined in the Leipzig glossing rules (https://www.eva.mpg.de/lingua/resources/glossing-rules.php).

Data availability

The anonymized learner speech corpus used for the analysis is openly available on Zenodo at the following DOI: https://doi.org/10.5281/zenodo.15917963.

Ethics and consent

Data collection for this study was approved by the Institutional Review Board of Wonkwang Digital University (protocol number: 2207-5-001). The study involved the collection of learner speech corpus data excluding any personally identifiable or sensitive information. All participants provided written informed consent and voluntarily agreed to participate.

Acknowledgements

We would like to express our sincere gratitude to Editage (www.editage.com) for their assistance with English language editing and journal submission. This manuscript was submitted with authorization from Editage. We are also grateful to the three anonymous reviewers for their insightful comments and valuable suggestions. In addition, we extend our appreciation to Andrew Sangpil Byon of the University at Albany for his thorough linguistic review and constructive feedback. Any remaining errors are solely our responsibility.

Competing interests

The authors have no competing interests to declare.

References

Cho, Yong-joon. 2012. L2 acquisition of Korean relativization by adult Chinese learners. *Journal of Korean National Language and Literature* 49. 353–386.

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

Doughty, Catherine. 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.

Du Bois, John W. 1981a. *Ergativity and preferred argument structure in Sacapultec discourse*. Santa Barbara: University of California. DOI: https://doi.org/10.1017/S0272263100010287

Du Bois, John W. 1981b. *The Sacapultec language*. Berkeley: University of California.

Du Bois, John W. 1985. Competing motivations. In Haiman, John (ed.), *Iconicity in syntax*, 343–365. Amsterdam/Philadelphia: John Benjamins Publishing Company. DOI: https://doi.org/10.1075/tsl.6.17dub

Du Bois, John W. 1987. The discourse basis of ergativity. *Language* 63(4). 805–855. DOI: https://doi.org/10.2307/415719

Du Bois, John W. 2003. Discourse and grammar. In. Du Bois, John W. & Kumpf, Lorraine E. & Ashby, William (eds.), *Preferred argument structure: Grammar as architecture for function,* 47–85. Amsterdam: Benjamins. DOI: https://doi.org/10.1075/sidag.14

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

Fox, Barbara A. 1987. The noun phrase accessibility hierarchy reinterpreted: Subject primacy or the absolutive hypothesis? *Language* 63(4). 856–870. DOI: https://doi.org/10.2307/415720

Fukuda, Shin & Choi, Soonja. 2009. The acquisition of transitivity in Japanese and Korean children. *Japanese and Korean Linguistics* 17. 613–624.

Gass, Susan M. 1979a. *An investigation of syntactic transfer in adult second language acquisition.* Ph.D. Dissertation. Indiana University, Indiana.

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

Gibson, Edward. 1998. Linguistic complexity: Locality of syntactic dependencies. *Cognition* 68. 1–76. DOI: https://doi.org/10.1016/S0010-0277(98)00034-1

Gibson, Edward. 2000. The dependency locality theory: A distance-based theory of linguistic complexity. In Marantz, Alec & Miyashita, Yasushi & O'Neil, Wayne (eds.), *Image, language, brain*, 95–126. Cambridge, MA: MIT Press. DOI: https://doi.org/10.7551/mitpress/3654.003.0008

Gibson, Edward & Wu, H.-H. Iris. 2013. Processing Chinese relative clauses in context. *Language and Cognitive Processes* 28(1-2). 125–155. DOI: https://doi.org/10.1080/01690965.2010.536656

Han, Song Hwa. 2015. A study on the acquisition of Korean relative clauses and Korean adnominal endings. *Grammar Education* 24. 211–243.

Hawkins, Roger. 1989. Do second language learners acquire restrictive relative clauses on the basis of relational or configurational information? The acquisition of French subject, direct object and genitive restrictive relative clauses by second language learners. *Interlanguage Studies Bulletin (Utrecht)* 5(2). 156–188. DOI: https://doi.org/10.1177/026765838900500204

Hogbin, Elizabeth & Song, Jae Jung. 2007. The accessibility hierarchy in relativisation: The case of eighteenth- and twentieth-century written English narrative. *Sky Journal of Linguistics* 20. 203–233.

Hsiao, Franny & Gibson, Edward. 2003. Processing relative clauses in Chinese. *Cognition* 90(1). 3–27. DOI: https://doi.org/10.1016/S0010-0277(03)00124-0

Hu, Shenai & Gavarró, Anna & Guasti, Maria T. 2016a. Children's production of head-final relative clauses: The case of Mandarin. *Applied Psycholinguistics* 37(2). 323–346. DOI: https://doi.org/10.1017/S0142716414000587

Hu, Shenai & Gavarró, Anna & Vernice, Mirta & Guasti, Maria T. 2016b. The acquisition of Chinese relative clauses: Contrasting two theoretical approaches. *Journal of Child Language* 43(1). 1–21. DOI: https://doi.org/10.1017/S0305000914000865

Ishizuka, Tomoko & Nakatani, Kentaro & Gibson, Edward. 2006. *Processing Japanese relative clauses in context*. Paper presented at the 19th annual CUNY conference on human sentence processing, CUNY, New York.

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

Jeon, K. Seon & Kim, Hae-Young. 2007. Development of relativization in Korean as a foreign language: The noun phrase accessibility hierarchy in head-internal and head-external relative clauses. *Studies in Second Language Acquisition* 29(2). 253–276. DOI: https://doi.org/10.1017/S0272263107070131

Keenan, Edward L. & Comrie, Bernard. 1977. Noun phrase accessibility and universal grammar. *Linguistic Inquiry* 8(1). 63–99.

Keenan, Edward L. & Comrie, Bernard. 1979. Data on the noun phrase accessibility hierarchy. *Language* 55(2). 333–351. DOI: https://doi.org/10.2307/412588

Kim, Chang-Goo. 2010. *The acquisition of Korean relative clauses as a foreign language.* Ph.D. Dissertation. Pukyong National University, Pusan.

Kim, Shinyoung & Kim, Youngju. 2016. Aspects of Korean relative clauses in L2 Korean learners' production. *The Korean Language and Literature* 70. 185–213. DOI: https://doi.org/10.18628/urimal.70..201609.185

Kweon, Soo-Ok & Lee, Dami. 2008. Processing constraints on relative clauses in Korean as a foreign language. *Contemporary Grammar Studies* 51. 129–152.

Lee, Hee-Ran & Choi, Yu-Li & Chang-Song, You-Kyung & Lee, Seung-Bok. 2008. The Early Verb Acquisition in Korean Children. *Journal of Speech-Language & Hearing Disorders* 17-3. 65-77. DOI: https://doi.org/10.15724/jslhd.2008.17.3.005

Lee, Sun-Young. 2005. Relative clauses and subject-drop in KSL learners' writing: Sentence processing approach. *Language Research* 41(2). 405–435.

Lee, Sunyoung & Choi, Eunji. 2020. Longitudinal study on the development of relative adnominal clause of beginner-level Korean learners. *DONAM Language and Literature* 38. 419–446. DOI: https://doi.org/10.17056/donam.2020.38..419

Matsumoto, Yoshiko. 1990. The role of pragmatics in Japanese relative clause constructions. *Lingua*. 82(2-3). 111–129. DOI: https://doi.org/10.1016/0024-3841(90)90059-T

Miyamoto, Edson T. & Nakamura, Michiko. 2003. Subject/object asymmetries in the processing of relative clauses in Japanese. In Garding, Gina & Tsujimura, Mimu (eds.), *Proceedings of the 22nd West Coast Conference on Formal Linguistics*, 342–355. Somerville MA: Cascadilla Press.

O'Grady, William & Lee, Miseon & Choo, Miho. 2003. A subject-object asymmetry in the acquisition of relative clauses in Korean as a second language. *Studies in Second Language Acquisition* 25(3). 433–448. DOI: https://doi.org/10.1017/S0272263103000172

Ozeki, Hiromi & Shirai, Yasuhiro. 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. DOI: https://doi.org/10.1017/S0272263107070106

Pavesi, Maria. 1986. Markedness, discoursal modes, and relative clause formation in a formal and an informal context. *Studies in Second Language Acquisition* 8(1). 38–55. DOI: https://doi.org/10.1017/S0272263100005829

Tarallo, Fernando & Myhill, John. 1983. Interference and natural language processing in second language acquisition. *Language Learning* 33(1). 55–76. DOI: https://doi.org/10.1111/j.1467-1770.1983.tb00986.x

Xia, Vera Y. & White, Lydia & Guzzo, Natália B. 2020. Intervention in relative clauses: Effects of relativized minimality on L2 representation and processing. *Second Language Research* 38(2). 347–372. DOI: https://doi.org/10.1177/0267658320958742

Xu, Yi. 2014. Evidence of the accessibility hierarchy in relative clauses in Chinese as a second language. *Language and Linguistics* 13(3). 435–464. DOI: https://doi.org/10.1177/1606822X14520666