

A corpus-based analysis of Chinese relative clauses produced by Japanese and Thai learners

Yike Yang

Department of Chinese and Bilingual Studies,
The Hong Kong Polytechnic University
Hong Kong SAR

yi-ke.yang@connect.polyu.hk

Abstract

Concerning the acquisition of relative clauses (RCs), studies on head-initial languages consistently reported a preference for subject-gapped RCs, but the issue of subject-object asymmetry is still a controversial one in research on the acquisition of RCs in head-final languages. Using written corpus data, this study investigated the second language production of RCs in Mandarin Chinese (Chinese) by Japanese-speaking and Thai-speaking learners with various proficiency levels. We first extracted the RCs produced by Japanese and Thai learners from the HKS Dynamic Composition Corpus, and coded head type and gap type for further analyses. The learners from the intermediate-level groups produced a significant number of error-free RCs, which suggests that the intermediate learners have already mastered Chinese RCs. Both Japanese and Thai learners exhibited a strong preference for the subject RCs, which is consistent with predictions that follow from the Noun Phrase Accessibility Hierarchy (NPAH) and the results of studies on head-initial languages. Our data also provided partial support for the Subject-Object Hierarchy (SOH). However, the size of the corpus was insufficient to exhaustively investigate the tested theories. More data are needed to examine the applicability of the NPAH and SOH hypotheses in L2 Chinese and in general.

Keywords: corpus; second language acquisition; Chinese relative clauses; subject-object asymmetry

1 Introduction

The acquisition of relative clauses (RCs) in both first language (L1) and second language (L2) has

been widely studied for decades. However, there remain controversial issues in previous research, such as that of subject-object asymmetry, especially in the context of Asian languages. Subject-object asymmetry has been the focus of inquiry in RC acquisition since the Noun Phrase Accessibility Hierarchy was proposed in the 1970s (Keenan & Comrie, 1977), and studies on post-nominal Indo-European languages generally accept that there is a subject preference (Izumi, 2003; Keenan & Hawkins, 1987), but the data from pre-nominal languages yield mixed results (Chan et al., 2011; Zhang & Yang, 2010). Using written corpus data, this study examined the production of RCs in Mandarin Chinese (henceforth, Chinese) by Japanese and Thai learners to investigate the subject-object asymmetry. Our data will also shed light on whether and how typological variations affect the acquisition process.

This paper begins with a brief introduction of RCs in Chinese, Japanese and Thai languages, and then reviews issues in the acquisition of RCs, particularly in the acquisition of Chinese RCs. The research questions and methods of the study will be described in Section 2. The results of the study will be presented in Section 3. General discussion and concluding remarks will appear in the final section.

1.1 Relative constructions

A relative construction comprises a nominal (i.e. head) and a subordinate clause (i.e. relative clause) that attributively modifies the nominal (Lehmann, 1986). An ‘under-represented’ element of the RC is co-indexed with the head, and this element picks up its interpretation from the head (O’Grady, 2011), as shown in (1), where the green apple has the same referent as the apple eaten by the addressee:

(1) The apple_i [that you ate t_i] was green.

(O’Grady, 2011: 19)

Conventionally, RCs have been divided into two types: 1) the pre-nominal type, where the head appears to the right of the RC (RC-head), and 2) the post-nominal type, where the head occurs to the left of the RC (head-RC) (Keenan & Comrie, 1977). Dryer (2013) conducted a typological investigation on the relationship between the basic word order and the order of the RC and head. The results are presented in Table 1. He found that almost all the world’s VO languages qualify as the second type, the post-nominal type; only five out of 421 VO languages have pre-nominal RCs, and Mandarin Chinese is one of those rare cases. Among the OV languages, there are more languages that have pre-nominal RCs than those that have post-nominal RCs (132 vs. 113). Given the special case of Chinese RCs, it is theoretically interesting to examine the acquisition of Chinese RCs by speakers of more ‘typical’ languages to determine whether and how typological variations affect the acquisition process.

Basic word order	Order of RC and head	Number	Example
Verb-object	RC-head	5	Mandarin
Verb-object	Head-RC	416	Thai
Object-verb	RC-head	132	Japanese
Object-verb	Head-RC	113	Persian
Languages that do not fall into the four types		213	Kutenai

Table 1: Typology of word order and order of RC and head (adapted from Dryer, 2013)

Chinese uses a relative marker *de* to link the matrix clause and the embedded RC (Li & Thompson, 1981). The RC always precedes the head, and *de* precedes the head to mark the RC, as illustrated in (2). (2a) is a subject-gapped (SU) RC, where the subject of the RC is co-indexed with the relativised head, and (2b) is an example of a direct object-gapped (DO) RC. With the basic word order of SOV, Japanese shares the pre-nominal property with Chinese RCs, as in (3). However, in Japanese, the head is directly modified by the RC, and there lacks any overt relative marker or relative pronoun (Yabuki-Soh, 2007), so the addressee will only notice the RC when the head is heard or read. The structure of Thai RCs resembles that of English RCs. Both Thai and English are VO languages with post-nominal RCs, and their RCs are introduced by a relative marker

that follows the head, as illustrations in (1) for English and in (4) for Thai (Sornhiran, 1978). There are three relative markers in Thai (i.e. *thūi*, *sýŋ* and *an*), among which *thūi* is the most productive one because it can be used under all circumstances (Phoocharoensil, 2014). It is clear that the SU RCs have a longer gap-filler dependency and that the DO RCs share their ordering with a canonical SVO clause in Chinese. Like those in Chinese, the SU RCs in Japanese also have a longer gap-filler dependency. In Thai, the SU RCs show a shorter gap-filler dependency. The contrast between Japanese and Thai RCs may lead to some differences in Japanese and Thai learners’ acquisition of Chinese RCs and help us better understand cross-linguistic influence (Yang, 2020).

- (2) a. [*t_i mai shu de*] *nanhai_i*
 buy book REL boy
 ‘the boy who bought a book’
 b. [*ta mai t_i de*] *shu_i*
 he buy REL book
 ‘the book that he bought’
- (3) a. [*t_i watashi ni hon o kure-ta*] *hito_i*
 me DAT book ACC give-PST person
 ‘the person who gave me a book’
 b. [*watashi ga kinoo t_i at-ta*] *hito_i*
 I NOM yesterday meet-PST person
 ‘the person whom I met yesterday’
 (Yabuki-Soh, 2007: 228)
- (4) a. *phét_i [thūi t_i mii khāa mahāasāan]*
 diamond REL have value tremendous
 ‘the diamond that has tremendous value’
 b. *dèki [thūi chān liay ti maa]*
 child REL I bring up come
 ‘the child that I brought up’
 (Sornhiran, 1978: 177)

1.2 Subject-object asymmetry

Several hypotheses have been adopted to account for the subject-object asymmetry in the acquisition and processing of RCs; among these, the Noun Phrase Accessibility Hierarchy (NPAH) and the Subject-Object Hierarchy are of particular relevance to the current study. The NPAH, a typological generalisation proposed by Keenan and Comrie (1977), has been employed to predict the difficulty order in the acquisition of various types of RCs over recent decades (Shirai & Ozeki, 2007; Song, 2002). According to the NPAH, there exists a universal hierarchy of psychological ease of relativisation of the RC head, as suggested in (5). This means that, the subject position in a sentence is always the most accessible to undergo the process of relativisation,

Type of RC	Illustration	No. of discontinuity
SS	<i>[_{TP} t_i xuyao gongzuo] de nanren_i bangzhu-le ta.</i> need job REL man help-PST her 'The man who needed a job helped her.'	1
SO	<i>[_{TP} mama [_{VP} yang t_i]] de tuzi_i chi-le huluobo.</i> mom feed REL rabbit eat-PST carrot 'The rabbit that mom feeds ate the carrot.'	2
OS	<i>wo kanjian [_{DP} [_{TP} t_i mai shu] de nanhai_i].</i> I see buy book REL boy 'I saw the boy who bought a book.'	2
OO	<i>ta mai-le [_{DP} [_{TP} nver [_{VP} xiangyao t_i]] de hua_i].</i> he buy-PST daughter want REL flower 'He bought the flower that his daughter wanted to have.'	3

Table 2: Illustration of discontinuity in Chinese RC according to SOH

whereas the object of comparative position is the least accessible to be relativised. SU and DO RCs have been the focus of subsequent studies testing the NPAH, and, consistent with the prediction of the NPAH, studies on post-nominal Indo-European languages have discovered that SU RCs are easier to acquire and process than DO RCs in both L1 (Diessel & Tomasello, 2001; Keenan & Hawkins, 1987) and L2 (Gass, 1979; Izumi, 2003) contexts. Inquiries into pre-nominal languages, however, have not yet reached an agreement, because some results favour the SU type (Lau, 2016; Lee, 1992), whereas others support the DO type (Chen & Shirai, 2015; Hsiao & Gibson, 2003).

(5) A universal hierarchy of ease of relativisation: Subject > Direct Object > Indirect Object > Oblique > Genitive > Object of comparative

Furthermore, Hamilton (1994) postulated the Subject-Object Hierarchy (SOH) to rank four types of RCs according to the position of the head in the matrix clause and the role of the relativised head in the subordinate RC: OS > OO/SS > SO, where the first code refers to the position of the head in the matrix clause and the second indicates the role of the relativised head in the RC. The SOH followed the idea of processing discontinuity (O'Grady, 1987) and proposed two levels of processing discontinuity: 1) if an RC is embedded in the middle of the main clause, it will create a processing discontinuity in the main clause, and 2) an SU RC establishes one discontinuity within the RC, whereas a DO RC sets

up two discontinuities (as illustrated in Table 2 with examples in Chinese). The Chinese RCs are typologically different from their English counterparts in the order of the head and the RC, so we adapted the SOH to accommodate Chinese RCs in Table 3, which also lists the rankings and the number of discontinuity for Japanese and Thai RCs. The information from the updated SOH will help us determine whether there is any transfer from the learners' L1s to their L2 Chinese.

Languages	Updated SOH
Chinese	SS (1) > SO/OS (2) > OO (3)
Japanese	SS (1) > SO/OS (2) > OO (3)
Thai	OS (1) > OO/SS (2) > SO (3)

Table 3: An updated SOH

Note: The parentheses after each RC type indicate the number of discontinuity.

1.3 Acquisition of Chinese relative clauses

Although a consensus on the subject preference in Indo-European languages has been reached, previous studies on the acquisition and processing of Chinese RCs exhibited complex results on subject-object asymmetry. For example, a preference for subject RCs in Chinese has been supported from L1 child comprehension data (Hu et al., 2016), L1 adult comprehension data (Lin & Bever, 2006), L2 comprehension data (Li et al., 2016) and L2 production data (Xu, 2014). In contrast, a preference for object

RCs in Chinese has also been suggested in L1 child production data (Chen & Shirai, 2015), L1 adult comprehension data (Chen et al., 2008) and L2 comprehension data (Packard, 2008). Moreover, Lam (2017) showed a slight subject preference in the production but an object preference in the comprehension of Chinese RCs by typical Cantonese-speaking children. However, consistency has been found in corpus-based studies on the spoken and written Chinese of native speakers, which report a preference for the subject RCs over the object RCs (e.g. 73.8% vs 26.2% in Pu (2007) and 60.8% vs 39.2% in Wu et al. (2011)).

Conflicting results were also found in Chinese for the hierarchy generated from the SOH. Only one comprehension study by Cheng (1995) exactly followed the hierarchy in Table 3. A corpus study of native written data (Wu et al., 2011) and an act out task performed by Chinese children (Lee, 1992) roughly supported the postulate of SOH, as shown by the following ranking: SS > OS > SO > OO. Pu (2007) analysed oral and written data gathered from native speakers, and both types of data followed the same pattern: SS > OS > OO > SO. Chang (1984) tested the comprehension of RCs by Chinese children and found the following hierarchy: SS/SO > OO > OS.

As reviewed above, although data from native Chinese speakers corroborated the generally accepted subject preference, acquisition studies on Chinese RCs presented more inconsistency. Besides, the results concerning the SOH hierarchy did not converge with each other, which needs to be further examined in this study.

2 The current study

2.1 Research questions

To fill gaps in the field, this study addresses the following research questions:

- Is there subject-object asymmetry in the production of RCs by Japanese and Thai learners with different levels of proficiency?

- Is the updated version of the SOH consistent with the written data of Japanese and Thai learners?

2.2 Research methods

A corpus-based approach was adopted for the current study. Correspondence between ease of processing and frequency of occurrence has been suggested in literature (Hawkins, 2004; Wu et al., 2011). This is why distributional data from a learner corpus can inform us of the possible ease of processing for L2 learners and help us answer our research questions.

The data were extracted from the Version 1.1 of the HSK Dynamic Composition Corpus¹ developed by the Beijing Language and Culture University. The Version 1.1 of the corpus contains more than 11,500 essays (more than four million Chinese characters) written by L2 learners of Chinese who took the standardised Chinese proficiency test for non-native speakers (Hanyu Shuiping Kaoshi, HSK). The L2 learners varied in their language background, so it is possible to make cross-linguistic comparisons. We chose Japanese and Thai learners because of the properties of Japanese and Thai RCs and also because of the number of Japanese and Thai learners in the corpus. In addition, there were a variety of topics, based on which the test takers were required to write essays during the test.

In this study, data from learners at the upper intermediate (henceforth, intermediate) level and advanced level were collected. Because there were much more intermediate learners than advanced learners, we first exhausted all the essays of advanced Japanese and Thai learners and manually identified typical SU and DO RCs². Then, we reviewed the essays of intermediate learners and randomly selected the essays to match the number of typical RCs produced by intermediate and advanced learners who shared the same language background. We also considered the topic of the selected essays when choosing the essays. In total, there were 80 essays collected and analysed in this study, with 47

¹ The Version 1.1 of the HSK Dynamic Composition Corpus is not available. The current version of the corpus is Version 2.0, which can be accessed at <http://hsk.blcu.edu.cn/>.

² By 'typical SU and DO RCs', we are referring to the real Chinese RCs where the head noun is co-indexed with the subject or object of the RC. We did not include adjunct or possessive RCs (Lin, 2018).

An adjunct RC:

you kunnan de shihou, women yinggai huxiang bangzhu.
have trouble REL time we should each_other help
'We should help each other when we are in trouble.'

A possessive RC:

xuesheng najiang de laoshi jintian mei lai.
student win_prize REL teacher today NEG come
'The teacher whose student won the prize did not come today.'

essays produced by Japanese learners and 33 essays by Thai learners.

The role of the head in the main clause, the role of the gap in the RC, the type of the verb in the RC and the animacy of the head noun were manually coded. The analyses of the first two properties are reported in this paper. Since our data concern the frequency of occurrence of RCs, we adopted the chi-square test and the binomial test in the statistical analysis. When the sample size was insufficient for a chi-square test (as reported in Section 3.3), we used the Fisher’s exact test instead, which calculates the deviation from a null hypothesis directly and is more appropriate for small-scale data.

3 Results

3.1 An overview of the corpus data

An overview of the corpus data can be found in Table 4, which shows the number and percentage of typical RCs produced by each group. There was a tendency that Japanese and Thai learners produced a larger proportion of RCs as their Chinese proficiency increased. When analysing the data, we not only examined the intermediate and advanced groups separately but also considered all the learners with the same language background as one group, so there were six groups in the table. Each group produced a considerable number of RCs, and no errors in RC usage were spotted throughout our careful investigation, which suggested that the intermediate learners had already mastered Chinese RCs fairly well. Sentences in (6) and (7) were produced by Japanese and Thai learners, respectively. Sentences (6a) and (7a) are examples of SU RCs, and (6b) and (7b) represent DO RCs.

Group	No. of essays	No. of sentences	No. of typical RCs	
JP_IN	28	430	48	(11.16%)
JP_AD	19	270	47	(17.41%)
JP_Sum	47	700	95	(13.57%)
TH_IN	21	266	26	(9.77%)
TH_AD	11	174	25	(14.37%)
TH_Sum	33	440	51	(11.59%)

Table 4: An overview of the corpus data

Note: JP_IN = intermediate Japanese learner group; JP_AD = advanced Japanese learner group; JP_Sum = all Japanese learners as a group; TH_IN = intermediate Thai learner group; TH_AD = advanced Japanese learner group; TH_Sum = all Thai learners as a group.

- (6) a. *wo jiushi yi ge [t_i bu xiyan de] ren_i*
 I be one CL NEG smoke REL person
 ‘I am a person who does not smoke.’
 b. *women yao zuo yixie [women yingdang zuo t_i de] shiqing_i*
 we need do some we should do
 REL thing
 ‘We need to do the things that we should do.’
 (7) a. *[t_i cizhi huijia de] funv_i ye yinggai zhuyi yixie wenti*
 resign go_home REL woman also should care some issue
 ‘The women who resigned and went back home should also pay attention to some issues.’
 b. *wo kaishi ganxie [muqin dangchu suo zuo t_i de] jue ding_i*
 I start thank mother then ACC make REL decision
 ‘I started to feel grateful for the decision that my mother made at that time.’

Chi-square tests of independence were employed to examine the relationship between occurrence of RCs or non-RCs and language background as well as between occurrence of RCs or non-RCs and proficiency level. The proportion in the production of RCs or non-RCs did not differ by language background ($\chi^2(1) = .780, p = .377$), indicating that the two language groups produced similar proportions of RCs. There was significant association between the occurrence of RCs or non-RCs and language proficiency in Japanese learners ($\chi^2(1) = 4.994, p = .025$) but not in Thai learners ($\chi^2(1) = 1.741, p = .187$), although an increase of the proportion of RC production is seen in both Japanese and Thai learners.

3.2 Subject-object asymmetry

To answer our first research question, we compared the SU RCs and DO RCs produced by each group. As shown in Figure 1, a clear subject preference was found in both Japanese and Thai learners at both proficiency levels, and the distribution pattern was very similar across the four groups. The intermediate Japanese learners produced 33 SU RCs and 15 DO RCs (68.75% vs 31.25%), and the advanced Japanese learners produced 35 SU RCs and 12 DO RCs (74.47% vs 25.53%). Similarly, the intermediate Thai learners produced 20 SU RCs and 6 DO RCs (76.92% vs 23.08%), and the advanced Thai learners produced 18 SU RCs and 7 DO RCs (72.00% vs 28.00%).

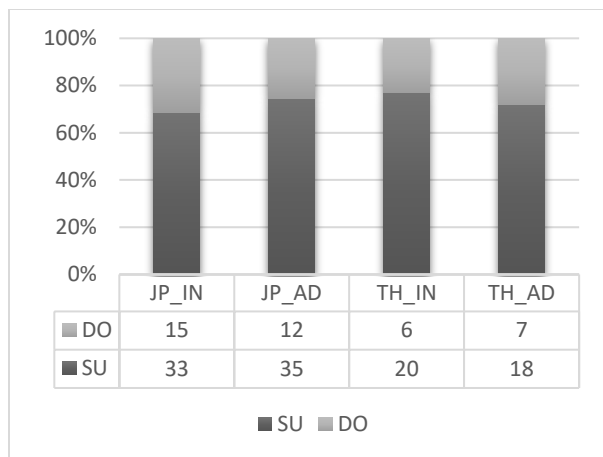


Figure 1: Production of two types of RCs by Japanese and Thai learners

Binomial tests with the type of RC as the test variable showed that there was a significant difference between the production of the two types of RCs within each group: $p = .013$ for the Japanese intermediate group, $p = .010$ for the Japanese advanced group, $p < .001$ for all Japanese learners, $p = .009$ for the Thai intermediate group, $p = .043$ for the Thai advanced group, and $p = .001$ for all Thai learners, revealing an obvious subject RC preference in all the groups. Chi-square tests of independence were then employed to test whether there was any association between RC type and language background and between RC type and proficiency level. The occurrence of SU or DO RC was unaffected by language background ($\chi^2(1) = .034$, $p = .854$) or proficiency level (Japanese learners: $\chi^2(1) = .152$, $p = .696$; Thai speakers: $\chi^2(1) = .007$, $p = .935$), which further confirmed our findings from the binomial tests.

3.3 The Subject-Object Hierarchy

To test the ranking of the four types of RCs compared in the SOH, we further filtered out some RCs from our data and kept only the RCs in which the heads are in the subject or object positions³. As Figure 2 and Table 5 suggest, the six groups did not conform to the same ranking of hierarchy, and none of them followed the hierarchies as predicted by the

SOH in Table 3. However, a bias toward the OS and SS types and a dispreference for the OO and SO types can be observed among the six groups, which required further interpretations.

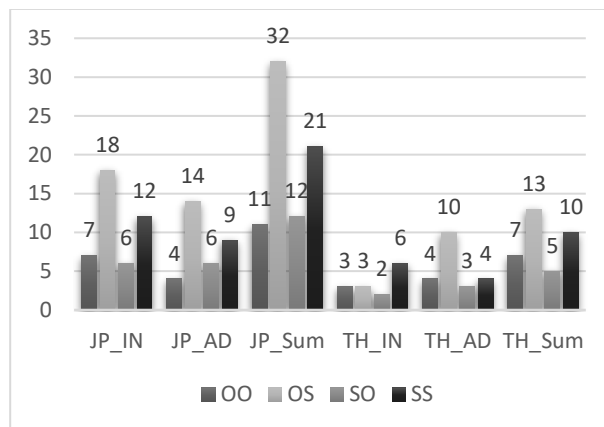


Figure 2: Production of four types of RCs by Japanese and Thai learners

Group	Hierarchy
JP_IN	OS > SS > OO > SO
JP_AD	OS > SS > SO > OO
JP_Sum	OS > SS > SO > OO
TH_IN	SS > OO/OS > SO
TH_AD	OS > SS/OO > SO
TH_Sum	OS > SS > OO > SO

Table 5: Hierarchy of the four types of RCs

Because the sample size was too small for a chi-square test, we used the Fisher's exact test to examine the relationship between the language background and the RC type as well as the relationship between the proficiency level and the RC type. No significant association between language background and type of RC was observed ($p = .893$). Nor did we find any statistically significant association between proficiency level and the RC type ($p = .937$ for the Japanese groups; $p = .365$ for the Thai groups). The results suggested that, in general, the hierarchies of the language and proficiency groups in our data did not distinguish from each other.

³ Headless RCs and RCs in which the RC head is not in the subject or direct object position of the main clause were removed from our analysis.

A headless RC:

xie-wan zuoye de dou hen nuli.
write-PST assignment REL all very diligent
'Those having done the assignment are diligent.'

A case of a RC head modifying the subject noun of the main clause:

jiating bu fuyu de ren de yiliaofei dui jia ren fudan tai da
situation not rich REL people POSS medical_cost for family burden too big
'The medical cost of the poor is a heavy burden for the family.'

4 Discussion and conclusions

This paper provided an analysis of Chinese RCs produced by Japanese and Thai learners using written corpus data. As suggested in Table 4, the Japanese and Thai learners at different proficiency levels produced a significant number of error-free Chinese RCs (ranging from 9.77% to 17.41%), which is even higher than the proportion of Chinese RCs produced by native speakers reported in Pu (2007)⁴. The high frequency of RCs in L2 learners' production indicates that Chinese RCs are acquirable by L2 learners at the intermediate level and that the L2 learners have the mental representations of RCs similar to those of native speakers, although RCs are complex in nature and are not very frequently used by native speakers. The productivity of RCs in our data can be explained by the modality of the data. RCs are embedded in a main clause, and a Chinese RC always interrupts the processing of the main clause because the RC precedes the head noun in the main clause. For example, in Sentence (9b), the readers will take 'mother' as the object of the sentence when they encounter the first four words '*wo kaishi ganxie muqin* (I start thank mother)', but as they read on, they will recognise that they should re-analyse this sentence with 'mother' as the subject of the RC (and retain the incomplete information of the main clause at the same time). The processing of RCs requires a heavy cognitive load and is thus not preferred in oral communication. Also, our data were from L2 learners of Chinese attending the HSK test, during which the learners must have been striving for higher scores by writing sentences with diverse and complex structures. It is plausible that they tried to use RCs as much as possible so that their essays may look more professional to the examiners.

This study explored the issue of subject-object asymmetry in the L2 Chinese RCs of Japanese and Thai learners. Despite the fact that Japanese and Thai are typologically different in terms of basic word order (SOV vs SVO) and the order of RC and head (RC-head vs head-RC), the distribution of the SU and DO RCs in the Japanese and Thai learners' production aligns with that of native Chinese speakers (Pu, 2007). Specifically, a clear preference for the subject RCs over the object RCs has been identified across all the groups in the production data.

We thus provided some support for the NPAH with evidence from L2 Chinese. Due to the limitation of the corpus size, however, only SU and DO RCs were extracted and compared in our analysis. A detailed analysis of all the possible RC types in Chinese is needed to yield a fuller picture of the difficulty order predicted by NPAH in L2 Chinese, which requires either a more large-scale L2 corpus or a carefully-designed elicited production test.

Since SU RCs have a longer filler-gap dependency than DO RCs in Chinese, and the DO RCs also mirror the word order of Chinese sentences, it is logical that DO RCs should be more preferred than SU RCs in Chinese. But corpus data suggest that SU RCs occurs much more frequently than DO RCs, and the pattern is consistent both in L1 (Pu, 2007) and L2 speakers of Chinese. Pu (2007) borrowed the notion of 'markedness' from Givón (1993) to provide some explanations. According to Givón (1993: 178), there are three criteria for markedness: structural complexity, discourse distribution and cognitive complexity. The marked categories are more complex in structure, less frequent in distribution and harder to process than the unmarked categories. As a pro-drop language, Chinese allows a null subject and a null object, and a null subject is much more likely to occur than a null object (Pu, 1997; Xu, 1986), which should have resulted from the subject usually being the topic of a sentence in Chinese and is thus allowed to be an empty category (Tsao, 1990). If the null subject is unmarked, it is reasonable to assume that SU RC is also unmarked. Then the higher frequency of SU RC than DO RC can be accounted for.

As for the SOH, our results in Table 5 did not exactly follow the predicted hierarchies in Table 3. This may be attributed to the relatively small number of RCs collected from the corpus, especially after we further removed the headless RCs and the RCs whose head was not in the subject or object position of the main clause. But there is a tendency of development from the intermediate groups to the advanced groups that fits approximately with the predicted hierarchies. For example, it has been proposed that the SO type should be easier to process and thus should appear more frequently than the OO type in Japanese (in Table 3). The intermediate Japanese learners produced more OO RCs than SO RCs,

⁴ Pu (2007) collected both written and oral data from native Chinese speakers. The proportion of RCs is 4.66% for the written data and 0.77% for the oral data.

but a reversed hierarchy is observed for the advanced Japanese learners. The same case is also found for the Thai learners, which indicates a developmental pattern in the learners' L2 Chinese. However, this pattern is surprising, because the data of advanced learners showed more similarities to the learners' native languages than the data of intermediate learners. Further studies are needed to provide satisfactory explanations to this phenomenon.

Also, the difference in the hierarchy of the Japanese and Thai learners might be a case of cross-linguistic influence where participants' L1s are playing a role (Kidd et al., 2015), and we can hypothesise that the L1s are affecting the learners' L2 acquisition. However, no statistical difference was found between the Japanese and Thai learners, suggesting that the Japanese group and the Thai group generally resembled each other in the production data. The question then arises of why there was little cross-linguistic influence. One possible explanation comes from the typological differences documented by Dryer (2013). Chinese RCs are typologically scarce, and they diverge from Japanese and Thai RCs. Consequently, the learners may treat RCs in Chinese as a new structure that is not equivalent to the RCs in their L1s (Comrie, 2007). If this is the case, not finding any obvious transfer from the L1s would be reasonable. And this may also explain the phenomenon we discussed in the previous paragraph.

In conclusion, both Japanese and Thai learners produced native-like Chinese RCs at the intermediate level, and our findings partially supported the hypotheses of the NPAH and the SOH. Moreover, no obvious cross-linguistic influence was found in our data. However, the size of the corpus was insufficient to exhaustively investigate the tested theories. More data are needed to examine the applicability of the NPAH and SOH hypotheses in L2 Chinese and in general.

Acknowledgements

Part of the data from the Japanese-speaking learners has been presented at the Pacific Second Language Research Forum 2016 in Chuo University, Tokyo, Japan. The author would like to thank the audience for their suggestions. We also acknowledge the useful comments from the PACLIC reviewers that have improved the quality of this paper.

References

- Chan, Angel, Stephen Matthews, and Virginia Yip, 'The Acquisition of Relative Clauses in Cantonese and Mandarin', in *The Acquisition of Relative Clauses: Processing, Typology and Function*, ed. by Evan Kidd (Amsterdam: John Benjamins Publishing Company, 2011), pp. 197–225 <<https://doi.org/10.1075/tilar.8.10cha>>
- Chang, Hsing-Wu, 'The Comprehension of Complex Chinese Sentences by Children: Relative Clause', *Chinese Journal of Psychology*, 26 (1984), 57–66
- Chen, Baoguo, Aihua Ning, Hongyan Bi, and Susan Dunlap, 'Chinese Subject-Relative Clauses Are More Difficult to Process than the Object-Relative Clauses', *Acta Psychologica*, 129 (2008), 61–65 <<https://doi.org/10.1016/j.actpsy.2008.04.005>>
- Chen, Jidong, and Yasuhiro Shirai, 'The Acquisition of Relative Clauses in Spontaneous Child Speech in Mandarin Chinese', *Journal of Child Language*, 42 (2015), 394–422 <<https://doi.org/10.1017/S0305000914000051>>
- Cheng, Sherry Ya-Yin, 'The Acquisition of Relative Clauses in Chinese' (National Taiwan Normal University, 1995)
- Comrie, Bernard, 'The Acquisition of Relative Clauses in Relation to Language Typology', *Studies in Second Language Acquisition*, 29 (2007), 301–9 <<https://doi.org/10.1017/S0272263107070155>>
- Diessel, Holger, and Michael Tomasello, 'The Development of Relative Clauses in Spontaneous Child Speech', *Cognitive Linguistics*, 11 (2001), 131–51 <<https://doi.org/10.1515/cogl.2001.006>>
- Dryer, Matthew S, 'Relationship between the Order of Object and Verb and the Order of Relative Clause and Noun', in *World Atlas of Language Structures Online*, ed. by Matthew S. Dryer and Martin Haspelmath (Leipzig: Max Planck Institute for Evolutionary Anthropology, 2013)
- Gass, Susan M., 'Language Transfer and Universal Grammatical Relations', *Language Learning*, 29 (1979), 327–44 <<https://doi.org/10.1111/j.1467-1770.1979.tb01073.x>>
- Givón, Talmy, *English Grammar: A Function-Based Introduction* (Amsterdam/Philadelphia: John Benjamins Publishing Company, 1993)
- Hamilton, Robert L., 'Is Implicational Generalization Unidirectional and Maximal? Evidence from Relativization Instruction in a Second Language', *Language Learning*, 44 (1994), 123–57 <<https://doi.org/10.1111/j.1467-1770.1994.tb01451.x>>
- Hawkins, John A., *Efficiency and Complexity in Grammars* (Oxford: Oxford University Press, 2004)
- Hsiao, Franny, and Edward Gibson, 'Processing Relative Clauses in Chinese', *Cognition*, 90 (2003), 3–27

- <[https://doi.org/10.1016/S0010-0277\(03\)00124-0](https://doi.org/10.1016/S0010-0277(03)00124-0)>
 Hu, Shenai, Mirta Vernice, Maria Teresa Guasti, and Anna Gavarró, 'The Acquisition of Chinese Relative Clauses: Contrasting Two Theoretical Approaches', *Journal of Child Language*, 43 (2016), 1–21
 <<https://doi.org/10.1017/S0305000914000865>>
- Izumi, Shinichi, 'Processing Difficulty in Comprehension and Production of Relative Clauses by Learners of English as a Second Language', *Language Learning*, 53 (2003), 285–323
 <<https://doi.org/10.1111/1467-9922.00218>>
- Keenan, Edward L., and Bernard Comrie, 'Noun Phrase Accessibility and Universal Grammar', *Linguistic Inquiry*, 8 (1977), 63–99
 <<https://doi.org/10.2307/4177973>>
- Keenan, Edward L., and Sarah Hawkins, 'The Psychological Validity of the Accessibility Hierarchy', in *Universal Grammar: 15 Essays*, ed. by Edward L. Keenan (London: Croom Helm, 1987), pp. 60–85
- Kidd, Evan, Angel Chan, and Joie Chiu, 'Cross-Linguistic Influence in Simultaneous Cantonese–English Bilingual Children's Comprehension of Relative Clauses', *Bilingualism: Language and Cognition*, 18 (2015), 438–52
 <<https://doi.org/10.1017/S1366728914000649>>
- Lam, Scholastica Wai Sze, 'Acquisition of Chinese Relative Clauses by Deaf Children in Hong Kong', *Language and Linguistics*, 18 (2017), 72–115
 <<https://doi.org/10.1075/lali.18.1.03lam>>
- Lau, Elaine, 'The Role of Resumptive Pronouns in Cantonese Relative Clause Acquisition', *First Language*, 36 (2016), 355–82
 <<https://doi.org/10.1177/0142723716648840>>
- Lee, Thomas Hun-tak, 'The Inadequacy of Processing Heuristics: Evidence from Relative Clause Acquisition in Mandarin Chinese', in *Research on Chinese Linguistics in Hong Kong*, ed. by Thomas Hun-tak Lee (Hong Kong: Linguistic Society of Hong Kong, 1992), pp. 47–85
- Lehmann, Christian, 'On the Typology of Relative Clauses', *Linguistics*, 24 (1986), 663–80
 <<https://doi.org/0024-3949/86/0024-0663>>
- Li, Charles N., and Sandra A. Thompson, *Mandarin Chinese: A Functional Reference Grammar* (Oakland: University of California Press, 1981)
- Li, Qiang, Xiaoyu Guo, Yiru Yao, and Müller Nicole, 'Relative Clause Preference in Learners of Chinese as a Second Language', *Chinese Journal of Applied Linguistics*, 39 (2016), 199–215
 <<https://doi.org/10.1515/cjal-2016-0013>>
- Lin, Chien-Jer Charles, 'Subject Prominence and Processing Dependencies in Prenominal Relative Clauses: The Comprehension of Possessive Relative Clauses and Adjunct Relative Clauses in Mandarin Chinese', *Language*, 94 (2018), 758–97
 <<https://doi.org/10.1353/lan.2018.0053>>
- Lin, Chien-Jer Charles, and Thomas G. Bever, 'Subject Preference in the Processing of Relative Clauses in Chinese', in *Proceedings of the 25th West Coast Conference on Formal Linguistics*, ed. by Donald Baumer, David Montero, and Michael Scanlon (Somerville, MA: Cascadilla Proceedings Project, 2006), pp. 254–60
 <<http://www.iub.edu/~lacl/att/paper1456.pdf>>
- O'Grady, William, *Principles of Grammar and Learning* (Chicago: University of Chicago Press, 1987)
- , 'Relative Clause: Processing and Acquisition', in *The Acquisition of Relative Clause: Processing, Typology and Function*, ed. by Evan Kidd (Amsterdam: John Benjamins Publishing Company, 2011), pp. 13–38
- Packard, Jerome L., 'Relative Clause Processing in L2 Speakers of Mandarin and English', *Journal of the Chinese Language Teachers Association*, 43 (2008), 107–46
- Phoocharoensil, Supakorn, 'Errors on the Relative Marker WHERE: Evidence from an EFL Learner Corpus', *3L: Language, Linguistics, Literature*, 20 (2014), 1–20
 <<https://doi.org/10.17576/3L-2014-2001-01>>
- Pu, Ming-Ming, 'Zero Anaphora and Grammatical Relations in Mandarin', in *Grammatical Relations: A Functionalist Perspective*, ed. by Talmy Givón (Amsterdam/Philadelphia: John Benjamins Publishing Company, 1997), pp. 283–322
- , 'The Distribution of Relative Clauses in Chinese Discourse', *Discourse Processes*, 43 (2007), 25–53
 <https://doi.org/10.1207/s15326950dp4301_2>
- Shirai, Yasuhiro, and Hiromi Ozeki, 'Introduction', *Studies in Second Language Acquisition*, 29 (2007), 155–67
 <<https://doi.org/10.1017/S027226310707009X>>
- Song, Jae Jung, 'Linguistic Typology and Language Acquisition: The Accessibility Hierarchy and Relative Clauses', *Language Research*, 38 (2002), 729–56
- Sornhiran, Pasinee, 'A Transformational Study of Relative Clauses in Thai' (University of Texas at Austin, 1978)
- Tsao, Feng-Fu, *Sentence and Clause Structure in Chinese: A Functional Perspective* (Taipei: Student Book, 1990)
- Wu, Fuyun, Elsi Kaiser, and Elaine Andersen, 'Subject Preference, Head Animacy and Lexical Cues: A Corpus Study of Relative Clauses in Chinese', in *Processing and Producing Head-Final Structures*, ed. by Hiroko Yamashita, Yuki Hirose, and Jerome Packard (Dordrecht: Springer Netherlands, 2011), pp. 173–93
 <<https://doi.org/10.1007/978-90-481-9213-7>>

- Xu, Liejiong, 'Free Empty Category', *Linguistic Inquiry*, 17 (1986), 75–93
- Xu, Yi, 'Evidence of the Accessibility Hierarchy in Relative Clauses in Chinese as a Second Language', *Language and Linguistics*, 15 (2014), 435–64
<<https://doi.org/10.1177/1606822X14520666>>
- Yabuki-Soh, Noriko, 'Teaching Relative Clauses in Japanese: Exploring Alternative Types of Instruction and the Projection Effect', *Studies in Second Language Acquisition*, 29 (2007), 219–52
<<https://doi.org/10.1017/S027226310707012X>>
- Yang, Yike, 'Acquisition of the Mandarin Ba-Construction by Cantonese Learners', *Macrolinguistics*, 8 (2020), 88–104
<<https://doi.org/10.26478/ja2020.8.12.6>>
- Zhang, Qiang, and Yiming Yang, 'Object Preference in the Processing of Relative Clause in Chinese: ERP Evidence', *Linguistic Sciences*, 9 (2010), 337–53

Appendix A. List of abbreviations

ACC	Accusative case
DAT	Dative case
DO	Direct object-gapped
JP_AD	Advanced Japanese learner group
JP_IN	Intermediate Japanese learner group
JP_Sum	All Japanese learners as a group
L1	First language
L2	Second language
NEG	Negative
NOM	Nominative case
NPAH	Noun Phrase Accessibility Hierarchy
OO	DO RC appearing in object position
OS	SU RC appearing in object position
OV	Object-verb
POSS	Possessive marker
PST	Past
RC	Relative clause
REL	Relative marker
SO	DO RC appearing in subject position
SOH	Subject-Object Hierarchy
SOV	Subject-object-verb
SS	SU RC appearing in subject position
SU	Subject-gapped
SVO	Subject-verb-object
TH_AD	Advanced Thai learner group
TH_IN	Intermediate Thai learner group
TH_Sum	All Thai learners as a group
VO	Verb-object