Animacy Effects in Chinese Relative Clause Processing

Fuyun Wu, Elsi Kaiser and Elaine Andersen
University of Southern California

1. Introduction

Previous studies on relative clauses (henceforth RCs) have shown a correlation between RC type and head noun animacy. That is, subject-extracted RCs frequently occur with animate head nouns (1a), and object-extracted RCs tend to occur with inanimate head nouns (1b):

(1a) [the journalist] who _ bypassed the boulder (subject-extracted RC)

(1b) [the boulder] that the journalist bypassed _ (object-extracted RC)

This correlation pattern has been replicated in sentence completion tasks (Gennari and MacDonald, 2008), and in corpus analyses of different languages, including Dutch and German (Mak, Vonk, & Schriefers, 2002, 2006), English (Roland, Dick, & Elman, 2007), and Chinese (Pu, 2007).

Research on the processing of RCs has found differences between subject-extracted and object-extracted RCs, as well as effects of head-noun animacy. It has been reported that subject-extracted RCs are easier to process than object-extracted RCs in languages with head-initial RCs (e.g., in English: Just & Carpenter, 1992; Ford, 1983; King & Just, 1991; King & Kutas, 1995; in French: Frauenfelder, Segui, & Mehler, 1980; Holmes & O’Regan, 1981; in German: Schriefers, Friederici, & Kuhn, 1995, inter alia). Furthermore, existing research has found that object-extracted RCs with animate heads are harder to process than object-extracted RCs with inanimate heads (e.g., Traxler, Morris, & Seely, 2002, 2005; Gennari & MacDonald, 2008 on English and Mak et al., 2002, 2006 on Dutch). These findings suggest an interaction between 1) syntactic complexity arising from extraction type and 2) head animacy in head-initial RCs.
This paper aims to investigate whether such an interaction also affects the processing of Mandarin RCs. Mandarin RCs are head-final: The head noun occurs at the very end of the RC, preceded by the relativizer (RC marker) DE. However, other phrasal categories in Mandarin are head-initial, and thus clauses have basic subject-verb-object word order. The Chinese versions of (1a-b) are given in (2a-b) below.

(2a) [ei, raokai jushi] de jizhe (subject-extracted)
[ boulder bypass ] DE journalist
‘the journalist who _ bypassed the boulder’

(2b) [jizhe raokai ei] de jushii (object-extracted RC)
[journalist bypass ] DE boulder
‘the boulder that the journalist bypassed _’

In contrast to research on head-initial languages which has shown that subject-extracted RCs are easier to process overall than object-extracted RCs, the findings for Mandarin are more mixed. Some researchers have found that subject-extracted RCs are easier to process (e.g. Lin & Bever, 2006; Lin, 2006; Kuo & Vasishth, submitted), whereas others have found the opposite (Hsiao & Gibson, 2003; B. Lin & Garnsey, 2008; Wu & Gibson, 2008; Chen, Ning, Bi, & Dunlap, 2008). It is worth noting that nearly all existing experiment-based studies on Chinese used RCs that had a human head noun as well as a human embedded noun inside the RC. However, this configuration is rare in natural language corpora. For instance, Hsiao’s 2003 analysis of the Chinese Treebank 3.0 corpus found only 6 RCs with two animate NPs (out of 882 RCs). Similarly, Kuo and Vasishth’s analysis of the Taiwan Sinica Corpus 3.0 revealed only 16 RCs with two animate NPs (out of 164 RCs). It is further worth noting that RCs with two animate NPs may potentially induce similarity-based interference (e.g., Gordon, Hendrick & Johnson, 2001, 2004; Vasishth & Lewis, 2006).

Given the correlations between animacy and RC type found in existing corpus work, most existing psycholinguistic research on Chinese RCs has not adequately explored the effect of animacy. To investigate this issue more closely, in this paper we examine (1) the frequency of different animacy configurations in Chinese RCs in a corpus; and (2) the contribution of different animacy configurations to real-time processing of subject- and object-extracted RCs in Chinese. We first present the results of our corpus study, followed by three self-paced reading experiments.

2. Corpus Study

A total of 331 RCs with transitive action verbs were extracted from the Chinese Treebank 5.0 corpus (Palmer, Chiou, Xue & Xia, 2005). The head noun and the
embedded noun were coded for two animacy categories: animate and inanimate. Nouns categorized as animate included humans, institutions and organizations (e.g., the Pentagon, WTO, Washington), and animals. Following the standard notation, we examined the animacy patterns for the four RC types, defined on the basis of the grammatical role of the head in the matrix clause and the extraction site in the relative clause itself: (i) subject-modifying, subject-extracted RCs (SS), (ii) object-modifying, object-extracted RCs (SO); (iii) object-modifying, subject-extracted RCs (OS), and (iv) object-modifying, object-extracted RCs (OO).

Let us first consider the results for head noun animacy. Of the 148 animate head nouns in the corpus, 88.51% (131/148) are in subject-extracted RCs (SS and OS). In contrast, among the 183 inanimate heads in the corpus, 60.65% (111/183) occur in object-extracted RCs (SO and OO). This result replicates the findings of prior studies: (1) animate heads generally tend to occur in subject-gapped RCs, and (2) inanimate heads tend to occur in object-extracted RCs.

When we looked at the animacy of both the head noun and the embedded nouns (Figure 1), we observed a striking pattern: RCs with two non-contrastive NPs (i.e., two animates or two inanimates; black bars and grey bars in Figure 1) occur only rarely, except for OS RCs. This suggests that RCs in general tend not to have two NPs with identical animacy.

![Figure 1: Animacy distribution of head nouns and embedded nouns across four types of RCs.](image-url)
RCs with NPs which contrast in animacy also exhibit an asymmetrical distribution. Object-extracted RCs (SO and OO) tend to have inanimate head nouns and animate embedded nouns (101/128, 78.91%). In contrast, SS RCs tend to have animate head nouns and inanimate embedded nouns (78/137, 56.93%). OS RCs do not show a clear preference for a specific animacy pattern.

On the basis of these corpus patterns and existing work, the following animacy preference constraints can be formulated: (i) head nouns that are RC-subjects tend to be animate; (ii) head nouns that are RC-objects tend to be inanimate; (iii) in both cases, the animacy of the head tends to contrast with the animacy of the embedded noun.

We tested these animacy constraints in three self-paced reading experiments, using subject-modifying RCs. We aimed to address the following questions: Are these animacy preference patterns reflected in the ease of processing subject-extracted and object-extracted RCs? If we take these animacy patterns into account, will that help to resolve the controversy regarding subject- vs. object-gapped RCs in Chinese?

3. Experiment 1

In the first self-paced reading study, we examined the real-time processing of subject-gapped, subject-modifying RCs (SS RCs) by testing 36 Mandarin native speakers. The experiment contained 20 target items and 46 filler items. The crucial manipulation was the animacy of the RC-object (i.e., the embedded noun) and the animacy of the RC-subject (i.e., the head), resulting in 4 animacy configurations: (i) Oi-Sa = inanimate RC-object and animate RC-subject/head, (ii) Oa-Sa = animate RC-object, animate RC-subject/head, (iii) Oa-Si = animate RC-object, inanimate RC-subject/head, (iv) Oi-Si = inanimate RC-object, inanimate RC-subject/head. Thus, by crossing the animacy of the RC-object and RC-subject (head), we created two contrastive configurations, one being preferred (animate head ‘journalist’ // inanimate embedded noun ‘gate’), the other being reversed (inanimate head ‘egg’ // animate embedded ‘guard’), as well as two matched configurations (two animate nouns ‘journalist’/‘guard’; two inanimate nouns ‘egg’/‘gate’). An example of the preferred animacy configuration (animate head, inanimate embedded noun, Oi-Sa) is given in (3). We predicted that (i) SS RCs with animate RC-subjects (i.e., head nouns) would be read faster than SS RCs with inanimate RC-subjects; and that (ii) SS RCs with the preferred contrastive animacy configuration would be processed faster than the other three RC types.

(3) SS RCs with animate head and inanimate embedded noun

ei, raokai    damen de jizhe   chenggongde liule    jinqu
  circumvent gate    DE journalist successfully    slip-ASP inside

‘The journalist that ___ circumvented the gate slipped in successfully.’
We focus on seven positions, starting with the RC-internal verb (position 1): the RC-internal object (position 2), the relativizer DE (position 3), the head of the RC (position 4), the post-head adverb (position 5), the matrix verb (position 6) and the postverbal matrix object (position 7).

As can be seen in Figure 1, participants’ RTs show a main effect of subject animacy at the embedded noun position (the RC-object ’gate’, denoted with ‘O’ on the x-axis): RC-objects following verbs that signal an upcoming animate subject (‘journalist’) are read faster than those following verbs signaling an inanimate subject (‘egg’). An interaction was also found at the relativizer DE: DE was read faster in sentences that contained inanimate RC-objects than in sentences that contained animate RC-objects; and in sentences with animate subjects (RC heads), DE was read faster than in sentences with inanimate subjects (RC heads). At the last two word regions, there was a marginal interaction: at the main verb, conditions with animate subjects were read faster than conditions with inanimate subjects; at the main object, inanimate subjects
were read slower than animate objects. In addition, these two regions also showed a contrastive animacy effect, that is, contrastive animacy configurations were read faster than matched animacy configurations.

Overall, the results of Experiment 1 provide evidence for the animacy preference constraints: in subject-extracted RCs, head nouns that are animate RC-subjects are easy to process, and processing is also facilitated when the animacy of the head contrasts with the animacy of the embedded noun.

4. Experiment 2

Interestingly, Experiment 1 did not find facilitatory effects for inanimate objects (except at the DE region), perhaps because the objects nouns were not the heads of the RCs. Experiment 2 used object-extracted RCs (SO RCs) to test whether object head nouns that are inanimate are easier to process than object heads that are animate. Except for extraction site, the logic of the design is the same as in Experiment 1, with the same animacy manipulation and the same predictions. But now the head noun is the RC-object, and the embedded noun is the RC-subject. A new set of native speakers of Mandarin (n=36) participated in the experiment. The experiment contained 18 critical items and 46 fillers. An example of target items in the preferred animacy configuration is provided in (4).

(4) SO RC with inanimate head and animate embedded noun
jizhe raokai cì de damen yaoyande shuazhe hongqi
gate circumvent DE gate glaringly paint-ASP red-paint
‘The gate that the journalist circumvented __ was painted glaringly red.’

The reading time patterns are shown in Figure 2 (below). Again as in Experiment 1, we focus on seven positions, starting with the RC-internal subject (position 1), and also including, the RC-internal verb (position 2), the relativizer DE (position 3), the head of the RC (position 4), the post-head adverb (position 5), the matrix verb (position 6) and the postverbal matrix object (position 7).

We found no main effects of object (head) animacy, except for an interaction between subject animacy and object animacy at the main verb region (e.g. ‘paint’ in ex.(4)). The interaction suggests that the [object=inanimate] preference has less of a facilitation effect than the [subject=animate] preference, and an animate object also has a less disruptive effect than an inanimate subject. Although we found no main effect of object animacy, there was a pervasive effect of RC-subject animacy across all seven regions. In addition, there was a delayed contrastive animacy effect at the matrix object (e.g. ‘red paint’ in (4)): the matched animacy conditions were read slower than the contrastive animacy conditions.

The results of Experiment 2 replicated the subject-animacy preference effects and the delayed contrastive animacy effect that were observed in Experiment 1.
However, we again found no clear evidence for an [object=inanimate] preference.

![Graph](image)

**Fig 2.** Mean reading times per word position in Experiment 2 (SO RCs)

To summarize, both experiments show that processing of RCs where the two nouns contrast in animacy is easier than processing of RCs where the two nouns match in animacy. This finding fits with previous work in other domains on similarity-based interference (Gordon et al., 2001; 2006). Crucially, in addition to showing facilitatory effects of dissimilar animacy, our results also indicate that this ‘mismatch preference’ interacts with the well-known observation that subjects tend to be animate and objects tend to be inanimate (also see Traxler et al. (2002) for a similar line of reasoning).

5. **Experiment 3**

The findings from the first two experiments laid the foundation for the third experiment. Experiment 3 aimed to further test whether distinct animacy status,
mapped onto the appropriate syntactic position, facilitates online processing of RCs. In Experiment 3, we compared subject-extracted and object-extracted RCs (SS vs. SO RCs). The two NPs always differed in animacy, but the animacy configuration was either preferred (animate subject, inanimate object) or reversed (inanimate subject, animate object). If contrastive animacy is kept constant, we predict that if RC-subject animacy and extraction type are manipulated, we should find that RCs with preferred animacy configurations are processed faster than RCs with reversed animacy configurations. Crucially, if we find a difference in the ease of processing subject-extracted vs. object-extracted RCs even when the animacy configurations for both are of the preferred type, that could help clarify the controversy regarding the processing asymmetry in Mandarin RCs.

We tested 40 new native speakers of Mandarin. The experiment contained 24 critical items and 46 fillers. We manipulated extraction type (subject-extracted, object-extracted) and animacy of RC-subject (animate, inanimate), yielding four conditions as in (5).

(5) a. **SS RC, Animate RC-Subject (preferred Oi-Sa)**
   
   ei, raokai jushi de jizhe, jingtide huangu sizhou
   circumvent boulder DE journalist vigilantly look-about surrounding
   ‘The journalist that __ circumvented the boulder cautiously looked about his surroundings.’

   b. **SS RC, Inanimate RC-Subject (reversed Oa-Si)**
   
   ei, zaizhong jizhe de jushii mimide zhangzhe qingtai
   pound journalist DE boulder thickly grow-ASP moss
   ‘The boulder that __ pounded upon the journalist thickly grew moss.’

   c. **SO RC, Animate RC-Subject (preferred Sa-Oi)**
   
   jizhe raokai e, de jushii mimide zhangzhe qingtai
   journalist circumvent DE boulder thickly grow-ASP moss
   ‘The boulder that the journalist circumvented __ thickly grew moss.’

   d. **SO RC, Inanimate RC-Subject (reversed Si-Oa)**
   
   jushi zaizhong e, de jizhe, jingtige huangu sizhou
   boulder pound DE journalist vigilantly look-about surrounding
   ‘The journalist that the boulder pounded upon __ cautiously looked about his surroundings.’

The reading time patterns are shown in Figure 3 (below). At the second word position (RC-object ‘boulder’ for SS RCs and RC-verb ‘circumvent’ for SO RCs), reading times in SS RCs with inanimate subjects (reversed Oa-Si) were significantly slower than in the other three types of RCs.
At the relativizer DE region, SS RCs with reversed animacy (Oa-Si) were significantly slower than SS RCs with preferred animacy (Oi-Sa). At the head noun region, SS RCs with preferred animacy (SS Oi-Sa) were processed faster than SO RCs with reversed animacy (SO Si-Oa). At the adverb region, SS RCs were read faster than SO RCs; in addition, SO RCs with preferred animacy were read faster than SO RCs with reversed animacy. At the main verb region, SO RCs with preferred animacy were read faster than SO RCs with reversed animacy. At the main object region, both kinds of SS RCs were read faster than SO RCs with reversed animacy.

The results of Experiment 3 confirm that preferred animacy configurations (animate RC subjects, inanimate RC objects) are indeed processed faster than reversed animacy configurations. Crucially, the results showed an overall processing advantage for subject-extracted RCs over object-extracted RCs, modulated by the animacy configuration. That is, subject-extracted RCs with the preferred animacy configuration (SS Oi-Sa) were processed significantly faster than object-extracted RCs with the non-preferred animacy configuration (SO Si-Oa).
When RCs of both extraction types satisfied the preferred animacy configurations (SS Oi-Sa, SO Sa-Oi), sentences were equally easy to process, although numerically subject-extracted RCs (i.e., SS Oi-Sa) were processed faster than object-extracted RCs (i.e., SO Sa-Oi). When RCs of both extraction types had the reversed animacy configurations (SS Oa-Si, SO Si-Oa), a processing advantage associated with subject-extracted RCs emerged: reading times in the SS Oa-Si condition were faster than in the SO Si-Oa conditions in all post RC-head regions.

6. Conclusions

In this paper, we presented a corpus study and three experiments that investigated the role of animacy on the processing of head-final relative clauses in Mandarin. The first two self-paced reading experiments manipulated the animacy (animate/inanimate) of the head-noun and the embedded noun. The results show that RCs with animate subjects are easier to process than RCs with inanimate subjects, but that the animacy of the object is not as important — regardless of whether we are dealing with subject-extracted or object-extracted RCs. We also found a late facilitatory effect of contrastive animacy. Also, subject-extracted RCs are processed fastest when the head (RC-subject) is animate and the embedded noun (RC-internal object) inanimate, whereas object-extracted RCs are processed fastest when the head (RC-object) is inanimate and the embedded noun (RC-internal subject) is animate. This facilitatory effect of contrastive animacy correlates with the frequency patterns observed in the corpus.

To clarify whether contrastive animacy alone facilitates RC processing, and to shed light on the controversy regarding the processing asymmetry of subject-versus object-extracted RCs in Chinese, Experiment 3 investigated effects of extraction site in RCs with preferred and dispreferred animacy configurations. Our results show that when RCs had animate subjects and inanimate objects (i.e., the preferred animacy configuration), subject-extracted and object-extracted RCs were equally easy to process. However, when RCs had the reversed animacy configuration, object-extracted RCs were more difficult than subject-extracted RCs.

In all three experiments, we found a consistent facilitatory effect for animate subjects (main effect of subject animacy), but no facilitation for inanimate objects (no main effect of object (in)animacy), except at the DE region in Experiment 2. We offer two possible reasons as to why the animacy of subject is more important than the (in)animacy of the object. First, there is a closer association between Subject and the [+animate] feature than between Object and the [-animate] feature. It has been a long-standing observation that human referents tend to be realized in subject position (e.g., Croft, 1990; Givon, 1983;
See also Gennari and MacDonald, 2008; Traxler et al., 2002; Clifton et al., 2003; Just and Carpenter, 1992).

The second possible reason for why subject animacy is so important has to do with the fact that information about subject animacy is available to the language comprehension system earlier than information about object animacy in Chinese. In Mandarin object-extracted RCs, the subject is the first element inside the RC and thus information about its animacy is available before information about the object’s (head’s) animacy. Even in subject-extracted RCs, information about the subject’s animacy is available early: the sentence-initial verb provides information about the animacy of its upcoming subject. In contrast, information about object animacy only becomes available post-verbally when the object itself is encountered. In general, the object’s animacy is less dependent on the verb than the subject’s animacy; the computation of object animacy needs to take into consideration the information about both the subject and the verb.

As a whole, the results presented here highlight the important role of animacy in RC processing, and suggest that the relation between the animacy of the head noun and that of the embedded noun plays an important role in modulating the processing ease of subject-extracted vs. object-extracted RCs in Mandarin Chinese.

References


*Fuyun Wu, Elsi Kaiser & Elaine Andersen*  
*University of Southern California*  
*Department of Linguistics*  
*Los Angeles, CA 90089-1693*  
*fuyunwu@usc.edu*  
elsi.kaiser@usc.edu  
eanderse@usc.edu