Grammatical Roles of the Head Noun in Chinese Relative Clauses

Tao Ming
Concordia College

Combination of grammatical roles in relative clauses (RC) is determined by a multitude of factors (Fox and Thompson 1990, Pu 2007, Ming and Chen 2009). This study shows that four factors (discourse functions of RCs, grounding mechanism, information status, and animacy of head nouns) interact with one another to determine which combination is favored in a Chinese discourse.

1. Introduction

The past two decades or so have witnessed an increasing number of analyses on relative clauses (RC) demonstrating that distribution and structural properties of RCs can be attributed to a multitude of factors such as information flow in discourse, information status of head nouns, humanness of head nouns, discourse functions of RCs etc (Givon 1993; Fox 1984; Fox and Thompson 1990; Chen; 1995; Chu 1998; Tao 2002; Pu 2007, Ming and Chen 2009). Functional-pragmatic analyses on linguistic structures in general and on RCs in particular demonstrate that linguistic structures arise from the discourse need and there is an isomorphic relation between syntax and pragmatics. The object of this study is to show that the distribution of Chinese RCs and the combination patterns of grammatical roles of their head nouns can be approached from a semantic-pragmatic perspective by examining the interaction between the semantic properties of the head noun and discourse functions of their modifying RCs. It also seeks to explain what motivates language users to use a particular type of relative causes (RC) to modify a particular type of head nouns (NP).

Chinese RCs, unlike their English counterparts, always precede their head NPs. The following are several examples where the head noun is italicized and the relativized noun inside the RC is spelled out as a zero (i.e.∅). For our purpose, three grammatical roles are distinguished: subject (S), object (O), and others (X). We first discuss the grammatical of the relativized head NP inside the RC. Subject RC is used to name RCs where the relativized head noun is the subject of the RC, object RCs to name RCs in which the relativized head NP functions as the object of the RC and X RCs refer to a RC whose head NP does not serve as the core argument of the RC, i.e. neither subject role nor object role. Object RC, subject RC, and X RC are exemplified in (1a), (1b), and (1c) respectively.

(1)   a. Object RC: zero relativized NP as the object of the RC
          上面 抄录着我们都十分喜欢的一首小诗
      b. Subject RC: zero relativized NP as the subject of the RC
          这位 急于 离京 出走 的 男子 终于 低下 了 头.
      c. X RC: zero relativized NP as a non-core argument
          植树造林发展林业的确是关系国民经济和社会发展全局的一件大事。

The grammatical role of the head NP in the main clause is also coded. The relativized head NP which functions as the subject of the main clause is called subject head. In the same vein, a head NP which takes object role in the main clause is called object head. X head is utilized to name a head noun which is not a core argument in the main clause. Subject head, object head, and X head are illustrated in (2a), (2b), and (2c) respectively.

(2)   a. subject head: head NP as the subject of the main clause
          焦裕禄用生命绘制的那张蓝图今天已经成为兰考 大地 的 现实 。
      b. object head: head NP as the object of the main clause
          我 不会 忘记 那些 令 他 老人家 饮恨 千古 的 人。
      c. X Head: head NP as a non-core argument of the main clause
          在 国内念大学 时对那些当过兵再来念书的男生总是佩服得要命

Besides discussing the grammatical roles of the relativized head noun in the main clause and RC, we will also discuss their combination patterns. The combinatory patterns of the grammatical role in the main clause and that in the RC is represented as AB. For example, SS refers to a combination in which the relativized head noun is the subject in the main clause and the object in the RC. Three examples are presented in the following to illustrate some combinatory patterns of grammatical roles.

(3)   SS: subject head modified by a subject RC
          这位 急于 离京 出走 的 男子 终于 低下 了 头.
      SO: subject head modified by an object RC
          焦裕禄 用生命绘制的那张蓝图今天已经成为兰考大地 的 现实 。
      XS: X head modified by a subject RC
          在国内念大学时,对那些当过兵再来念书的男生总是佩服得要命.

Following previous studies (Fox 1984; Fox and Thompson 1990; Pu 2007), we only focus on core arguments in this study, that is, subject, object and their four combinatory patterns: SS, OS, SO, and OO.

2. Prior studies

Study of the impact of grammatical roles on relativization can be traced back to Keenan and Comrie (1977). According to the Noun Phrase Accessibility Hierarchy proposed by them, all languages conform to the following scale:

(4) Subject > Direct Object > Indirect Object > Obliques > Genitives > OComps

On this scale, if a language can relativize on a grammatical role lower on the scale, it can also relativize on the grammatical role/roles higher on the scale. For example, if a language allows for the relativization on an indirect object, it also permits the relativization on the direct object and subject higher in this scale. The focus of Keenan and Comrie (1977) is mainly on the structural properties of RCs and semantic properties associated with the head noun are not integrated in their study.

Fox (1984) argues against the Noun Phrase Accessibility Hierarchy and comes up with Absolute Hypothesis which states that if a language permits relativization, it at least must allow for relativization on P and S. The Absolute Hypothesis is based on her findings that subject RCs and object RCs tend predominantly to outnumber A (Agent) RCs where the relativized NP is the agent of the RC. The rarity of A RCs, according to her, can be attributed to the fact that A which in general is realized as a pronoun in conversational data is a better anchor than P which tends not to carry given information. In other words, grammatical roles alone can not account for the distribution of RCs in the discourse. Semantic properties associated with a grammatical role should be called into service to provide a better explanation of the distribution of RCs.

Givon (1993) discusses the role of semantic properties of the head noun. According to him, all referents must be grounded to make it relevant to the current discourse and RCs serve to ground the head NP. The information status of the head noun plays significant role in explaining the discourse function of the RC. If the head noun is definite and codes given information, the RC grounds the head NP anaphorically into preexisting mental structure. By contrast, if the head noun is indefinite and carries new information, the RC serves to cataphorically ground it to the subsequent discourse.

Fox and Thompson (1990) found that there is a remarkable skewed distribution of syntactic types of RCs in their conversation data. Their data shows that for nonhuman head noun the combination pattern SO is the dominant one and the combination pattern OO is seldom observed. For human head nouns, their discussion is limited to existential sentences where the head noun of the RC is mainly human and indefinite. The finding on human head noun is that subject RCs overwhelmingly exceed object RCs. According to Fox and Thompson, human head NP and nonhuman head NP are grounded differently. The former are mainly grounded by "being grounded by their own activities" whereas the latter by a given referent in the modifying RC. Their study convincingly shows that the distribution of syntactic types of RCs can be explained by exploring the discourse where they occur and that information flow in discourse plays a crucial role in determining the

syntactic types of RCs. Their discussion, however, does not cover the interaction between the information status of the head NP and the discourse function of the RC. In other words, they do not discuss the differences between given head NP and new head NP in terms of grounding and distributional patterns of RCs. What is more, human head NP in other position other than existential construction is not discussed.

Studies on RCs in Chinese have also attracted lots of attention from Chinese functional linguists (Chu 1998; Chen 1995; Chen 1997; Tao 2002; Pu 2007; Ming and Chen 2009). Among them, of particular interest to our study are Chen (1997) and Pu (2007) because both studies focus on the distributional patterns of Chinese RCs. Three patterns emerge from Chen's study. For nonhuman head nouns, SO is a favored choice in discourse. The second finding is that OO structure is also a preferred pattern for nonhuman head nouns. The third finding is that for human head nouns, subject RCs exceed object RCs regardless of the grammatical role of the head noun in the main clause.

Pu (2007) Studies various combination patterns of grammatical roles in Chinese RCs. She found that of the four possible combinations between S and O, SS is the most dominant one and SO is least frequent. OO combination is seldom observed in human subject heads and OS combination is rare in discourse. Three factors are reported to influence the choice of RCs. The first factor is a cognitive one, which states that object RCs are more marked than subject RCs. The second factor which influences the distributional patterns of RCs is a discourse-pragmatic one, i.e. the information status of head NP and the discourse function of RCs, and the third factor is the semantic properties of the head NP such as humanness, agentivity, saliency, and so on. Of the three factors, according to her, the first factor, i.e. markness, is the most important factor.

Previous studies on distributional patterns of RCs in Chinese have made great contribution to our understanding of the factors underlying the deployment of different syntactic types of RCs. The influence of information status of the head noun on the distributional patterns, however, has not been clearly spelled out and more researches are needed to provide a better understanding of the distributional patterns of Chinese RCs. This study hopes to make some contribution toward this end.

3. This study
3.1. Data

The data for this paper are extracted from a publicly available Chinese language corpus the Lancaster Corpus of Modern Chinese (McEnery et al. 2003). The Lancaster Corpus of Modern Chinese (LCMC), a one-million-word balanced corpus of written Mandarin Chinese, consists of five hundred 2,000-word samples of written Chinese texts selected from fifteen text categories published in Mainland China around 1991. LCMC provides web-based concordance search functionality, which greatly facilitates this research. The concordance results from LCMC always come with a complete sentence where the searched word occurs. The complete context where a RC occurs is examined

when it comes to determine the information status of the head noun and discourse functions of the RC.

3.2. Coding

Discourse oriented studies of RCs (e.g., Fox 1987; Fox and Thompson 1990; Givon 1993; Pu 2007) have identified various factors influencing the distribution of RCs. Of particular interest here are information status of the head noun, the animacy of head nouns containing a RC, grounding mechanisms, and discourse functions of RCs. In the following subsections, we will discuss the coding along the four dimensions:

- Information status of head nouns
- Animacy of head nouns
- Grounding mechanisms
- Discourse functions of RCs

3.2.1. Information status

The focus of this study is on the influence of the information status of the head noun on the distributional pattern of Chinese RCs. Therefore, it is not necessary to extract all RCs from the large corpus LCMC. Although Chinese does not have articles to index information status of a noun, it does provide linguistic clue as to where to find the head noun with different information status. As a result, we use a text analysis software Concordance (Watt, 1999) to extract all sentences where a demonstrative occurs and then eliminate all sentences where there lacks a RC. By doing so, we are able to extract head nouns which carry given information. In the same vein, with the help of Concordance, we extract all sentences where the numeral *yi* 'one' occurs and get rid of all sentences where there is no co-occurring RC. As a result, we succeed collecting RCs where the head noun encodes new information. Although it is a well established fact that the numeral *yi* 'one' is to index an indefinite noun which tends to be new and demonstratives such as *zhe* 'this' is to signal a definite noun which in most cases encodes given information, there is no absolute correlation between the information status of the head noun and their co-occurring linguistic units denoted by the numeral and the demonstrative. Sometimes it is possible to observe a mismatch between them.

3.2.2. Humanness

Following previous studies (Fox 1987; Fox and Thompson 1990; Pu 2007), humanness of the head noun containing a RC is also coded. Humanness of a referent has been shown to play a significant role in various studies. For example, Fox and Thompson (1990) observe that humanness of the head noun plays an important role in explaining the distribution of syntactic types of RCs in their conversation data in English. Pu (2007) also acknowledges the important role of humanness play in explaining the distributional patterns of Chinese RCs.

3.2.3. Grounding

We add grounding as one factor because grounding is closely related to animacy of head nouns and information status. Both Fox and Thompson (1990) and Pu (2007) include grounding as a crucial factor in accounting for the skewed distribution of different types of RCs. What is more, it is shown in Fox and Thompson (1990) that there is a positive correlation between grounding and discourse functions of RCs. To achieve effective communication, a speaker/writer presents new referents into the discourse in such a way as to make them relevant for the listener/reader at the point where they are introduced; and grounding is the primary way of making relevant NPs whose relevance is not clear from prior mention or situation (Fox and Thompson 1990). Following Fox and Thompson (1990), we focus on three kinds of grounding: anchoring, main clause grounding, and proposition linking. First, a new referent can be grounded through linking itself to a known referent in its modifying RC. The first way of grounding a new referent, according to Fox and Thompson (1990), is *anchoring,* is illustrated in example (5).

(5) 桌上留着朋友的一个留言，上面抄录着我们都十分喜欢的一首小诗。

In (5), the new referent 一首小诗 is grounded by the human subject 我们 "we" in the RC. The referent of the first person pronoun 我们 is a given one by virtue of the speaker's role as speech participant, and thus the RC containing the pronoun anchors the new referent 一首小诗, which is then made relevant to the current discourse through its connection to the given referent.

When the RC provides no grounding, the main clause can ground a new NP referent by relating it to a given discourse referent. In other words, a new referent can also be grounded by known information contained in a main clause. Fox and Thompson (1990) refer to this second way of grounding a new referent in the same main clause as the given referent as *main clause grounding*. This can be illustrated with example (6).

(6) １９８３年２月，张申府还以９０高龄写了一篇怀念罗素的文章。

The excerpt in (6) is preceded by a discussion of 张申府, which is a given referent in discourse. The new referent 文章 is grounded by the known subject 张申府 in the main clause and the modifying RC serves to characterize the head noun and provide additional information regarding the head noun.

The third way to ground a new referent is by means of *proposition-linking*, which, according to Fox and Thompson (1990), is to link an entity to given referents "by means of frames invoked in earlier discourse" or by way of a modifying RC which is linked to earlier proposition. Fox and Thompson (1990) provide the following as an example of proposition-linking (Fox and Thompson 1990, P. 109).

(7) B: Y'know I've been reading about people very old people lately,
     A: Yea//:h?
   B: Like they had an article in the Rolling Stone with this guy who's supposed to be a hundred and thirty. The oldest American. He is a black who lives in Florida and they interviewed him,…
   B: and one thing they said in the article that was really intriguing was, in the United States at this point, there are over a hundred thousand people [who are over a hundred years old]

In this example, the entire head NP referent *a hundred thousand people* is grounded by the RC *who are over a hundred years old* by means of *proposition linking*: the new referent *a hundred thousand people* is made relevant to the current discourse by the established link between the RC and the earlier proposition *I've been reading about people very old people lately*.

3.2.4. Discourse functions of RCs

Having introduced different grounding mechanisms, we will proceed to discuss the discourse functions of RCs in the two constructions. Fox and Thompson (1990, p. 301) identify two major types of RCs according to their functional roles: characterization and identification. In the first type, the RC provides a characterizing assertion or description of a new head NP referent in a particular discourse situation to supply additional descriptive information regarding the head noun. In the second type the RC makes the referent of a head NP relevant at a point in a particular discourse situation when it is first introduced. They use the contrast in (8) to illustrate the two discourse functions.

(8) a. *This man* [who I have for linguistics] is really too much.
    b. There's *a woman* in my class [who's a nurse].

While the RC in (8a) is used to ground the referent by virtue of providing a given referent *I* to anchor the new head referent *this man* and the RC in (8b) does not ground the referent; rather, it makes a characterizing assertion because the RC does not provide any *anchoring* given referent to identify the new referent *a woman*.

3.3. Statistics on Chinese RCS

Careful examination of the LCMC generates a total of 587 RCs. Of them, the subject RC overwhelmingly outnumbers the object RCs, by a ratio of 2 to 1. The following table presents different types of RCs

Table 1. Types of RCs

Types of RCs	Subject RC	Object RC	X RC
Percentage	360 (61%)	155 (26%)	72 (13%)

The preponderance of subject RCs over object RCs is compatible with previous studies on Chinese RCs. The ratio of subject RCs to object RCs in Chen (1997) and Pu (2007) approximates 3 to 1. In this study, the distribution of X RCs will not be discussed to get comparable data with previous studies. In the same vein, head NPs which do not take subject or object role is not considered either. After eliminating all those RCs whose head NPs occur in X position (i.e. neither subject nor object position), we get 434 RCs whose head nouns assume either a subject role or object role in the main clause as well as in the RCs.

We now present the distribution of RCs which co-occurs with a given head NP. Table 2 indicates that for a given head NP, regardless it is human (H) or not, subject RCs overwhelmingly exceeds object RCs. For a given human head NP, subject RCs are predominantly used to modify a subject head as evidenced by the high occurrence of SS (69.8% or 60 tokens). Object heads modified by a subject RC also makes up a sizable portion of the data (22.1% or 19 tokens). However, the combinations of OO and OS are rare in the data. For a given nonhuman (NH) head NP, a similar tendency is observed although the number of SS is decreased and that of OS is boosted. It is shown in table 2 that for a given nonhuman head NP, the most dominant combination pattern is OS which is slightly higher than SS.

Table 2. Given Information, Humanness, and Grammatical Roles of Chinese RCs

	SS	SO	OS	OO	TOTAL
H	60 (69.8%)	4 (5.8%)	19 (22.1%)	3 (3.4%)	86 (100)
NH	15 (30%)	7 (14%)	22 (44%)	6 (12%)	50 (100%)

Investigation of head NPs which carry new information shows different distributional patterns. For a new human head NP, subject RCs (45+54) overwhelmingly outnumber object RCs (3+2). The same tendency is not observed on new nonhuman head NPs. As shown in table 3, the number of subject RCs (15+81) is more or less the same with that of object RCs (25+73).

Table 3. New Information, Humanness, and Grammatical Roles of Chinese RCs

	SS	SO	OS	OO	TOTAL
H	45 (43.3%)	3 (2.9%)	54 (51.9%)	2 (1.9%)	104 (100%)
NH	15 (7.7%)	25 (12.9%)	81 (41.8%)	73 (37.6%)	194 (100%)

Table 3 shows that of the four possible combinations of grammatical roles, for new human head NPs, SS and OS predominantly exceeds SO and OO. However, for new nonhuman head, besides OS, OO also makes up a sizable proportion of the whole data and the combination pattern SS only accounts for a small portion of the data.

The data in table 2 and table 3 suggests that the combination of grammatical roles depends on the information status as well as humanness of the head noun. For human head NPs, subject RCs are favored regardless of the information status. By contrast, for nonhuman head NPs, information status plays a significant role in determining the use of

a particular type of RCs in discourse. If the information status is given, subject RCs are chosen over object RCs; if the information status is new, the preponderance of subject RCs over object RCs is not observed.

4. Discussion

The Findings presented in section 3 challenges some findings in previous studies (Fox and Thompson 1990, Pu 2007). The main focus of this section is show how information status interacts with humanness and discourse functions to determine the possible combination of grammatical roles.

4.1. Distribution of RCs modifying new nonhuman heads

According to Fox and Thompson (1990, P. 304), the nonhuman object heads do not tend to occur with object RCs. In other words, the combination OO is not expected for nonhuman head NPs.

Table 4. New Information, Humanness, and Grammatical Roles of Chinese RCs

	SS	SO	OS	OO	TOTAL
NH	15 (7.7%)	25 (12.9%)	81 (41.8%)	73 (37.6%)	194 (100%)

For nonhuman head nouns which encode given information, our data supports their observation as shown in Table 4. For nonhuman head NPs which carry new information, however, OO is one of the favored patterns (37.6% of the data). Our data also challenges the finding in Pu (2007) with regard to the OO combination. On her account, OO pattern is mainly observed on nonhuman head NPs and the information status of the head NP in OO is mainly given (see table 6 in Pu). To resolve the conflicting findings, we need to examine how a new nonhuman head NP is grounded in discourse. Functional linguists such as Fox and Thompson (1990) and Givon (1993) propose that all referents should be grounded to warrant its relevance to the current discourse. Givon (1993) further argues that a new referent differs from a given referent in terms of the way how it is grounded. For a new referent, it is grounded by the current text location because it can not be grounded by a previous mention or situation.

A careful study of the distribution of RCs containing a new nonhuman head shows that they predominantly occur in the object position of the main clause (see Table 3). The tendency for a new nonhuman head to occur in object position instead of in subject position is not surprising. The tendency for new nonhuman heads to be grammatical object has been well-established in several studies (DuBois 1987; Givon 1993; Fox 1984; Fox and Thompson 1990; Pu 2007). Non-humanness, newness, nontopicality are reported to be prototypical features associated with object position which predicts that a new nonhuman head containing a RC mainly occurs in object position. Unlike new nonhuman heads in English, which are mainly grounded by a given referent in the main clause, new nonhuman heads are grounded almost equally by RCs as well as by the main clauses. The observed difference of grounding for new nonhuman

heads can be attributed to the different positioning of RCs in English and Chinese. RCs in English differ from their Chinese counterparts in that modifying clauses precede head nouns whereas the opposite is true for their Chinese RCs.

The different positioning of RCs in the two languages has repercussions on the way how a new head is grounded. RCs in English do not tend to provide grounding for the object in on-line discourse processing because they are positioned after their modifying head nouns. As a result, new nonhuman head nouns are mainly grounded by a given referent in the main clause (Fox and Thompson 1990). What is more, the major way for a nonhuman head to be grounded is by virtue of human beings who own them, use them, and manipulate them (Du Bois 1980; Fox and Thompson 1990). Therefore in English the positioning of RCs and nonhumanness of the head noun conspire to prevent the occurrence of OO combination. Chinese RCs, by contrast, can serve to ground a new human head noun because they precede their modifying RCs. In on-line discourse processing, if the main clause does not provide a grounding referent, a RC can still fulfill the role of grounding by providing a grounding given referent. The sentence in (9) is to illustrate how a Chinese RC serves to ground a new nonhuman referent.

(9) 这 是 当时 张 作相 无法 解决 的 一 大 难题 。

The subject of the main clause 这 in (9) is a demonstrative which can not serve as a grounding referent. As a result, the RC serves the function of grounding by providing a given human referent. Investigation of the discourse shows that the OO combination is desired one for new nonhuman head noun because object RCs modifying object heads provide the necessary grounding for them to warrant their relevance to the current discourse. Chen (1997) also made the similar observation that OO combination mainly associates with nonhuman head nouns which carry new information. Her explanation, however, stand in striking contrast with the explanation provided in this paper. On her account, the new nonhuman head nouns in OO structure are mainly grounded by the subject of the main clause, i.e. main clause grounding, and the RC in OO structures mainly serves the discourse function of characterization.

Examination of the data extracted from the large corpus LCMC shows that the new nonhuman head noun is mainly grounded by the given referent in its modifying RC, therefore the main discourse function of the RC is to ground instead of characterizing the new nonhuman head noun. In other words, our finding regarding the deployment of OO structure is compatible with Chen (1997) although the explanation is different. The frequent occurrence of OO structure in the discourse is also reported in Pu (2007) and she explains the prevalence of OO structure in terms of humanness of the head noun. The argument is that the head noun in OO structure is mainly nonhuman and that nonhuman heads are mainly grounded by human beings who own them, use them, and manipulate them. As a result, the passive role played by the nonhuman being in relation to central human being is naturally realized by an object RC. Apparently, information status of the

head noun is not a factor to account for the occurrence of OO structure in her account. Investigation of our data suggests that information status is crucial for the prevalence of the OO structure for the nonhuman head noun and that for given nonhuman head nouns, the OO structure is not a favored choice in the discourse.

Having explained why the OO combination is a preferred one for new nonhuman head nouns, we now turn to answer why the combination OS (41.8% of the data or 81 tokens) is also a favored combination for them. Object heads occur in the later part of a sentence, besides being grounded by their modifying RCs, they may also be grounded by the given referent in the main clause, as in (10).

(10) １９８３年２月，张申府还以９０高龄写了一篇怀念罗素的文章。

The example in (10) is preceded by discussion of *张申府* (proper name), which is a given referent in discourse. The new nonhuman head referent *文章* is grounded by the known subject *张申府* in the main clause. As a result, there is no discourse need to have the RC *怀念罗素* to ground the new nonhuman object head *文章* because the grounding has been taken care of by the given human referent *张申府* in the main clause and the RC turns out to serve the discourse function of characterization by providing additional, descriptive information regarding the new nonhuman head noun. According to Fox and Thompson (1990), characterization is mainly realized by a subject RC, therefore the combination OS is also a preferred choice in the deployment of Chinese RCs.

Compared with the occurrence of OS and OO which are preferred structures in discourse, for new nonhuman head nouns, the occurrence of SO is rare (12.9% of the data or 25 tokens) and that of SS is even rarer (7.7% of the data or 15 tokens). We first answer why the combination SO is not a desired one for new nonhuman head nouns. In terms of grounding, SO is a favored combination (Fox and Thompson 1990), as illustrated in (11).

(11) 昆明 地区 彝族 桑尼帕 支系 １９８５年 搞 的 一 次 宗教 活动 很 能 说 明 问题。

The new nonhuman head *一次宗教活动* occurs in the subject position of the main clause. By the time it is introduced into the discourse, it is not grounded by the main clause because of its clause-initial position. As a result, the RC *昆明 地区 彝族 桑尼帕 支系 １ ９８５年 搞* serves to ground it by providing a given human beings *彝族 桑尼帕 支系* to warrant its discourse relevance. The question arises as to why SO is seldom observed in the data although it is a preferred combination in terms of grounding. We believe that the answer to this question lies in the information status and humanness of the head noun. It is well established on previous studies that humanness, giveness, saliency, and topicality are typical features associated with the subject position (DuBois 1987; Givon 1993; Fox 1987; Fox and Thompson 1990; Pu 1997; Pu 2007). It is not surprising that

nonhumaness and newness of the head noun discourages the occurrence of SO. The scarcity of SO for new nonhuman head nouns is consistent with previous studies (Chen 1997; Pu 2007). Chen observed that the SO structure is prevalent in the discourse and they mainly associate with nonhuman head nouns which carry given information. Pu (2007) also made the similar observation in her study.

Lastly we answer the question why the combination of SS is disfavored in the discourse. We believe that for new nonhuman head nouns, the rarity of SS can be attributed to two factors: 1) the mismatch between subject position and the newness and nonhumaness of the head noun; 2) the discourse functions of subject relatives. It is reported in previous studies that new, nonhuman head nouns are discouraged to occur in the subject position because it is mainly reserved for human, given, salient referents. Therefore new nonhuman head nouns are not expected to occur in the subject position of the main clause. On the other hand, According to Fox and Thompson (1990), the main discourse function of subject RCs is to characterize its head noun. A new nonhuman head noun occurring in subject position of the main clause needs to be grounded by its modifying RC to justify its discourse relevance to the current discourse. However, subject RC can not fulfill such discourse requirement.

4.2. Distribution of RCs modifying given nonhuman heads

The data with regard to RCs modifying given nonhuman head shown in Table 5 shows that for given nonhuman head nouns, the favored structures are OS (44% of the data or 22 tokens) and SS (30% of the data or 15 tokens) and the disfavored ones are SO (14% of the data or 7 tokens) and OO (12% of the data or 6 tokens).

Table 5. Given Information, Humanness, and Grammatical Roles of Chinese RCs

	SS	SO	OS	OO	TOTAL
NH	15 (30%)	7 (14%)	22 (44%)	6 (12%)	50 (100%)

The findings from our data challenges the finding in Chen (1997) where it is reported that for given nonhuman head nouns, OO and SO are the dominant patterns but supports Pu (2007) where it shows that SO is not a preferred choice. We believe that the different grounding mechanisms in relation to head nouns with different information status help to resolve the conflicting findings. According to Fox and Thompson (1990) and Givon (1993), all referents must be grounded to make it relevant to the current discourse. Givon (1993) further argues that a new referent differs from a given referent in terms of the way how it is grounded. For a new referent, it is grounded by the **current text location** because of the fact that it can not be grounded by a previous mention or situation. If a new referent is modified by a RC, the modifying RC tends to provide the grounding information because the RC occurs in the current text location of its modifying head owing to its proximity with it. By contrast, for a given referent, it is mainly grounded by **other text location** in previous discourse by virtue of its previous mention

or a frame established in prior discourse and the RC in general does not serve to ground the given head referent because it is already established in previous discourse. As a result, RCs modifying a given head referent are not deployed to provide grounding information but to characterize the given head noun by providing additional, descriptive information (Fox and Thompson 1990). For example,

(12) 这 场 涉及 到 家家户户 切身利益 的 重大 改革 ， 深深 牵动 着 每个 职工 和 家属 的 心。

The prior discourse in (12) centers around the discussion of the reform, therefore, by the time the head noun *重大改革* which occurs in the subject position of the RC is introduced in the discourse, it has been grounded by the previous discourse through the frame established. Consequently, there is no discourse need for the RC to ground it and the RC turns out to serve the discourse function of characterization.

Having shown that the SS combination is a preferred choice in discourse for given nonhuman head nouns, we now turn to the other favored combination OS. We believe that the explanation of the prevalence of SS can also apply to account for the favored choice of SO. The sentence in (13) shows an example where a given nonhuman object head *报道* is modified by a subject RC *赞扬 日本 老师.* The discourse prior to (13) is about a report where a Japanese teacher tries to boost his students' national pride by counting the number of Japanese cars passing a street intersection in a China. In other words, the head noun *报道* carries given information and its relevance to the current discourse has been well established in the previous discourse. Therefore there is no discourse motivation for the modifying clause to ground it. As a result, the RC is used to characterize the head noun and that is the reason why a subject RC is used.

(13) 我看 了 那 篇 赞扬 日本 老师 的 报道。

In short, subject RCs are chosen object RCs for a given nonhuman head noun owing to the information flow and the different discourse functions the two types of RCs play in the discourse. A related question arises as to why the number of SO approximates that of OS. We believe that the answer to this question is related to the interaction of the semantic properties the head noun and discourse requirements on the grammatical roles of a sentence. Subject position tend to associates with identifiable, given, specific human beings whereas object slot is reserved for new, nonspecific, unidentifiable nonhuman referent (DuBois 1987; Givon1993; Fox 1987; Fox and Thompson 1990; Pu 1997; Pu 2007). A given nonhuman head noun does not fit either of the two roles. Its givenness makes it a less than prototypical object while its humanness makes it a less than subject. Therefore it is expected that given nonhuman heads straddle across both grammatical roles.

Our attention now turns to the rarity of SO and OO in the discourse. Grounding function, according to Fox and Thompson (1990), tend to fulfilled by object RCs where a given human subject mainly serves to ground the head noun. As shown in previous discussion, a given nonhuman head need not to be grounded by the RC because it is already grounded by its previous mention or a frame established in prior discourse. For this reason, its modifying RC tends to serves the discourse function of characterization and characterization is mainly fulfilled by subject RCs (Fox and Thompson 1990). The combination of SO and OO, therefore, is not expected to be preferred choices in the discourse. For given nonhuman head nouns, the scarcity of SO is also corroborated in Pu (2007) where it is reported that OO structure mainly associates with new nonhuman head nouns and the combination of SO mainly used in conjunction with given nonhuman heads is seldom observed. Pu's explanation of the rarity of SO differs from the explanation offered by us. She argues that four factors conspire to the rare occurrence of SO. Firstly, object RCs are marked in the sense that it produces marked structures [S V ∅]; secondly the modifying RC is less informative because it provides repeated and redundant information; thirdly 'the definiteness, givenness, and topicality not only discourages a modifying RC but also disfavor the coding of a full NP' (Pu 2007, P. 49); lastly a subject slot is not a preferred position for the coding of nonhuman heads. According to Pu (2007), the first factor is the most important one. We, however, do not believe the object RCs are more marked than subject RCs. The reason is that the zero form resulting from relativization is different from the zero form in the main clause. The zero form in the main clause is to substitute frequently occurring referent conforming to "the light subject constraint" proposed by Chafe (1994) and that zero form is seldom observed in object position in Chinese discourse (Chu 1998; Pu 1997).

Zero forms resulting from relativization, however, are definitely different from zero forms in the RCs in terms of frequency of occurrence because relativization engenders zero forms regardless of the grammatical role of the relativized noun. For example, if the relativized noun is the object inside the RC, a zero form in object position is obligatory. In the same vein, a zero form in subject position is also mandatory if the subject position is relativized. In other words, it is hard to say that the zero form in the subject position is more frequent than that in the object position in RCs. Pu proposes that subject RCs are easier to process than object RCs in Chinese and that is the reason why there is a preponderance of subject RCs over object RCs. The research by Hsiao and Gibson (2003), however, showed that "contrary to the patterns found in all other languages, Chinese RCs showed a processing preference for object extractions" What is more, according to the Markedness principle, subject RCs are always more frequent than object RCs regardless of information status of the head noun. The data in Table 3, however, shows that for new nonhuman head nouns, subject RCs (81+15) do not exceed object RCs (73+25).

4.3. Distribution of RCs modifying given human head nouns

The data in Table 6 shows that for given human head nouns, the favored patterns are SS (69.8% of the data or 60 tokens) are (22.1% of the data or 19 tokens) and the two disfavored patterns are SO (5.8% of the data or 4 tokens) and OO (3.4% of the data or 3 tokens).

Table 6. Given Information, Humanness, and Grammatical Roles of Chinese RCs

	SS	SO	OS	OO	TOTAL
H	60 (69.8%)	4 (5.8%)	19 (22.1%)	3 (3.4%)	86 (100)

The other way to interpret it is that subject RCs predominantly outnumber object RCs. For given human head nouns, the preponderance of subject RCs over object RCs can be attributed to two factors. Firstly a given human head nouns does not need to be grounded by its modifying RC, which greatly reduces the occurrence of object RCs because object RCs are mainly used to ground their head nouns (Fox and Thompson 1990). Secondly, given human nouns tend to be deployed in subject positions of main clauses as well as subject RCs. Therefore, the pattern of SO and OO are disfavored for a given human head. Excerpt in (14) presents an example to illustrate how a given human head is grounded.

(14) 母亲 则 于 心灵 深处 对 幼子 怀 着 羞怯 而 不可 明 言 的 指望 ， 相
　　　信 这个 *不 说话 而 贪 食 的* **孩子** 终究 会 大 有 前途 。

The head noun 孩子 in (14) functions as the subject of the subordinate clause. It codes old information because it is introduced into the previous discourse as 幼子. What is more, its previous mention 幼子 is immediately adjacent to the head noun 孩子 of the RC. It is apparent that there is no discourse need to ground the given head referent at the moment it is reintroduced into the discourse (Givon 1993) because its identity and relation to the current discourse is well established in the prior discourse. As a result, the RC serves the discourse function of characterization to provide additional descriptive information. Therefore the association of a given human head with a subject RC is an expected tendency.

The next question is why the SS patterns are greater than the OS patterns although both of them are favored choice in discourse. The answer, we believe, lies in the semantic properties of the head noun. As a given human head noun, it is supposed to occur in subject position of the main clause because subject slot is mainly reserved for identifiable, given, human referent. Therefore the OS combinations are expected to lower than the SS patterns because the OS pattern results in a mismatch between the semantic properties of object head and the discourse requirements of the object slot which mainly associates with new, nonhuman referent.

4.4. Distribution of RCs modifying new human head nouns

Table 7 summarizes the distributional patterns of RCs and shows that the two most favored patterns are OS and SS. In contrast, SO and OO are strongly discouraged to occur in the discourse.

Table 7. New Information, Humanness, and Grammatical Roles of Chinese RCs

	SS	SO	OS	OO	TOTAL
H	45 (43.3%)	3 (2.9%)	54 (51.9%)	2 (1.9%)	104 (100%)

Contrary to Pu (2007) where SS is predicted to be the most dominant pattern for new human head nouns, SS is not found to be the most dominant one although it is a preferred it one. As shown in Table 7, OS structures are slightly greater than SS structures. The question arises as to why new human head nouns behave in a way similar to given human head nouns considering the fact that SS and OS are also the two favored combinations for given human head nuns. Can we apply the same explanation to account for the behavior of new human head nouns? The answer, we believe, lies in the way how a new human being is grounded. Givon (1993) argues that a new referent is grounded differently from a given referent. The former is cataphorically linked to the subsequent discourse via the modifying RC whereas the latter is grounded by its previous mention or a frame established in earlier discourse. Fox and Thompson propose that a human being is grounded from a nonhuman referent. Nonhuman referents are in general grounded by a given human referent either in the main clause (i.e. main clause grounding) or a given human referent in the RC (i.e. anchoring) who own, use, manipulate it. Human being referents, by contrast, do not need to be grounded by other human beings. Instead they, according to Fox and Thompson (1990), tend to be grounded by their own activities, which naturally produces subject RCs. That is the reason why for a new human head noun, subject RCs (45+54) predominantly outnumber object RCs (3+2), which naturally disfavors the occurrence of SO and OO structures. For example

(15)   一 位 架子工 出身 的 局长 接受 记者 的 采访。
(16)   一 位 头 戴 箬 帽 、 挂 着 竹竿 的 老汉 ， 噙 着 热泪 紧握 副 省长 的
       手 说。
(17)   一个 穿 绿色 T恤 的 矮个子 插 到 前面 了

The human head nouns in these three examples are typical of the SS structures in our data. They are introduced into the discourse for the first time and thence can not be anaphorically grounded by their previous mention or a frame established in prior discourse. As a result, their modifying RCs function to provide grounding or distinguishing information to help the language receiver to single out the new referent and establish its relevance to the current discourse (Pu 2007). The human referents are grounded by their modifying RCs depicting its identity in (15), describing its appearance

in (16), and stating its dressing style in (17). All those grounding RCs describing a human being's activity or properties are naturally subject RCs. The other way to interpret it is that object RCs can not fulfill the function of describing a human being's behavior or properties. Once again, object RCs are not predicted to be normal pattern in the discourse, as concluded in previous researches (Fox and Thompson 1990; Chen 1997; Pu 2007).

We now revert to answer why SS structure parallel OS structures in terms of frequency of distribution. The following present two examples of OS structures.

(18)   采访 了 几 个 从 数十 里外 赶到 这里 参加 秋播 的 农机 专业户
(19)   笔者 问 一 位 被 挤 下 阵 来 的 中年 妇女 抢 什么

The human head referents in these two sentences carry new information because they are introduced into the discourse for the first time. Their modifying RCs ground them by describing their activities. The reason for the equal distribution of SS and OS structures, we believe, is related to the prototypical associations of a grammatical role. It is well established that subject role tend to associate with giveness and humanness and object role tend to associate with newness and nonhumanness. A new human referent, however, fit neither of them. As a human referent, it is expected to occur in subject position; as a new referent, it is supposed to occur in object position. The mismatch between the new human head noun and its associating grammatical role dictates that there is no single strategy to deploy it (Fox and Thompson 1990). That is the reason why a new human head referent is more or less equally across the subject position and the object position.

5. Conclusion

We hope that we have succeeded in our efforts to explain various distributional patterns of RCs. It has shown in this paper that information flow, semantic properties of the head noun such as information status and humanness, grounding, and discourse functions of RCs all play a role in explaining the distribution of RCs. The paper has made several important findings which challenge previous studies on the same topic: 1) OO is favored pattern for new nonhuman heads; 2) SS is mainly associated with given human head nouns; 3) Information status is of vital important in the explanation of combinatory patterns of grammatical roles. This study also shows that syntactic constructions are motivated in large part by functional considerations.

REFERENCES

Chafe, Wallace. (1987). Cognitive constraints on information flow. In R. S. Tomlin (Ed.), Coherence and grounding in discourse (pp. 21–51). Amsterdam: Benjamins.

Chafe, Wallace.1993. Discourse, Consciousness and Time: The Flow and Displacement of Conscious Experience in Speaking and Writing, U Chicago Press

Chen, Ping. (1995). Pragmatic interpretation of structural topics and relativization in Chinese. Journal of Pragmatics, 26, 389–406.

Chen, Shu-Ching Susan. (1997). A discourse analysis of RCs in spoken Chinese: A study based on grammatical reflexes and information structure. Unpublished master's thesis, National Taiwan University.

Chu, Chauncey. (1998). A discourse grammar of Mandarin Chinese. New York: Peter Lang.

DuBois, John. W. (1987). The discourse basis of ergativity. *Language*, 63, 805–855.

Keenan, Edward. & Comrie, Bernard. (1977). Noun Phrase Accessibility and Universal Grammar. Linguistic Inquiry, 8:63-99.

Barbara A. Fox. 1987. The noun phrase accessibility hierarchy reinterpreted: subject primacy or the absolutive hypothesis?. *Language*, Vol. 63, 856-870.

Fox,BarbaraA., Thompson, Sandra A., 1990.Adiscourse explanation of the grammar of relative clauses inEnglish conversation. Language 66, 297–316.

Givón, Talmy. (1983). Topic continuity and word order pragmatics in Ute. In T. Givón (Ed.), Topic continuity in discourse: Quantitative cross-language studies (pp. 343–363). Amsterdam: Benjamins.

Givón, Talmy. 1993. English Grammar: a Function-Based Introduction. vol. 1 & 2. John Benjamins Publishing Company.

Hsiao, Franny., & Gibson, Edward. (2003). Processing RCs in Chinese. Cognition, 90, 3-27.

McEnery, Anthony, Xiao, Zhonghua, Mo, Lili, 2003. Aspect marking in English and Chinese: using the Lancaster Corpus of Mandarin Chinese for contrastive language study. Literary and Linguistic Computing 18, 361–378.

Ming, Tao and Chen, Liang. 2009. A discourse-pragmatic study of the word order variation in Chinese relative clauses. *Journal of Pragmatics*. doi:10.1016/j.pragma.2009.05.023.

Pu, Mingming. 2007. The distribution of Chinese relative clauses. *Discourse processes* 43(1): 25–53.

Tao, Hongyin. 2002. Hanyu kouyu xushiti guanxi congju de yuyi he pianzhang shuxin[The semantic and textual characteristics of RCs in Chinese spoken narratives]. *Contemporary Research in modern Chinese* 4: 47-57.

Watt, Rob J.C., 1999. Concordance-text analysis software (1st version). http://www.concordancesoftw