

## RESEARCH

## Open Access



# Grammatical roles, Coherence Relations, and the interpretation of pronouns in Chinese

Andrew Simpson<sup>1\*</sup>, Zoe Wu<sup>2</sup> and Yan Li<sup>3</sup>

\* Correspondence:

[andrew.simpson@usc.edu](mailto:andrew.simpson@usc.edu)<sup>1</sup>Department of Linguistics,  
University of Southern California,  
Los Angeles 90069, CA, USAFull list of author information is  
available at the end of the article

## Abstract

This paper reports on an experimental study of the interpretation of pronouns in Chinese which provides additional support for the proposal in Yang et al. (1999, 2003) that the resolution of pronominal reference in Chinese is more influenced by syntactic information than often assumed in approaches to discourse anaphora in Chinese such as Li and Thompson (1979), Givon (1983), Chen (1986), Christensen (2000), and Pu (2011), where the interpretation of such elements is solely attributed to semantic, pragmatic, and discourse structure-related factors. The paper makes use of a series of sentence completion tasks, adapted from Kehler and Rohde (2013) for Chinese, to try to tease apart the often complex roles played by syntactic position, Coherence Relations, and discourse structure.

**Keywords:** Pronouns, Coherence Relations, Discourse anaphora

## 1 Background

This paper reports on an experimental study of the interpretation of pronouns in Chinese which provides additional support for the proposal in Yang et al. (1999, 2003) that the resolution of pronominal reference in Chinese is more influenced by syntactic information than often assumed in approaches to discourse anaphora in Chinese such as Li and Thompson (1979), Givon (1983), Chen (1986), Christensen (2000), and Pu (2011), where the interpretation of such elements is solely attributed to semantic, pragmatic, and discourse structure-related factors. The paper makes use of a series of sentence completion tasks, adapted from Kehler and Rohde (2013) for Chinese, to try to tease apart the often complex roles played by syntactic position, Coherence Relations, and discourse structure. In strengthening the case for a facilitating role of syntax among other discourse-related processes of anaphora resolution in Chinese, the paper extends Kehler and Rohde's experimental approach, initially developed with English, to a pro-drop language which has substantial use of pronouns, and shows how realized pronouns in such a language also exhibit a subject bias in their interpretation, in addition to being governed in their distribution by their relation to pro, full noun phrases/ NPs and structural aspects of discourse.

The structure of the paper is as follows. Section 2 reviews previous work carried out on principles governing the interpretation of overt pronouns in Chinese, highlighting salient ideas which have emerged in prominent functional and experimental approaches to the

use of pronouns in Chinese. Sections 3, 4, 5, and 6 then each describe one of three sentence completion experiments carried out to further probe the interpretation of pronouns. Experiment 1 first identifies a regular subject bias on the location of antecedents for pronouns in connected discourse. Experiment 2 then investigates the interpretation of pronouns in a particular context, using transfer of possession verbs, and notes changes in the subject antecedent bias caused by the manipulation of verb choice and aspect. This leads on to Experiment 3, which specifically controls for the influence of discourse structure/Coherence Relations on pronoun interpretation, and shows how the results of the experiment, in conjunction with those of the previous experiments, offer a way to distinguish the influence of discourse structure from syntactic role in the interpretation of realized pronouns in Chinese. Section 7 summarizes and further discusses how the findings of the paper add to ongoing debates on the use and interpretation of discourse anaphora in Chinese.

## 2 Previous studies of zero and overt pronoun use in Chinese

An early prominent paper on the use of discourse anaphora, Li and Thompson (1979), offers two potentially important generalizations concerning the distribution of zero and realized pronouns in Chinese, which have relevance here. First, making use of data from two historical texts of Chinese (argued to display patterns also found in modern Mandarin), Li and Thompson suggest that the interpretation of discourse anaphora is not constrained by syntactic factors such as the syntactic roles of referent and antecedent NP, but rather driven by purely semantic and pragmatic forces, and that zero pronouns may identify an antecedent in any kind of syntactic position when this is licensed by an appropriate semantic relation. Second, it is proposed that realized/overt pronouns are rather exceptional in their occurrence in Chinese and that the use of zero pronouns is the rule, rather than any overt counterparts. Li and Thompson thus present Chinese as a pro-drop language in which anaphoric relations are determined by purely non-syntactic factors, and where there is apparently only marginal use of overt pronouns.

While Li and Thompson (1979) has become a much-referenced work on Chinese discourse anaphora, subsequent studies of referential forms have questioned its depiction of Chinese as a language not making any significant use of overt pronouns and have reported substantial use of such forms, particularly so in spoken Chinese. Both Chen (1986) and Christensen (2000) found an almost equal occurrence of overt and zero anaphora/pro in oral narratives produced by recounting the pear stories in Chinese (Chafe 1980), and the robust use of overt pronouns has been described in many other works (e.g., Tai 1978; Pu 2011; Tao 1996; Tomlin and Pu 1991), which have set about identifying discourse conditions which may cause alternations between the use of zero and overt pronouns<sup>a,b</sup>.

One general, functional account of the choice of referential forms in discourse (zero anaphora, overt pronouns, or full NPs) which has been considered for Chinese is the approach outlined in Givón (1983), sometimes referred to as the “Distance Model” or the “Recency Model.” This approach assumes that the more accessible a referent is within a discourse, the less overt coding it will be given, hence that highly accessible antecedents will be referenced with zero anaphora, less accessible antecedents with (overt) pronouns, and very weakly accessible referents with the use of a full NP. Importantly, the accessibility of a referent is seen to be determined by the simple distance which exists in a discourse between mentions of the referent, as well as the possible

intervention of other potential antecedents. Chen (1986) presents experimental data which initially appears to provide good support for the Distance Model in Chinese, but then rejects such an approach on the grounds that it fails to capture other important aspects of the linking of anaphora to antecedent NPs. Chen (1986: 279–280) (and later Pu 2011) argues that the Distance Model over-emphasizes the linear nature of discourse and in doing so is unable to account for instances of “long-distance pronominalization” (where there is a considerable distance between pronoun and antecedent, unexpected in the Distance Model), and occurrences of “short-distance nominalization” (where an explicit full NP form is used to refer to an antecedent even though there is little distance separating the two referents, and these may even occur in adjacent sentences).

A different direction of investigation is explored in Chen (1986), Tai (1978), Pu (1995, 2011), and, to some extent, in Li and Thompson (1979), following a general structural model of discourse and the use of anaphora developed in Hinds (1977, 1979) and Fox (1987, 1996). These works suggest that major and minor discontinuities in discourse structure cause alternations in the different anaphoric forms of reference that speakers regularly use. Zero anaphora/zero pronouns are described as typically being used when a topical referent remains in focal attention across many sentences in succession in a “topic chain,” in which the sentences all “cohere” well and describe a closely related sequence of events (Chen 1986), or are otherwise semantically closely linked with each other (Tai 1978)<sup>c</sup>. Overt pronouns are argued to occur when anaphor-antecedent relations are structurally interrupted by minor breaks and discontinuities in discourse structure, caused by changes in temporal, spatial, or action continuity, from transitions in description of a referent’s physical activities to his/her mental activity, and from switches in narration from background to foreground information (Pu 2011). Such disruptions are suggested to affect speakers’ attention on a referent, lowering its activation level and favoring the use of a more explicit form of reference—an overt pronoun rather than zero anaphora. More major breaks in the episodic structure of a discourse/narrative, often corresponding with paragraph breaks in writing, are described as resulting in the use of repeated full NPs to refer back to topical referents in a story line. The selection of elements from within the hierarchy of forms *zero pronouns* > *overt pronouns* > *NPs*, taken to reflect a referent’s current activation level, is therefore suggested to directly result from the *structural* rather than purely *linear* organization of discourse, allowing for a more successful account of instances of long-distance pronominalization and short-distance nominalization<sup>d</sup>. What remains challenging in such an approach is an accurate characterization of the kinds of narrative conditions which result in a sequence of clauses/sentences being perceived to be sufficiently “semantically closely linked” to form a single unit of discourse. The following examples from Tai (1978) are an illustration of some of the subtlety observed in choice of anaphoric forms as caused by the connection between clauses/sentences. In example 1, a zero pronoun/pro is suggested to be preferred to an overt pronoun (though the latter is not unacceptable), as the two clauses are felt to have a tight connection with each other, whereas in 2 there is less of a close connection perceived in the juxtaposed clauses, and an overt pronoun is suggested to be preferred to the use of pro/a zero pronoun:

- (1) 小梅结了婚,√*pro*?她生了两个孩子。(Tai 1978)  
 xiao\_mei\_jie-le\_hun\_ta\_sheng-le\_liang-ge\_haizi  
 little\_mei\_connect-ASP\_marriage\_pro/she\_bear-ASP\_two\_CL\_child  
*Little Mei got married and had two children.*
- (2) 老张昨天来看我,√他/?*pro*带了一本杂志给我。(Tai 1978)  
 lao\_zhang\_zuotian\_lai\_kan\_wo\_ta\_dai-le\_yi-ben\_zazhi\_gei\_wo  
 old\_zhang\_yesterday\_come\_see\_me\_he/pro\_bring-ASP\_one\_  
 CL\_magazine\_for\_me  
*Yesterday old Zhang came to see me and he brought me a magazine.*

It is also acknowledged that certain speaker variation occurs in the choice of anaphoric reference form, as the result of different perceptions of the semantic connectedness of clauses and sentences within a narrative.

The above-mentioned approaches to the analysis of discourse anaphora in Chinese all have in common that the linking of discourse referents with their antecedents is driven by semantics/meaning in tandem with discourse structure and the centrality/topicality of referents within a discourse, and is not a direct function or product of syntactic structure in any way. The potential influence of syntax on the choice and interpretation of referential forms is, however, investigated and argued for in Yang et al. (1999, 2003), two experimental studies of zero and overt pronouns in Chinese. The first of these works, Yang et al. (1999), considers how reading time is affected by the manipulation of different types of referential forms—repeated names, overt pronouns, and zero pronouns/*pro*. The study showed that a slow-down occurs when a name is repeated as an anaphoric referential form, and this effect is determined by the syntactic position of the repeated name, occurring when the repeated name occurs in subject position but not object position. A sample stimulus set from Yang et al. is illustrated in 3 below:

- (3) Sample stimuli from Experiment 1 (Yang et al. 1999: 720)
- Initial sentence  
 小明告诉小玲这一带住宅区不安全。  
 xiao\_ming\_gaosu\_xiao\_ling\_zhe\_yi\_dai\_zhuzhaiqu\_bu\_anquan  
 little\_ming\_tell\_little\_ling\_this\_place\_residential-area\_not\_safe  
*Xiaoming told Xiaoling that the residential area in this neighborhood is not safe.*
- Second sentence  
 小明/他警告小玲/她平常随时提高警觉。  
 xiao\_ming/ta\_jinggao\_xiao\_ling/ta\_pingchang\_suishitigao\_jingjue  
 little\_ming/he\_warn\_little\_ling/her\_regularly\_always\_heighten\_vigilance  
*Xiaoming/He warned Xiaoling/her to keep vigilant all the time.*
- Passage-final sentence  
 提高警觉是预防意外的不二法门。  
 tigao\_jingjue\_shi\_yufang\_yiwai\_de\_bu'er\_famen  
 heighten\_vigilance\_be\_prevent\_accident\_DE\_only\_way  
*Keeping alert is the one and only way to avoid accidents.*

Yang et al. note that this result suggests that syntactic information does indeed play an important role in the processing of co-reference at a local discourse level, providing potential evidence against the “context-and-pragmatics-only” approach proposed by Li and Thompson (1979). A follow-up experiment focused on whether the slow-down effect was related to the position of the antecedent NP. It was found that the repeated name slow-down occurred only when the name itself was in a subject position and it referred back to an antecedent expression that was also in a subject position in a preceding clause/sentence. There was no slow-down when the repeated name referred to an object in a preceding clause/sentence. Yang et al. argue that this again emphasizes that there is a syntactic component to the use of referential forms in discourse and that processing is sensitive to the syntactic positions of antecedents and also that of anaphoric elements.

Yang et al. (2003) describes an experiment that involved self-paced reading, where it was found that participants read their way faster through a short narrative when the pronoun subject of one sentence matched the subject rather than the object of a preceding sentence, even when the meaning of the sentences naturally favored co-reference with the object. Yang et al. argue that this indicates that the prominence of referents in a discourse, and hence the assignment of antecedents to discourse anaphora, is “more sensitive to syntactic factors than was previously recognized” (Yang et al. 2003:286), and it is not simply meaning that guides the processing of discourse anaphora, contra Li and Thompson (1979)<sup>e</sup>.

With all of the above as background on work previously carried out on the distribution and processing of zero and overt pronouns in Chinese, the present paper will now present a series of experiments focused on the interpretation of overt pronouns (henceforth simply referred to as pronouns) which provides further support of a different experimental type for Yang et al.’s position that syntactic information does play a role in the resolution of pronominal anaphoric reference relations in Chinese, in addition to other properties of discourse structure. The present study considers specifically configurations of *ambiguity* with multiple compatible antecedents for a pronoun in a discourse, which are not focused on in other studies, and show how the ways this ambiguity is typically resolved reveal biases of interpretation for pronouns that relate to syntactic position. It will in fact be shown that both discourse structure, in the form of Coherence Relations (Hobbs 1979), and syntactic structure combine to strongly bias the actual interpretation of pronouns and that both components of language are important for the determination of anaphoric relations. Hence, while we acknowledge the important insights of Chen (1986), Pu (2011), and others concerning the potential role that discourse structure and semantics may have in shaping the use and interpretation of pronouns, we argue, along with Yang et al. (1999, 2003), that the processing of pronouns also accesses syntactic information and not simply aspects of discourse structure and meaning.

### 3 Experiment 1: an initial probe into pronoun-antecedent relations

The first experiment initiated in this investigation of pronoun interpretation in Chinese was a sentence completion task, modeled on a similar experiment described in Kehler and Rohde (2013) (henceforth regularly referred to as KR). The 27 participants in the experiment, who were all adult native speakers of Mandarin Chinese from northern China, were asked to write completions for a set of short discourses, in which the first word of the second sentence was provided as a prompt, as seen in examples 4 and 5

below. In all of the 56 discourses, the first sentence contained two referents which could potentially serve as antecedents for pronouns appearing in the second sentence, one in subject position and one in a non-subject position (direct object, indirect object, or object of preposition/co-verb/把 *ba* “BA”). In the 34 target discourses, the second sentence regularly began with a third person pronoun prompt, which in principle could be interpreted as referring to either of the two noun phrase referents in the preceding sentence, or possibly some other unmentioned individual, as illustrated in examples 4a–b. The remaining 22 filler/distractor discourses had prompts that were either first/second pronouns or adverbials, as exemplified in 5a–b. The focus of the experiment was to see how often the third person pronoun in the completed target sentences might refer to the subject or the non-subject noun phrase in the preceding sentence. In the instructions given to participants, participants were not asked to interpret any instance of the third person pronoun 他 *ta* “he” as necessarily referring to an individual in the preceding sentence, and participants were asked to simply provide any narrative completion that naturally came to mind, with no restriction.

- (4) a. 小李去见老马了。他...  
 xiao\_li\_qu\_jian\_lao\_ma\_le\_ta  
 little\_Li\_go\_see\_old\_Ma ASP\_he  
*Little Li went to see Old Ma. He...*
- b. 老板刚刚把厨师开除了。他...  
 laoban\_ganggang\_ba\_chushi\_kaichu\_le\_ta  
 owner\_just\_BA\_chef\_fire-ASP\_he  
*The owner just fired the chef. He...*
- (5) a. 这件衬衫太小了。你...  
 zhe\_jian\_chenshan\_tai\_xiao\_le\_ni  
 this\_CL\_shirt\_too\_small ASP\_you  
*This shirt is too small. You...*
- b. 昨天我参加了一个朋友的婚礼。没想到...  
 zuotian\_wo\_canjia-le\_yi\_ge\_pengyou\_de\_hunli\_meixiangdao  
 yesterday\_I\_attend-  
 ASP\_one\_CL\_friend\_DE\_wedding\_unexpectedly  
*Yesterday I went to a friends' wedding. To my surprise...*

Once participants had finished the sentence completion task, each target sentence was coded by two native speakers of Mandarin, who considered how the meaning of the completed sentences formed a coherent discourse with the preceding (provided) sentence, and noted whether the pronouns seemed to be being used to refer back to the subject noun-phrase in the preceding sentence, or the non-subject noun-phrase, or some other individual. Target sentences where coding from both native speakers coincided were maintained in the database, and instances where there were disagreements between the native speaker coders over which referent was the intended antecedent for the pronoun were discarded. From an initial total of 918 target sentence pairs, this produced a total of 875 tokens of sentence pairs in which the native speaker coders agreed on which noun phrase in the first sentence was likely to be the intended antecedent for



the pronoun beginning the second sentence, as exemplified in 6a below. In the discarded 43 sentences, there was either no agreement between the two coders on which of the referents in the first sentence was the intended antecedent, as for example in 6b below, or participants' completions began with the pluralizing morpheme 们 *men* which converted the element 他 *ta* "he" provided in the completion into a plural pronoun 他们 *tamen* "they," so that this subject referred back jointly to both referents in the preceding sentence, as in 6c.

- (6) a. 小华替大明工作了两个钟头。他一分钱也没收。  
 xiao\_hua\_ti\_da\_ming\_gongzuo-le\_liang\_ge\_zhongtou  
 \_ta\_yi\_fen\_qian\_ye\_mei\_shou  
 little\_Hua\_for\_big\_Ming\_work-ASP\_two\_CL\_hour  
 \_he\_one\_CL\_money\_even\_not\_receive  
*Little Hua worked two hours for big Ming. He didn't even get \$1.*
- b. 老张昨天遇到小陈了。他低着头没说话。  
 lao\_zhang\_zuotian\_yudao\_xiao\_chen-le\_  
 ta\_di-zhe\_tou\_mei\_shuohua  
 old\_Zhang\_yesterday\_met\_little\_Chen-ASP\_  
 he\_lower-ASP\_head\_NEG\_speak  
*Yesterday old Zhang met little Chen. He kept his head lowered and didn't say anything.*
- c. 小高昨天跟小王吵了一架，他们俩打起来了。  
 xiao\_gao\_zuotian\_gen\_xiao\_wang\_chao-le\_yi\_jia\_tamen\_lia\_  
 da-qi-lai\_le  
 little\_Gao\_yesterday\_with\_little\_Wang\_argue-ASP\_one\_round\_  
 they\_two\_hit-ASP-ASP ASP  
*Yesterday little Gao argued with little Wang. They began fighting with each other.*

Considering the 875 sentence pairs in which the pronoun *ta* in the second sentence did appear to refer back to just one of the possible antecedents in the preceding sentence in an unambiguous way, the observation emerged that this pronoun was being used much more frequently to refer back to antecedents in subject position than in all other grammatical relations, with close to 80 % of pronoun reference involved linking to a preceding subject, vs. a little over 20 % for all other grammatical relations combined, as tabulated in 7.

- (7) Experiment 1: results of agreed codings  
 Pronoun reference to subject of preceding sentence: 79.89 % (699 tokens)  
 Pronoun reference to non-subject of preceding sentence: 20.11 % (176 tokens)

What this first experiment therefore shows is a very significant asymmetry in the spontaneous interpretation of pronouns, and speakers appear to very strongly favor linking the interpretation of a pronoun to a preceding subject in Chinese<sup>f</sup>. Such patterns suggest that the grammatical relation of a potential antecedent is indeed important for the resolution of the reference of a pronoun in Chinese, as has been suggested for certain other languages such as English (Crawley et al. 1990; Fredericksen 1981;

Smyth 1994). A corpus study considering the use of pronouns in Chinese which was carried out after completion of the experimental work bears this initial conclusion out, as will be noted in Section 7, and a similar high proportion of pronouns was found to be used to refer back to previous subjects vs. non-subjects (see Section 7).

Next, shadowing Kehler and Rohde's (2013) investigation of pronoun interpretation in English, we focused attention on a particular subset of target sentences in Experiment 1, those with predicates that depict an act of "transfer of possession," where some entity undergoes relocation into the possession of a second individual Goal argument, for example, verbs like "give/send/throw something to someone." Transfer of possession (TOP) verbs have been identified as having a special effect on the ways that pronouns are interpreted in English, causing a decrease in dominant reference to the subject of a preceding sentence, and a significant increase in reference to a post-verbal Goal argument. KR posit that TOP verbs cause a natural mental focus to occur on the Goal recipient, as the act of transfer is conceptualized as being completed once the object arrives in the possession of the Goal. A Goal argument thus becomes a salient, likely antecedent for further mention (by pronouns or other referential devices) in continuations of a dialog.

Considering the patterning of pronoun interpretation in Experiment 1 in sentences with TOP verbs (8 of the 34 target sentences) vs. those with non-TOP verbs (26 out of 34), a significant decrease in reference to subjects was noted when a TOP verb was present in the first sentence, as for example, in 8a–b:

- (8) a. 朱光给了丁力 两个礼物。他...  
 zhuguang\_gei-le\_dingli\_liang\_ge\_liwu\_ta  
 Zhuguang\_give-ASP\_Dingli\_two\_CL\_present\_he  
*Zhuguang gave Dingli two presents. He...*
- b. 吕建踢了一个球给周平。他...  
 lüjian\_ti-le\_yi\_ge\_qiu\_gei\_zhouping\_ta  
 Lüjian\_kick-ASP\_one\_CL\_ball\_to\_Zhouping\_he  
*Lüjian kicked a ball to Zhouping. He...*

In sentences with non-TOP verbs, reference to the subject was 83.48 %, and reference to non-subjects was just 16.52 %, whereas in sentences with TOP verbs, reference to the subject fell to 68.83 %, and reference to the non-subject argument present, the Goal with TOP verbs, was 31.16 %, almost twice the frequency of the reference to non-subjects in the non-TOP sentences.

- (9) Pronoun reference to subjects and non-subjects in non-TOP sentences (660 tokens)  
 Reference to subjects: 83.48 % (551 tokens)  
 Reference to non-subjects: 16.52 % (109 tokens)  
 Pronoun reference to subjects and non-subjects in TOP sentences (215 tokens)  
 Reference to subjects: 68.83 % (148 tokens)  
 Reference to non-subjects: 31.16 % (67 tokens)

This difference in patterning clearly suggests that a default subject reference bias for pronouns present in Chinese may be reduced *to some extent* by a secondary bias toward



referencing a Goal argument when present. This tension of pressures will continue to be explored in the further experiments carried out in the study.

#### 4 Experiment 2: Goal reference and predicate/aspectual structure

The secondary Goal bias revealed by a more refined look at the results of Experiment 1 raises the question of whether Goal arguments always show increased occurrences of reference to pronouns in subsequent sentences simply in virtue of being Goals, or whether this increase occurs only in certain circumstances, as Goals are constructed into larger discourse structures. In English, KR found evidence that it is principally when Goals represent the endpoint of a *completed* action of transfer that they show higher proportions of reference with following pronouns, as in 10, and that when a TOP verb occurs in an aspectually imperfective environment, and the action of transfer is not completed, there is no similar increase in the linking of Goal arguments to pronouns in the discourse, as for example in 11. The natural explanation KR give for this observation is that it is only in completed occurrences of transfer of possession that hearers build a salient mental picture of the action terminating at the Goal, favoring the latter as an activated, focused antecedent for a following pronoun.

(10) John passed a comic to Bill. He....

(11) John was passing a comic to Bill. He...

Experiment 2 attempted to establish whether a similar effect also occurs in Chinese, by means of a sentence completion task patterned on the experiment carried out by KR on English. The experiment compared TOP verbs in Chinese in perfective aspect, where transfer of the object to the Goal had occurred, with TOP verbs in imperfective aspect and with future markers, where no transfer had been completed or initiated. TOP verbs used in the experiment in different tense/aspect forms included 给 *gei* “give,” 寄 *ji* “send,” 送 *song* “send,” 发 *fa* “send,” 踢 *ti* “kick,” and 扔 *reng* “throw.” Examples of the sentence completion tasks are given in 12–14. Example 12 is an example of a TOP verb in perfective aspect, 13 of a TOP verb in imperfective aspect, and 14 of a TOP verb with a future marker<sup>8</sup>.

(12) 王春强给了赵玉柱一件衬衫。他...

wangchunqiang\_gei-le\_zhaoyuzhu\_yi\_jian\_chenshan\_ta  
Wangchunqiang\_give-ASP\_Zhaoyuzhu\_one\_CL\_shirt\_he  
*Wangchunqiang gave Zhaoyuzhu a shirt. He...*

(13) 赵来福在给李浩恩发短信。他...

zhaolaifu\_zai\_gei\_lihaoen\_fa\_duan\_xin\_ta  
Zhaolaifu\_ASP\_to\_Lihaoen\_send\_short\_letter\_he  
*Zhaolaifu is sending Lihaoen a text message. He...*

(14) 邹师长要送一瓶酒给赵团长。他...

zou\_shizhang\_yao\_song\_yi\_ping\_jiu\_gei\_zhao\_tuanzhang\_he  
Zou\_major\_will\_send\_one\_CL\_wine\_to\_Zhao\_captain\_he  
*Major Zou will give a bottle of wine to Captain Zhao. He...*

Experiment 2 incorporated 26 target sentences (13 with perfective verbs, 9 with future markers, and 4 with imperfective aspect markers) among 24 filler sentences and was completed by 28 adult native speaker participants from northern China, none of whom took part in Experiment 1. As in Experiment 1, participants were asked to write spontaneous continuations for each mini-dialog, making use of the initial pronoun prompt in the second sentence. From the total of 728 target sentences produced by participants, 20 sentences were eliminated from the data set by the two coders for the same reasons sentences were eliminated from the data set in Experiment 1 (participant use of the pluralizing morpheme—*men*, and ambiguity in the intended antecedent of *ta*). The results of the experiment, computed from the remaining 708 token sentences, showed a significant difference in the frequency of pronoun-Goal reference in the perfective and imperfective/future sentence pairs. When the TOP verb in the first sentence occurred in an imperfective or future-time predicate, reference to the Goal only occurred at an average of 9 %, and the remaining 91 % of pronoun references were to the subject. However, when the TOP verb occurred in a perfective environment, the rate of reference to the Goal was 34 %, and reference to the subject 66 %.

(15) Pronoun reference in perfective aspect sentence pairs (total 353 tokens)

Reference to subject: 66 % (233 tokens)

Reference to post-verbal Goal: 34 % (120 tokens)

Pronoun reference in imperfective aspect and future-marked sentence pairs (total 355 tokens)

Reference to subject: 91 % (323 tokens)

Reference to post-verbal Goal: 9 % (32 tokens)

This distinction between perfective and imperfective/future contexts demonstrates that a TOP event must be completed for it to make the Goal heavily salient for subsequent pronominal reference. Goal arguments are therefore not intrinsically salient in virtue of having the Goal thematic relation, but acquire salience for later pronominal reference when they constitute the realized endpoints for TOP events as new foci of attention.

### 5 The influence of Coherence Relations on referent-antecedent relations

In their study of referential devices used in discourse in English, Kehler and Rohde highlight the important role that appears to be played by the ways in which sub-parts of a discourse relate to each other—the “Coherence Relations” that exist between sentences in connected discourse. KR and others (see, for example, Hobbs 1979; Sanders and Noordman 2000, Kaiser 2011, and the collection of papers in Renkemer 2009) show how the different kinds of connections between sentences may result in a next mention bias for a particular referent—an expectation that a certain referent in one part of the discourse will figure strongly in a subsequent part given the way the discourse unfolds, sentence by sentence. A range of Coherence Relations have been posited and argued to effect the structuring of discourse. Examples of a number of common Coherence Relations are presented below, with illustrative examples from English adapted from KR, along with descriptions of types of discourse connection which exist between two sequential sentences, S1 and S2.

*Explanation:* S2 gives a reason why S1 happens—it explains S1

(16) Mary bought a new Spanish-English dictionary. She is going to Mexico this weekend.

*Occasion:* A temporal sequencing of events. S1 precedes S2 in time. The endpoint of the event described in S1 provides the starting point for what happens in S2.

(17) David took a taxi to the station. He then boarded the train for San Francisco.

*Result:* S1 makes it possible for S2 to occur.

(18) Chelswu just got a visa for North Korea. Therefore, she can now visit her distant relatives in Pyongyang.

*Elaboration:* S2 provides additional information about what happens in S1.

(19) John got accepted into Cal Tech. He is going to study physics.

*Violated Expectation:* S2 represents an unexpected outcome of the event in S1.

(20) Mary graduated summa cum laude. However, her parents still weren't satisfied.

TOP verbs with realized actions focus attention onto the individual (the Goal) who receives the transferred object, and this salience makes the Goal a natural potential starting point/new salient focus of attention for use as the subject of a continuation of the discourse, resulting in increased use of pronoun prompts to refer back to the Goal. In the theory of Coherence Relations, this linking of the subject position of the second sentence to the Goal of the first sentence might naturally be expected to correspond to situations in which one of two types of Coherence Relation exist between S1, which depicts the act of transfer, and S2 which describes a new eventuality/state arising as a result of the transfer. These are the Occasion and Result relations, both of which make use of the endpoint of the event in S1 as the starting point for S2. If there is a focus on the Goal argument as the endpoint of the act of transfer in S1, this may make the Goal a natural starting point for S2 if Coherence Relations do indeed create next mention biases in the way argued for in the literature.

Returning to consider the results of Experiment 1, the target sentences in this experiment were coded by two native speakers for the Coherence Relations which existed between each pair of sentences. This resulted in the observation of seven distinct types of Coherence Relations: (a) Explanation, (b) Occasion, (c) Elaboration, (d) Result, (e) Violated Expectation, (f) Parallelism, and (g) Topic Comment. The sub-set of TOP-verb sentences in which the pronoun in S2 referred back to the Goal in S1 was then separated out to examine the patterning of Coherence Relations that occurred in these sentence pairs. As anticipated, a very high proportion of the Goal-referring sentence pairs with TOP verbs involved either an Occasion or a Result relation existing between S1 and S2: 80.45 %.

(21) Coherence Relations in Experiment 1 with pronoun-Goal reference (67 tokens)

Proportion of Occasion/Result relations: 80.45 % (54 tokens)

The initial conclusion from this consideration of Coherence Relations in Experiment 1 is therefore that pronouns refer back to preceding Goals in TOP sentences with a

very high rate of frequency when there is an Occasion/Result relation between S1 and S2, and consequently that the Coherence Relation maintained between sentences in connected discourse can facilitate a next mention bias for processes of interpretation (i.e., an expectation is created that reference will be made to a certain NP in a preceding sentence). This role of Coherence Relations in influencing the interpretation of pronouns is examined further in Experiment 3.

### 6 Experiment 3: explicitly controlling for Coherence Relations

Experiment 3 set out to specifically control for the potential effect of Coherence Relations on pronoun interpretation in Chinese, again utilizing the experimental paradigm developed in KR. The experiment set out to see whether a manipulation of the Coherence Relation existing between two sentences could affect pronominal reference patterns. In order to do this, Experiment 3 presented the same set of TOP sentences to two different groups of native speakers of Chinese in a sentence completion task, with different instructions. Group 1 were asked to complete sentences beginning with a pronoun prompt that would be answers to the question “What happened next?” as shown in 22.

- (22) A: 吕建扔了一根香蕉给周平。  
 lüjian\_reng-le\_yi\_gen\_xiangjiao\_gei\_zhouping  
 Lüjian\_throw-ASP\_one\_CL\_banana\_to\_Zhouping  
*Lüjian tossed a banana to Zhouping.*
- B: 接下来发生了什么事?  
 jixialai\_fasheng-le\_shenme\_shi  
 next\_occur\_ASP\_what\_thing  
*What happened next?*
- A: 他.....  
 ta  
 he  
*He...*

Group 2, by way of contrast, were asked to complete sentences that would be answers to the question “Why?” as seen in 23.

- (23) A: 吕建扔了一根香蕉给周平。  
 lüjian\_reng-le\_yi\_gen\_xiangjiao\_gei\_zhouping  
 Lüjian\_throw-ASP\_one\_CL\_banana\_to\_Zhouping  
*Lüjian tossed a banana to Zhouping.*
- B: 为什么?  
 weishneme  
 why  
*Why?*
- A: 他.....  
 ta  
 he  
*He...*

Completions answering the question “What happened next?” were expected to result in Occasion relations, or Result relations, which would favor the occurrence of Goal co-reference. Occasion relations could also in theory give rise to the subject of the preceding sentence being construed as the topic/focus of the next sentence and the antecedent of a subject pronoun in S2, but this was expected to be less of a favored interpretation with TOP sentences than reference to the Goal, as the Goal of a realized TOP verb can be naturally taken to be the focal starting point for a continuation of the discourse.

Completions answering the question “Why?” were expected to result in Explanation relations, which would favor the subject-Agent of S1 being interpreted as the subject of S2. If S2 explains why the action of S1 is carried out, and the subject of S1 is the instigator/Agent of this action, it would be natural for speakers to reference the subject of S1 in the S2 continuation.

The predictions of Experiment 3 were therefore as follows. If Coherence Relations play an important role in affecting the interpretation of pronouns in discourse, the pronouns present in the continuations created by Group 2 might be more regularly expected to refer to the subject in S1, whereas pronouns present in the Group 1 continuations were expected to refer to the Goal in S1 more frequently than in the Group 2 continuations—even though the same realized TOP sentences were used as the S1 for both groups.

Experiment 3 required participants to make 40 sentence completions, which included 20 target sentences, all with TOP main verbs<sup>h</sup>, randomly mixed with 20 fillers. There were 37 adult native speaker participants from northern China in Group 1, producing 715 data tokens (following elimination of 25 sentences for reasons of ambiguity and use of the plural morpheme—*men*), and 40 in Group 2, producing 762 data tokens (following elimination of 38 sentences from the data set). None of the participants in either group had taken part in Experiment 1 or 2.

The results of the experiment were as follows. In Group 2, where participants provided completions answering the “Why?” question, which was expected to cause Explanation relations favoring subject reference, there was indeed a very high rate of reference to the subject of S1 with the pronoun in S2 (as coded by two native speakers): 83 %, vs. just 17 % reference to the Goal in S1. However, in Group 1, whose continuations answered the “What happened next?” question favoring Occasion/Result continuations and greater facilitation of reference to the Goal, there was significantly less reference to the subject of S1—46 %—and a much higher rate of reference to the Goal—54 %, as tabulated in 24.

#### (24) Results of Experiment 3

Group 1 “What happened next?” continuations (total, 715 tokens)

Reference to subject: 46 % (329 tokens)

Reference to Goal: 54 % (386 tokens)

Group 2 “Why?” continuations (762 tokens)

Reference to subject: 83 % (632 tokens)

Reference to Goals: 17 % (130 tokens)

The conclusions from Experiment 3 therefore seem to be that Coherence Relations do indeed significantly affect the way that pronouns are interpreted in Chinese. A forced Explanation relation causes a very strong reference to the subject, even in

realized TOP sentences which might otherwise favor reference to the Goal. However, when an Occasion or Result relation occurs, there is much more frequent co-reference of pronouns with the Goal. Additionally, it can be noted that the rate at which pronominal subjects in S2 reference the Goal in S1 clearly increases when S2 is naturally manipulated to be an Occasion/Result relation through the introduction of a “What happened next?” question. In Experiment 2, where TOP verbs occur as realized events but there was no contextual forcing of S2 to represent an Occasion/Result (and participants were free to write any kind of continuation), there was a 34 % rate of reference to the Goal of S1, whereas when the contextual cue provided by the question following S1 causes speakers to create Occasion/Result continuations, as in Experiment 3, the rate of Goal reference climbs very clearly, up to 54 %.

Such a conclusion about the influence of Coherence Relations on pronoun interpretation raises a further interesting and important question about the results found in Experiment 3, namely, why is it the case that the proportion of pronominal references to the Goal in Group 1 is actually not higher still, resembling the rate of references to subjects in Group 2 (83 %)? If an Occasion/Result relation combined with a realized TOP event provides a very natural next mention bias for the Goal argument of S1, why do we not find that the pronoun prompt is used to refer back to the Goal with the same high frequency that pronouns are used to refer to the subject of S1 in situations of Explanation relations? With Group 1, where all factors apparently seem to favor reference to the S1 Goal, there is 54 % reference to the Goal of S1, while in Group 2, where the Coherence Relation favors reference to the subject, this causes the subject to be referenced at the much higher rate of 83 %.

Following KR’s interpretation of similar patterns in English, we suggest that the particular distribution of reference found in Group 1 in fact results from the combination and interaction of *two potentially conflicting pressures*, whose existence is independently supported by other aspects of the patterns found in Experiments 1–3. The first of these pressures is that an Occasion or Result relation with a realized TOP verb heavily favors reference to the Goal of S1—it creates a next mention bias for the Goal due to the structuring of the discourse, as the Goal becomes a new focus of attention. The clear effect of this next mention bias has just been noted in the way that the rate of reference to the Goal increases dramatically when an Occasion/Result relation is forced, distinguishing the patterns found in Experiments 2 and 3. The second pressure stems from the actual type of referential device explicitly being used here—a *pronoun*. It can be posited that the presence of a pronoun in the continuation prompt favors reference to a preceding subject rather than a non-subject/Goal. A general high rate of use of pronouns to refer to subjects rather than other grammatical relations was independently attested in Experiment 1, with nearly 80 % of the pronouns in S2 continuations being linked to NPs in the preceding sentences’ subject position. In Experiment 3, the combination of these two conflicting pressures affecting the creation of continuations in Group 1 can be argued to result in the observed close-to-even split in referential use of the pronoun in S2 continuations created by this group, with a linking of the pronoun to the Goal in 54 % of the continuations, and to the subject in 46 % of the completed sentences<sup>1</sup>.

Such an interpretation of the asymmetrical patterning of results in Experiment 3 suggests that there are two different factors at work affecting the interpretation of pronouns and other referential devices in Chinese, as summarized in 25.



- (25) Principles guiding discourse continuity and the interpretation of pronouns
- (i) Coherence Relations influence hearer's expectations about what entities/individuals are likely to be talked about next. Based on the Coherence Relation linking two adjacent sentences, S1 and S2, hearers establish an expectation that the subject of S2 will correspond to a certain entity in S1. Such an expectation among hearers is referred to as a "next mention bias," following terminology used in Kehler and Rohde (2013), in parallel experiments on English<sup>l</sup>.
  - (ii) When pronouns occur, there is a bias to interpreting such elements as referring back to subjects of preceding sentences, rather than non-subjects.

Section 7 now contextualizes and discusses these conclusions further.

## 7 Further discussion and conclusions

The goal of the current experimental study of (overt) pronouns in Chinese has been to establish how such elements are interpreted in discourse, focusing specifically on contexts of potential ambiguity with two compatible antecedents, and how the experimental manipulation of various factors might both affect and reveal forces guiding the resolution of pronominal interpretation. Before reviewing the results of the experiments and the potential significance they have for analyzing processes of interpretation with discourse anaphora in Chinese, it is important to emphasize again, as discussed in Section 2, that pronouns constitute an important referential device in Chinese, despite the fact that Chinese is a pro-drop language and permits zero anaphora. While Li and Thompson (1979) have suggested that pronouns may be rather rare in their occurrence in Chinese and have a low frequency of distribution due to the presence of and preference for zero anaphora/pro, this characterization of Chinese has been argued by many other works to be incorrect as a general picture of the language, and the use of pronouns has been shown to be both common and frequently more natural than zero anaphora, in a broad range of discourse contexts. Concerning the design of the experiments in the current study, in which a pronoun prompt occurs in subject position of a sentence which is preceded by a sentence with two compatible antecedent NPs, pronouns in such positions are highly natural wherever there is even a very minor shift in the action described in the discourse (e.g., minor change in location, time, or trajectory of action) or a slight change in perspective (e.g., from foreground to background description), or if no minor discontinuity of any type occurs and speakers assume that the pronoun refers to the object of the first sentence (Chen 1986; Pu 2011). The multi-sentential design of the three experiments thus makes available a very broad range of contexts in which pronouns are the natural subjects of continuations supplied by participants in the experiments<sup>k</sup>.

Experiment 1 first set out to establish a base-line of pronoun use in Chinese, and whether any distinct patterns of use could be observed in the linking of pronouns to referents linguistically present in the preceding discourse. The sentence completion task showed that participants used pronouns provided in the prompts to reference elements in preceding subject positions at a rate of nearly 80 %, and linked pronouns to non-subjects much less frequently. This provided initial signs of a clear strong subject preference for pronominal interpretation in Chinese. An additional corpus study of Chinese has confirmed this dominant use of pronouns to refer to subjects found in

Experiment 1. A film script with extensive dialog was selected in order to check that the basic patterns identified in the experimental study were in line with pronoun use in other non-experimental environments<sup>1</sup>. The authors accessed the script for the film 颐和园 *yiheyuan* “Summer Palace” written by Lou Ye and Mei Feng<sup>m</sup> and tracked the positions that all third person pronouns in the text occurred in, as well as the syntactic positions of their antecedents. The script contained 1,361 occurrences of third person pronouns, within a total script length of 202,332 characters, and hence showed a very robust usage of such pronouns. Of these pronouns, 54.15 % occurred in subject positions, 18.1 % in possessor positions, and the remaining 27.75 % in other non-subject, non-possessor positions. Examining the position of the antecedents of pronouns present in the script, it was found that 75.18 % of these were in previous subject positions and only 24.82 % in non-subject positions. The use of pronouns in the film script thus provides additional, independent support for the conclusion from Experiment 1 that the interpretation of pronouns in Chinese exhibits a clear subject bias. Examples 26a–b below are representative examples from the movie script:

- (26) a. 陈军从舞会上回来了。他兴奋地和小林说话。  
 chenjun\_cong\_wuhui\_shang\_hui-  
 ai\_le\_ta\_xingfende\_he\_xiaolin\_shuohua  
 Chenjun\_from\_dance\_on\_return-  
 come\_ASP\_he\_excited\_with\_Xiaolin\_speak  
*Chenjun returned from the dance. He spoke excitedly with Xiaolin.*
- b. 周伟对这个消息有些意外。他尽量保持平静。  
 zhouwei\_dui\_zhe\_ge\_xiaoxi\_you\_xie\_yiwai\_ta\_jinliang\_baochi\_  
 pingjing  
 Zhouwei\_to\_this\_CL\_news\_have\_some\_surprise\_he\_try\_keep\_  
 calm  
*Zhouwei was a bit surprised about this news. He tried to remain calm.*

Having observed that reference to non-subject Goals in Experiment 1 was higher with transfer of possession/TOP verbs than other predicates, as in KR’s study of English, Experiment 2 demonstrated that the increase in reference to Goals and the decrease in reference to subjects with TOP verbs occur only in perfective contexts, where the action of the verb has been completed, resulting in a focus on the Goal and the occurrence of dialog continuations that make use of the Goal referent as a starting point. These continuations were noted to primarily instantiate Occasion or Result Coherence Relations. Experiment 3 then specifically controlled for the use of certain types of Coherence Relation with TOP verbs, and it was found that when Occasion and Result continuations are naturally stimulated by means of question equivalents to “What happened next?” pronoun reference to Goals climbed to 54 %. Remarking that the common subject bias for pronouns is heavily depressed in these conditions, we also asked why the frequency of reference to Goals was actually not considerably higher when Occasion/Result relations occurred with realized TOP events, as this combination of features might be expected to cause an even higher rate of pronoun use to link to preceding Goals. We suggested that the patterning across Experiments 1–3 can be interpreted as indicating that there are in fact conflicting pressures on the interpretation of

pronouns in Chinese, as KR have proposed for English, and that a natural subject antecedent bias interacts with and sometimes competes with the force of discourse structure and Coherence Relations present in Chinese. In cases in which specific Coherence Relations create a next mention bias for a non-subject argument, such as a Goal in a realized TOP predicate, this favors mention of the non-subject as the anchor/starting point of the continuation of the dialog. However, the presence of a pronoun provided as an initial prompt in the continuation favors interpretation of the pronoun as referencing a preceding subject. Such a conflict of pressures was suggested to result in a relatively even split of reference, and the pronoun prompt in such conditions is interpreted as referring to the Goal and to the subject with almost equal frequency. In other instances, where the next mention bias created by the discourse structure does not conflict with the preference to interpret pronouns as referring back to previous subjects, pronoun reference resumes its heavy subject bias and there is infrequent linking of pronouns to non-subject antecedents.

The results of the present study, and their suggested interpretation, contribute to ongoing debates on the factors that may either determine or influence the interpretation of discourse anaphora in Chinese. On the one hand, there is a tradition of work that emphasizes the role of semantics and pragmatics and the structuring of discourse on the interpretation of discourse anaphora, both in Chinese and in other languages, and recognizes no independent role for syntactic structure in the shaping of referential dependencies between pronouns and their antecedents (Li and Thompson 1979; Fox 1987; Givon 1983 among others). On the other hand, there are certain recent suggestions, most notably in Yang et al. (1999, 2003), that the syntactic roles occupied by discourse anaphora and their antecedents may indeed play a role in the referential linking of such elements and exert influences on this that are independent of discourse structure (see also Tai 1978). The present experimental investigation of pronominal reference provides an additional new paradigm of data in support of the view presented in Yang et al. (1999, 2003) that syntax does have the potential to exert itself as a force in the computation of anaphoric relations in Chinese and that a default subject bias exists in the linking of pronouns with their antecedents. This is revealed both directly in the high frequency of anaphoric relations between pronouns and elements in subject positions, and in the way that this subject bias can be deliberately manipulated and affected by certain specific changes made to aspects of the discourse structure. The general position supported by the present study and the experimental investigations in Yang et al. (1999, 2003) (and Kehler and Rohde (2013) for English) is that the interpretation of discourse anaphora such as pronouns is a hybrid operation involving the interaction of both syntactic factors and discourse structure, and therefore that earlier semantics/pragmatics-only approaches need to be revised to allow for certain (malleable) influences of syntax.

More experimental work will certainly be important to further confirm or refute such conclusions, and a number of issues need to be investigated in future work. The current project has focused heavily on the interpretation of pronouns relative to preceding subjects and Goal NPs, and we have not attempted to examine potential contrasts between subjects and NPs in other syntactic relations, such as direct object position, or pre-verbal objects in the 把 *ba* “BA” construction. It will be useful to see whether sentence completion tasks similar to those made use of here but focused on contrasting antecedent NPs in subject and direct/把 *ba* “BA” object positions will lead

to parallel results to those of the present study or possibly different conclusions. We also believe that other experimental methods can be usefully employed to further probe the influence of Coherence Relations on pronominal interpretation, such as those made use of in Yang et al. (1999, 2003) and the measuring of reading time with coherent and incoherent sentences when Coherence Relations are manipulated. Finally, in the present study, we have not attempted to investigate the referential preferences of zero anaphora/pro, and whether these can be systematically manipulated by adjustment of the discourse structure, as suggested in a range of works (Li and Thompson 1979; Tai 1978; Chen 1986; Pu 2011), and how syntactic factors may potentially play a role in constraining antecedent choice for pro (Huang 1984; Huang et al. 2009). While there is consequently much more to be done in investigating the use and interpretation of realized and zero pronouns in Chinese, it is hoped that the present study has broken some additional ground in this area of Chinese and will provoke further experimental and corpus studies of forces influencing patterns of pronominal reference in the language.

## 8 Endnotes

<sup>a</sup>All such studies recognize that overt pronouns in Chinese are, however, commonly used to refer to human entities and only rarely occur referring to inanimate objects. The latter are most frequently referenced with either full NPs or zero anaphora/pro.

<sup>b</sup>Chen (1986) also points out that the 37.7 % rate of occurrence of zero pronouns/pro in the Chinese version of the pear stories he investigated is much lower than the 73.2 % rate of occurrence of pro in Japanese pear stories reported in Clancy (1980), hence that Chinese might seem to make much less use of zero pronouns and more use of overt pronouns than the neighboring pro-drop language, Japanese.

<sup>c</sup>Li and Thompson (1979) use the phrase “conjoinable” to refer to the same basic idea that certain clauses or sentences may be perceived to be tightly connected with each other and form a single unit of discourse for the purposes of anaphoric reference.

<sup>d</sup>In such an approach, the high proportion of zero anaphora reported for Chinese in Li and Thompson (1976) will be due to a high frequency of topic chains in the particular texts they selected for their study, rather than due to any strong dispreference for overt pronouns, which are entirely appropriate for use in many discourse structures.

<sup>e</sup>Tai (1978) also notes various restrictions on the use of zero and overt pronoun forms which reference the syntactic positions occupied by such elements and their antecedents in the discourse, though Tai embeds such observations within a broader discourse structure perspective.

<sup>f</sup>Note that the NPs in subject position in Experiment 1 bore a variety of semantic roles, due to the choice of predicates in the examples, and that in addition to an Agent role, subjects bearing Experiencer, and Goal roles also occurred when the first sentence contained verbs of perception, psych verbs, verbs encoding accidental contact with an individual (running into an individual), and verbs of object motion toward an NP (e.g., receiving an object). The regular interpretation of pronouns as referencing the subject of the first sentence occurred across all subject types, Agent, Experiencer, and Goal, and was therefore the interpretative association of pronouns with an NP in a preceding *syntactic position*, not a specific *semantic role* (such as Agent).

<sup>e</sup>Note that the instance of 给 *gei* which occurs in 12 is a main verb meaning “give,” while the instances of *gei* in 13 and 14 are the non-verbal use of *gei* to mark indirect object arguments of the main verbs *fa* “send” and *song* “send.” In such occurrences, *gei* is commonly assumed to have grammaticalized as a preposition and is no longer a verb, as determined by a range of syntactic tests which distinguish verbs and prepositions in Chinese (inability of prepositions to attach the aspectual suffix *le*, inability of prepositions to combine with a zero pronoun, inability of many prepositions to occur as a short answer-form to a yes-no question, see Huang et al. 2009).

<sup>h</sup>In examples such as 22, the main verb of the sentence is 扔 *reng* “throw” and the occurrence of 给 *gei* is as a preposition marking the indirect object, as noted in footnote a for similar examples in Experiment 6.

<sup>i</sup>The fact that there is still a high proportion of pronoun reference to the subject when the Goal is activated as the focus of attention in realized TOP sentences is further support for the view that one force influencing the interpretation of pronouns is *syntactic*, and the regular linking of a pronoun to an NP in a particular syntactic position (the subject position), rather than the construal of pronouns with NPs having specific semantic roles (e.g., an Agent role—see footnote f), or a particular pragmatic role such as topic. The topic of a sentence is taken in Tomlin and Pu (2011) to be the entity which is the current focus of attention (see also Pu 2011), and in Li and Thompson (1979) to be the most focused referent in a discourse. With realized TOP predicates, the Goal argument regularly instantiates the (end) focus of attention in the event of object transfer, and so may naturally be interpreted as the ongoing, new topic of the discourse sequence. The observation that there is still a significant degree of interpretation of pronouns as referring to the subjects in such sentences, where the Goal and not the subject NP functions as the primary focus of attention/new topic, suggests that the association of pronouns with subject NPs is a *syntactic effect* which is separate and independent from any frequent *discourse effect* of linking pronouns with topic NPs.

<sup>j</sup>It is important to stress that the experiments reported in the present study all investigate patterns of speakers’ *interpretation* of pronouns in Chinese, and not their *production*. The next mention bias in 25i and the pronoun interpretation bias described in 25ii both refer to pressures that affect the way speakers *interpret* pronouns embedded in discourse, and we do not have any specific claims about the *production* of pronouns, though we would expect that there are close connections between processing and production. In order to be specific, testable predictions about when speakers will elect to use pronouns rather than other referential forms (zero anaphora, full NPs), a different set of experiments would need to be designed, allowing participants the freedom to make choices between referential forms in different discourse environments (see Pu 2011 for discussion and certain experiments of this type).

<sup>k</sup>Tai (1978) additionally shows that pronouns can occur across sentential boundaries regardless of the grammatical functions of the antecedent and the pronoun and that pronouns in subject or object positions can in principle link to antecedents in either the subject or object position of a preceding sentence.

<sup>l</sup>Film scripts are, of course, created linguistic products. However, to the extent that they attempt to mirror natural speech, they present a useful means of comparison to see whether the results of the experiments might be close to (or far away from) patterns assumed to be natural in regular speech.

<sup>m</sup>Accessed online at. Script version, September 26, 2003, Beijing. Currently available at: (i) <http://tieba.baidu.com/p/3537101341>, and (ii) <http://www.txt2016.com/ebook/7913/ebookread.html>

#### Competing interests

The authors declare that they have no competing interests.

#### Authors' contributions

All authors contributed towards the research design, the planning of the experiment, and the analysis of the results. ZW and YL administered the experiments. All authors read and approved the final manuscript.

#### Author details

<sup>1</sup>Department of Linguistics, University of Southern California, Los Angeles 90069, CA, USA. <sup>2</sup>Languages Division, Pasadena City College, Pasadena 91106, CA, USA. <sup>3</sup>Department of East Asian Languages and Cultures, Lawrence 66045, KS, USA.

Received: 20 June 2015 Accepted: 24 January 2016

Published online: 23 February 2016

#### References

- Chafe, Wallace (ed.). 1980. *The pear stories: Cognitive, cultural, and linguistic aspects of narrative production*. Norwood: Ablex.
- Chen, Ping. 1986. *Referent introducing and tracking in Chinese narratives*. Ph.D. dissertation. Los Angeles: University of California, Los Angeles.
- Christensen, Matthew. 2000. Anaphoric reference in spoken and written Chinese narrative discourse. *Journal of Chinese Linguistics* 28(2): 303–336.
- Clancy, Paricia. 1980. Referential choice in English and Japanese narrative discourse. In *The pear stories: Cognitive, cultural, and linguistic aspects of narrative production*, ed. Wallace Chafe, 127–202. Norwood: Ablex.
- Crawley, Rosalind, Rosemary Stevenson, and David Kleinman. 1990. The use of heuristic strategies in the interpretation of pronouns. *Journal of Psycholinguistic Research* 4: 245–264.
- Fox, Barbara. 1987. *Discourse structure and anaphora*. Cambridge: Cambridge University Press.
- Fox, Barbara. 1996. *Studies in anaphora*. New York: John Benjamins.
- Fredericksen, John. 1981. Understanding anaphora: Rules used by readers in assigning pronominal referents. *Discourse Processes* 4: 323–347.
- Givon, Talmy. 1983. *Topic continuity in discourse: Quantitative cross-linguistic studies*. New York: John Benjamins.
- Hinds, John. 1977. Paragraph structure and pronominalization. *Papers in Linguistics* 10: 77–99.
- Hinds, John. 1979. Organizational patterns in discourse. In *Syntax and semantics 12: Discourse and syntax*, ed. Talmy Givon, 135–157. New York: Academic.
- Hobbs, Jerry. 1979. Coherence and coreference. *Cognitive Science* 3: 67–90.
- Huang, C.-T. James. 1984. On the distribution and reference of empty pronouns. *Linguistic Inquiry* 15: 531–574.
- Huang, C.-T. James, Y.-H. Audrey Li, and Yafei Li. 2009. *The syntax of Chinese*. Cambridge: Cambridge University Press.
- Kaider, Elsi. 2011. On the relation between coherence relations and anaphoric demonstratives in German. In *Proceedings of Sinn & Bedeutung 15*, ed. Ingo Reich et al., 337–351. Saarbrücken, Germany: Saarland University Press.
- Kehler, Andrew, and Hannah Rohde. 2013. A probabilistic reconciliation of coherence-driven and centering-driven theories of pronoun interpretation. *Theoretical Linguistics* 39: 1–37.
- Li, Charles, and Sandra Thompson. 1979. Third person pronouns and zero anaphora in Chinese discourse. In *Syntax and semantics 12: Discourse and syntax*, ed. Talmy Givon, 311–335. New York: Academic.
- Pu, Ming-Ming. 1995. Anaphoric patterning in English and Mandarin narrative production. *Discourse Processes* 19: 279–300.
- Pu, Ming-Ming. 2011. *Discourse anaphora: A cognitive functional approach*. Munich: Lincom Europa.
- Renkemer, Jan (ed.). 2009. *Discourse, of course*. Amsterdam: John Benjamins.
- Rohde, Hannah. 2008. *Coherence-driven effects in sentence and discourse processing*. Ph.D. dissertation. San Diego: University of California, San Diego.
- Sanders, Ted, and Leo Noordman. 2000. The role of coherence relations and their linguistic markers in text processing. *Discourse Processes* 29: 37–60.
- Simpson, Andrew. 2014. Verbal answers to yes/no questions, focus and ellipsis. In *Chinese syntax in a cross-linguistic perspective*, ed. Audrey Li, Andrew Simpson, and Dylan Tsai, 300–333. Oxford: Oxford University Press.
- Smyth, Ron. 1994. Grammatical determinants of ambiguous pronoun resolution. *Journal of Psycholinguistic Research* 23: 197–229.
- Tai, James H.-Y. 1978. Anaphoric restraints in Mandarin Chinese narrative discourse. In *Anaphora in discourse*, ed. John Hinds, 279–338. Edmonton: Linguistic Research.
- Tao, Liang. 1996. Topic discontinuity and zero anaphora in Chinese discourse: Cognitive strategies in discourse processing. In *Studies in anaphora*, ed. Barbara Fox, 485–511. New York: John Benjamins.
- Tomlin, Russell, and Ming-Ming Pu. 1991. The management of reference in Mandarin discourse. *Cognitive Linguistics* 2(1): 65–93.
- Yang, Chin Lung, Peter Gordon, Randall Hendrick and Jei Tun Wu. 1999. Comprehension of referring expressions in Chinese. *Language and Cognitive Processes* 14: 715–743.
- Yang, Chin Lung, Peter Gordon, Randall Hendrick, and Chih Wei Hue. 2003. Constraining the comprehension of pronominal expressions in Chinese. *Cognition* 86: 283–315.