Gender interference in processing Chinese compound reflexive: evidence from reading eye-tracking

Wenshuo Chang , Yunyan Duan , Jingjing Qian , Fuyun Wu , Xiaoming Jiang & Xiaolin Zhou

Published online: 24 Jun 2020.

Submit your article to this journal ⟋

Article views: 182

View related articles ⟋

View Crossmark data ⟋

Routledge
Taylor & Francis Group

Check for updates

REGULAR ARTICLE

Gender interference in processing Chinese compound reflexive: evidence from reading eye-tracking

Wenshuo Chang [a], Yunyan Duan [a,b], Jingjing Qian[c], Fuyun Wu [d], Xiaoming Jiang[e,f] and Xiaolin Zhou[a,f]

[a]School of Psychological and Cognitive Sciences, Peking University, Beijing, People's Republic of China; [b]Department of Linguistics, Northwestern University, Evanston, IL, USA; [c]School of International Education, Shanghai University of Sport, Shanghai, People's Republic of China; [d]School of Foreign Languages, Shanghai Jiao Tong University, Shanghai, People's Republic of China; [e]Department of Psychology, Tongji University, Shanghai, People's Republic of China; [f]Institute of Linguistics, Shanghai International Studies University, Shanghai, People's Republic of China

ABSTRACT

In Chinese, the compound reflexive *ta-ziji* ("him/her-self") has the gender marking pronoun *ta*, hence presenting a good test case for interference effects from structurally illicit antecedents predicted by cue-based retrieval models. Using reading eye-tracking, we manipulated the gender of *ta-ziji* that (mis)matches that of matrix- and local-subject. Results showed no interference whatsoever when *ta-ziji* matched local subjects. Only when *ta-ziji* mismatched local subjects did we find an *inhibitory* interference on first fixation duration and gaze duration at the verb immediately preceding *ta-ziji*, but a *facilitatory* interference on gaze duration at *ta-ziji*. Furthermore, at *ta-ziji*, total reading times were longer for gender-mismatching local subjects than for gender-matching ones. These findings are partially predicted by the standard cue-based retrieval model, but are mostly consistent with the structure-favoring cue-based retrieval model, suggesting that the structural cue plays a dominant role in the antecedent retrieval process, with interference occurring only in highly constrained situations.

1. Introduction

In parsing sentences, a reader needs to identify, construct and complete non-local syntactic dependencies between linguistic elements. One such dependency is the reflexive-antecedent dependency, as in "*John knows Tom believes himself.*" where upon encountering the reflexive *himself*, the reader will trigger a search process for the appropriate antecedent. A purely structure-based account, based on Principle A of Chomsky's (1981) Binding Theory (henceforth BT-A), would posit that the parser only considers the noun within a local governing domain, which in this case must be the local subject *Tom*. Note however, that the matrix subject *John*, though outside the governing domain of the reflexive, shares the gender feature with the reflexive and the local subject. Then a question immediately arises: Compared to the sentence "*Mary knows Tom believes himself*" where the matrix subject *Mary* does not share the gender feature, will the gender-matching noun *John* be considered as an antecedent candidate, thereby interfering in the retrieval process of the local noun *Tom*?

To address this question, various parsing models have been proposed, and some might be subsumed under one influential framework, namely cue-based memory retrieval (Jäger et al., 2017; Lewis & Vasishth, 2005). One essential claim is that upon encountering a reflexive (or more generally, an anaphor), the parser necessarily initiates a retrieval process from working memory that is content-addressable, searching for its appropriate antecedent (McElree et al., 2003; Van Dyke & McElree, 2006). Memory representations with features that match relevant properties of the reflexive (called retrieval cues) can be directly accessed. As a result, there can be more than one candidate antecedent, among which the best one is selected depending on the strength of activation. This cue-based retrieval framework has obtained ample evidence from psycholinguistic experiments and computational modelling on reflective-antecedent dependency and other linguistic dependencies as well (for a comprehensive review, see Jäger et al., 2017).

While the structural cue (i.e. BT-A) is clearly important in the antecedent retrieval process, it remains unclear whether and how it interacts with other retrieval cues. One variant of cue-based retrieval models, known as the structure-favoring account (Dillon et al., 2013; Frazier et al., 2015; Parker & Phillips, 2017), posit that

the structural information has priority over other cues in the retrieval process of antecedents. In contrast, the standard cue-based retrieval account (Jäger et al., 2017; Lewis & Vasishth, 2005) contends that all cues – structural (e.g. BT-A) and nonstructural (e.g. gender, number, focus-position) – are treated equally, jointly guiding the antecedent-search process for reflexive resolution. Thus, while both accounts agree that a structurally illicit noun that matches the reflexive in nonstructural features may be (mis)retrieved as a candidate antecedent, leading to the so-called interference effect, they differ in how and when the interference effect appears.

Given that the structure-favoring cue-based retrieval model views BT-A as a primary cue in retrieving the antecedent of a reflexive, a structurally illicit candidate will minimally interfere in the retrieval of the structurally licit antecedent. Rather, interference from a structurally illicit candidate, if any, would only be observed in highly constrained situations, especially when the structurally licit antecedent fails to match any non-structural cue(s). Perhaps the earliest evidence for the structure-favoring cue-based retrieval model comes from Sturt (2003), who, using materials as in (1), manipulated gender congruence between the reflexive and the matrix subject (*Jonathan/Jennifer*), with the local subject (*the surgeon*) being a stereotypical male or female.

(1) Jonathan (Jennifer) was pretty worried at the City Hospital. He/she remembered that the surgeon had pricked himself (herself) with a used syringe needle. There should be an investigation soon.

Sturt (2003) found that at the reflexive, sentences with gender-mismatching local subjects (i.e. *surgeon … **herself***) were read longer than sentences with gender-matching local subjects (i.e. *surgeon … **himself***), as reflected in first fixation duration, gaze duration, and second-pass reading time, suggesting that BT-A was immediately applied in antecedent search. The interference from the gender-matching matrix subject was only observed on second-pass reading time.

Using similar stimuli, Cunnings and Sturt (2014) replicated the early (i.e. gaze duration) effect of gender congruence between the reflexive and the local subject, but failed to replicate the late interference effect. Instead, they only found a gender congruence effect of local subjects on the second-pass reading time, indicating that reflexive resolution in English is strongly biased towards the local subject, with little interference from matrix congruence.

In addition to the gender cue, recent evidence has shown that the structural constraint can override other non-structural cues as well. In Dillon et al. (2013), prolonged reading times were observed at the reflexive when it mismatched in number with the structurally licit antecedent, regardless of the number feature of structurally illicit antecedents. Similarly, while Parker and Phillips (2017) found interference from the (structurally illicit) matrix subject when the reflexive mismatched the local subject in both gender and animacy, this interference disappeared when the reflexive matched the local subject in either feature. Taken together, these results suggest that the structural cue plays a dominant role in antecedent search for reflexives in English. Only in cases where the reflexive-antecedent dependency cannot be completed by the structural cue alone would non-structural cues play a role.

In contrast, the standard cue-based retrieval model assigns an equal status to the structural cue as other cues, and makes the following predictions: When a perfectly matched antecedent is available, its activation would be reduced in the presence of a structural illicit antecedent that matches one or more non-structural cue(s) (e.g. gender, number, or animacy) due to activation spreading from the retrieval cue to both antecedents, leading to a slowdown in retrieval latencies or increased reading times (*inhibitory interference*, also known as *similarity-based interference* or *cue overload*). When a perfectly matched antecedent is unavailable, that is, the structurally licit candidate mismatches a certain non-structural cue, then a candidate antecedent matching in that non-structural cue would race with the structurally licit candidate until the one with a higher activation is retrieved, resulting in a speed-up in the average retrieval latencies or decreased reading times (*facilitatory interference*) (Logačev & Vasishth, 2016).

In addition, while the standard cue-based retrieval model also predicts a main effect of local congruence (i.e. a penalty of violation of the locality constraint), the underlying reason is motivated by faster decay of activation of non-local versus local antecedents, without even using the structural cue (Jäger et al., 2015, p. 11).

Evidence for interference effects has been reported in English reflexives, but the results are mixed (Badecker & Straub, 2002; Kwon & Sturt, 2014; Parker & Phillips, 2017; Sturt & Kwon, 2015). Using self-paced reading, Badecker and Straub (2002) found inhibitory interference effects in the post-reflexive region when a perfectly matched local subject was available, reflected by longer reading times when the reflexive matched the matrix subject (*John thought that Bill owed himself another opportunity to solve the problem*) than when *John* was

replaced by *Jane*. In contrast, in a Bayesian meta-analysis of 30 published studies (eye-tracking and self-paced reading) on reflexive/reciprocal-antecedent dependency, Jäger et al. (2017) found no interference when a perfectly matched antecedent was available, but inhibitory (instead of *facilitatory*) interference when the structurally licit antecedent candidate mismatched the nonstructural cue.

Despite a wealth of research on English reflexive processing, few studies have focused on reflexive processing in Chinese. Reflexives in Chinese can take either bare (*ziji*, "self") or compound (*ta-ziji*, "himself/herself") forms, with the latter requiring its antecedent to match the pronoun *ta* in gender and number features (Huang et al., 2009). While the bare reflexive *ziji* has long been known to allow long-distance binding from antecedents in a matrix clause, the compound reflexive *ta-ziji* is claimed by theoretical linguists to strictly obey BT-A (Huang et al., 2009). Previous work on reflexive processing in Chinese has mainly focused on *ziji*, with converging results verifying the theoretical claims that *ziji* can be affected by nonstructural information and be bound by a non-local antecedent. Specifically, using reading eye-tracking, Jäger et al. (2015) found an inhibitory interference from the structurally illicit antecedent (i) when *ziji* mismatched the local antecedent in animacy (Experiment 1), and (ii) when *ziji* matched the local antecedent if an additional memory load was enforced on the participants (Experiment 2). In addition, specific referential properties of local verbs, which may serve as nonstructural cues in processing, can also directly affect how the bare reflexives *ziji* are interpreted (Jin, 2003; Li & Zhou, 2010).

Given that *ta-ziji* resembles English reflexives more closely than *ziji* (Huang et al., 2009), one might assume that its online resolution process should be more likely to follow BT-A. To our knowledge, only a few experimental studies have investigated real-time processing of *ta-ziji*. Using self-paced reading, Dillon et al. (2016) compared *ziji* with *ta-ziji* using complex sentences as in (2a-b), where the animate feature of reflexives matched with either the local/head noun (*seamstress*) or the distant/embedded noun (*Mrs. Zhang*) of a prenominal relative clause.

(2) Sample stimulus set from Dillon et al. (2016)
  (a) 媒体报道的那个女裁缝上个星期把她自己（自己）不小心弄 伤了。
    Media/ report-on/ that/ seamstress/ last-week/ BA/ *ta-ziji* (*ziji*)/ carelessly/ harm.
    The seamstress that the media reported on carelessly harmed herself last week.

  (b) 张太太经常光顾的那个时装店上个星期把她自己（自己）不小心弄伤了。
    Mrs. Zhang/ often/ visit/ that/ boutique/ last-week/ BA/ *ta-ziji* (*ziji*)/ carelessly/ harm.
    The boutique that Mrs. Zhang often visits carelessly harmed herself last week.

Dillon et al. (2016) found that reflexives were read longer when their animacy features matched with the distant/embedded noun (2b) than with the local/head noun (2a), and the effect size (i.e. the difference in reading times between the two types of sentences) was larger for *ziji* than *ta-ziji* in the spillover region (不小心 "*carelessly*"). They took these findings as evidence for *ta-ziji* being less subject to the locality constraint than *ziji*, and attributed this difference in the parsing profile of the two reflexives to the presence of an additional retrieval cue *ta* in *ta-ziji* that encoded human, number and gender features. This suggests that when the local antecedent does not match with the reflexives in animacy, *ziji* will undergo a stronger inhibitory interference from the distant animacy-matching antecedent than *ta-ziji*. Note, however, in (2b) the distant noun phrase "Ms. Zhang" is actually the only appropriate antecedent for both *ziji* and *ta-ziji*, making its status as a "distractor" less clear in the cue-based retrieval theories.

In a self-paced reading experiment, Qian and Wu (2016) manipulated gender congruence of *ta-ziji* with the local subject and the matrix subject in sentential complement sentences as in (3). All proper nouns are typical Chinese names for males (e.g. *Lijun*) or females (e.g. the matrix subject *Limei* or the local subject *Huwei*). The verbs in the subordinate clauses were equally biased regarding which noun can serve as the antecedent of *ta-ziji*, that is, the patient of the local verb could be interpreted as referring to either the local subject or the matrix subject.

(3) 李美（李骏）说胡薇早就推荐她自己（他自己）以及镇上的其他 干部参加市里的辩论赛。
    Limei (Lijun)/ says/ Huwei/ already/ recommended/ *ta-ziji*/ and/ other/ in the town/ officials/ participate/ held by the city hall/ the debate competition.
    Limei (Lijun) says Huwei already recommended herself (himself) and other officials in the town to participate the debate competition held by the city hall.

Qian and Wu (2016) found that at the reflexive (*ta-ziji*) and in the post-reflexive region (*yiji* "and"), reading times

were longer when *ta-ziji* mismatched the local subject than when it matched the local subject, regardless of the gender congruence between *ta-ziji* and the matrix subject. Only at the fourth word following *ta-ziji* (i.e. *qita* "other") did they find longer reading times when *ta-ziji* mismatched the matrix subject than when *ta-ziji* matched the matrix subject. These results were largely analogous to the finding from English reflexives (Cunnings & Sturt, 2014), suggesting that Chinese compound reflexive is governed by BT-A, with potential interference from the matrix subject taking effects in such a delayed fashion as to be negligible. Note however, that self-paced reading is less fine-grained than reading eye-tracking in temporal resolution (Cunnings & Felser, 2013; Cunnings & Sturt, 2014; Jäger et al., 2015; Sturt, 2003). Thus, it might not be surprising that they failed to detect early interference effects from the matrix subject.

Building on Qian and Wu (2016), we aim to further investigate potentially early interference effects in Chinese compound reflexive processing, using reading eye-tracking. According to the structure-favoring cue-based retrieval account, we would expect to observe (i) a main effect of gender congruence when *ta-ziji* matches the local subject, reflected by prolonged reading time on eye movement measures in local incongruence conditions than local congruence conditions, and (ii) a potential interference effect from the matrix subject that would be shown only when *ta-ziji* mismatches the local subject. In other words, interference from gender-matching matrix subjects is predicted to occur only in ungrammatical conditions where no structurally licit antecedent is available, with the condition containing gender-matching matrix subjects being processed faster (i.e. facilitatory interference or "illusionary grammaticality") than the condition containing gender-mismatching matrix subjects. According to the standard cue-based retrieval model, however, since both structural and (non-structural) gender cues are assigned equal weight, we would expect to observe an interaction depending on the availability of a perfectly matched antecedent, specifically, (i) an *inhibitory* interference from the gender-matching matrix subject relative to the gender-mismatching matrix subject when *ta-ziji* matches the local subject, and (ii) a *facilitatory* interference from the gender-matching matrix subject when *ta-ziji* mismatches the local subject.

In addition, we expect to see early interference effects prior to the reflexive region, given the strong parafoveal-on-foveal effect reported in the literature on Chinese reading (Tsai et al., 2012; Yan et al., 2010; Yan & Sommer, 2015; Yan, Zhou, et al., 2012; Yang et al., 2009; Yen et al., 2008). The perceptual span in Chinese

reading extends from 1 character on the left of the fixated character to up to 4 characters on its right (Inhoff & Liu, 1998; Pan et al., 2017; Yan et al., 2010, 2015), depending on the frequency (Yan et al., 2010) and predictability (Zang et al., 2016) of parafoveal words. Given that *ta* and *ziji* are highly frequent (11853.78 words/million for masculine ta, 6600.96 words/million for feminine *ta*, and 1728.29 words/million for *ziji*; Cai & Brysbaert, 2010) and that these two words form natural collocations (i.e. *ta-ziji*), the early congruence effect(s) on processing *ta-ziji* might be detected when the eyes are fixated to the local verb preceding the reflexive.

2. Method

2.1. Participants

Forty-five undergraduate and graduate students (18 males; age range: 19–25 years) from Peking University participated in this experiment. They were all native speakers of Chinese and had normal or corrected-to-normal vision. This study was conducted in accordance with the Declaration of Helsinki and was approved by the Ethics Committee of the School of Psychological and Cognitive Sciences, Peking University.

2.2. Materials and design

Adapting the experimental design of Qian and Wu (2016), we created 48 sets of experimental stimuli with a compound reflexive *ta-ziji* embedded in the subordinate clause. The structure of each sentence was "$NP_1 + V_1 + NP_2 + ADV + V_2 + ta\text{-}ziji + CONJ + NP_3 + VP_3$", as shown in Table 1. The matrix subject (NP_1) and the local subject (NP_2) are proper names with stereotypical gender (See *Norming tests* below), each consisting of two characters. The matrix verb (V_1) is always mono-morphemic *shuo* ("say"). A verb phrase follows NP_2, consisting of a two-character adverb (e.g. "already", ADV) and a two-character local verb (V_2). The local verb (V_2) is always transitive with an equal bias towards the matrix or local subject as the patient (see *Norming tests* below). The local verb (V_2) takes an embedded clause, consisting of *ta-ziji* (three characters), a two-character conjunction phrase *yiji* ("and"; CONJ), and a collective noun referring to a group of people (four characters; NP_3), followed by a predicate structure (VP_3) having four to six characters.

We manipulated gender congruencies of *ta-ziji* to the matrix subject NP_1 and to the local subject NP_2, yielding four conditions: (1) *ta-ziji* matched both NP_1 and NP_2 (Local-match, Matrix-match), (2) *ta-ziji* matched NP_1

Table 1. Experimental conditions, exemplar sentences and interest regions.

Condition	Example Matrix subject		Local subject		Local verb		Compound reflexive		Conjunction			
Matrix-match, local-match	李美 Lijmei	说 says	胡薇 Huwei	早就 already	推荐 recommended		她自己 herself		以及 and	其他 other	干部 officials	参加辩论赛。 participate in the debate competition.
Matrix-mismatch, local-match	李骏 Lijun	说 says	胡薇 Huwei	早就 already	推荐 recommended		她自己 herself		以及 and	其他 other	干部 officials	参加辩论赛。 participate in the debate competition.
Matrix-match, local-mismatch	李骏 Lijun	说 says	胡薇 Huwei	早就 already	推荐 recommended		他自己 himself		以及 and	其他 other	干部 officials	参加辩论赛。 participate in the debate competition.
Matrix-mismatch, local-mismatch	李美 Limei	说 says	胡薇 Huwei	早就 already	推荐 recommended		他自己 himself		以及 and	其他 other	干部 officials	参加辩论赛。 participate in the debate competition.

Translation: *Limei (Lijun) says Huwei has already recommended herself (himself) and other officials to participate in the debate competition.*

only, but mismatched NP$_2$ (Local-mismatch, Matrix-match), (3) *ta-ziji* matched NP$_2$ only, but mismatched NP$_1$ (Local-match, Matrix-mismatch), and (4) *ta-ziji* mismatched both NP$_1$ and NP$_2$ (Local-mismatch, Matrix-mismatch). Within each set of experimental stimuli, NP$_2$ was identical across all four conditions. Half of the NP$_2$ were feminine names, and the other half were masculine names, with comparable visual complexity as indexed by the number of strokes ($t_{(47)} = 0.75$, $p = 0.457$). Half of the materials contained "他自己" ("himself"), and the other half contained "她自己" ("herself").

Experimental sentences were assigned into four lists, using Latin-square procedure, such that each list had 48 sentences, 12 from each condition. Another 92 sentences were added into each list as fillers. All filler sentences were well-formed with a variety of syntactic structures. Stimuli in each list were pseudo-randomized such that the first five sentences were fillers and no more than three sentences from the same condition appeared consecutively.

2.3. Norming tests

2.3.1. Gender stereotypicality of proper names
The proper names used for NP$_1$ and NP$_2$ were assessed for gender typicality to ensure the strength of the manipulation of gender congruence. We created 304 proper names consisting of a common Chinese family name (1 character) and a given name (1 character) which typically contained the gender information. Eight participants who did not participate in the eye-tracking experiment were asked to rate on a 5-point scale how typical these Chinese names were associated with a particular gender (1 for extremely masculine and 5 for extremely feminine). Based on the mean ratings, 144 names were selected for the experimental stimuli, half being masculine (mean = 1.72, $SD = 0.28$, all lower than or equal to 2.5), and the other half being feminine (mean = 4.59, $SD = 0.23$, all higher than or equal to 3.88).

2.3.2. The property of local verb phrases (VP)
To ensure that only gender and syntactic features of antecedents were relevant in interpreting *ta-ziji*, we conducted a norming test for selectional properties of verb phrases.

Twenty-six participants who did not participate in the eye-tracking experiment were asked to finish a multiple-choice cloze test consisting of 108 items. Each item consisted of an embedded clause taken from the experimental stimuli, with the direct object missing, as in "NP$_2$ (proper name) + ADV + V$_2$ + [BLANK] + CONJ + NP$_3$ + VP$_3$". Participants were instructed to read the sentence and to fill in the blank by choosing one of the three options: "*ta-ziji*", "*A proper name other than NP$_2$ (e.g. Wangyao)*", or "*Both*". Participants were told that they should choose "*Both*" only when they believed that both *ta-ziji* and *Wangyao* were acceptable to fill in the blank. The three choices were encoded as 0 for *Both*, 1 for *ta-ziji*, and −1 for *Wangyao*. Averaging across participants, an acceptability score close to zero would mean both *ta-ziji* and a proper name other than NP$_2$ were acceptable. In the end, 48 pairs of adverb and verb were selected with acceptability scores close to zero (mean = −0.03, $SD = 0.23$), and one-sample t-test further indicated that the acceptability scores were comparable to zero ($t_{(27)} = -0.93$, $p = 0.357$). These 48 verb phrases consisting of adverb and verb (VP) pairs were then used to construct experimental stimuli. Of these 48 pairs, 20 verbs were used twice each (paired with two different adverbs), and 8 verbs were used only once.

2.4. Apparatus

Eye-movements were recorded with an Eyelink 2 K system at a sampling rate of 2000 Hz. Each sentence was presented in one line at the middle vertical position of a 21-inch CRT screen (1024*768 resolution; frame rate 120 Hz). The font Song-28 was used, with each character subtending 1 degree of visual angle. Participants read

each sentence with their head positioned on a chin rest 60 cm from the screen. All recordings and calibrations were based on the left eye but viewing was binocular.[1] Experimental presentation was programmed with Psychtoolbox-3 (Brainard, 1997; Pelli, 1997) and Eyelink toolbox (Cornelissen et al., 2002).

2.5. Procedure

Before the formal experiment, participants underwent a practice block of 10 trials. After this, participants were calibrated with a nine-point grid. A fixation cross was presented at the position on the screen where the first character of the sentences would appear. The fixation cross was presented for 300, 400 or 500 ms randomly, followed by visual presentation of the whole sentence. Participants were requested to read the sentence silently and to press the space bar on a keyboard when finishing reading. Upon pressing the bar, the sentence disappeared and a comprehension question was presented. The question was related to the meaning of parts of the sentence but was irrelevant to how the participants resolved the reflexive. Participants were instructed to answer the question by pressing the "yes" or the "no" button. Half of the participants were instructed to press the "yes" button with their left index finger and the "no" button with the right one, and the other half made their responses with a reversed button-hand assignment scheme. Half of the trials required a "yes" answer and half required a "no" answer.

After the eye-tracking experiment, participants were asked to complete an antecedent choice questionnaire, in which they made a forced choice between the matrix subject and the local subject as the antecedent of *ta-ziji*, using the same material used in the eye-tracking experiment.

2.6. Data analyses

We focused on three regions of interest for eye movement analysis: the local verb immediately preceding *ta-ziji* (e.g. *tuijian* "recommended"), the compound reflexive *ta-ziji*, and the conjunction (CONJ) immediately following *ta-ziji* (*yiji* "and"). We computed four reading duration measures that index the time course of antecedent-search process triggered by *ta-ziji*: first fixation duration, gaze duration, regression path duration, and total reading time, with the first two reflecting early processing, and the latter two signifying late processing (Rayner, 1998; Yang et al., 2009). First fixation duration is the duration of the initial fixation in a region during first-pass reading. Gaze duration is the sum of all first-pass fixation durations in a region before eyes make a

saccade to another region. Regression path duration is the sum of fixation durations from when a region was first fixated up until the eyes first moved to the right of the region, including the time regressed to regions prior to this region. Total reading time includes the sum of all fixation durations in a region, regardless of whether fixations occurred during the first pass or later. We also checked the skipping rate on each region of interest to make sure that the effects observed on reading time measures could not be attributed to differences in skipping rate between conditions (Cunnings et al., 2015).

First fixation durations shorter than 60 ms or longer than 800 ms, and gaze duration longer than 1000 were excluded from analyses for all measures except total reading time, leaving 90% of observations for these analyses. Trials with no fixation in a region of interest were excluded from analyses for all duration measures, including total reading time analyses.

2.6.1. Linear mixed models and generalised linear mixed models

We ran linear mixed models (LMMs) fitted by maximum likelihood for reading duration analyses, and generalised linear mixed models (GLMMs) for skipping rate and the post-experiment antecedent choice analyses (Baayen et al., 2008). Duration measures were log-transformed, but untransformed durations also yielded the same pattern. The LMMs and GLMMs were fitted using the *lmer* function and the *glmer* function of the *lme4 1.1-13* package (Bates et al., 2015) respectively. The *p*-values for *t*-statistics for LMMs were computed using the *lmerTest 3.0-1* package (Kuznetsova et al., 2017) in the R environment for statistical computing (R-Core Development Team, 2017). We used sum contrasts to code two fixed factors, namely Matrix (NP$_1$) Congruence (the matrix subject-mismatch vs. the matrix subject-match) and Local (NP$_2$) Congruence (the local subject-mismatch vs. the local subject-match). All models were fitted with fixed effects structure consisting of the two critical fixed factors and their interaction, and one additional fixed factor, stimulus order (centred and scaled), aiming to control for adaptation effect during the experiment (Fine et al., 2013). Random effects structure included by-subject and by-item random intercepts, and random slopes were determined by backward model selection according to the Akaike information criterion (AIC, Akaike, 1974) and a significance test based on the χ^2-distributed likelihood-ratio test (Matuschek et al., 2017).

The model selection procedure started with the full model with random effects structure including by-subject and by-item random slopes for all fixed factors

(Barr et al., 2013). We then defined a set of reduced models by excluding one of the random slopes. The AICs of reduced models were compared, and the one with the minimum AIC was selected to compare with a more complex model by a likelihood-ratio test. The more complex model would be selected when the p-value of a χ^2-statistic was smaller than 0.2 (Matuschek et al., 2017). Otherwise, the model reduction procedure would be repeated until a model was selected or all the random slopes were excluded. Models that failed to converge were not considered in this procedure. We report in the *Results* section estimated effect sizes (*b* values), standard errors (*SE*), *t*-statistics for LMM (or *z*-statistics for GLMM), and statistical significance (*p*-values) of the best models selected by the backward model selection. The results reported below hold with or without the stimulus order fixed factor in the analyses. All the data (including the gender norms in Mandarin) and statistical analyses reported in this paper are available at https://github.com/ChangWenshuo/Chinese_reflexive_R

2.6.2. Additional analyses with Bayesian linear mixed models

We also conducted additional Bayesian analyses, in order to further verify our LMM results and to better understand potential "interference" effects from the matrix subject in local-match and local-mismatch conditions, which are critical for distinguishing the structure-favoring cue-based retrieval model from the standard cue-based retrieval model.

Following Jäger et al. (2020),[2] we fitted Bayesian linear mixed models with fixed effects' contrasts coded as in Table 2: The first contrast tested the main effect of Local Congruence; the second contrast tested the effect of Matrix Congruence when *ta-ziji* matched the local subject; and the third contrast tested the effect of Matrix Congruence when *ta-ziji* mismatched the local subject.

The Bayesian LMMs used a log normal likelihood on reading time measures in milliseconds (ms), and were fitted by the *brm* function of *brms 2.8.0* package (Bürkner, 2017). All models included full random effects

and were consistently converged. For prior distributions, we used a standard normal distribution prior $N(0, 1)$ for all slopes of fixed effects and a $N(0,10)$ for intercepts, and a Cauchy distribution prior $Cauchy(0,5)$ for standard deviation parameters. Within the variance-covariance matrices of the by-subject and by-items random effects, priors were defined for the correlation matrices using a LKJ prior with parameter η 2.0 (Lewandowski et al., 2009). For each model, the joint posterior distribution was sampled by running four Monte-Carlo Markov Chains (MCMCs) at 2000 iterations, with the first half of the samples discarded as warm-up samples. Convergence was checked using $\hat{R}$ convergence diagnostic and by visual inspection of the chains; posterior estimates were considered as convergent if $\hat{R}$ is lower than 1.1 (Gelman & Rubin, 1992). We report in the *Results* section estimated means and 95% credible intervals (*CrI*s) of posterior distributions. A 95% *CrI* indicates that the true effect size lies within this range with a probability of 95%. All posterior estimates reported below have $\hat{R}$s lower than 1.1.

3. Results

3.1. Comprehension question

Data from one participant was excluded due to low accuracy rate (10%) of comprehension questions. For the remaining 44 participants, their mean accuracy rates ranged from 88% to 100% (*mean* = 95%, *SD* = 3%), indicating that they were attentive to the task.

3.2. The off-line antecedent choice questionnaire

Table 3 shows the means of the probabilities of choosing the local subject (i.e. NP$_2$) vs. the matrix subject (i.e. NP$_1$) as a function of Matrix Congruence and Local Congruence in the post-experiment questionnaire on antecedent choice.

Consistent with the claim of theoretical linguistics, our participants were more likely to choose the local subject as the antecedent of *ta-ziji* than the matrix subject in

Table 2. Contrast coding of the fixed effect in Bayesian linear mixed models.

	Local-match		Local-mismatch	
	Matrix-match	Matrix-mismatch	Matrix-match	Matrix-mismatch
Main effect of local congruence	0.5	0.5	−0.5	−0.5
Matrix Congruence (local match)	0.5	−0.5	0	0
Matrix Congruence (local mismatch)	0	0	0.5	−0.5

Table 3. Means and Standard Errors (in parenthesis) of the probabilities of choosing the local subject and the matrix subject in the off-line antecedent choice questionnaire.

	Local-match		Local-mismatch	
	Matrix-match	Matrix-mismatch	Matrix-match	Matrix-mismatch
Probability of choosing the local subject	71% (2%)	81% (2%)	49% (2%)	67% (2%)
Probability of choosing the matrix subject	29% (2%)	19% (2%)	51% (2%)	33% (2%)

nearly all conditions[3], except for the local-mismatch, matrix-match condition, where participants were slightly more likely to choose the matrix subject than the local subject to be the antecedent (51% vs. 49%).

Results of the GLMM showed a main effect of Local Congruence ($b = 0.57$, $SE = 0.1$, $z = 5.79$, $p < 0.001$), a main effect of Matrix Congruence ($b = 0.47$, $SE = 0.08$, $z = 6.02$, $p < 0.001$), but no interactions between Matrix Congruence and Local Congruence ($b = 0.1$, $SE = 0.06$, $z = 1.66$, $p = 0.097$). Gender-matching local subjects were more likely than mismatching ones to be chosen as the antecedent of ta-ziji (76% vs. 58%). While gender-matching matrix subjects were also more likely than mismatching ones to be chosen as the antecedent (40% vs. 26%), neither of these two mean ratings was above chance.

3.3. Online eye-movement measures

Results of LMM on eye movement measures showed no effects in the Conjunction region. Thus, we do not report results for this region. Table 4 shows the untransformed means of eye movement measures for each condition in the remaining two critical regions.

3.3.1. Local verb region

In the verb region, totals of 1867, 1867, 1867, 2007, and 2112 observations were available for first fixation duration, gaze duration, regression path duration, total reading time, and skipping rate analyses, respectively. As shown in Table 4, the skipping rates were comparable across conditions (11% on an average, $ps > 0.2$).

We found no main effect of Local Congruence and no main effect of Matrix Congruence on any measure, but

significant interactions on first fixation duration ($b = 0.01$, $SE = 0.01$, $t = 2.23$, $p = 0.026$) and gaze duration ($b = 0.02$, $SE = 0.01$, $t = 2.06$, $p = 0.039$). Further planned tests examining the effect of Matrix Congruence as a function of Local Congruence suggested that gender-matching matrix subjects induced longer first fixation duration and gaze duration than gender-mismatching matrix subjects only when ta-ziji mismatched local subjects (first fixation duration: 231 vs. 220 ms, $b = 0.02$, $SE = 0.01$, $t = 2.39$, $p = 0.017$; gaze duration: 264 vs. 245 ms, $b = 0.03$, $SE = 0.01$, $t = 2.48$, $p = 0.01$), but not when ta-ziji matched local subjects (first fixation duration: 226 vs. 228 ms; gaze duration: 259 vs. 258 ms, $ps > 0.4$).

Similar patterns were also shown on total reading time, but the interaction did not reach significance ($ps > 0.09$): Total reading time was numerically longer for gender-matching matrix subjects than for gender-mismatching matrix subjects (448 vs. 407 ms) only when ta-ziji mismatched local subjects in gender, but not when ta-ziji matched local subjects (399 vs. 407 ms).

3.3.2. Reflexive region

In the reflexive region, totals of 1925, 1925, 1925, 2051, and 2112 observations were available for first fixation duration, gaze duration, regression path duration, total reading time, and skipping rate analyses, respectively. As shown in Table 3, the skipping rates were comparable across conditions (8% on average, $ps > 0.2$).

On first fixation duration, we found a main effect of Local Congruence ($b = 0.02$, $SE = 0.01$, $t = 2.33$, $p = 0.02$), a significant interaction between Local Congruence and Matrix Congruence ($b = 0.01$, $SE = 0.01$, $t = 2.04$, $p = 0.042$), but no main effect of Matrix Congruence. Further planned tests examining the effect of Matrix Congruence as a function of Local Congruence showed that when ta-ziji mismatched local subjects, the gender-matching matrix subject induced significantly shorter first fixation duration than gender-mismatching matrix subjects (228 vs. 236 ms, $b = -0.02$, $SE = 0.01$, $t = -2.31$, $p = 0.021$), but no significant difference was observed when ta-ziji matched local subjects (226 vs. 224 ms, $p = 0.504$).

The main effect of Local Congruence was also observed on gaze duration (313 vs. 297 ms, $b = 0.02$, $SE = 0.01$, $t = 2.36$, $p = 0.019$), regression path duration (372 vs. 347 ms, $b = 0.02$, $SE = 0.01$, $t = 1.97$, $p = 0.049$), and total reading time (549 vs. 487 ms, $b = 0.06$, $SE = 0.01$, $t = 4.17$, $p < 0.001$), with gender-mismatching local subjects being read longer than gender-matching local subjects. However, we found no main effect of Matrix Congruence (p-values > 0.07) and no interaction

Table 4. Means and Standard Errors (in parenthesis) of eye-movement measures for the four experimental conditions.

	Local-match		Local-mismatch	
	Matrix-match	Matrix-mismatch	Matrix-match	Matrix-mismatch
Local verb				
First-fixation duration	226 (3)	228 (3)	231 (4)	220 (3)
Gaze duration	259 (5)	258 (5)	264 (6)	245 (5)
Regression path duration	292 (9)	285 (8)	300 (10)	278 (8)
Total reading time	399 (10)	407 (10)	448 (12)	407 (11)
Skipping rate	10% (1%)	12% (1%)	11% (1%)	11% (1%)
Reflexive				
First-fixation duration	226 (3)	224 (4)	228 (4)	236 (4)
Gaze duration	296 (7)	300 (7)	306 (7)	320 (7)
Regression path duration	352 (12)	342 (10)	369 (16)	374 (12)
Total reading time	472 (12)	489 (14)	547 (14)	543 (14)
Skipping rate	8% (1%)	8% (1%)	9% (1%)	7% (1%)

between the two factors (*p*-values > 0.1) on any of these measures.

3.3.3. Bonferroni-corrected significance

As one reviewer suggested, one potential issue with the LMM analyses reported above is increased Type I errors as a result of testing multiple eye-tracking measures (von der Malsburg & Angele, 2017). Given that four eye movement measures were analysed in each region, we applied a Bonferroni correction to the LMM analyses, yielding a corrected significance threshold of 0.0125 for alpha, such that the Type I error probability is kept at 0.05. With this conservative significance threshold, nearly all the effects were gone, and the only effect survived was the main effect of Local Congruence on total reading time at *ta-ziji*.

3.3.4. Additional analyses using Bayesian linear mixed model

To provide a quantitative assessment of the effects detected by LMMs and to further examine potential interference effects, we followed existing conventions by running Bayesian analysis. Table 5 and Figure 1 summarise the estimates of the Bayesian analyses, including the means of the posterior distribution of each planned contrast (back-transformed to ms) and corresponding 95% CrIs.

When *ta-ziji* matched local subjects in gender, 95% *Crl*s of posterior distributions of the effect of Matrix Congruence included 0 on all eye movement measures in both positions, specifically, first fixation duration (local verb: *mean* = −3 ms, *Crl*: [−12, 7]; *ta-ziji*: *mean* = 3 ms, *Crl*: [−6, 13]), gaze duration (local verb: *mean* = −2 ms, *Crl*: [−17, 14]; *ta-ziji*: *mean* = −2 ms, *Crl*: [−17, 16]), regression path duration (local verb: *mean* = 1 ms, *Crl*: [−16, 20]; *ta-ziji*: *mean* = 5 ms, *Crl*: [−17, 29]), and total reading time (local verb: *mean* = −17 ms, *Crl*: [−43, 13]; *ta-ziji*: *mean* = −9 ms, *Crl*: [−40, 27]). Close examination reveals that most of these 95% CrIs were centred around 0, except for the total reading time, where the

means of the posterior distributions were −17 ms and −9 ms and the uncertainty as reflected by the CrIs were rather high. We will return to this point in the Discussion section. But overall, the fact that the 95% CrIs contained 0 on all these measures suggest that Bayesian analyses are consistent with the results of LMM, confirming a lack of evidence for reliable difference between the Local-match, Matrix-match condition and the Local-match, Matrix-mismatch condition.

When *ta-ziji* mismatched local subjects in gender, 95% *Crl*s of posterior distributions of the effect of Matrix Congruence excluded 0 on (i) first fixation duration (*mean* = 9 ms, *Crl*: [1,19]) and gaze duration (*mean* = 13 ms, *Crl*: [2,26]) at the local verb, and (ii) on gaze duration at *ta-ziji* (*mean* = −16 ms, *Crl*: [−30,−1]), with the effects in opposite directions, suggesting that differences were found on these measures between the Local-mismatch, Matrix-match condition and the Local-mismatch, Matrix-mismatch condition. Specifically, when local subjects could not serve as perfect antecedents for *ta-ziji*, gender-matching matrix subjects induced *inhibitory* interference (reflected by prolonged early eye movement measures) at the local verb, and *facilitatory* interference at the reflexive (reflected by shortened gaze duration). It is worth noting that while the interference effects revealed by the two methods of statistical analysis are consistent at the verb, specific eye-movement measures varied in the critical region of the reflexive, where the interaction was revealed on first fixation duration by the LMM, but on gaze duration by the Bayesian analyses. As the Bayesian analysis is known to be more reliable compared with the LMM, we focus on the converging results in the Discussion section, and in the case of inconsistency, we only discuss the results yielded by the Bayesian analysis.

Note also at *ta-ziji* (see Table 5), the 95% *Crl*s of posterior distributions of main effect of Local Congruence excluded 0 on first fixation duration (*mean* = −7 ms, *Crl*: [−13, −1]) and total reading time (*mean* = −51 ms, *Crl*: [−69, −29]), confirming that the "local-mismatch"

Table 5. Posterior means (ms) and credible intervals of Bayesian linear mixed models.

	Main effect of Local Congruence		Matrix Congruence (local match)		Matrix Congruence (local mismatch)	
	Mean	95% Crl	Mean	95% Crl	Mean	95% Crl
Local verb						
First-fixation duration	2	[−4,8]	−3	[−12,7]	9	[1,19]
Gaze duration	4	[−6,15]	−2	[−17,14]	13	[2,26]
Regression path duration	3	[−10,18]	1	[−16,20]	15	[−2,33]
Total reading time	−8	[−29,17]	−17	[−43,13]	18	[−4,46]
Reflexive						
First-fixation duration	−7	[−13,−1]	3	[−6,13]	−9	[−18,1]
Gaze duration	−12	[−23,1]	−2	[−17,16]	−16	[−30,−1]
Regression path duration	−13	[−29,5]	5	[−17,29]	−15	[−33,7]
Total reading time	−51	[−69,−29]	−9	[−40,27]	1	[−33,42]

Note: The posterior means were back-transformed to ms; Crl: credible interval.
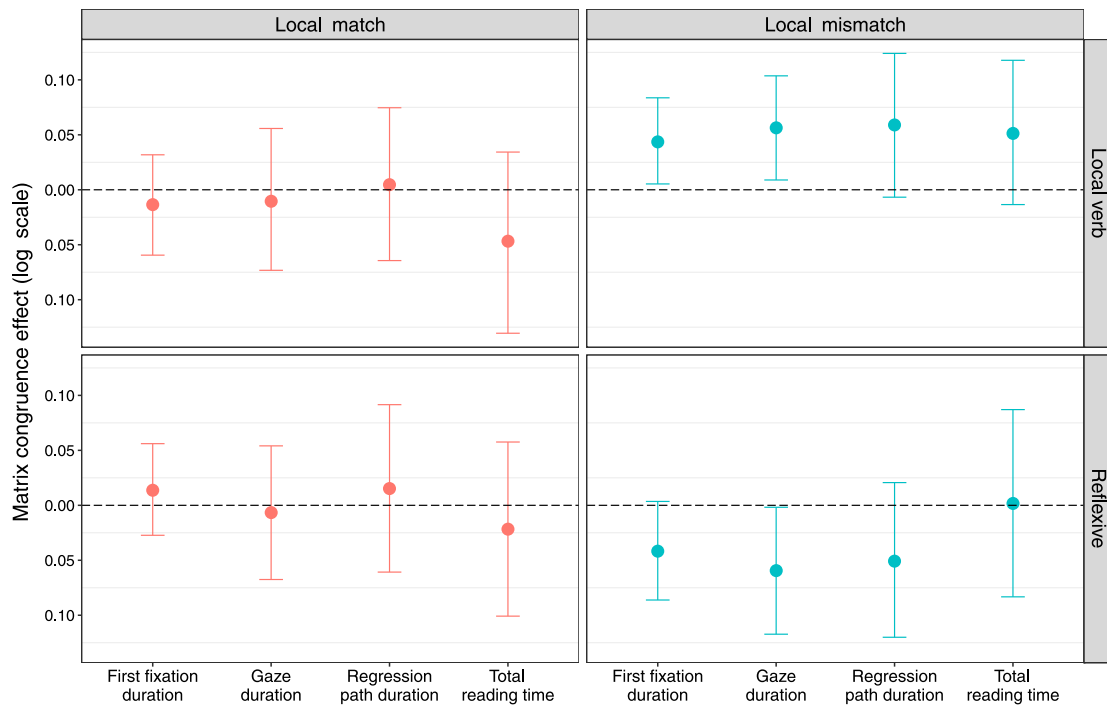
Figure 1. Estimates obtained from Bayesian linear mixed models of eye movement duration measures (in log-scale). Circular points denote the posterior means, and error bars denote the corresponding 95% credible intervals. The upper penal shows the estimates of eye movement measures at the local verb, whereas the lower penal shows the estimates of eye movement measures at the reflexive (i.e. ta-ziji). Red lines on the left penal denote estimates of Matrix congruence effect in the Local-match conditions, whereas turquoise lines on the right penal denote estimates of Matrix congruence effect in the Local-mismatch conditions.

penalty effect is due to violation of the locality constraint (reflected by both early and late eye movement measures).

4. Discussion

In this study, we examined online processing of Chinese compound reflexive *ta-ziji*. Our eye-tracking data yielded four main findings. First, at the local verb preceding *ta-ziji*, we consistently found in both LMM and Bayesian analyses *inhibitory* interference effects from the matrix subject when *ta-ziji* mismatched the local subject, as reflected by longer first fixation duration and gaze duration for the gender-matching matrix subject sentences than for the gender-mismatching matrix subject sentences. Second, at *ta-ziji*, we found evidence for *facilitatory* interference effects from the matrix subject when *ta-ziji* mismatched the local subject, as indicated by shorter gaze duration (by the Bayesian analysis) for gender-mismatching matrix subject sentences relative to gender-matching matrix subject sentences. Third, at *ta-ziji*, a main effect of Local Congruence was found in both early (first fixation duration by the Bayesian analysis) and late (total reading time by both Bayesian and LMM analyses) measures. Fourth, converging results from both Bayesian and LMM analyses show a lack of

evidence for interference when local subjects matched reflexives. As discussed below, while the latter two findings are consistent with the structure-favoring cue-based retrieval model, the first two findings are predicted by either model.

Recall that while both structure-favoring and standard cue-based retrieval models predict a Local Congruence × Matrix Congruence interaction reflected by interference effects from gender-matching matrix subjects in our experiment, the standard cue-based retrieval model explicitly predicts that the nature of such interference is *inhibitory* when the reflexive matches the local subject, and *facilitatory* when the reflexive mismatches the local subject. However, in our experiment, interference effects from matrix subjects were observed only when *ta-ziji* mismatched the local subject but were barely detected when a perfect local antecedent was available. Furthermore, in the ungrammatical conditions the types of such interference were exactly opposite to each other in two adjacent regions, that is, *inhibitory* at the local verb and *facilitatory* at the reflexive. These patterns of result cast doubts on the predictions of the standard cue-based retrieval model.

Note also that virtually no evidence for interference effects from the matrix subject were found from both LMM and Bayesian analyses when *ta-ziji* matched the

local subject, contrary to what the standard cue-based retrieval model would predict. While the absence of an effect does not necessarily mean the effect does not exist, hence it might not be deemed as strong evidence falsifying the standard cue-based retrieval model, we nevertheless would like to point out that this lack of inhibitory interference has been reported by both experimental studies (Cunnings & Sturt, 2014; Dillon et al., 2013; Kush & Phillips, 2014; for additional references, see Jäger et al., 2015, p. 3) and meta-analysis (Jäger et al., 2017) on antecedent-reflexive dependency processing, and has been taken as evidence for the parsing theories claiming that the structural cue guides antecedent search process for reflexives (Dillon et al., 2013; Kush & Phillips, 2014). That being said, on the total reading time measure, there is a fairly large degree of uncertainty in the estimates of Bayesian analyses, and the 95% CrIs are not tightly centred around zero. This potentially allows the observed data from this measure to be compatible with a small interference effect when *ta-ziji* matched the local subject, which is in the right direction as predicted by the standard cue-based retrieval model.[4] In this sense, the possibility for *inhibitory* interference to arise in the local-match conditions cannot be unequivocally ruled out, and hence it remains possible to detect such interference effects in future work. However, assuming the overall pattern we observed in the current data can be generalised to *ta-ziji* processing in the presence of a perfect-matching antecedent and a long-distance candidate antecedent, then our results favour the structural cue-based retrieval model.

It is also worth noting that while facilitatory interference was found by early eye-movement measures at *ta-ziji*, this finding is compatible with both the structure-favoring cue-based retrieval and the standard cue-based retrieval model. Furthermore, while both models predict a slowdown in reading times associated with gender-mismatching local subjects relative to gender-matching local subjects, this main effect of Local Congruence is claimed to be "unrelated to the cues used for retrieval" under the standard cue-based retrieval model, but directly follows from "assumption[s] of decay" (Jäger et al., 2015, p. 11). Thus, these findings could not be used to support the standard cue-based retrieval model.

Instead, most data patterns better fit the structure-favoring cue-based retrieval model, which, in addition to a main effect of Local Congruence due to the dominant role of the structural cue in processing reflexive-antecedent dependencies, also predicts facilitatory interference effects from gender-matching matrix subjects only in highly constrained situations, where the structurally licit local subject fails to serve as a perfect

antecedent due to gender mismatch. In fact, this effect of illusionary grammaticality has also been reported in antecedent-reflexive dependency in Hindi (Kush & Phillips, 2014), and has been reliably found in other types of linguistic dependency including subject-verb agreement (Dillon et al., 2013; Wagers et al., 2009) and negative polarity item licensing (Vasishth et al., 2008; Xiang et al., 2009). Thus, our finding that interference occurred when *ta-ziji* mismatched structurally licit subjects in gender, coupled with the virtual absence of interference in the presence of perfect antecedents for *ta-ziji*, provides supporting evidence for the structure-favoring cue-based retrieval model. These patterns also echo the results of our off-line antecedent choice test, which indicated that the local subject was more likely to be selected as the antecedent of *ta-ziji*; only when the local subject mismatched *ta-ziji* in gender was the gender-matching matrix subject likely, albeit below the chance level, to be chosen as the antecedent.

In sum, while more work needs to be done to further explore potential interference in the presence of perfectly matching antecedent, our behavioural data – both online and offline – lend support to the structure-favoring cue-based retrieval model, suggesting that the structural cue (i.e. BT-A) is likely to ultimately take precedence over the nonstructural cues (i.e. the gender cue), guiding the antecedent search for *ta-ziji*.

4.1. Early inhibitory interference effects at the verb

One noteworthy finding in the current study is the *inhibitory* interference effect detected on the verb immediately preceding the reflexive. This type of interference, again, was observed only in ungrammatical conditions where the structurally licit local subject mismatched the reflexive in gender, but was virtually absent in grammatical conditions where a perfectly matched antecedent was available. The effects were robust in both LMM and Bayesian analyses, revealed by first fixation duration and gaze duration that index early eye-movement measures.

The inhibitory nature and the pattern of the observed effect – though strictly speaking not predicted by either of cue-based retrieval models under discussion, are not exceptional in the existing literature. For instance, in the meta-analysis of reflexives/reciprocal-antecedent dependencies including studies of English, Chinese and Hindi, Jäger et al. (2017) found inhibitory interference when the reflexives mismatch local antecedents and no interference when they match each other. In terms of experimental evidence, our results are similar to the findings of Jäger et al. (2015) on Chinese bare reflexive

ziji: When the structurally licit antecedent candidate was inanimate and thus mismatched the animacy cue of *ziji*, the structurally illicit antecedent induced longer first fixation duration and gaze duration at the reflexive (i.e. *ziji*) when it was animate compared to when it was inanimate. To explain this unpredicted effect, Jäger et al. (2015) added cue confusion to the standard cue-based retrieval model. Cue confusion arises when retrieval cues are not perfectly distinguishable from each other. Since (*ta*-)*ziji* usually requires its antecedent to be both local and animate, the structural cue and a non-structural cue (i.e. the "animate" cue) frequently co-occur, and thus can have crossed associations with more than one features, leading to similarity-based interference as a result of high similarity between these two cues. This cue confusion account may initially appear appealing for the inhibitory interference observed in our study: All matrix subjects in our sentences were proper names, and thus were animate, boosting their likelihood of being considered as antecedent candidates when the structural cue was retrieved. The activation of the structural cue would spread to both local and matrix subjects, resulting in inhibitory interference. However, given that both gender-matching and gender-mismatching matrix subjects were animate in our design, and that there is no reason to assume that animacy differed in any of these proper names, the gender cue should presumably be rarely confused with the animacy cue or the structural cue. We therefore do not consider cue confusion as a viable account to fully explain the inhibitory effect observed in the current study.

Here we would like to offer a post-hoc explanation by assuming memory "encoding" (Dillon et al., 2014; Laurinavichyute et al., 2017) in the cue-based retrieval model. As will be argued in Section 4.2, due to parafoveal processing, our participants were able to preview the following reflexive *ta-ziji* when fixating the verb, making three types of processing at work, probably in parallel:

(i) Upon the verb, the local subject will be retrieved as the agent, which might compete with the memory representation, possibly decaying, of the distant subject as the agent of the matrix verb.
(ii) Meanwhile, due to its first presence, the previewed *ta-ziji* needs to be encoded into memory representation and be activated. Specifically, in order to get its referential interpretation, a reflexive requires its antecedent to be local ({locality}) and be matching with the gender conveyed by *ta* ({gender}).
(iii) With the previewed reflexive, both the local and matrix subjects might be activated as appropriate patients of the verb, with a roughly equal level of activation (as the verb was equi-biased in our design).

Together, lexical processing of the verb and parafoveal processing of the reflexive would mean that representations of these lexical items, along with their syntactic and/or semantic associations (e.g. Agent, Theme, locality), might overlap in feature content, potentially causing confusions when our participants attempted to initiate antecedent search for the reflexive. Thus, the slower reading times in the Local-mismatch, Matrix-match condition perhaps reflect confusion or additional time necessary for our participants to interpret the sentence up to that point. In Local-mismatch, Matrix-mismatch condition, however, our participants could quickly recognise that what they had seen so far just was not correct for referential interpretation of the reflexive; therefore, they could continue reading the rest of the sentence. Hence, the "no-match" condition was read a bit more quickly. In the case of two local-match conditions, no processing difference would ensue as long as the local subject satisfied the highly-ranked locality constraint, regardless of gender features. Thus, the overall profile would surface as inhibitory interference in the two Local-mismatch conditions at the verb.

Speculative as it seems, this explanation might be deemed as similar to the noisy-channel model[5] (Gibson et al., 2013; Levy et al., 2009), where an ideal (rational) reader would take input noises into account – as in the case of the Local-Mismatch/ Matrix-Match condition, while reducing parsing uncertainty by inferring a plausible (but non-literal) reading, leading to prolonged reading times. This line of research is worth further investigation, yet formalising it involves a non-trivial amount of work, which is beyond the scope of our paper.

4.2. Parafoveal effect with the presence of gender-mismatching local subject

Another novel finding of the inhibitory effect at the verb lies in its (temporal or spatial) locus, which is surprisingly early. We attribute this effect to a parafoveal effect with the presence of the gender-mismatching local subject. To our knowledge, Jäger et al. (2015) reported interference effect (or a main effect of locality) at the verb immediately preceding the bare reflexive *ziji*, but they ruled out the possibility of parafoveal preview because a significant effect was also found at the verb by Chen et al. (2012) using self-paced reading, where preview was implausible. But in our case, no effect was found at the verb in the self-paced reading experiment using similar experimental stimuli (Qian & Wu,

2016), validating that the effect on the verb can be attributed to preview in our eye-tracking experiment. Preview effect is the influence of the information presented in the parafoveal region (in this case, the reflexive *ta-ziji*) on reading the current region (here, the verb). We suggest that upon previewing the compound reflexive, Chinese readers were able to activate/encode relevant features ({locality, gender}) that *ta-ziji* requires for an appropriate antecedent in order to get its referential interpretation, or process the gender information contained in *ta-ziji*, leading to the observed early inhibitory effects.

It is worth noting that when *ta-ziji* mismatched the local subject, the facilitatory interference at *ta-ziji* was opposite to the inhibitory interference at the preceding verb. This pattern might be due to a tradeoff between the time spent at the local verb and that at the reflexive, and is consistent with the previous findings that the parafoveal gaze duration is inversely correlated with the foveal gaze duration (Kennedy, 1998; Kennedy et al., 2002).

One might question the plausibility of interpreting the effects on early eye movement measures at the local verb as reflecting high-level processing of the following word, *ta-ziji*. Yet ample evidence on Chinese reading has shown preview benefits (Yan et al., 2009; Yan, Risse, et al., 2012) and parafoveal-on-foveal effects (Yan et al., 2009; Yan & Sommer, 2015) at the semantic level. Using gaze-contingent boundary paradigm (Rayner, 1975) where the word in the preview is manipulated before the reader's eyes crossed an invisible boundary and changes to the target word after the boundary is crossed, Yan and colleagues have obtained clear evidence for semantic preview benefit on early eye movement measures, such that first fixation duration and gaze duration at the target word were shorter in the semantically related preview than in the unrelated preview (Yan et al., 2009; Yan, Risse, et al., 2012). Even in natural reading without an invisible boundary, Yan and Sommer (2015) also found preview effects influenced by high-level processing, such that first fixation duration and gaze duration at the pre-target words were longer when the upcoming target words conveyed positive emotional meanings than when they conveyed neutral meanings. Increasing evidence has shown that Chinese readers also integrate the meaning of parafoveal words with contextual information (Li et al., 2018; Yang et al., 2012, 2014). We therefore believe that the effect at the local verb is likely due to parafoveal processing of *ta-ziji*.

Despite that high-level (especially, semantic) preview effect in English reading is rare and barely available with highly constrained context (Schotter et al., 2015),

semantic parafoveal processing is not only found in reading Chinese, but also in reading German (Hohenstein & Kliegl, 2014) and Korean sentences (Yan et al., 2019). How preview benefits influence the results pattern observed in antecedent retrieval in languages with large preview benefits is beyond the scope of this paper, but may be worth further investigation in the future.

4.3. An alternative account

While it is very likely that *ta-ziji* as a whole was parafoveally processed when the preceding local verb was fixated upon, there is an alternative account: The parafoveal-on-foveal effect observed at the verb was not resulted from the processing of *ta-ziji* (i.e. reflexive), but simply from the processing of the first character *ta* (i.e. pronoun) alone. It can be said that when the local verb was in the attentional focus, only *ta* (with the accompanying gender information) was activated parafoveally. In this case, because Principle B of the Binding theory (Chomsky, 1981) assumes that the pronoun refers to the antecedent outside the local domain, the enhanced reading time would be observed for gender-matching matrix subject relative to the gender-mismatching one, regardless of the gender congruence between the reflexive and the local subject. However, in the present study, the interference effect from the matrix subject was robustly observed only when the local subject mismatched *ta* in gender. Therefore, it is unlikely that only *ta* is available in parafoveal vision; instead, *ta-ziji* as a whole is visible in the parafovea.

5. Conclusion and future directions

Our eye-tracking experiment on Chinese compound reflexive revealed early interference effects from gender-matching matrix subjects, though such effects occurred only under highly constrained conditions where the local subject mismatched *ta-ziji*. Our findings are mostly consistent with the structure-favoring cue-based retrieval model, where both structural and non-structural (gender) cues guide antecedent search, with the structural cue outweighing the gender cue in the resolution of *ta-ziji*.

An open question for future work has to do with other non-structural retrieval cues, such as animacy/humanness information of nouns and properties of verbs. In our experimental sentences, only human nouns were used as candidate antecedents for *ta-ziji*, and local verbs equally restrain the probability of *ta-ziji* referring to the local subject or the matrix subject. But animacy/humanness is likely to be a stronger retrieval cue than

gender (cf. Dillon et al., 2016), and unequally biased verbs might provide extra information for the interpretation of the reflexive (Jin, 2003; Li & Zhou, 2010). Thus, future work might explore potential effects of the property of verbs and their potential interaction with gender or animacy information on the processing of *ta-ziji*.

Notes

1. Existing work has shown that the two eyes are well-aligned most of the time, and fixation durations recorded from two eyes are highly correlated ($r = 0.98$) (Kliegl et al., 2006; Nuthmann & Kliegl, 2009).
2. We thank Brian Dillon for suggesting this.
3. One thing worth noting is in all the four conditions, the probability of choosing structurally licit antecedents (i.e., local subjects) never reached 100%. In other words, there were retrieval errors (i.e., misretrievals of matrix subjects) across the board, even in the Local-mismatch, Matrix-mismatch condition where the rate of misretrieval was 33%. As we speculate, this might be in part due to the nature of the forced-choice questionnaire and the use of equi-biased verbs in our design.
4. Note that our 95% *CrI* estimation of the interference effect on total reading time is similar to the 95% *CrI* estimate in Jäger et al. (2020, p. 13, Table 4): the present paper (back transform to ms): [–40, 27], mean = –17 ms; Jager et al.: [–43, 16], mean = –12 ms.
5. We thank Brian Dillon for suggesting this.

Acknowledgements

Disclosure statement

Funding

ORCID

Wenshuo Chang http://orcid.org/0000-0002-3873-235X
Yunyan Duan http://orcid.org/0000-0003-3730-1133
Fuyun Wu http://orcid.org/0000-0001-9995-918X

References

Akaike, H. (1974). A new look at the statistical model identification. *IEEE Transactions on Automatic Control*, 19(6), 716–723. https://doi.org/10.1109/TAC.1974.1100705

Baayen, R. H., Davidson, D. J., & Bates, D. M. (2008). Mixed-effects modeling with crossed random effects for subjects and items. *Journal of Memory and Language*, 59(4), 390–412. https://doi.org/10.1016/j.jml.2007.12.005

Badecker, W., & Straub, K. (2002). The processing role of structural constraints on interpretation of pronouns and anaphors. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 28(4), 748–769. https://doi.org/10.1037/0278-7393.28.4.748

Barr, Dale J., Levy, Roger, Scheepers, Christoph, & Tily, Harry J. (2013). Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language*, 68(3), 255–278. https://doi.org/10.1016/j.jml.2012.11.001

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, 67(1), 1–48. https://doi.org/10.18637/jss.v067.i01

Brainard, D. H. (1997). The Psychophysics Toolbox. *Spatial Vision*, 10(4), 433–436. https://doi.org/10.1163/156856897X00357

Bürkner, P.-C. (2017). *Brms*: An *R* package for Bayesian multilevel models using *Stan*. *Journal of Statistical Software*, 80(1), 1–28. https://doi.org/10.18637/jss.v080.i01

Cai, Q., & Brysbaert, M. (2010). SUBTLEX-CH: Chinese word and character frequencies based on film subtitles. *PLoS ONE*, 5(6), e10729. https://doi.org/10.1371/journal.pone.0010729

Chen, Z., Jäger, L., & Vasishth, S. (2012). How structure-sensitiveis the parser? Evidence from Mandarin Chinese. In B. Stolterfoht & S. Featherston (Eds.), *Empirical approaches to linguistic theory: Studies of meaning and structure (studies in generative grammar)* (pp. 43–62). Mouton de Gruyter.

Chomsky, N. (1981). *Lectures on government and binding: The Pisa Lectures*. Foris Publications.

Cornelissen, F. W., Peters, E. M., & Palmer, J. (2002). The Eyelink Toolbox: Eye tracking with MATLAB and the Psychophysics Toolbox. *Behavior Research Methods, Instruments, & Computers*, 34(4), 613–617. https://doi.org/10.3758/BF03195489

Cunnings, I., & Felser, C. (2013). The role of working memory in the processing of reflexives. *Language and Cognitive Processes*, 28(1-2), 188–219. https://doi.org/10.1080/01690965.2010.548391

Cunnings, I., Patterson, C., & Felser, C. (2015). Structural constraints on pronoun binding and coreference: Evidence from eye movements during reading. *Frontiers in Psychology*, 6(840). https://doi.org/10.3389/fpsyg.2015.00840

Cunnings, I., & Sturt, P. (2014). Coargumenthood and the processing of reflexives. *Journal of Memory and Language*, 75, 117–139. https://doi.org/10.1016/j.jml.2014.05.006

Dillon, B., Chow, W.-Y., Wagers, M., Guo, T., Liu, F., & Phillips, C. (2014). The structure-sensitivity of memory access: Evidence from Mandarin Chinese. *Frontiers in Psychology*, 5 (1025). https://doi.org/10.3389/fpsyg.2014.01025

Dillon, B., Chow, W.-Y., & Xiang, M. (2016). The relationship between anaphor features and antecedent retrieval: Comparing Mandarin Ziji and Ta-Ziji. *Frontiers in Psychology*, 6(1966). https://doi.org/10.3389/fpsyg.2015.01966

Dillon, B., Mishler, A., Sloggett, S., & Phillips, C. (2013). Contrasting intrusion profiles for agreement and anaphora:

Experimental and modeling evidence. *Journal of Memory and Language*, *69*(2), 85–103. https://doi.org/10.1016/j.jml.2013.04.003

Fine, A. B., Jäger, T. F., Farmer, T. A., & Qian, T. (2013). Rapid expectation adaptation during syntactic comprehension. *PLoS ONE*, *8*(10), e77661. https://doi.org/10.1371/journal.pone.0077661

Frazier, M., Ackerman, L., Baumann, P., Potter, D., & Yoshida, M. (2015). Wh-filler-gap dependency formation guides reflexive antecedent search. *Frontiers in Psychology*, *6*(1504). https://doi.org/10.3389/fpsyg.2015.01504

Gelman, A., & Rubin, D. B. (1992). Inference from Iterative Simulation using multiple Sequences. *Statistical Science*, *7*(4), 457–472. https://doi.org/10.1214/ss/1177011136

Gibson, E., Bergen, L., & Piantadosi, S. T. (2013). Rational integration of noisy evidence and prior semantic expectations in sentence interpretation. *Proceedings of the National Academy of Sciences*, *110*(20), 8051–8056. https://doi.org/10.1073/pnas.1216438110

Hohenstein, S., & Kliegl, R. (2014). Semantic preview benefit during reading. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *40*(1), 166–190. https://doi.org/10.1037/a0033670

Huang, C.-T. J., Li, Y. A., & Li, Y. (2009). *The syntax of Chinese*. Cambridge University Press.

Inhoff, A. W., & Liu, W. (1998). The perceptual span and oculomotor activity during the reading of Chinese sentences. *Journal of Experimental Psychology: Human Perception and Performance*, *24*(1), 20–34. https://doi.org/10.1037/0096-1523.24.1.20

Jäger, L. A., Engelmann, F., & Vasishth, S. (2015). Retrieval interference in reflexive processing: Experimental evidence from Mandarin, and computational modeling. *Frontiers in Psychology*, *6*(617). https://doi.org/10.3389/fpsyg.2015.00617

Jäger, L. A., Engelmann, F., & Vasishth, S. (2017). Similarity-based interference in sentence comprehension: Literature review and Bayesian meta-analysis. *Journal of Memory and Language*, *94*, 316–339. https://doi.org/10.1016/j.jml.2017.01.004

Jäger, L. A., Mertzen, D., Van Dyke, J. A., & Vasishth, S. (2020). Interference patterns in subject-verb agreement and reflexives revisited: A large-sample study. *Journal of Memory and Language*, *111*, 1–21. https://doi.org/10.1016/j.jml.2019.104063

Jin, Z. (2003). Verb restraint function to Ziji (自己) long-distance binding. *Chinese Language Learning*, *4*, 9–12. (in Chinese).

Kennedy, A. (1998). The influence of parafoveal words on foveal inspection time: Evidence for a processing trade-off. In *Eye guidance in reading and scene perception* (pp. 149–179). Elsevier Science. https://doi.org/10.1016/B978-008043361-5/50008-0

Kennedy, A., Pynte, J., & Ducrot, S. (2002). Parafoveal-on-foveal interactions in word recognition. *The Quarterly Journal of Experimental Psychology Section A: Human Experimental Psychology*, *55*(4), 1307–1337. https://doi.org/10.1080/02724980244000071

Kliegl, R., Nuthmann, A., & Engbert, R. (2006). Tracking the mind during reading: The influence of past, present, and future words on fixation durations. *Journal of Experimental Psychology: General*, *135*(1), 12–35. https://doi.org/10.1037/0096-3445.135.1.12

Kush, D., & Phillips, C. (2014). Local anaphor licensing in an SOV language: Implications for retrieval strategies. *Frontiers in Psychology*, *5*(1252). https://doi.org/10.3389/fpsyg.2014.01252

Kuznetsova, A., Brockhoff, P. B., & Christensen, R. H. B. (2017). lmerTest package: Tests in linear mixed effects models. *Journal of Statistical Software*, *82*(13), 1–26. https://doi.org/10.18637/jss.v082.i13

Kwon, N., & Sturt, P. (2014). The use of control information in dependency formation: An eye-tracking study. *Journal of Memory and Language*, *73*, 59–80. https://doi.org/10.1016/j.jml.2014.02.005

Laurinavichyute, A., Jäger, L. A., Akinina, Y., Roß, J., & Dragoy, O. (2017). Retrieval and encoding interference: Cross-linguistic evidence from anaphor processing. *Frontiers in Psychology*, *8*(965). https://doi.org/10.3389/fpsyg.2017.00965

Levy, R., Bicknell, K., Slattery, T., & Rayner, K. (2009). Eye movement evidence that readers maintain and act on uncertainty about past linguistic input. *Proceedings of the National Academy of Sciences*, *106*(50), 21086–21090. https://doi.org/10.1073/pnas.0907664106

Lewandowski, D., Kurowicka, D., & Joe, H. (2009). Generating random correlation matrices based on vines and extended onion method. *Journal of Multivariate Analysis*, *100*(9), 1989–2001. https://doi.org/10.1016/j.jmva.2009.04.008

Lewis, R. L., & Vasishth, S. (2005). An activation-based model of sentence processing as skilled memory retrieval. *Cognitive Science*, *29*(3), 375–419. https://doi.org/10.1207/s15516709cog0000_25

Li, N., Wang, S., Mo, L., & Kliegl, R. (2018). Contextual constraint and preview time modulate the semantic preview effect: Evidence from Chinese sentence reading. *Quarterly Journal of Experimental Psychology*, *71*(1), 241–249. https://doi.org/10.1080/17470218.2017.1310914

Li, X., & Zhou, X. (2010). Who is ziji? ERP responses to the Chinese reflexive pronoun during sentence comprehension. *Brain Research*, *1331*, 96–104. https://doi.org/10.1016/j.brainres.2010.03.050

Logačev, P., & Vasishth, S. (2016). A multiple-channel model of task-dependent ambiguity resolution in sentence comprehension. *Cognitive Science*, *40*(2), 266–298. https://doi.org/10.1111/cogs.12228

Matuschek, H., Kliegl, R., Vasishth, S., Baayen, H., & Bates, D. (2017). Balancing Type I error and power in linear mixed models. *Journal of Memory and Language*, *94*, 305–315. https://doi.org/10.1016/j.jml.2017.01.001

McElree, B., Foraker, S., & Dyer, L. (2003). Memory structures that subserve sentence comprehension. *Journal of Memory and Language*, *48*(1), 67–91. https://doi.org/10.1016/S0749-596X(02)00515-6

Nuthmann, A., & Kliegl, R. (2009). An examination of binocular reading fixations based on sentence corpus data. *Journal of Vision*, *9*(5), 31–31. https://doi.org/10.1167/9.5.31

Pan, J., Yan, M., & Laubrock, J. (2017). Perceptual span in oral reading: The case of Chinese. *Scientific Studies of Reading*, *21*(3), 254–263. https://doi.org/10.1080/10888438.2017.1283694

Parker, D., & Phillips, C. (2017). Reflexive attraction in comprehension is selective. *Journal of Memory and Language*, *94*, 272–290. https://doi.org/10.1016/j.jml.2017.01.002

Pelli, D. G. (1997). The VideoToolbox software for visual psychophysics: Transforming numbers into movies. *Spatial Vision*, *10*(4), 437–442. https://doi.org/10.1163/156856897X00366

Qian, J., & Wu, F. (2016). Immediate effects of Binding Principle A on complex reflexive processing in Chinese. *Modern Foreign Languages*, *39*(4), 495–506. (in Chinese).

Rayner, K. (1975). The perceptual span and peripheral cues in reading. *Cognitive Psychology*, *7*(1), 65–81. https://doi.org/10.1016/0010-0285(75)90005-5

Rayner, K. (1998). Eye movements in reading and information processing: 20 years of research. *Psychological Bulletin*, *124*(3), 372–422. https://doi.org/10.1037/0033-2909.124.3.372

Schotter, E. R., Lee, M., Reiderman, M., & Rayner, K. (2015). The effect of contextual constraint on parafoveal processing in reading. *Journal of Memory and Language*, *83*, 118–139. https://doi.org/10.1016/j.jml.2015.04.005

Sturt, P. (2003). The time-course of the application of binding constraints in reference resolution. *Journal of Memory and Language*, *48*(3), 542–562. https://doi.org/10.1016/S0749-596X(02)00536-3

Sturt, P., & Kwon, N. (2015). The processing of raising and nominal control: An eye-tracking study. *Frontiers in Psychology*, *6*(331). https://doi.org/10.3389/fpsyg.2015.00331

Tsai, J.-L., Kliegl, R., & Yan, M. (2012). Parafoveal semantic information extraction in traditional Chinese reading. *Acta Psychologica*, *141*(1), 17–23. https://doi.org/10.1016/j.actpsy.2012.06.004

Van Dyke, J. A., & McElree, B. (2006). Retrieval interference in sentence comprehension. *Journal of Memory and Language*, *55*(2), 157–166. https://doi.org/10.1016/j.jml.2006.03.007

Vasishth, S., Brüssow, S., Lewis, R. L., & Drenhaus, H. (2008). Processing polarity: How the ungrammatical intrudes on the grammatical. *Cognitive Science: A Multidisciplinary Journal*, *32*(4), 685–712. https://doi.org/10.1080/03640210802066865

von der Malsburg, T., & Angele, B. (2017). False positives and other statistical errors in standard analyses of eye movements in reading. *Journal of Memory and Language*, *94*, 119–133. https://doi.org/10.1016/j.jml.2016.10.003

Wagers, M. W., Lau, E. F., & Phillips, C. (2009). Agreement attraction in comprehension: Representations and processes. *Journal of Memory and Language*, *61*(2), 206–237. https://doi.org/10.1016/j.jml.2009.04.002

Xiang, M., Dillon, B., & Phillips, C. (2009). Illusory licensing effects across dependency types: ERP evidence. *Brain and Language*, *108*(1), 40–55. https://doi.org/10.1016/j.bandl.2008.10.002

Yan, M., Kliegl, R., Shu, H., Pan, J., & Zhou, X. (2010). Parafoveal load of word N+1 modulates preprocessing effectiveness of word N+2 in Chinese reading. *Journal of Experimental Psychology: Human Perception and Performance*, *36*(6), 1669–1676. https://doi.org/10.1037/a0019329

Yan, M., Richter, E. M., Shu, H., & Kliegl, R. (2009). Readers of Chinese extract semantic information from parafoveal words. *Psychonomic Bulletin & Review*, *16*(3), 561–566. https://doi.org/10.3758/PBR.16.3.561

Yan, M., Risse, S., Zhou, X., & Kliegl, R. (2012). Preview fixation duration modulates identical and semantic preview benefit in Chinese reading. *Reading and Writing*, *25*(5), 1093–1111. https://doi.org/10.1007/s11145-010-9274-7

Yan, M., & Sommer, W. (2015). Parafoveal-on-foveal effects of emotional word semantics in reading Chinese sentences: Evidence from eye movements. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *41*(4), 1237–1243. https://doi.org/10.1037/xlm0000095

Yan, M., Wang, A., Song, H., & Kliegl, R. (2019). Parafoveal processing of phonology and semantics during the reading of Korean sentences. *Cognition*, *193*, 104009. https://doi.org/10.1016/j.cognition.2019.104009

Yan, M., Zhou, W., Shu, H., & Kliegl, R. (2012). Lexical and sublexical semantic preview benefits in Chinese reading. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *38*(4), 1069–1075. https://doi.org/10.1037/a0026935

Yan, M., Zhou, W., Shu, H., & Kliegl, R. (2015). Perceptual span depends on font size during the reading of Chinese sentences. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *41*(1), 209–219. https://doi.org/10.1037/a0038097

Yang, J., Li, N., Wang, S., Slattery, T. J., & Rayner, K. (2014). Encoding the target or the plausible preview word? The nature of the plausibility preview benefit in reading Chinese. *Visual Cognition*, *22*(2), 193–213. https://doi.org/10.1080/13506285.2014.890689

Yang, J., Wang, S., Chen, H.-C., & Rayner, K. (2009). The time course of semantic and syntactic processing in Chinese sentence comprehension: Evidence from eye movements. *Memory & Cognition*, *37*(8), 1164–1176. https://doi.org/10.3758/MC.37.8.1164

Yang, J., Wang, S., Tong, X., & Rayner, K. (2012). Semantic and plausibility effects on preview benefit during eye fixations in Chinese reading. *Reading and Writing*, *25*(5), 1031–1052. https://doi.org/10.1007/s11145-010-9281-8

Yen, M.-H., Tsai, J.-L., Tzeng, O. J. L., & Hung, D. L. (2008). Eye movements and parafoveal word processing in reading Chinese. *Memory & Cognition*, *36*(5), 1033–1045. https://doi.org/10.3758/MC.36.5.1033

Zang, C., Wang, Y., Bai, X., Yan, G., Drieghe, D., & Liversedge, S. P. (2016). The use of probabilistic lexicality cues for word segmentation in Chinese reading. *Quarterly Journal of Experimental Psychology*, *69*(3), 548–560. https://doi.org/10.1080/17470218.2015.1061030