Journal of Memory and Language 79–80 (2015) 97–120

Contents lists available at ScienceDirect

Journal of Memory and Language

journal homepage: www.elsevier.com/locate/jml

The subject-relative advantage in Chinese: Evidence for expectation-based processing

Lena Jäger [a,*], Zhong Chen [b], Qiang Li [c,d], Chien-Jer Charles Lin [e], Shravan Vasishth [a,f]

[a] Dept. of Linguistics, University of Potsdam, Germany
[b] Dept. of Modern Languages and Cultures, Rochester Institute of Technology, USA
[c] School of Foreign Languages, Dalian University of Technology, China
[d] Dept. of Communicative Disorders, University of Louisiana at Lafayette, USA
[e] Dept. of East Asian Languages and Cultures, Indiana University Bloomington, USA
[f] School of Mathematics and Statistics, University of Sheffield, UK

ARTICLE INFO

Article history:
Received 27 September 2013
revision received 10 October 2014

Keywords:
Sentence processing
Relative clause
Structural expectation
Working-memory
Surprisal
Chinese

ABSTRACT

Chinese relative clauses are an important test case for pitting the predictions of expectation-based accounts against those of memory-based theories. The memory-based accounts predict that object relatives are easier to process than subject relatives because, in object relatives, the distance between the relative clause verb and the head noun is shorter. By contrast, expectation-based accounts such as surprisal predict that the less frequent object relative should be harder to process. In previous studies on Chinese relative clause comprehension, local ambiguities may have rendered a comparison between relative clause types uninterpretable. We designed experimental materials in which no local ambiguities confound the comparison. We ran two experiments (self-paced reading and eye-tracking) to compare reading difficulty in subject and object relatives which were placed either in subject or object modifying position. The evidence from our studies is consistent with the predictions of expectation-based accounts but not with those of memory-based theories.

Tracking expectations is of central importance in language comprehension: as we read or hear a sentence, we constantly generate predictions about upcoming material at every linguistic level. According to one view, these predictions are a result of our production system being deployed as an "emulator" during language comprehension (Pickering & Garrod, 2007); a related assumption is that the comprehender maintains and uses linguistic knowledge probabilistically to incrementally parse a sentence (Hale, 2001; Levy, 2008). Common to these views is the idea that the frequency of the structures we produce in a language at least partly determines what we incrementally predict while engaged in a comprehension task.

Once an expectation has been generated, it is either met or not met. When an expectation is met, generally no disruptions occur; but when the expectation is dashed, we often experience processing difficulty. Although the role of predictive processing is well-established in the psycholinguistic literature, in sentence comprehension research, the work of Hale (2001) and Levy (2008) has considerably sharpened our understanding of what it means to have a syntactic expectation, and, more importantly, how to quantify the effect of a dashed expectation. The key idea here is that building a rarer syntactic structure, with which comprehenders have less experience, is more difficult than building a relatively more frequent structure. As a simple

* Corresponding author at: Department of Linguistics, University of Potsdam, Karl-Liebknecht-Str. 24/25, 14476 Potsdam, Germany.
E-mail address: lena.jaeger@uni-potsdam.de (L. Jäger).

http://dx.doi.org/10.1016/j.jml.2014.10.005
0749-596X/

example, in English, reading the relative pronoun in a sentence like *The man who* ... raises an expectation for a subject relative because it is more frequent; if the sentence continues with an object relative (e.g., *The man who the woman* ...), a slowdown is predicted (relative to an appropriate baseline sentence). Staub (2010) has shown that this prediction is correct. The surprisal metric introduced by Hale (2001) thus formalizes the well-established idea of reanalysis that dates back to the early days of psycholinguistics (Frazier, 1979). Surprisal assumes a ranked-parallel parser that ranks the available parses by their conditional probabilities. When the input favors a parse that has not been ranked highest, a re-ranking occurs. This re-ranking due to dashed expectations is computationally costly and formalizes the idea of reanalysis.

Over the last few years, evidence has started to accumulate in favor of such an expectation-based account of sentence comprehension. For example, Levy, Fedorenko, and Gibson (2013) have recently shown, using Russian relative clauses, that dashed expectations lead to slowdowns in reading; in other words, building a rarer structure is more difficult than building a more frequent structure. However, it is currently unknown how generally applicable the expectation-based account is cross-linguistically; as Pickering and Van Gompel (2006) have documented, even robust theories of sentence processing have floundered in the face of a cross-linguistic investigation. It is particularly important to stress-test the expectation-based account given that its predictions sometimes go directly against another, well-established class of explanation, the memory-based accounts of sentence comprehension. Such models attribute processing difficulty to limitations of memory resources (Clifton & Frazier, 1989; Frazier, 1979; Frazier & Fodor, 1978; Gibson, 1998; Gibson, 2000; Just & Carpenter, 1992; Lewis & Vasishth, 2005; Lewis, Vasishth, & Van Dyke, 2006; Miller & Chomsky, 1963). A prominent example for this class of account is the Dependency Locality Theory (DLT) (Gibson, 1998; Gibson, 2000). DLT assumes that processing difficulty depends on so-called integration cost and storage cost. Structural integration cost is defined as a linearly increasing function of the number of new discourse referents that intervene between the constituent that is currently being processed and the constituent(s) with which a syntactic dependency has to be built. Storage cost assumes that processing difficulty linearly increases as a function of the number of predicted heads.

Chinese relative clauses (RCs) are a critical test of the opposing predictions of the expectation-based account and memory-based accounts like the DLT. Due to the syntactic properties of Chinese, memory-based accounts and expectation-based accounts make diametrically opposed predictions about the processing difficulty associated with subject relative clauses (SRs) compared to object relative clauses (ORs). Because of the great importance of Chinese relative clauses in unpacking the relative contributions of expectations and memory cost (among other theoretical explanations), they have drawn considerable attention in recent psycholinguistic research. In contrast to other languages such as Japanese and Korean with pre-nominal relative clauses, which display a consistent SR advantage in all published studies (Japanese: Miyamoto & Nakamura, 2003; Ueno & Garnsey, 2008; Korean: Kwon, Kluender, Kutas, & Polinsky, 2013; Kwon, Lee, Gordon, Kluender, & Polinsky, 2010; Kwon, Polinsky, & Kluender, 2006; Yun, Whitman, & Hale, 2010), the evidence from Chinese relatives is not conclusive. As discussed below, the conflicting results are likely due to local ambiguities in the experimental materials.

(1) a. **Subject-modifying SR**

		邀请	男孩	的		女孩	认识	老师。
[$_{RC}$	t_i	yaoqing	nanhai	de	]	nühai$_i$	renshi	laoshi.
		invite	boy	REL		girl	know	teacher

'The girl who invites the boy knows the teacher.'

b. **Subject-modifying OR**

	男孩	邀请		的		女孩	认识	老师。
[$_{RC}$	nanhai	yaoqing	t_i	de	]	nühai$_i$	renshi	laoshi.
	boy	invite		REL		girl	know	teacher

'The girl who the boy invites knows the teacher.'

c. **Object-modifying SR**

老师	认识			邀请	男孩	的		女孩。
laoshi	renshi	[$_{RC}$	t_i	yaoqing	nanhai	de	]	nühai$_i$.
teacher	know			invite	boy	REL		girl

'The teacher knows the girl who invites the boy.'

d. **Object-modifying OR**

老师	认识		男孩	邀请		的		女孩。
laoshi	renshi	[$_{RC}$	nanhai	yaoqing	t_i	de	]	nühai$_i$.
teacher	know		boy	invite		REL		girl

'The teacher knows the girl who the boy invites.'

Chinese relative clauses (1) are pre-nominal, i.e., the head noun (here, *nühai*, 'girl') appears after the relative clause. This has the interesting consequence that the distance between head noun and the gap inside the relative clause (or, equivalently, the relative clause verb) is greater in SRs than in ORs. By contrast, in English, the distance between the gap (or verb) and head noun is greater in ORs than SRs. As a result of this reversal in gap-head distance in Chinese, memory-based accounts like the DLT, which index processing cost in terms of the distance between co-dependent elements, predict longer reading times at the head noun in SRs than ORs. In other words, memory-based accounts predict an object relative advantage.[1]

By contrast, the expectation account outlined above, which assumes that rare structures are harder to process, predicts that in ORs, *longer* reading times should be seen compared to SRs (i.e., a subject relative advantage); this is because, similar to English, SRs are more frequent than object relatives in Chinese (Chen, Grove, & Hale, 2012; Hsiao & Gibson, 2003; Vasishth, Chen, Li, & Guo, 2013; Wu, 2009; Wu, 2011; Wu, Kaiser, & Andersen, 2010). The expectation-based account would predict an SR advantage as soon as the relative clause is built; and this SR advantage could plausibly spill over to the head noun and beyond as well. But, crucially, the predicted onset of the SR advantage depends on the point at which the structural properties of the experimental materials allow the comprehender to predict a relative clause. In the materials used in previous experiments, several local ambiguities might have confounded the results (for a discussion of local ambiguities in Chinese RCs also see Lin & Bever, 2006; Lin & Bever, 2011; Qiao, Shen, & Forster, 2012; Vasishth et al., 2013). Indeed, Hsiao and MacDonald (2013) and Hsiao, Li, and MacDonald (2014) have argued that the results of previous studies on Chinese RCs can be largely explained by the local ambiguities in the stimuli. Thus, although the theoretically interesting aspect of Chinese relative clauses lies in the diametrically opposed predictions of the expectation-based account vs the memory-based accounts, it is vital to bring the local ambiguities under experimental control before we can investigate these opposing predictions. In the present paper, we investigate the predictions of these two classes of account using an experiment design where we strongly constrain the local ambiguities that have confounded previous studies.

We will first give an overview of the various ambiguities present in Chinese RCs and then describe how they might have confounded previous studies and how we deal with these ambiguities in our experimental materials.

[1] Note that these predictions do not depend on whether we assume the psychological reality of empty categories (the gap inside the relative clause) or not (Pickering & Barry, 1991; Traxler & Pickering, 1996). For Chinese relative clauses the same predictions hold if we assume that the head noun is directly associated with the verb – not via a two-step mechanism involving a gap — because the distance between the head noun and the relative clause verb is shorter in object relatives compared to subject relatives.

Local ambiguities in Chinese relative clauses

Fig. 1 provides an overview of how the various local ambiguities differently affect subject- and object-modifying SRs and ORs.[2]

Local ambiguity 1 in Fig. 1 shows an alternative parse that is due to the lexical ambiguity of the relativizer *de*: when *de* directly follows a noun phrase, it can also be read as a possessive marker. In this case, the NP preceding *de* is interpreted as a possessor of a following NP. Therefore, in SRs, when reading the RC object followed by the RC head noun, the reader can interpret them as a complex NP (resulting in 'the boy's girl' in example 1a and 1c). This parse becomes impossible at the main clause verb in subject-modifying SRs and at the end of the clause in object-modifying SRs.

Local ambiguity 2 in Fig. 1 shows alternative parses that involve null subjects (for a discussion of null subjects in Chinese see Huang, 1989). In SRs, a dropped subject (*pro*) whose reference is determined by the context of the utterance is postulated before the RC verb. Thus, the RC verb together with the RC object can be read as a main clause in subject-modifications (resulting in '*pro* invites the boy' in example 1a) and as sentential object in object-modifications (resulting in 'The teacher knows that *pro* invites the boy' in example 1c).[3] These parses become impossible upon reaching the relativizer *de*. (If the relativizer is interpreted as a possessive marker, the parse becomes impossible at the main clause verb in subject-modifications and the end of the clause in object-modifications.) Note that in object-modifications, *Local ambiguity 2* is only present if the main clause verb allows a sentential object.

Local ambiguity 3 in Fig. 1 shows an alternative possible parse in ORs. The RC subject is interpreted as the main clause subject (in subject-modifications) or the main clause object (in object-modifications). In object-modifying ORs, this main clause reading becomes impossible already at the following word, the RC verb. In subject-modifying ORs, the RC verb can still be integrated into this parse as main clause verb. This parse becomes impossible at the relativizer.

Local ambiguity 4 in Fig. 1 shows how, in object-modifying ORs, the RC subject together with the RC verb can be interpreted as a sentential object (in case the main clause verb allows a sentential object). This parse becomes impossible at the relativizer.

Thus, it is clear that in order to fairly compare SRs and ORs, we need a syntactic configuration that allows us to

[2] Note that the list of ambiguities in Chinese RCs can never be exhaustive due to the recursive nature of syntax. E.g., all sentential nodes could be embedded below another sentential node as sentential subject or sentential object. The local ambiguities which result from the recursive application of a production rule on a sub-tree where the same local ambiguity is already present are therefore not discussed here. Crucially for the experimental design described below, the elimination of an alternative parse automatically eliminates all alternative parses that result from this parse via recursion.

[3] In the concrete example given in 1c, this parse is not possible since the verb *renshi* ('know' or 'recognize') does not take sentential complements. If *renshi* is replaced by a verb like *zhidao* ('know' or 'acknowledge'), which allows sentential complements, this local ambiguity is present.

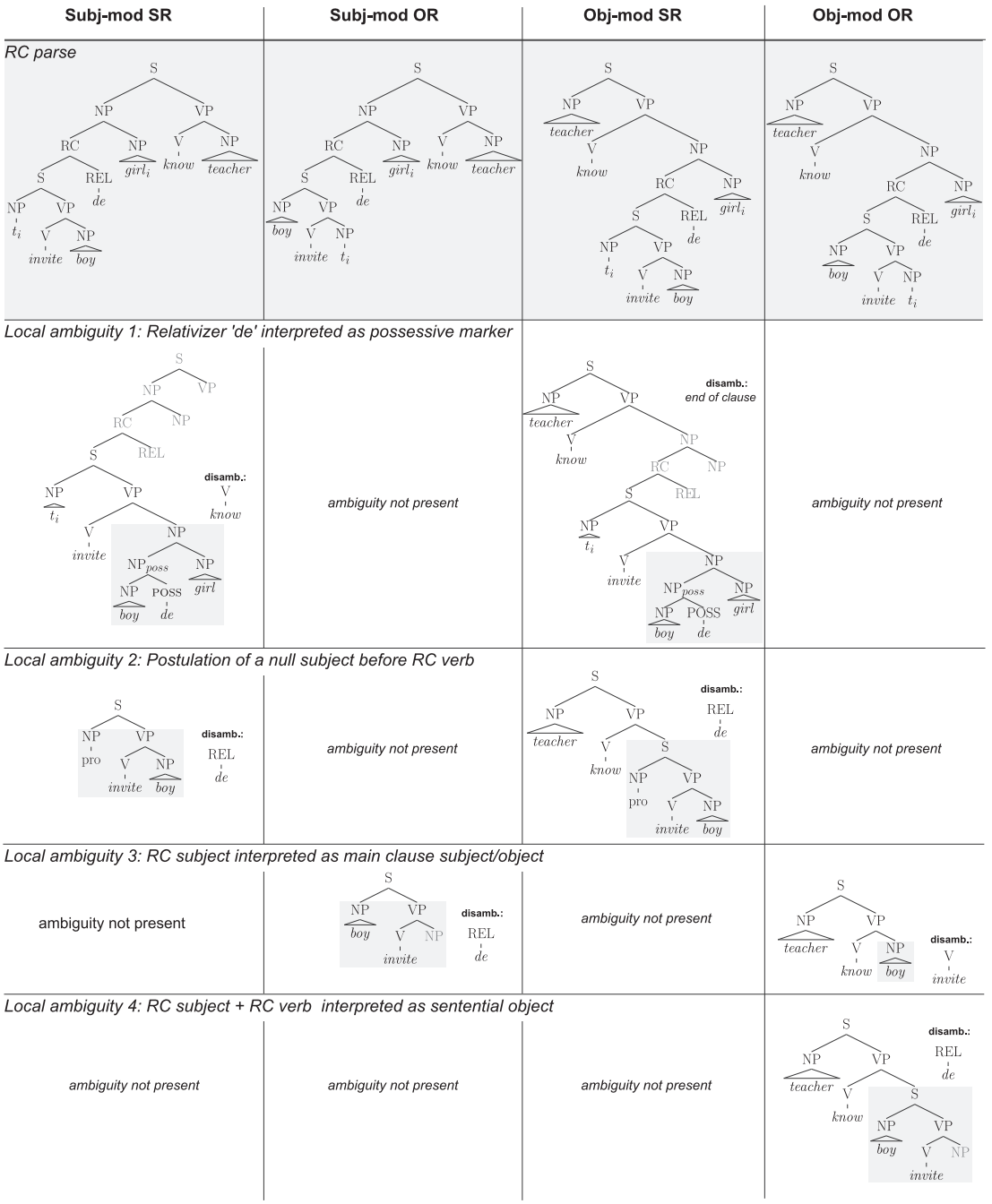

Fig. 1. Local syntactic ambiguities in Chinese relative clauses. The first line shows the relative clause parses for the subject- and object modifying SRs/ ORs presented in example 1. Lines 2–5 show the alternative parses up to the disambiguating word labeled as 'disamb'. The mis-analysed part of the tree that differs from the correct RC parse is highlighted with gray shading. Predicted nodes whose head has not been encountered yet are depicted in gray color.

eliminate the local ambiguities described above such that they cannot confound the comparison.

Previous work on Chinese relative clauses

Research on Chinese relative clause processing has dealt with the local ambiguities discussed in the last section in different ways. Earlier studies of Chinese relative clause processing (e.g., Hsiao & Gibson, 2003; Lin & Bever, 2006) directly presented Chinese relative clauses such as those in (1) in isolated sentences without taking confounding ambiguities into account. Hsiao and Gibson (2003) conducted a 2×2 self-paced reading study in which they manipulated RC type (SR vs OR) and number of embeddings (single vs double). In single embeddings, the authors found faster reading times in ORs compared to SRs at the relative clause region, namely the region containing the RC verb and the RC object in SRs, and the region containing

the RC subject and the RC verb in ORs. At the head noun, no effect was found. In double embeddings, the summed up reading time of the region containing the two relative clauses up to the second relativizer *de* was shorter in ORs than in SRs. The authors interpreted this OR advantage as evidence for a memory-based account and explained their results in terms of the Dependency Locality Theory (Gibson, 1998; Gibson, 2000). Note that, as pointed out by Lin and Bever (2006), the differences between SRs and ORs found in the pre-relativizer regions can be explained by the fact that the NP + V sequence of an OR, but not the V + NP sequence of an SR, is actually the canonical order of a main clause (Chinese is an SVO language).

Lin and Bever (2006) carried out two studies which suggest that SRs are processed faster than ORs; however, the SR advantage was only found in cases where the object of the main clause was modified by a relative clause. Subject-modifying relative clauses did not show any difference between RC types. Gibson and Wu (2013) pointed out that the Lin and Bever (2006) finding can be explained by another local ambiguity confounding the materials, namely that the initial NP of an OR may have been mistaken as the object of the matrix clause, as can be seen in (1d).

Subsequent studies report processing facilitation for ORs whereas others report an SR advantage. An SR advantage has been reported by Wu (2009); Vasishth et al. (2013) and Liu, Zhou, and Yang (2011). By contrast, several self-paced reading studies (e.g., Chen & Ning, 2008; Chen, Ning, Bi, & Dunlap, 2008; Lin & Garnsey, 2011) as well as experiments using maze tasks (e.g., Qiao et al., 2012) report an OR advantage. Wu, Kaiser, and Andersen (2012) show that the subject relative advantage is further modulated by the animacy configuration of the head noun and the embedded noun. They found an SR advantage only in case the subject was inanimate and the object was animate. In the reverse animacy configuration, the SR advantage disappeared. Contrary to these results, Zhang and Jiang (2010) report an ERP study in which they found that SRs were harder to process than ORs in subject-modifying relatives while in object-modifying relative clauses, a processing facilitation was observed in SRs. In none of the studies mentioned so far has the issue of local ambiguity been taken into account.

One attempt to overcome the local ambiguity issue was undertaken by Gibson and Wu (2013), who examined SRs vs ORs with disambiguating preceding context; they investigated only subject-modifying RCs. In their design, sentences similar to the materials of Hsiao and Gibson (2003) were preceded by a context that introduced the action described in the relative clause. The target sentence itself was the answer to a question that required the reader to identify either the agent or the patient of the RC action. The assumption was that the presence of such a context ensures that the reader expects either an SR or an OR and therefore will not be garden pathed. At the relativizer *de*, Gibson and Wu (2013) found slower self-paced reading times in SRs than in ORs. This comparison reached significance in the by-participant analysis but not in the by-item analysis. At the head noun, SRs were read significantly slower than ORs. No other comparisons reached significance. In contrast to Hsiao and Gibson (2003), who did not find an effect at the head noun in single embeddings, Gibson and Wu (2013)'s results are consistent with the predictions of DLT's structural integration cost. Moreover, at the relative clause region preceding the relativizer, Gibson and Wu (2013) did not find any statistically significant effects. This absence of an effect is consistent with the DLT storage cost metric under the assumption that the reader is aware of an upcoming relative clause from the very outset due to the preceding context, so that the number of predicted heads is identical in SRs vs ORs.

Vasishth et al. (2013) also replicated the OR advantage at the head noun that Gibson and Wu found. However, the effect was already significant at the relativizer preceding the head noun, which is not predicted by either DLT integration or storage cost metrics. In recent work, Lin (2014) has argued that the OR advantage observed in Gibson and Wu (2013)'s materials might reflect a syntactic priming from the context rather than a lower DLT integration cost in ORs. Using materials similar to Gibson and Wu (2013)'s, Lin (2014) additionally manipulated the different thematic orders in the context and found a stronger OR advantage when the thematic order in the preceding context was similar to the one of an OR.[4] Thus, the Gibson and Wu (2013) context manipulation may have introduced a new confound, thematic order of the context and target sentences, that rendered ORs easier to process than SRs.

As the above summary shows, the evidence about the processing of Chinese RCs is inconsistent across studies, and this is due at least in part to various confounds introduced by local syntactic ambiguities in the stimuli. Importantly, these ambiguities differ between SRs and ORs, as well as between subject-modifications and object-modifications.

Using syntactic context to eliminate local ambiguities in Chinese relative clauses

We present an experimental design which leads the reader to strongly predict that a relative clause is coming up. Unlike Gibson and Wu (2013), we eliminate the local ambiguities not by providing a biasing context but rather by creating a syntactic configuration in which the comprehender is highly likely to predict a relative clause. Several previous studies have used structural cues to generate a

[4] There is also an inconsistency in the predictions of the DLT as presented in Gibson (2000) and the DLT predictions derived for the experimental materials used in Gibson and Wu (2013). The structural integration cost metrics of DLT are defined as a function of the number of *new* discourse referents intervening between the two elements of the dependency that is being built. The following is the definition in Gibson (2000, p. 125) (the emphasis is ours): *The structural integration cost associated with connecting the syntactic structure for a newly input head h_2 to a projection of a head h_1 that is part of the current structure for the input is dependent on the complexity of the computations that took place between h_2 and h_1. For simplicity, it is assumed that one EU is consumed for each **new** discourse referent in the intervening region.* However, in Gibson and Wu (2013)'s materials, the discourse referents in question are already introduced in the context sentence and hence are no new discourse referents anymore when processing the relative clause. Thus, in fact, for the Gibson and Wu materials the DLT should predict no processing difference at the head noun in SRs vs ORs.

prediction for an upcoming RC parse in Chinese. For instance, Hsu, Phillips, and Yoshida (2005) showed that the presence of the OR marker *suo* led to a higher proportion of OR continuations in a sentence completion task and to shorter reading times at the relativizer and the head noun in self-paced reading. Other studies have used the semantic clash caused by local classifier-noun mismatches as an indicator of embedded clauses (Hsu, Hurewitz, & Phillips, 2006; Hsu et al., 2005; Wu, Luo, & Zhou, 2014; Yoshida, Aoshima, & Phillips, 2004). Note, however, that while a classifier-noun mismatch indicates RC boundaries in Japanese and Korean (Yoshida & Yoon, 2014), it alone may not be strong enough to predict an RC in Mandarin (Hsu et al., 2005; cf. Wu et al., 2014). Wu and colleagues also used the passive marker *bei* at the onset of the relative clause to lead comprehenders to expect an upcoming passive SR. Participants indeed showed faster self-paced reading times in the presence of this passive marker (Wu, 2013; Wu et al., 2014). Crucially, the above-cited studies differ from the present study in that they did not examine the processing asymmetry between SRs and ORs but the predictive strength of certain syntactic markers. Note also that the use of *suo* and the classifier-noun mismatch only indicated the existence of an OR, not an SR (see Lin & Bever, 2011) for a comprehensive review; by contrast, the passive marker *bei* can only mark a passivized SR. Thus, the syntactic markers *suo* and *bei* allow us to predict either an SR or an OR but not an RC in general and hence are not sufficient for our purposes.

We created a 2×2 factorial design that, across experimental conditions, leads readers to strongly predict a relative clause from the first word of the relative clause on. In addition to the manipulation of RC type, we included the manipulation of modification type for two reasons. First, the conflicting pattern reported in the literature might reflect different RC type preferences depending on modification type. For example, Hsiao and Gibson (2003) and Gibson and Wu (2013), who both report an OR advantage, only tested subject modifications. Lin and Bever (2006) report an SR advantage, but this effect reached significance only in object modifications. Second, for our materials, the different accounts not only differ in their predictions with respect to RC type but also with respect to modification type, as will be discussed below.

The experimental sentences consisted of a main clause whose subject or object was modified by an SR or an OR. This main clause was followed by another coordinate main clause as the spillover material; see (2) for an example item and Fig. 2 for the underlying syntactic structures.

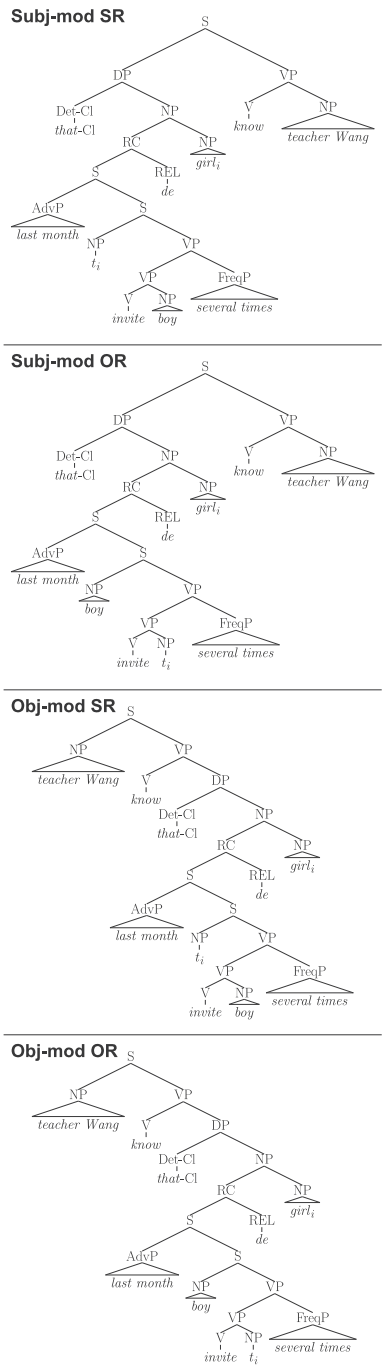

Fig. 2. Syntactic structure of the experimental materials. For the Chinese wording see example 2.

(2) a. **Subject-modifying SR**

那个			上个月	邀请了	男孩	几次	的		女孩	认识	王老师	因为	上过	她的	课。
Na-ge	[$_{RC}$	t_i	shanggeyue	yaoqing-le	nanhai	ji-ci	de	]	nühai$_i$	renshi	Wang laoshi	yinwei	shang-guo	ta-de	ke
DET-CL			last month	invite-ASP	boy	several-CL	REL		girl	know	teacher Wang	because	attend-ASP	her	class

'The girl who invited the boy several times last month knows teacher Wang because [she] has attended her class.'

b. **Subject-modifying OR**

那个		上个月	男孩	邀请了		几次	的		女孩	认识	王老师	因为	上过	她的	课。
Na-ge	[$_{RC}$	shanggeyue	nanhai	yaoqing-le	t_i	ji-ci	de	]	nühai$_i$	renshi	Wang laoshi	yinwei	shang-guo	ta-de	ke
DET-CL		last month	boy	invite-ASP		several-CL	REL		girl	know	teacher Wang	because	attend-ASP	her	class

'The girl who the boy invited several times last month knows teacher Wang because [she] has attended her class.'

c. **Object-modifying SR**

王老师	认识	那个			上个月	邀请了	男孩	几次	的		女孩	因为	教过	她的	课。
Wang laoshi	renshi	na-ge	[$_{RC}$	t_i	shanggeyue	yaoqing-le	nanhai	ji-ci	de	]	nühai$_i$	yinwei	jiao-guo	ta-de	ke
teacher Wang	know	DET-CL			last month	invite-ASP	boy	several-CL	REL		girl	because	teach-ASP	her	class

'Teacher Wang knows the girl who invited the boy several times last month because [she] has taught her class.'

d. **Object-modifying OR**

王老师	认识	那个		上个月	男孩	邀请了		几次	的		女孩	因为	教过	她的	课。
Wang laoshi	renshi	na-ge	[$_{RC}$	shanggeyue	nanhai	yaoqing-le	t_i	ji-ci	de	]	nühai$_i$	yinwei	jiao-guo	ta-de	ke
teacher Wang	know	DET-CL		last month	boy	invite-ASP		several-CL	REL		girl	because	teach-ASP	her	class

'Teacher Wang knows the girl who the boy invited several times last month because [she] has taught her class.'

First, we ensured that the relativizer *de* cannot be interpreted as a genitive marker (*Local ambiguity 1* in Fig. 1); this was done by inserting a frequency phrase (FreqP, *ji-ci* 'several times' in 2) consisting of a numeral adjective and a verbal classifier (Cl) before the relativizer. This FreqP attaches to the relative clause VP (see Fig. 2).

Second, we inserted a determiner (Det) (*zhe* 'this' or *na* 'that') followed by a nominal classifier (Cl) at the onset of the relative clause.[5] In Chinese, a Det + Cl sequence predicts an NP.[6] Therefore, the head noun is already predicted before the onset of the RC. This predicted NP, which will become the RC head, eliminates *Local ambiguity 2* (see Fig. 1): it ensures that in subject- and object-modifying SRs, a *pro* subject is impossible since the subject position of the clause is already occupied by Det + Cl and the predicted NP.

Third, we inserted an adverbial phrase at the onset of the relative clause, i.e., between the Det + Cl sequence and the RC verb (in SRs) or the RC subject (in ORs). Together with the inserted Det + Cl, the AdvP ensures that in ORs, the RC subject cannot be interpreted as a main clause subject/object (*Local ambiguity 3* in Fig. 1). Moreover, it ensures that in object-modifying ORs, the RC subject and RC verb cannot be interpreted as a sentential object (*Local ambiguity 4* in Fig. 1). This is because the AdvP predicts a clausal node to which it can attach (see Fig. 2). It cannot directly attach to the main clause because it appears between the Det + Cl and the NP predicted by the Det + Cl combination. Hence, the only possible phrase for the adverb to attach to is an RC that modifies the predicted NP.

Although the local ambiguities presented in Fig. 1 are eliminated in the present design, one new ambiguity has been introduced by the insertion of Det + Cl. The NP predicted by Det + Cl could be covertly realized as an elided NP, given a suitable context. In SRs, a parse structurally similar to *Local ambiguity 2* where the [$_{NP}$ pro] constituent is replaced by [$_{DP}$ Det Cl [$_{NP}$ ~~NP~~]] is still possible (*Na-ge ~~NP~~ shanggeyue yaoqing-le nanhai ji-ci* 'That one invited the boy several times last month'). This parse becomes impossible at the relativizer. In contrast to *pro*, Det + Cl + ~~NP~~ is also allowed in object positions. Therefore, in object-modifications, a parse like *Wang laoshi renshi na-ge ~~NP~~* 'Teacher Wang knows that one' is possible until encountering the AdvP. However, the interpretation of Det + Cl as Det + Cl + ~~NP~~ needs a highly constraining discourse context. Indeed, such a reading is highly infrequent, as we will show with a corpus search and a sentence completion task reported below.

In sum, the present design ensures that, in both object and subject-modifications, the most frequent structure that comprehenders expect is a relative clause. This avoids the potential garden path effects that might have confounded the results of earlier studies on Chinese RCs.

In order to empirically validate our assumptions about the experimental materials and derive quantitative predictions of the expectation-based account, we conducted two pre-tests: a corpus analysis and a sentence completion study. These are described next.

Pre-tests

Corpus analyses

We conducted corpus analyses in order to (i) empirically validate the assumption that the local ambiguities presented in Fig. 1 are eliminated in the experimental materials and that the most frequent structure that occurs after a determiner + classifier + adverb sequence is a relative clause and (ii) derive empirically grounded region-by-region predictions of the expectation-based account for the present materials. Predictions of the modification type and RC type factors can be derived from the probabilities of the respective structure at the point in the sentence where this structure is built given the left context of the sentence. We carried out four corpus searches on the Chinese Treebank 7.0 (Xue, Xia, Chiou, & Palmer, 2005), which consists of 51,447 fully parsed sentences (1,196,329 words) using the pattern matching tool Tregex (Levy & Andrew, 2006).

In order to verify that the local ambiguities presented in Fig. 1 are eliminated, we extracted all tokens containing Det + Cl + Adv sequences to obtain counts for the different structural types that follow this structure. Since the experimental items only used the determiner *zhe* or *na* ('this' or 'that', respectively), we restricted the determiner to these lexical items. The raw corpus counts are provided in Table A1 in the appendix. The results suggest that the zhe/na + Cl + Adv combination in Chinese indeed predicts a relative clause: 98.6% of all tokens of zhe/na + Cl + Adv are followed by a relative clause. Only a single token of a Det + Cl + Adv chunk (1.4% of all tokens) is not followed by a relative clause. In this sentence, an elided NP is postulated between the Det + Cl and the adverb (i.e., the newly introduced ambiguity in the present materials). These results confirm our syntactic analysis that the insertion of the Det + Cl + Adv sequence indeed eliminates all the local ambiguities presented in Fig. 1 and that the possibility of a configuration involving an elided NP is highly infrequent and thus unlikely to lead to a garden path later in the sentence. Here, we make the assumption, uncontroversial in psycholinguistics, that incremental parsing consists of a ranking of alternative parses using their relative frequencies, with a preference to predict the structure that has the highest frequency (Gibson, 2000; Pickering & Garrod, 2007). Thus, no non-RC parses leading to garden paths confound the comparison between SRs and ORs.

Next, in order to assess the region-by-region conditional probabilities associated with each experimental condition of the present materials, we conducted three more corpus analyses. The conditional probabilities derived from these corpus analyses are presented in Table 1.

[5] Note that the disambiguating character of the Det + Cl sequence in our materials only makes use of the general syntactic prediction that a classifier attaches to an NP; unlike Hsu et al. (2006) or Wu et al. (2014), our reasoning does not depend on any semantic match/mismatch of the classifier and a following noun.

[6] In case this NP is modified by a relative clause, the Det + Cl can be either located between the relative clause and the NP that is modified (i.e., the head of the RC), or it can precede the relative clause. In our materials, the Det + Cl sequence always precedes the relative clause.

Table 1
Summary of the corpus analyses based on the Chinese Treebank 7.0 (see Tables A2 and A3 for the underlying corpus counts). Conditional probabilities Prob (Target|Prefix) associated with the upcoming structure (Target) at the various regions (Prefix) of the experimental materials grouped by modification type.

Mod. type	Prefix	Target	Prob(Target | Prefix)
subj-mod	Det + Cl	RC	0.14
	Det + Cl + Adv	SR	0.78
	Det + Cl + Adv	OR	0.20
	Det + Cl + Adv + SR + *de*	Overt RC head	1.00
	Det + Cl + Adv + OR + *de*	Overt RC head	1.00
obj-mod	NP + V + Det + Cl	RC	0.12
	NP + V + Det + Cl + Adv	SR	0.85
	NP + V + Det + Cl + Adv	OR	0.15
	NP + V + Det + Cl + Adv + SR + *de*	Overt RC head	1.00
	NP + V + Det + Cl + Adv + OR + *de*	Overt RC head	1.00

First, to estimate the conditional probability of a subject-modifying vs an object-modifying relative clause given that the subject/object NP (i.e., the RC head) has been predicted by a preceding Det + Cl, we carried out a corpus search on structures following a Det + Cl chunk that either attached to a subject NP or to an object NP. When the Det + Cl modified a subject, it was followed by an RC (i.e., a subject-modifying RC) in 13.9% of the tokens. When Det + Cl modified an object, it was succeeded by an RC (i.e., an object-modifying RC) in 12.3% of the tokens (see Table A3 in the appendix). Thus, after having read a Det + Cl, an RC parse is similarly likely to follow in subject- and object-modifying contexts.

Second, in order to estimate conditional probabilities of an SR or OR appearing after Det + Cl + Adv for subject and object-modifications separately, we further categorized all tokens of the very first corpus search according to whether Det + Cl modified the subject NP or the object NP (or another constituent) of the sentence (see Table A2 in the appendix). In subject-modifications, 77.5% of the Det + Cl + Adv-tokens are followed by an SR and 20% by an OR.[7] In object-modifications, the Det + Cl + Adv chunk is followed by an SR in 85% of the tokens and by an OR in 15% of the tokens.[8] This analysis shows that SRs are predicted overwhelmingly more frequently than ORs in the present materials.

Third, we conducted a corpus analysis to derive the conditional probability of an overt RC head noun given the Det + Cl + Adv + OR/SR prefix. We therefore checked all tokens of SRs/ORs preceded by a Det + Cl + Adv sequence (i.e., the tokens obtained in the first corpus search) for whether they are headed or headless RCs. In all tokens, the RC head was overtly realized (headed RC).

In sum, the corpus analyses show that:

(i) The Det + Cl + Adv sequence in the experimental items leads readers to strongly predict an RC parse. This ensures that readers do not follow a non-RC parse leading them into garden path, i.e., the present design rules out that non-RC parses confound the comparison between SRs and ORs.
(ii) The conditional probability of an RC appearing in subject-modifying position is similar to the conditional probability of an RC appearing in object-modifying position.
(iii) The conditional probability of an SR appearing after Det + Cl + Adv is higher than the conditional probability of an OR.
(iv) The conditional probability of an overt RC head noun appearing after an RC that is preceded by a Det + Cl sequence and starts with an adverbial phrase is very high (1.0) across all conditions.

Sentence completion task

Despite the clear evidence in the corpus investigation that the local ambiguities discussed earlier have been eliminated, it is nevertheless possible that the relatively small size of the corpus may have yielded incomplete information. We carried out a direct test of the possible completions at the critical regions in the sentence by conducting a sentence completion study. The main aim of the sentence completion study was to cross-methodologically validate the corpus findings with the exact materials to be used in the reading experiments. In particular, we wanted to further validate the experimental design by testing whether participants indeed predict a relative clause when reading the Det + Cl + Adv sequence in the experimental materials and are not garden pathed by other non-RC-parses. Moreover, we wanted to obtain sentence continuations from participants so that we could derive region-by-region predictions of the expectation-based account for the factors RC type and modification type.

In contrast to the corpus counts, these predictions will be based on the exact materials to be used in the reading experiments. We therefore conducted a sentence completion task with all experimental items. The 32 experimental items to be used in Experiment 2 (i.e., a superset of the items used in Experiment 1) were presented with a truncation either after Det + Cl or after the adverb. This resulted in sentence fragments consisting of Det + Cl or Det + Cl + Adv in subject-modifying conditions and NP + VP + Det + Cl or NP + V + Det + Cl + Adv in object-modifying conditions. We included 78 filler sentences (the same sentences as in Experiment 2) that were pseudo-randomly truncated at different points. Twenty Mandarin speakers from Taiwan currently living in the US (12 participants) or in Germany (8 participants) participated in

[7] In all of these ORs, the RC subject is covertly realized as a *pro*.
[8] 5% of these ORs have an overt RC subject and 10% are ORs with a *pro* as RC subject.

Table 2
Summary of the results of the sentence completion experiment (see B2, B1 in the appendix for the underlying raw counts). Conditional probabilities Prob (Target|Prefix) associated with the upcoming structure (Target) at the various regions (Prefix) of the experimental materials grouped by modification type. Canonical RCs are RCs with the canonical SVO word order and an overt RC subject in ORs. The experimental items are canonical SRs/ORs.

Mod. Type	Prefix	Target	Prob(Target | Prefix)
subj-mod	Det + Cl	RC	0.23
	Det + Cl + Adv	SR	0.72
	Det + Cl + Adv	Canonical SR	0.68
	Det + Cl + Adv	OR	0.14
	Det + Cl + Adv	Canonical OR	0.02
	Det + Cl + Adv + SR + *de*	overt RC head	0.99
	Det + Cl + Adv + canonical SR + *de*	overt RC head	0.99
	Det + Cl + Adv + OR + *de*	overt RC head	0.95
	Det + Cl + Adv + canonical OR + *de*	overt RC head	1.00
obj-mod	NP + V + Det + Cl	RC	0.43
	NP + V + Det + Cl + Adv	SR	0.86
	NP + V + Det + Cl + Adv	Canonical SR	0.82
	NP + V + Det + Cl + Adv	OR	0.11
	NP + V + Det + Cl + Adv	Canonical OR	0.05
	NP + V + Det + Cl + Adv + SR + *de*	overt RC head	1.00
	NP + V + Det + Cl + Adv + canonical SR + *de*	overt RC head	1.00
	NP + V + Det + Cl + Adv + OR + *de*	overt RC head	1.00
	NP + V + Det + Cl + Adv + canonical OR + *de*	overt RC head	1.00

this experiment against payment of 18 USD or 13 EUR respectively. The sentence completion task was implemented as an online questionnaire, with the items being presented in Traditional Chinese characters and arranged according to a Latin Square in pseudo-randomized order, such that each experimental item was preceded by at least one filler sentence.

The sentence completion data provide us with a quantitative estimate of the conditional probability of an RC parse being adopted given the left context of the experimental items up to Det + Cl (i.e., the conditional probability of an RC parse modulated by modification type) and the conditional probability of an SR or OR parse being adopted given the left context of the experimental items up to Det + Cl + Adv. Moreover, participants' productions were analysed for whether the RC head noun was overtly or covertly produced. This provided us with an estimate of the conditional probability of an overt head noun appearing given a Det + Cl + Adv + SR/OR sequence.

We can compute the conditional probability at each region by taking all alternative parse completions into account. Thus, in our estimation procedure, we make the independently motivated assumption (Pickering & Garrod, 2007, 108) that the parser predicts a range of syntactic alternatives, each associated with a probability conditional on its prefix. For example, the conditional probability of an SR given left context Det + Cl + Adv is then the number of SRs produced divided by the total count of all constructions produced given the same left context. The sentence completion task therefore allows us to empirically estimate surprisal (Hale, 2001). An overview of the results of the sentence completion study is provided in Table 2.

Sentence completions after Det + Cl/NP + V + Det + Cl

A classification of all sentence completions produced when the sentence was truncated after the sequence Det + Cl (subject-modification) or NP + V + Det + Cl (object-modification) is provided in the appendix (see Table B1). Participants continued the Det + Cl sequence with an RC modifying an upcoming subject in 23% of the trials. In object-modifying contexts, the NP + V + Det + Cl fragment was continued with an RC modifying an upcoming object in 43% of the trials. Thus, the conditional probability of an RC appearing is higher in an object-modifying syntactic environment compared to a subject-modifying environment.

Sentence completions after Det + Cl + Adv/NP + V + Det + Cl + Adv

A classification of all sentence completions produced when the sentence was truncated after the sequence Det + Cl + Adv (subject-modification) or NP + V + Det + Cl + Adv (object-modification) is provided in the appendix (see Table B2). Participants completed the sentence fragment with an RC that started with the Adv and that modified an upcoming NP attaching to the Det + Cl in 88% of all trials in subject-modifications, and in 98% of the trials in object-modifying conditions. In the trials in which no RC continuation was produced, participants either postulated an elided NP between the Det + Cl and the Adv (i.e., the newly introduced ambiguity in the present design) (10% in subject-modifications, none in object-modifications)[9] or assigned a different part-of-speech to the adverb: In three trials of each modification type, they interpreted the adverb

[9] The higher proportion of productions with an elided NP compared to the corpus counts (one token) might be explained by the fact that truncating the experimental sentences after Det + Cl + Adv (three open clause boundaries) might have introduced a bias towards an elided NP reading. The reason for this is that the insertion of a covert NP before the adverb leaves the reader with only one open clause (only the main clause VP to be produced in order to complete the sentence) in contrast to three predicted heads in case no elided NP is postulated (the RC, the main clause subject and the main clause VP). This bias towards a reduction of open clauses by postulating an elided NP might have been even strengthened by the fact that the filler items in the sentence completion task were all cut between clause boundaries.

as an NP. (In Chinese, the mapping between lexical items and the part-of-speech categories is not always one-to-one; Kwong & Tsou, 2003. The temporal adverbs in our materials can function as genitive NPs in case they modify a second NP to their right, similar to *today's* + NP in English). In sum, the sentence completion study replicated the corpus finding that all the local ambiguities presented in Fig. 1 are indeed eliminated in the present materials. Low-probability continuations are a parse with an elided NP interposed between Det + Cl and Adv (the newly introduced ambiguity) or a parse with the adverb being read as a genitive NP. Thus, the most probable structure predicted upon encountering the Det + Cl + Adv sequence is a relative clause and the local ambiguities that might have confounded previous studies have been brought under experimental control.

In order to derive conditional probabilities associated with each RC type modulated by modification type, the RCs produced after the Det + Cl + Adv/NP + V + Det + Cl + Adv sequence were classified manually into SRs (72% of all trials in subject-modifications, and 86% in object-modifications), ORs (14% in subject-modifications and 11% in object-modifications), possessive RCs (1% in subject-modifications and 1% in object-modifications) and adjunct RCs[10] (1% in subject-modifications and 0 tokens in object-modifications). Since the SR/OR structures used in the experimental items are canonical SRs/ORs, i.e., RCs with the canonical SVO word order and, in the case of ORs, with an overt RC subject, the SRs and ORs observed in participants' sentence completions were further categorized into canonical SRs/ORs, passive RCs, topicalized RCs, RCs with an adjectival predicate and RCs with a dropped RC subject.

Canonical SRs were produced in 68% and 82% of the trials in subject- and object-modifications, respectively. Canonical ORs were produced in 2% and 5% of the trials in subject- and object-modifications, respectively. This pattern confirms the main finding of the corpus analyses: The probability of an SR continuation after Det + Cl + Adv is higher than the probability of an OR continuation. Note that in this highly constraining context, the difference in conditional probabilities of SRs vs ORs is even bigger than the general frequency difference between SRs and ORs reported in other studies.

In order to derive estimates for the probability of an overt RC head noun, as opposed to a covert NP, being produced after a subject/object-modifying Det + Cl + Adv + SR or Det + Cl + Adv + OR, we further categorized the SR and OR productions by whether their head noun was overtly produced or realized as a covert NP (headless RC). This estimation procedure is necessary to compute the conditional probability associated with the head noun in the experimental items, because the headless NP constitutes an alternative parse to the one in the experimental materials. In subject-modifications, one headless SR (canonical SR) (1% of all subject-modifying SRs) and one headless OR (OR with dropped subject) (5% of all subject-modifying ORs) were produced. In object-modifying contexts, no headless RCs were produced.

In sum, the sentence completion data show that:

(i) Consistent with the corpus findings, the Det + Cl + Adv sequence strongly predicts an RC. This effectively eliminates the possibility that non-RC continuations are predicted. In other words, the present design eliminates the confound present in previous experimental comparisons between ORs and SRs.
(ii) The conditional probability of an SR continuation after Det + Cl + Adv is higher than the conditional probability of an OR. This is also consistent with the corpus analyses.
(iii) The conditional probability of an RC appearing in object-modifying position is higher than the conditional probability of an RC appearing in subject-modifying position. This finding differs from the results of the corpus analyses where subject- and object-modifications did not differ in their conditional probability.
(iv) The conditional probability of an RC-head being overtly produced was close to 1.0 across conditions. This replicates the pattern observed in the corpus.

In summary, the corpus analyses and the sentence completion data allow us to derive predictions of the expectation-based account for the relative processing ease associated with subject- and object-modifying SRs and ORs for each regions of the sentence. Corpus counts and sentence completion data provide similar estimates of the conditional probability associated with RC type for all regions but differ with respect to the conditional probability of modification type and its interaction with RC type.

Experiment 1: Self-paced reading

Before discussing the experiment, we present the predictions of the two competing accounts by region of interest: the adverb at the onset of the relative clause, the relative clause region containing the V + N/N + V sequence, and the head noun.

Predictions

The expectation-based account

The expectation-based account assumes that the relative frequency of a certain structure influences processing time. Specifically, surprisal (Hale, 2001; Levy, 2008) predicts that the difficulty associated with the integration of a certain word is proportional to its negative log-probability given the preceding context. Higher surprisal values are thus predicted to be reflected in increased processing difficulty, i.e., longer self-paced reading times.

Table 3 gives an overview of the surprisal values associated with each experimental condition at the adverb, the RC region (the V + N sequence *yaoqing-le nanhai*, 'invited the boy' and N + V sequence *nanhai yaoqing-le*, 'the boy invited' in SRs and ORs respectively) and the RC head noun. These surprisal values are calculated from the conditional probabilities based on the sentence completion data presented in Table 2. We used the counts reflecting the probabilities of the subtypes of SRs and ORs that exactly match the structure of our experimental materials (canonical SRs

[10] Adjunct RCs are also called gap-less RCs.

Table 3
Predicted surprisal values derived from sentence completion data associated with each experimental condition by region of interest. Surprisal is calculated as the negative natural logarithm of the conditional probabilities presented in Table 2. The estimates are based on those numbers that take into account only the exact syntactic structure of the experimental materials (canonical SRs/ORs).

	Adv	V – N/N – V	RC head
$SR_{Subj\text{-}mod}$	1.47	0.39	0.01
$OR_{Subj\text{-}mod}$	1.47	3.91	0.00
$SR_{Obj\text{-}mod}$	0.84	0.20	0.00
$OR_{Obj\text{-}mod}$	0.84	3.00	0.00

Table 4
Effects predicted by surprisal at the adverb, the RC-region (V + N/N + V) and the RC head noun. The predicted effects are calculated from the sentence completion-based surprisal values presented in Table 3. A positive sign associated with an effect of modification type or RC type means that subject-modifications or SRs are easier to process than object-modifications or ORs, respectively.

	Adv	V – N/N – V	RC head
Main effect RC Type	*not applicable*	3.16	−0.01
Main effect Mod. Type	-0.63	−0.55	−0.01
RC Type × Mod. Type	*not applicable*	−0.36	0.01
RC Type [Subj-mod]	*not applicable*	3.52	−0.01
RC Type [Obj-mod]	*not applicable*	2.80	0.00

and ORs) as a basis for the calculation of the surprisal values.

From these surprisal values associated with each experimental condition, we derived predictions for the experimental factors (main effects of modification type and RC type, their interaction and pairwise effects of RC type nested within modification type). The surprisal value associated with the main effect of modification type was calculated by subtracting the mean surprisal value of the subject-modifying conditions from the mean surprisal value of the object-modifying conditions. The surprisal value associated with the main effect of RC type was calculated in a similar manner. The predicted interaction was calculated by subtracting the mean of the surprisal value of subject-modifying ORs and object-modifying SRs from the mean surprisal of subject-modifying SRs and object-modifying ORs. The predicted effects of RC type in subject-modification and object-modification were calculated by subtracting the surprisal value associated with ORs from the surprisal value associated with SRs within each modification type separately. The same contrasts are applied in the statistical analyses of the reading times. A summary of the surprisal values associated with each of these effects is provided in Table 4.

At the adverb, it becomes clear to the reader that an RC is being read. The sentence completion data show that the conditional probability of an RC appearing in this position (i.e., after a Det + Cl) is higher in object-modifying contexts. This leads to a lower surprisal value in object-modifying conditions. Therefore, a main effect of modification type with faster reading times in object-modifications is predicted.

At the RC-region (V + N/N + V), the conditional probability of an SR is much higher than the conditional probability of an OR, i.e., surprisal is higher in ORs than in SRs. Thus, the expectation-based account predicts a relatively strong main effect of RC type with faster reading times in SRs. This SR advantage is predicted to be stronger in subject-modifications, i.e., a small interaction between RC type and modification type is predicted. Moreover, the higher surprisal values in subject-modifications compared to object-modifications predict a main effect of modification type (faster reading times in object-modifications).

At the RC head noun, no effect is predicted since across all conditions the head noun is highly expected leading to no difference in surprisal between RC types and between modification types.

Calculating surprisal based on the conditional probabilities from corpus counts rather than on the sentence-completion data would result in similar, but not completely identical predictions.[11] No effect of modification type at any region and no interaction between RC type and modification type would be predicted. We will focus on the surprisal predictions derived from sentence completion data because (i) they more closely match the experimental materials, (ii) they allow us to directly link data from the production experiment to comprehension data, and (iii) the absence of an effect of modification type and its interaction with RC type in the corpus is inconclusive given the relatively small size of the corpus. We believe that, due to its experimental setting, the sentence completion study is likely to have higher sensitivity than the corpus analyses.

The working-memory based account

We turn next to the predictions of the DLT memory cost metrics (Gibson, 2000), beginning with the DLT's storage cost metric. In the pre-head region, DLT's storage cost predicts that ORs and SRs are equally hard to process; this is because in the experimental materials it is clear that the sentences are relative clauses, and therefore an equal number of incomplete dependencies remain when the V + N/N + V region is processed (see page 6 of Hsiao & Gibson, 2003). When reading the Det + Cl sequence, one nominal head is predicted in both SRs and ORs. When reading the subsequent adverb, a relative clause, i.e., a VP and the relativizer *de*, is predicted. When reading the next word (the RC verb in SRs and the RC subject in ORs), the RC object is predicted in SRs, while in ORs the RC verb (that was already predicted from the adverb on) is still predicted. In sum, at every word of the relative clause, an equal number of heads is predicted resulting in the same DLT storage cost predicted for SRs and ORs.

The storage cost metric also predicts that at the adverb, in the relative clause region and at the head noun, object-modifications should be easier to process than subject-modifications. This is because in object-modifications the main clause verb precedes the relative clause, whereas in subject-modifications the main clause verb follows the RC head, which leads to higher storage costs in the object-modifying conditions. The predictions of DLT storage cost are summarized in Table 5.

[11] See Chen et al. (2012) for an earlier attempt of using corpus frequency derived surprisal to account for the subject relative advantage in Chinese.

Table 5
Predicted syntactic heads with DLT storage cost for each regions of interest by experimental condition.

	N	V_{MC}	Det-Cl	Adv	V_{RC}/N_{RCsubj}	N_{RCobj}/V_{RC}	FreqP	de	N
$SR_{Subj\text{-}mod}$									
Predicted heads	–	–	N,V_{MC}	N,V_{MC}, V_{RC},de	N,V_{MC},de,N_{RCobj}	N,V_{MC},de	N,V_{MC},de	N,V_{MC}	V_{MC}
Storage cost			2	4	4	3	3	2	1
$OR_{Subj\text{-}mod}$									
Predicted heads	–	–	N,V_{MC}	N,V_{MC}, V_{RC},de	N,V_{MC}, V_{RC},de	N,V_{MC},de	N,V_{MC},de	N,V_{MC}	V_{MC}
Storage cost			2	4	4	3	3	2	1
$SR_{Obj\text{-}mod}$									
Predicted heads	V_{MC}	N	N	N,V_{RC},de	N,de,N_{RCobj}	N,de	N,de	N	–
Storage cost	1	1	1	3	3	2	2	1	0
$OR_{Obj\text{-}mod}$									
Predicted heads	V_{MC}	N	N	N,V_{RC}, *de*	N,V_{RC},de	N,de	N,de	N	–
Storage cost	1	1	1	3	3	2	2	1	0

In terms of the DLT's integration cost metric (Gibson, 2000), at the head noun and possibly beyond (due to spill-over), an OR advantage is predicted, because of the greater gap-head (or RC verb-head) distance in SRs compared to ORs. Integration cost is predicted to be higher in SRs due to a higher number of intervening discourse referents (the RC object and possibly the RC verb, depending on whether the verb or the gap preceding it is assumed to be retrieved). Similarly, the ACT-R based model of parsing (Lewis & Vasishth, 2005) predicts lower activation of the constituent that has to be retrieved (the gap or the relative clause verb) in SRs compared to ORs because the increased dependency length is predicted to lead to decay and interference.

Regarding the effect of modification type at the head noun, the integration cost metric predicts the opposite of the storage cost metric. Integration cost predicts that object-modification should be *harder* than subject-modification at the head noun, since in object-modifications there is one additional dependency to be built when reading the head noun: the dependency between the main clause verb and the main clause object (see Table 6). Similar to DLT integration cost, the ACT-R based model (Lewis & Vasishth, 2005) predicts a slowdown in object-modifications. An overview of the constituents retrieved at the head noun and their predicted DLT structural integration cost is provided in Table 6.

Participants

49 college students at the Dalian University of Technology, China participated in the experiment as volunteers. All participants were native speakers of Mandarin Chinese.

Table 6
Predicted DLT structural integration cost at the RC head noun with respect to each constituent that needs to be retrieved. The numbers in parenthesis refer to the structural integration cost under the assumption that not the gap site inside the relative clause but rather the RC verb itself is retrieved.

	GAP (V_{RC})	Det-Cl	V_{MC}	*total*
$SR_{Subj\text{-}mod}$	2 (1)	3	–	5(4)
$OR_{Subj\text{-}mod}$	0	3	–	3
$SR_{Obj\text{-}mod}$	2 (1)	3	3	8(7)
$OR_{Obj\text{-}mod}$	0	3	3	6

Stimuli

The stimuli were as described above (see example 2); there were 16 target items, which were adapted from the items used in the Gibson and Wu (2013) study. In addition, we presented 78 fillers consisting of various syntactic structures including eight relative clauses taken from Gibson and Wu (2013)'s filler items that were structurally different from the experimental sentences. All stimuli and fillers are available from the first author.

Procedure

The experiment used the non-cumulative self-paced moving window method (Just, Carpenter, & Woolley, 1982). We presented stimulus items using Douglas Rohde's Linger software, version 2.88 (http://tedlab.mit.edu/~dr/Linger/). The target items were presented in Simplified Chinese characters (font size 18) using the standard Latin square design. The target items and fillers were pseudo-randomized for each subject such that each item was succeeded by at least one filler sentence. Each trial was followed by a *yes/no*-comprehension question designed to probe the correct understanding of both the relative clause and the main clause.

Each experimental session started with six practice trials which allowed participants to get used to the procedure. At the beginning of each trial, the participant saw a mask of hyphens that covered the upcoming sentence. Each hyphen represented the space-delimited words shown in example 2. Participants were instructed to press the `f`-key in order to read the sentence; successive presses of this key displayed the next word of a sentence and masked the preceding word. In order to respond to the comprehension questions, the `k`-key was used for a 'yes' response, and the `l`-key for a 'no' response. Reading times or RTs (in milliseconds) were taken as a measure of relative momentary processing difficulty.

Results

All analyses were carried out using linear mixed models using the `lme4` package version 1.0–6 in R (Bates,

Maechler, Bolker, & Walker, 2014). The analyses of reading times were carried out on log-transformed values in order to stabilize variance and to achieve approximately normal residuals (Box & Cox, 1964). Question-accuracies were analysed using generalized linear mixed models with a binomial link function. We defined two sets of contrasts. First, we tested for main effects of RC type (SRs coded as −0.5 and ORs as +0.5) and modification type (subject-modifications coded as −0.5 and object-modifications as +0.5) and an interaction between the two. Second, two sum contrasts nested within modification type (coded as in the first model) were defined so that SRs vs ORs were each compared separately in the subject-modifying and object-modifying case. SRs were coded as −0.5 and ORs as +0.5. Residuals of linear mixed models were always checked to ensure that there were no serious deviations from the normality assumption. For linear mixed models, we took an absolute *t*-value equal to or above 2 to reach statistical significance at $\alpha = 0.05$. Throughout this paper, a full variance–covariance matrix was fit for participants and items (Gelman & Hill, 2007); whenever the model failed to converge or if the variance–covariance matrix of random effects was degenerate, we removed the relevant varying slopes for items or participants.

Question–response accuracies

In the analyses of accuracy, no differences were found between SRs and ORs in either the subject-modifying or the object-modifying cases, and no difference was found between subject and object-modification. In subject-modifying cases, SRs had accuracy 85%, and ORs 85%; for object-modifying cases, the corresponding accuracies were 84% and 86%.

Analyses of reading times

The regions of interest were the adverb at the onset of the relative clause, the pre-head region following the Det + Cl + Adv sequence, the head noun, and the regions following it. We therefore analysed the adverb (Adv) (only testing for a main effect of modification type since conditions are identical across RC type), the V + N/N + V-region (SRs and ORs respectively), the region of the frequency phrase (FreqP) that followed the V + N/N + V region, the relativizer *de* before the head noun, the head noun, and the two regions following the head noun. The regions following the head noun differ lexically and syntactically across modification type (but not across RC type), therefore we only report models with pairwise comparisons of RC type nested within each level of modification type for the post-head regions.

The reading times for the different regions are summarized in Figs. 3 and 4, and the results of the statistical analyses are shown in Tables 7 and 8.

At the adverb, no effect of modification type was observed. In the RC-region (V + N/N + V), a main effect of RC type with SRs being read faster than ORs was observed. In nested comparisons, this SR advantage reached significance in subject-modifications and was marginal in object-modifications. No effect of modification type and no interaction between modification and RC type were observed. No effects were observed at the frequency phrase (V + N/N + V), the relativizer, the RC head noun, and one word after the RC head. Two words after the head noun, we found faster reading times in SRs compared to ORs in subject-modifications, but not in object-modifications. No other effect reached significance in this region.

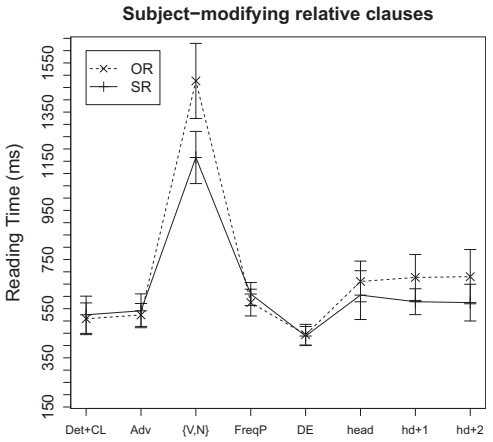

Fig. 3. Mean reading times of each region of interest in subject-modifying relatives, along with 95% confidence intervals.

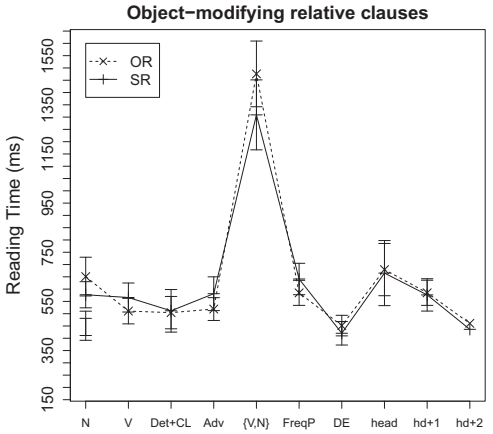

Fig. 4. Mean reading times of each region of interest in object-modifying relatives, along with 95% confidence intervals.

Table 7
Main effects of modification and RC type and their interaction by region of interest in Experiment 1. The dependent variable is log-transformed reading time.

Region	Contrast	Coef.	SE	*t*-value
Adv	Modification type	0.01	0.04	0.34
V + N/N + V	Modification type	0.04	0.05	0.84
	RC type	**0.11**	**0.04**	**2.51***
	Mod. type × RC type	0.00	0.03	−0.08
FreqP	Modification type	0.03	0.04	0.74
	RC type	−0.06	0.04	−1.57
	Mod. type × RC type	−0.01	0.04	−0.17
DE	Modification type	0.00	0.03	−0.05
	RC type	0.02	0.02	0.97
	Mod. type × RC type	0.00	0.03	0.17
head	Modification type	0.01	0.04	0.38
	RC type	0.06	0.05	1.34
	Mod. type × RC type	−0.03	0.04	−0.87

Table 8
The results of the RC type comparisons nested within each level of modification type by region of interest in Experiment 1. The dependent variable is log-transformed reading time.

Region	Contrast	Coef.	SE	*t*-value
V + N/N + V	RC Type [Subject mod]	**0.11**	**0.05**	**2.23***
	RC type [Object mod]	0.11	0.06	1.83
FreqP	RC type [Subject mod]	−0.06	0.06	−0.94
	RC type [Object mod]	−0.07	0.05	−1.42
DE	RC type [Subject mod]	0.02	0.03	0.52
	RC type [Object mod]	0.03	0.03	0.76
head	RC type [Subject mod]	0.1	0.06	1.56
	RC type [Object mod]	0.03	0.06	0.45
hd + 1	RC type [Subject mod]	0.09	0.05	1.56
	RC type [Object mod]	0.04	0.05	0.86
hd + 2	RC type [Subject mod]	**0.13**	**0.04**	**3.05***
	RC type [Object mod]	0.03	0.04	0.89

Discussion

The faster reading time in SRs in the pre-head region (V + N/N + V) is consistent with a sharpened expectation for the more frequently occurring SR. The data are consistent with the predicted higher surprisal (Hale, 2001; Levy, 2008) cost in ORs compared to SRs in both subject and object-modifications. Surprisal can also explain the stronger SR advantage observed in subject-modifications compared to object-modifications since the difference in surprisal associated with SRs vs ORs is smaller in object-modifications than in subject-modifications. Indeed, surprisal predicts a small interaction between RC type and modification type. In the reading data, however, this interaction is not significant. Surprisal might account for the absence of a significant interaction by the very small size of the predicted interaction compared to the size of the effect of RC type. The predictions of DLT storage cost (Gibson, 1998; Gibson, 2000) (i.e., no effect of RC type), in contrast, are not consistent with the SR advantage observed at the V + N/N + V region.

At the head noun, we do not see evidence for a difference between processing difficulty associated with SRs vs ORs. This absence of an effect is statistically inconclusive, but as predicted by surprisal and the DLT storage cost component. However, the observed null result cannot rule out retrieval-based memory accounts such as the DLT integration cost component (Gibson, 1998; Gibson, 2000) or the ACT-R based model proposed by Lewis and Vasishth (2005), both of which predict an OR advantage at the head noun.

The SR advantage in subject modifications in the spillover region after the head noun cannot be explained by either account under discussion. It is inconsistent with memory-based retrieval metrics like DLT integration cost (Gibson, 2000) and the ACT-R-based retrieval model by Lewis and Vasishth (2005), since they predict an effect in the opposite direction. It is also inconsistent with the storage cost component of DLT (Gibson, 2000) and surprisal (Hale, 2001; Levy, 2008) since both of them predict no effect of RC type at the head noun. Although one might interpret this effect as spillover from the RC region, in which case the effect would be consistent with surprisal, this explanation appears to be rather implausible given that the SR advantage had disappeared at the head noun and the following word. We suggest an explanation for this effect consistent with surprisal in the Discussion of Experiment 2.

The absence of an effect of modification type is inconclusive. We therefore do not discuss it any further here.

In sum, the evidence in Experiment 1 points in favor of the expectation-based account. However, given the conflicting results in the literature together with the SR advantage in subject-modifications two words after the head noun that cannot be explained by either account, and the null result with respect to the factor modification type, it is vital to attempt to replicate this result. Indeed, one of the biggest worries in psychology and linguistics today is the problem of non-replicable findings. Already Ronald Fisher, the founder of frequentist statistics, has advocated replication from the outset as the gold standard for science (Fisher, 1937, page 16). Today, a growing number of methodologically and statistically concerned researchers emphasize the need for replication in experimental psychology (e.g., Asendorpf et al., 2013; Simmons, Nelson, & Simonsohn, 2011).

Experiment 2: Eye-tracking study

This experiment extends Experiment 1 in two respects: first, the use of the eye-tracking method provides cross-methodological validation of the results in Experiment 1 and second, we doubled the number of items in order to increase statistical power. The same predictions hold as for Experiment 1.

Participants

This study was conducted at the eye-tracking lab of the Department of English at National Taiwan Normal University, Taipei. 49 students from that university participated in the experiment, each receiving payment of 250 NTD. All participants were native speakers of Mandarin and had normal or corrected-to-normal vision.

Design and materials

The experimental items, including comprehension questions, had the same design as in Experiment 1. We used all of the previous items and additionally created 16 new sets of items that had the same structure. A minor difference between the old and the new items was that in the original items the sentence final materials sometimes varied across modification type, whereas in the new items they were identical across all four conditions. In addition, we made very minor lexical changes to the self-paced reading items to adapt them to the cultural environment of Taiwan, where this experiment was conducted. In contrast to the first experiment, all items were written in Traditional Chinese characters, the script officially used in Taiwan. In order to ensure that any difference in processing difficulty associated with the experimental manipulations is not due to a difference in acceptability between ORs and SRs that is particular to the present materials (i.e., due to the insertion of the Det + Cl + AdvP), we conducted a web-based acceptability rating study on the experimental materials to be used in the eye-tracking experiment with twenty native speakers of Manda-

rin. We did not find any evidence for a difference in acceptability between SRs and ORs induced by the present stimulus design.

Procedure

Eye movements were recorded with an SR Research Eyelink 1000 eyetracker at a sampling rate of 1000 Hz using a desktop-mount camera system (leveled, illuminator on the right) with a 35 mm lens to track participants' right eye. The participants' head was stabilized with a chin-rest. The camera and the presentation screen were installed on a table of 74 cm height, the chair on which participants were seated as well as the chin rest were adjustable in height. The camera-to-eye distance measured 52 cm, the eye-to-screen distance 62 cm. Stimuli were presented on an 19" monitor with a resolution of 1440 × 900 pixels; the stimuli were written in Traditional Chinese characters (font type SimSum, font size 20) in a black font on a light gray background.

The experiment was run using Experiment Builder software provided by SR Research. 32 experimental items, each with four conditions were presented in a Latin square. 88 filler items with a variety of syntactic structures including the relative clause fillers of Experiment 1 were used. Items were pseudo-randomized such that each relative clause item was preceded by at least one filler item. Each trial was followed by a comprehension question to be answered with *yes* or *no* pressing a key on a response pad.

Results

Similar to Experiment 1, we used linear mixed effects models with a full variance–covariance matrix structure for participants and items when possible, applying two sets of contrasts: main effects and interaction and pairwise comparisons of RC type nested within modification type. Binomial dependent variables were analysed using generalized linear mixed effects models with a binomial link function.

Question–response accuracies

The mean accuracy for subject-modifying SRs was 79%, for subject-modifying ORs 80%, object-modifying SRs 77%, and object-modifying ORs 79%. As in Experiment 1, none of the comparisons showed any statistically significant differences.

Reading times

In eye-tracking data, the dependent measures can be partitioned into three broad classes: those that provide information about: (i) first-pass events; (ii) regression-related events (proportions of regressions and duration of regressive events); and (iii) second- and later pass events. Clifton, Staub, and Rayner (2007) have shown in a large scale review on experiments relying on eye-tracking to measure sentence processing difficulty that it is still unclear in which eye-tracking measure to expect effects of syntactic processing. It is therefore common practice to report a wide range of dependent variables. However, since many of the eye-tracking measures are by definition correlated, this is statistically problematic since it increases the probability of a Type I error. We try to strike a balance between this statistical concern and the aim of providing a comprehensive picture of the data by selecting the most commonly reported dependent variable from each class of eye-tracking measures. As a representative first-pass measure, we report first-pass reading time (FPRT) (often also referred to as gaze duration), i.e., the sum of all fixations on a region before leaving it if and only if this region is entered progressively. As a measure for proportions of regressions, we analysed first-pass regression probability (FPReg), i.e., the proportion of trials in which a regression was initiated from a region when first entering this region, and regression-path duration (RPD), i.e., the sum of the time of all fixations starting from the first fixation on this region until leaving this region to the right including all fixations to the left of this region that fall into this time window, as a representative measure for regressive reading events. As a later-pass measure, we report total fixation times (TFT) which is defined as the sum of all fixations on a region. For all dependent variables, trials in which the region under consideration was skipped (i.e., in case the dependent fixation measure was 0) were excluded from analyses. Figs. 5 and

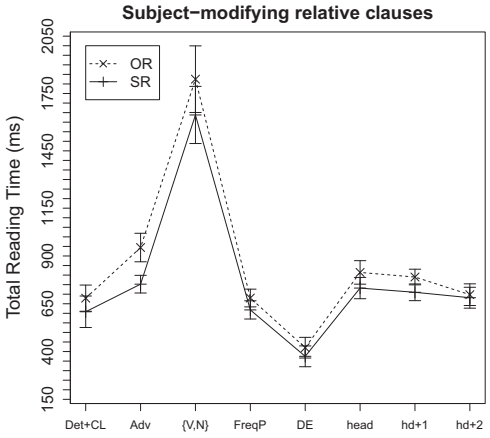

Fig. 5. Mean total fixation time of each region of interest in subject-modifying relatives, along with 95% confidence intervals in Experiment 2.

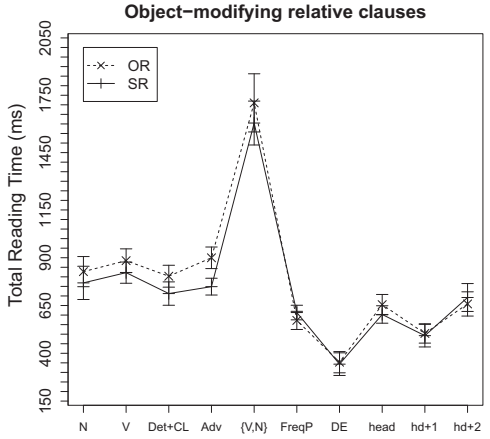

Fig. 6. Mean total fixation time of each region of interest in object-modifying relatives, along with 95% confidence intervals in Experiment 2.

6 provide a visual summary of relevant comparisons in TFT across all regions of the sentences up to two words after the head noun, along with 95% confidence intervals. The estimates, standard errors, and *t*-values of all linear mixed models for all regions of interest are shown in Tables 9 and 10.

At the adverb-region, a significant main effect of modification type was observed in both regression-based dependent variables (RPD and FPReg) with facilitated processing in object-modifications.

At the V + N/N + V region, we observed a significant main effect of RC type in FPReg, RPD and TFT with faster

Table 9
Main effects of RC type and modification type and their interaction in Experiment 2 by region of interest for the dependent measures log-first-pass reading time, first-pass regression probability, log-regression path duration and log-total fixation time.

Region		Modification type			RC type			Mod. type × RC type		
		Coef.	SE	*t* or *z*	Coef.	SE	*t* or *z*	Coef.	SE	*t* or *z*
Adv	FPRT	0.02	0.03	0.67	Not applicable			Not applicable		
	FPReg	**−0.73**	**0.18**	**−3.95***	Not applicable			Not applicable		
	RPD	**−0.24**	**0.05**	**−4.99***	Not applicable			Not applicable		
	TFT	−0.02	0.04	−0.6	Not applicable			Not applicable		
V + N/N + V	FPRT	−0.07	0.04	−1.79	−0.05	0.04	−1.39	0.04	0.04	0.97
	FPReg	0.01	0.14	0.05	**0.76**	**0.22**	**3.51***	−0.15	0.15	−1.02
	RPD	**−0.1**	**0.04**	**−2.73***	**0.11**	**0.04**	**2.81***	−0.02	0.04	−0.51
	TFT	0.03	0.04	−0.91	**0.08**	**0.03**	**2.82***	−0.03	0.03	−0.92
FreqP	FPRT	−0.03	0.02	−1.12	−0.03	0.03	−1.13	0	0.02	0.03
	FPReg	0.27	0.16	1.67	0.01	0.18	0.06	0.12	0.16	0.75
	RPD	0.03	0.04	0.67	−0.02	0.04	−0.58	−0.01	0.04	−0.25
	TFT	−0.06	0.04	−1.32	−0.02	0.04	−0.53	−0.07	0.04	−1.74
DE	FPRT	−0.03	0.03	−0.98	−0.01	0.03	−0.43	0.01	0.03	0.56
	FPReg	−0.04	0.2	−0.22	0.25	0.16	1.56	−0.29	0.22	−1.32
	RPD	0.04	0.05	0.81	0.01	0.05	0.22	−0.04	0.05	−0.87
	TFT	**−0.1**	**0.04**	**−2.34***	0.06	0.04	1.58	−0.04	0.04	−0.91
head	FPRT	−0.04	0.03	−1.63	0.01	0.02	0.23	**0.06**	**0.02**	**2.57***
	FPReg	−0.14	0.15	−0.95	0.2	0.16	1.28	−0.17	0.14	−1.23
	RPD	−0.07	0.04	−1.55	0.04	0.04	1	−0.01	0.04	−0.18
	TFT	**−0.18**	**0.03**	**−5.5***	0.07	0.04	1.81	−0.01	0.04	−0.22

Table 10
The results of the RC type comparisons nested within each level of modification type in Experiment 2 by region of interest for the dependent measures log-first-pass reading time, first-pass regression probability, log-regression path duration and log-total fixation time.

Region		RC Type [subj-mod]			RC Type [obj-mod]		
		Coef.	SE	*t* or *z*	Coef.	SE	*t* or *z*
V + N/N + V	FPRT	−0.08	0.05	−1.8	−0.01	0.04	−0.31
	FPReg	**0.91**	**0.28**	**3.27***	**0.61**	**0.24**	**2.51***
	RPD	**0.13**	**0.06**	**2.29***	0.09	0.05	1.81
	TFT	**0.1**	**0.04**	**2.62***	0.05	0.04	1.28
FreqP	FPRT	−0.03	0.04	−0.89	−0.03	0.03	−0.98
	FPReg	−0.11	0.26	−0.42	0.13	0.22	0.62
	RPD	−0.02	0.05	−0.29	−0.03	0.06	−0.57
	TFT	0.05	0.06	0.82	−0.09	0.06	−1.54
DE	FPRT	−0.03	0.04	−0.72	0	0.04	0.09
	FPReg	0.55	0.28	1.96	−0.04	0.27	−0.15
	RPD	0.06	0.06	0.87	−0.03	0.08	−0.42
	TFT	0.1	0.05	1.81	0.02	0.06	0.42
head	FPRT	−0.05	0.03	−1.6	0.06	0.03	1.97
	FPReg	0.37	0.22	1.71	0.03	0.2	0.14
	RPD	0.05	0.06	0.91	0.04	0.06	0.64
	TFT	0.08	0.05	1.47	0.06	0.05	1.22
hd + 1	FPRT	0.04	0.03	1.22	0.02	0.03	0.47
	FPReg	0.15	0.22	0.69	0.17	0.21	0.83
	RPD	0.09	0.06	1.64	0.04	0.07	0.56
	TFT	**0.13**	**0.05**	**2.88***	0.07	0.05	1.39
hd + 2	FPRT	0.04	0.04	1.05	−0.01	0.04	−0.17
	FPReg	0.24	0.19	1.22	−0.06	0.17	−0.35
	RPD	0.1	0.07	1.41	−0.1	0.07	−1.31
	TFT	0.07	0.05	1.26	−0.06	0.06	−1.08

reading times and less first-pass regressions in SRs compared to ORs. The pairwise comparisons nested within modification type showed that this main effect was driven by both, subject-modifying and object-modifying conditions, but the effect was stronger in subject-modifications: In subject-modifications, the SR advantage was significant in FPReg, RPD, and TFT. In object-modifications, the effect reached significance only in FPReg and was marginal in RPD. The main effect of modification type and the interaction between RC type and modification type did not reach significance in any measure.

At the frequency phrase (FreqP) that was inserted between the V + N/N + V region, none of the comparisons reached significance.

At the relativizer *de*, we found a marginal SR advantage in FPReg within subject-modifying conditions and a main effect of modification type in TFT (shorter fixations in object-modifications).

At the RC head noun, in FPRT, the interaction between modification type and RC type was significant. The pairwise comparisons revealed that this interaction was driven by a marginal SR advantage in object-modifications that was not present in subject-modifications. In TFT, we observed a main effect of modification type with shorter fixation times in object-modifications.

One word after the head noun, in subject-modifications only, a significant SR advantage was observed in TFT.

Discussion

The results of Experiment 2 confirmed the key findings of Experiment 1. We replicated the SR advantage at the RC region in both subject and object-modifications as well as the SR advantage in subject-modifications in the spillover region following the head noun. Moreover, we observed a main effect of modification type starting at the onset of the relative clause that was not present in Experiment 1.

The main effect of RC type (SR advantage) at the RC region (V + N/N + V) reached significance across all eye-tracking measures except for FPRT. This SR advantage is predicted by expectation-based accounts of parsing such as surprisal (Hale, 2001; Levy, 2008). As in the self-paced reading experiment, the effect was more pronounced in subject-modifications than in object-modifications. This can also be explained by surprisal, which predicts a bigger effect size in subject-modifications. Also similar to Experiment 1, the interaction between RC type and modification type did not reach significance in any dependent variable. Rather, the difference between the two modification types with respect to the effect of RC type manifested itself in the SR advantage reaching significance across more dependent variables in subject modifications than in object modifications. Note that although surprisal predicts an interaction driven by the stronger effect predicted for subject-modifications, the predicted size of this interaction is very small compared to the predicted size of the effect of RC type. Storage metrics (Gibson, 1998; Gibson, 2000) which assume that processing difficulty depends on the number of predicted heads are inconsistent with the SR advantage in the relative clause region V + N/N + V since they predict the absence of an effect of RC type.

At the head noun, similar to Experiment 1, no main effect of RC type was observed. This is predicted by surprisal and DLT storage cost but statistically inconclusive. Although the integration cost component of DLT (Gibson, 2000) as well as the ACT-R based parsing model (Lewis & Vasishth, 2005) predict an OR advantage at the head noun, the observed null result cannot be interpreted as evidence against these theories. In contrast to Experiment 1, the interaction between RC type and modification type reached significance at the head noun (in FPRT only). This interaction was driven by a marginal SR advantage present only in object-modifications. None of the accounts under discussion can account for this interaction. It is inconsistent with retrieval-based memory accounts (DLT integration cost and the ACT-R based model of sentence processing) as they predict the absence of an interaction (they predict an OR advantage that is not modulated by modification type). The expectation-based account and DLT storage cost also predict the absence of an interaction since they predict an SR advantage that is not modulated by modification type either. However, it should be noted that this effect might not be reliable since it was only observed in a single eye-tracking measure (in which no other effects were observed in the experiment), it did not reach statistical significance in pairwise comparisons and it was not observed in Experiment 1. We will therefore not discuss this effect more in detail.

The SR advantage seen in subject-modifications one word after the head noun replicates the effect observed in the spillover region in Experiment 1. Compared to Experiment 1, the effect appeared one word earlier in the sentence. This earlier appearance of the effect is likely to be due to the higher temporal sensitivity of eye-tracking compared to self-paced reading. This effect cannot be accounted for by either of the accounts under discussion. It is inconsistent with the predictions made by retrieval-based accounts, e.g., the structural integration cost metrics of the DLT (Gibson, 1998; Gibson, 2000) or the ACT-R based model (Lewis & Vasishth, 2005) as both predict an effect in the opposite direction. DLT storage cost (Gibson, 1998; Gibson, 2000) is also incompatible with the SR advantage one word after the head noun since it predicts no effect at the head noun. Expectation-based accounts cannot explain this effect either since the sentence completion data and the corpus counts incorrectly predict no difference in surprisal at the head noun. However, one plausible explanation in line with surprisal might be worth considering: The SR advantage in the spillover region might reflect a higher conditional probability of an overtly produced RC head noun (i.e., a headed RC) in SRs compared to ORs within subject-modifications. Our pre-tests might have failed to detect this difference in conditional probabilities due to some reason, e.g., low statistical power. The corpus analyses were restricted to RC tokens preceded by a Det + Cl + Adv sequence, which resulted in a total of 72 tokens considered in the estimation of the conditional probabilities of headless vs headed RCs. A larger corpus search in which all tokens of subject- and object-modifying SRs/ORs were taken into account and not only RCs preceded by a Det + Cl + Adv sequence showed that in subject-modifications, 18% of the SR tokens (221 of 1244 tokens) and

38% of the OR tokens (220 of 582 tokens) are headless RCs. In object-modifications, in contrast, headless RCs are very rare independently of RC type, 2% in SRs (15 of 845 tokens) and 3% in ORs (8 of 281 tokens). Crucially, these numbers are based on a much larger amount of data (2952 RC tokens in total) compared to the original corpus analyses presented in the Pre-tests section. Thus, it might be related to the low statistical power that in the original corpus analyses and the sentence completion test, we did not find any headless RCs.

In contrast to Experiment 1, the effect of modification type reached significance at various regions. This can be attributed to the higher sensitivity of the eye-tracking method compared to SPR. The processing facilitation in object-modifications starting at the adverb and continuing up to the head noun is in line with the predictions of DLT storage cost. For the adverb and the V + N/N + V region, surprisal also explains this effect. At the head noun, surprisal predicts no effect, and therefore can account for the observed pattern only under the assumption that the effect is due to spillover from the previous regions. An alternative explanation in line with surprisal would be that in object-modifications, the conditional probability of the head noun being overtly realized is indeed higher compared to subject-modifications but the corpus search restricted to the RC tokens following a Det + Cl + Adv sequence and the sentence completion task did not have enough statistical power to detect this difference. This argument is supported by the more general corpus counts on headless vs headed RCs that are not restricted to RC tokens preceded by Det + Cl + Adv (see above), where headless RCs occur more frequently in subject-modifications than in object modifications.

Retrieval-based metrics for processing difficulty such as DLT integration cost (Gibson, 1998; Gibson, 2000) or ACT-R (Lewis & Vasishth, 2005) are incompatible with the faster reading times in object-modifications at the head noun since they predict the opposite direction of the effect.

General discussion

We investigated whether subject relatives or object relatives (subject- and object-modifications) are easier to process in Chinese by creating experimental stimuli in which the left context strongly constrains the predicted structure to be a relative clause. Importantly, in the experimental materials, we eliminated several local ambiguities present in Chinese relative clauses that might have confounded previous studies. In Mandarin Chinese, SRs occur more frequently than ORs (Hsiao & Gibson, 2003; Vasishth et al., 2013). Therefore, if the left context leads the comprehender to posit a relative clause as the most likely continuation, expectation-based parsing accounts such as surprisal (Hale, 2001; Levy, 2008) predict an SR advantage at the RC region (the verb and the object within SRs and the subject and the verb within ORs respectively) in both subject- and object-modifications.

In order to derive precise predictions of the expectation-based account for the materials to be used in the reading experiments, we assessed the conditional probability of SRs vs ORs and subject- vs object-modifications in a sentence completion experiment. The sentence completion data showed that in the experimental materials, the conditional probability of an SR continuation is higher than that of an OR continuation and this difference is larger in subject-modifications. Moreover, the conditional probability of a subject-modifying RC is lower than the conditional probability of an object-modifying RC. (The sentence completion-based higher conditional probability of SRs over ORs replicated corpus findings, whereas the differences with respect to modification type were not observed in corpus data.) Thus, the expectation-based account, and surprisal in particular, predicts an SR advantage at the relative clause region which is slightly stronger in subject-modifications. Moreover, it predicts faster reading times in object modifications compared to subject-modifications starting at the onset of the relative clause.

In contrast to expectation-based accounts, storage-based working-memory accounts such as DLT storage cost (Gibson, 2000; Gibson, 1998) assume that processing cost increases as a function of predictions to be kept track of. For the experimental materials tested here, storage cost predicts no difference between SRs and ORs in the relative clause region and at the head noun because an equal number of upcoming heads is predicted (Hsiao & Gibson, 2003, p. 6). Moreover, storage cost predicts processing facilitation in object-modifications since a smaller number of predicted syntactic heads have to be kept in memory.

Retrieval-based working memory accounts such as DLT integration cost (Gibson, 2000; Gibson, 1998) or the ACT-R based model of sentence processing (Lewis & Vasishth, 2005) assume that processing difficulty is a function of the distance between the currently processed item and the item that needs to be retrieved for the successful integration of the current item into the parse constructed so far. For the experimental materials, retrieval-based working memory accounts predict an OR advantage at the head noun because of the greater gap-head distance in ORs compared to SRs. Moreover, they predict faster reading times in subject-modifications at the head noun since in object-modifications, an additional retrieval, namely the retrieval of the main clause verb, is triggered.

We conducted two experiments with similar materials but different methods (self-paced reading and eye-tracking). The evidence from both experiments was unequivocal: in both self-paced reading times and across eye-tracking measures, we found an SR advantage at the RC region V + N/N + V preceding the head noun. Nested comparisons showed that this effect was present in both subject- and object-modifications, but more pronounced in subject-modifications. Also in both experiments, in subject-modifications only, we found an SR advantage in the materials following the head noun. As for the factor modification type, while no effect was found in the self-paced reading experiment, the eye-tracking data showed clear evidence for a processing facilitation in object-modifications. This facilitation was strongest at the very onset of the relative clause (the adverb) and continued to the head noun.

The predictions of expectation-based accounts of parsing are mostly in line with these results. The SR advantage at the RC region and the fact that this SR advantage was stronger in subject-modifications compared to object-modifications is

predicted by surprisal. Moreover, surprisal can also account for the faster reading times in object-modifications starting at the very onset of the relative clause and continuing up to the head noun if we assume that the effect at the head noun is driven by spillover from the previous regions. In contrast, the SR advantage in the spillover region in subject-modifications is not in line with the surprisal predictions derived from sentence completion data or the corpus analyses restricted to relative clauses that have a similar structure as the experimental items. However, more general corpus counts that are not restricted to relative clauses appearing after a Det + Cl + Adv sequence indicate that the conditional probability of an overt head noun appearing after a relative clause is higher in subject-modifying SRs than in subject-modifying ORs while there is no such difference in object-modifications, where the RC head is almost always overtly produced. (This difference in conditional probabilities would also account for the effect of modification type at the RC head).

Memory-based accounts are not compatible with our results. Storage-based memory accounts such as the DLT storage cost metrics are compatible with the faster reading times in object modifications but are inconsistent with the SR advantage observed at the RC region and, in subject-modifications, in the materials after the head noun. Our results are also inconsistent with the retrieval component of memory-based accounts. The faster reading times in object-modifications at the head noun are incompatible with both DLT integration cost (Gibson, 2000) and Lewis and Vasishth (2005)'s memory-based retrieval architecture as both predict an effect into the opposite direction. The SR advantage in the materials following the head noun is also inconsistent with DLT integration cost or the Lewis and Vasishth (2005) model since they predict an OR advantage at the head noun.

Note that although the effect of modification type and the modulation of the effect of RC type by modification type, i.e., the stronger SR advantage in subject-modifications, can be accounted for by surprisal, it might also reflect an underlying difference in the processing of the relative clause depending on which syntactic part of the sentence is being modified. The present experimental design does not allow us to draw any conclusions in this respect, but we believe that this issue deserves investigation in future research.

The SR advantage in Chinese adds to the growing body of cross-linguistic evidence showing that expectations for upcoming structure play a crucial role in determining processing difficulty. We have already mentioned the recent evidence from English (Staub, 2010) and Russian (Levy et al., 2013). In earlier work, Konieczny (2000) tested German verb-argument dependencies in verb-final structures and found evidence for processing facilitation as a function of increasing distance between the verb and its arguments. Similar results are reported for English (Jaeger, Fedorenko, Hofmeister, & Gibson, 2008), German (Levy & Keller, 2013; Vasishth & Drenhaus, 2011), and for Hindi (Husain, Vasishth, & Srinivasan, 2014; Vasishth & Lewis, 2006). Such anti-locality effects have been explained by Levy (2008) in terms of an increasing conditional probability (which effectively translates to higher predictability) of the verb given preceding context. Related work by Vasishth, Suckow, Lewis, and Kern (2010) has shown that English native speakers exhibit a counter-intuitive grammaticality illusion in reading times but that Germans do not show this illusion; English speakers find double center embeddings with the middle verb missing easier to process than the grammatical counterpart, whereas German speakers find the ungrammatical version harder to process. This English grammaticality illusion effect has been argued to reflect the statistical infrequency of double embeddings in English; in German, double center embeddings occur relatively more often because all relative clauses are verb-final. In a follow-up study, Frank, Trompenaars, and Vasishth (submitted for publication) showed that German and Dutch native speakers reading English double center embeddings track the structural probabilities of upcoming material, and that more proficient non-native English speakers show a greater grammaticality illusion effect in English, suggesting that they are more closely approximating the native-speaker knowledge of structural probabilities of upcoming structure. Finally, several large scale eye-tracking corpora developed for English and German have also been analysed using surprisal as a predictor (Boston, Hale, Patil, Kliegl, & Vasishth, 2008; Boston, Hale, Vasishth, & Kliegl, 2011; Demberg & Keller, 2008); in all cases surprisal has been shown to be a statistically significant predictor of various eye-tracking measures. The present study adds to this broad range of evidence, and provides new support for the idea that we generate and track predictions based on our knowledge of grammar, and that the degree of difficulty that we experience during sentence comprehension is determined, at least in part, by the extent to which our expectations for upcoming structure are met.

At the same time, a large number of studies have shown that increased distance between the two elements of a dependency leads to an increase in processing difficulty. For example, Grodner and Gibson (2005) and Bartek, Lewis, Vasishth, and Smith (2011) present evidence that in English argument-verb dependencies, increasing the distance between the verb and the argument leads to slower reading times. A similar pattern is observed in argument-verb dependencies in Russian RCs (Levy et al., 2013), and this has been replicated in Hungarian (Kovács & Vasishth, 2013). These locality effects have been presented as evidence for decay or interference supporting memory-based accounts of sentence processing. In the light of this large body of evidence for memory-based accounts of sentence processing, it would be unreasonable to conclude that expectation rather than memory restrictions determine processing ease in general. However, it is possible that there is cross-linguistic variation in the extent to which one or the other factor dominates. Our data suggest that, at least in the case of Chinese relative clauses, it is expectation rather than memory cost that determines the relative processing ease of SRs vs ORs.

It is worth noting that our findings are partially consistent with accounts based on the accessibility hierarchy (Keenan & Comrie, 1977) and accounts based on phrase-structure distance (O'Grady, 2007), which have been previously adopted to account for the SR advantage in East

Asian languages (Kwon et al., 2013; Kwon et al., 2010; Kwon et al., 2006; Lin, 2008; Lin & Bever, 2006; Miyamoto & Nakamura, 2003; Ueno & Garnsey, 2008). According to the accessibility hierarchy, noun phrases of certain grammatical functions are easier to access (or extract) than noun phrases of other grammatical functions. The claim is that subject noun phrases, being highest on the hierarchy, are easier to process than object noun phrases. Regarding the phrase-structure distance of SRs and ORs, it has been proposed that processing SRs is less costly because a smaller number of phrasal nodes intervene between the head noun and a subject gap. Both accounts predict processing differences between SRs and ORs on and after the head noun rather than inside the RC regions. These accounts are therefore only consistent with one of the findings reported here, namely the SR advantage in the spillover region following the head noun in subject-modifications. However, the SR advantage inside the RC region is best accounted for by the expectation-based account.

Another account that can partially explain our results is the so-called Perspective Shift account proposed by MacWhinney (2005). According to this account, processing difficulty increases as a shift in perspective is required. In the experimental materials, the perspective of the main clause and the relative clause are determined by the main clause agent (i.e., the main clause subject) and the RC agent (i.e., the RC subject), respectively.[12] We will assume that a subject takes the perspective as soon as the head noun of the subject NP is predicted. In our subject-modifying conditions, the SR shares the perspective of the main clause while in ORs a perspective-shift from the main clause subject to the RC subject is required when reading the RC subject. This is in line with the SR advantage at the N + V/V + N region in our data. For our object-modifying RCs, in contrast, no difference between SRs and ORs is predicted because in both cases a perspective shift from the main clause agent to the RC agent is required. This shift is predicted to happen at the RC region (at the RC subject in ORs and at the RC verb, i.e., when the RC head is predicted, in SRs). This prediction is not in line with the SR advantage at the N + V/V + N region observed in our object-modifying conditions. If, alternatively, we do not assume that a subject takes perspective as soon as its head is predicted but rather when its head is being encountered in the input, the pattern observed in our subject-modifying conditions cannot be explained by the Perspective Shift account, while the pattern observed in object-modifications might be partially explained. In subject-modifying SRs, the perspective of the main clause subject takes scope over the whole sentence, while in ORs, perspective needs to be shifted from the first encountered RC subject to the main clause subject. Therefore, an SR advantage is predicted at the RC head noun. Thus, the SR advantage observed in the spillover region of the head noun can be explained by the Perspective Shift account, but the even stronger SR advantage at the RC region cannot. In object-modifying conditions, perspective is shifted from the main clause agent to the RC agent when reaching the RC subject in ORs and when reaching the RC head in SRs. Therefore, an SR advantage is predicted at the RC region and an OR advantage at the head noun. Our data are consistent with the former prediction, but not with the latter. However, one could argue that in object-modifications, we do not observe any effect at the head noun because spillover of the SR advantage from the RC region and an OR advantage at the head noun are canceling each other out.

In sum, the Perspective Shift account may explain either the pattern we observed in subject- or in object-modifying conditions depending on the nature of the additional assumptions we make to derive predictions. Crucially, it cannot explain both subject- and object-modifications under the same set of assumptions.

One open issue that remains to be addressed is the role of animacy of the RC subject/object and the head noun. Wu et al. (2012) showed in three self-paced reading studies that SRs were read faster when they had animate heads and inanimate objects compared to the reverse animacy configuration. ORs, in contrast, were processed faster when they had inanimate heads and animate subjects. Moreover, they found an SR advantage in sentences with an inanimate subject and an animate object. However, this difference disappeared when the animacy configuration was reversed (animate subject and inanimate object). These findings reflect animacy preferences found in corpus counts. Therefore, the Wu et al. (2012) results are also consistent with expectation-based accounts of relative clause processing. An informative test case would be to replicate our study with animacy of the head noun and the relative clause subject/object as additional factors.

Conclusion

We present the first study comparing Chinese subject- and object-modifying subject and object relative clauses with materials that use syntactic cues to lead the comprehender to predict a relative clause as the upcoming structure. Two experiments show that the differential conditional probabilities of subject vs object relative continuations, estimated using a sentence completion study, can predict the reading time difficulty experienced by readers. We found clear evidence for subject relatives being processed faster than object relatives at the region containing the relative clause verb and at the spillover region of the head noun in subject-modifying conditions. These data are consistent with a particular instantiation of an expectation-based account of processing, surprisal. We conclude from these results that expectation plays an important role in the processing of Chinese relative clauses. More generally, our data provide independent support for the idea that the human sentence comprehension system deploys its probabilistic knowledge of grammar to generate predictions about upcoming structure. When these predictions are met, processing is relatively easy, but when they are not, processing difficulty occurs. In sum, dashed expectations are costly.

[12] For accusative–nominative languages it is assumed that language users take the perspective of the thematic agent while for ergative-absolutive languages it is assumed that the perspective of the thematic patient is taken. To derive the predictions of the Perspective Shift account for our materials, we will assume that Chinese, which does not have overt case marking, clusters with accusative-nominative languages in taking the agent's perspective.

Acknowledgments

We thank the Department of English at National Taiwan Normal University for allowing us to conduct our experiment in their eye-tracking lab. We are grateful to their staff, and to Shiaohui Chan and Sun-Chieh Liang in particular, for their assistance. We thank three anonymous reviewers for their constructive suggestions and Anne Steinke, Yu-Yin Hsu, Dario Paape, Bruno Nicenboim and the audiences at CUNY 2012 and 2013 for valuable discussions. We are grateful to Ted Gibson for sharing his experimental items from a previous study on Chinese. Shravan Vasishth thanks the professors at the University of Sheffield's School of Mathematics and Statistics, in particular Jeremy Oakley and Kevin Walters, for valuable advice on statistical analysis using linear mixed models. Charles Lin thanks Indiana University's Office of the Vice Provost for Research for research support through its Faculty Research Support Program.

Corpus study of Chinese Treebank 7.0

Tables A1, A2 and A3.

Sentence completion data

Tables B1 and B2.

Table B1
Sentence completions produced after a Det + Cl sequence (subject-modification) and after a NP + V + Det + Cl (object-modification) sequence observed in the sentence completion study. Canonical RCs are defined as RCs with the canonical SVO word order and an overt RC subject in ORs.

Sentence completion			Count (subj-mod)	Count (obj-mod)
RC	SR	*canonical*	27	48
		adjectival	0	0
		passive (bei)	2	2
		topicalized	1	2
	OR	*canonical*	3	9
		subj. drop	2	2
	Possessive RC		0	2
	Adjunct RC		0	0
Not RC			117	87

Table A1
Corpus counts of structures following a zhe/na(*this/that*) + classifier + adverb sequence. All relative clause tokens are headed relative clauses. Headless relative clauses in this position have zero tokens in the corpus.

Structure following *zhe/na + Cl + Adv*		Count	Frequency (%)
SR	*V (N) de N*	58	79.5
OR (dropped subject)	*pro V de N*	13	17.8
OR (overt subject)	*N V de N*	1	1.4
Main clause VP	V N	1	1.4

Table A2
Categorization of the corpus tokens presented in Table A1 (structures following a zhe/na (*this/that*) + classifier + adverb sequence) by the syntactic role (subject, object or other) of the noun which the the *zhe/na + classifier* phrase modifies. All relative clause tokens are headed relative clauses. Headless relative clauses in this position have zero tokens in the corpus.

Noun modified by *zhe/na + Cl*	Structure following *zhe/na + Cl + Adv*		Count	Relative frequency (%)
Subject	SR	*V (N) de N*	31	77.5
	OR (dropped subject)	*pro V de N*	8	20.0
	OR (overt subject)	*N V de N*	0	0
	Main clause VP	V (N)	1	2.5
Object	SR	*V (N) de N*	17	85.0
	OR (dropped subject)	*pro V de N*	2	10.0
	OR (overt subject)	*N V de N*	1	5.0
	Main clause VP	V (N)	0	0
Other	SR	*V (N) de N*	10	76.9
	OR (dropped subject)	*pro V de N*	3	23.1
	OR (overt subject)	*N V de N*	0	0
	Main clause VP	V (N)	0	0

Table A3
Corpus counts of structures following a zhe/na(*this/that*) + classifier sequence, categorized by whether the zhe/na + classifier phrase modifies the subject or the object of the main clause.

Noun modified by *zhe/na + Cl*	Structure following *zhe/na + Cl*	Count	Relative frequency (5)
Subject	RC	163	13.9
	Other	1007	86.1
Object	RC	95	12.3
	Other	677	87.7

Table B2
Sentence completions produced after a Det + Cl + Adv sequence (subject-modification) and after a NP + V + Det + Cl + Adv sequence (object-modification) observed in the sentence completion study. Canonical RCs are defined as RCs with the canonical SVO word order and an overt RC subject in ORs.

Sentence completion			Count (subj-mod)	Count (obj-mod)
RC	SR	*canonical (headed)*	100	121
		canonical (headless)	1	0
		adjectival	1	3
		passive (bei)	5	3
		topicalized	0	1
	OR	*canonical (headed)*	3	8
		canonical (headless)	0	0
		subj. drop (headed)	16	8
		subj. drop (headless)	1	0
	Possessive RC		2	1
	Adjunct RC		1	0
Not RC	Diff. PoS for Adv		3	3
	Elided NP betw. Det + Cl and Adv		15	0

References

Asendorpf, J. B., Conner, M., De Fruyt, F., De Houwer, J., Denissen, J. J., Fiedler, K., et al. (2013). Recommendations for increasing replicability in psychology. *European Journal of Personality, 27*(2), 108–119.

Bartek, B., Lewis, R. L., Vasishth, S., & Smith, M. (2011). In search of on-line locality effects in sentence comprehension. *Journal of Experimental Psychology: Learning, Memory and Cognition, 37*(5), 1178–1198.

Bates, D., Maechler, M., Bolker, B., & Walker, S. (2014). lme4: Linear mixed-effects models using Eigen and S4 [Computer software manual]. Retrieved from http://CRAN.R-project.org/package=lme4 (R package version 1.0-6)

Boston, M. F., Hale, J. T., Patil, U., Kliegl, R., & Vasishth, S. (2008). Parsing costs as predictors of reading difficulty: An evaluation using the Potsdam Sentence Corpus. *Journal of Eye Movement Research, 2*(1), 1–12.

Boston, M. F., Hale, J. T., Vasishth, S., & Kliegl, R. (2011). Parallel processing and sentence comprehension difficulty. *Language and Cognitive Processes, 26*(3), 301–349.

Box, G. E. P., & Cox, D. R. (1964). An analysis of transformations. *Journal of the Royal Statistical Society, 26*, 211–234.

Chen, B., & Ning, A. (2008). The comparison of processing difficulty between Chinese subject-relative and object-relative clauses. *Chinese Journal of Applied Psychology (应用心理学), 14*(1), 29–34.

Chen, B., Ning, A., Bi, H., & Dunlap, S. (2008). Chinese subject-relative clauses are more difficult to process than the object-relative clauses. *Acta Psychologica, 129*(1), 61–65.

Chen, Z., Grove, K., & Hale, J.T. (2012). Structural expectations in Chinese relative clause comprehension. In J. Choi, E.A. Hogue, J. Punske, D. Tat, J. Schertz, & A. Trueman (Eds.), *Proceedings of the 29th West Coast Conference on Formal Linguistics* (pp. 29–37). Somerville, MA: Cascadilla Press.

Clifton, C., Jr., Staub, A. J, & Rayner, K. (2007). Eye movements in reading words and sentences. In M. Fisher R. Van Gompel & R. L. Hill W. Murray (Eds.), *Eye movements: A window on mind and brain* (pp. 341–372). Oxford, UK: Elsevier.

Clifton, C., Jr., & Frazier, L. (1989). Comprehending sentences with long distance dependencies. In G. Carlson & M. Tanenhaus (Eds.), *Linguistic structure in language processing* (pp. 273–317). Dordrecht, Netherlands: Kluwer.

Demberg, V., & Keller, F. (2008). Data from eye-tracking corpora as evidence for theories of syntactic processing complexity. *Cognition, 109*(2), 193–210.

Fisher, R. (1937). *The design of experiments* (2nd ed.). Edinburgh, UK: Oliver and Boyd.

Frank, S.L., Trompenaars, T., & Vasishth, S. (submitted for publication). Cross-linguistic differences in processing double-embedded relative clauses: Working-memory constraints or language statistics?

Frazier, L. (1979). On comprehending sentences: Syntactic parsing strategies (PhD thesis). University of Massachusetts, Amherst, MA.

Frazier, L., & Fodor, J. D. (1978). The sausage machine: A new two-stage parsing model. *Cognition, 6*(4), 291–325.

Gelman, A., & Hill, J. (2007). *Data analysis using regression and multilevel/hierarchical models*. New York, NY: Cambridge University Press.

Gibson, E. (1998). Linguistic complexity: Locality of syntactic dependencies. *Cognition, 68*(1), 1–76.

Gibson, E. (2000). Dependency locality theory: A distance-based theory of linguistic complexity. In A. Marantz, Y. Miyashita, & W. O'Neil (Eds.), *Image. language, brain* (pp. 95–126). Cambridge, MA: MIT Press.

Gibson, E., & Wu, H.-H. I. (2013). Processing Chinese relative clauses in context. *Language and Cognitive Processes, 28*(1-2), 125–155.

Grodner, D., & Gibson, E. (2005). Consequences of the serial nature of linguistic input for sentential complexity. *Cognitive Science, 29*, 261–291.

Hale, J. T. (2001). A probabilistic Earley parser as a psycholinguistic model. In *Proceedings of the 2nd Meeting of the North American Chapter of the Association for Computational Linguistics* (pp. 159–166). Pittsburgh, PA: Association for Computational Linguistics.

Hsiao, F., & Gibson, E. (2003). Processing relative clauses in Chinese. *Cognition, 90*(1), 3–27.

Hsiao, Y., Li, J., & MacDonald, M. (2014). Ambiguity affects Mandarin relative clause processing. In *27th Annual CUNY Conference on Human Sentence Processing*. Columbus, OH.

Hsiao, Y., & MacDonald, M. C. (2013). Experience and generalization in a connectionist model of Mandarin Chinese relative clause processing. *Frontiers in Psychology, 4*(767).

Hsu, C.-C.N., Hurewitz, F., & Phillips, C. (2006). Contextual and syntactic cues for head-final relative clauses in Chinese. In *19th Annual CUNY Conference on Human Sentence Processing*. New York, NY.

Hsu, C.-C.N., Phillips, C., & Yoshida, M. (2005). Cues for head-final relative clauses in Chinese. In *18th Annual CUNY Conference on Human Sentence Processing*. Tucson, AZ.

Huang, C. T. J. (1989). Pro drop in Chinese: a generalized control approach. In O. Jaeggli & K. Safir (Eds.), *The null subject parameter* (pp. 185–214). Dordrecht, Netherlands: Kluwer.

Husain, S., Vasishth, S., & Srinivasan, N. (2014). Strong expectations cancel locality effects: Evidence from Hindi. *PLoS ONE, 9*(7), e100986.

Jaeger, F.T., Fedorenko, E., Hofmeister, P., & Gibson, E. (2008). Expectation-based syntactic processing: Antilocality outside of head-final languages. In *21th Annual CUNY Conference on Human Sentence Processing*. Chapel Hill, NC.

Just, M. A., & Carpenter, P. A. (1992). A capacity theory of comprehension: Individual differences in working memory. *Psychological Review, 99*(1), 122–149.

Just, M. A., Carpenter, P. A., & Woolley, J. D. (1982). Paradigms and processes in reading comprehension. *Journal of Experimental Psychology: General, 111*(2), 228–238.

Keenan, E. L., & Comrie, B. (1977). Noun phrase accessibility and universal grammar. *Linguistic Inquiry, 8*(1), 63–99.

Konieczny, L. (2000). Locality and parsing complexity. *Journal of Psycholinguistic Research, 29*(6), 627–645.

Kovács, N., & Vasishth, S. (2013). The processing of relative clauses in Hungarian. In *Architectures and Mechanisms for Language Processing Conference*. Marseille, France.

Kwon, N., Kluender, R., Kutas, M., & Polinsky, M. (2013). Subject/object processing asymmetries in Korean relative clauses: Evidence from ERP data. *Language, 89*(3), 537–585.

Kwon, N., Lee, Y., Gordon, P., Kluender, R., & Polinsky, M. (2010). Cognitive and linguistic factors affecting subject/object asymmetry: An eye-tracking study of pre-nominal relative clauses in Korean. *Language, 86*(3), 546–582.

Kwon, N., Polinsky, M., & Kluender, R. (2006). Subject preference in Korean. In D. Baumer, D. Montero, & M. Scanlon (Eds.), *Proceedings of the 25th West Coast Conference on Formal Linguistics* (pp. 1–14). Somerville, MA: Cascadilla Press.

Kwong, O.Y., & Tsou, B.K. (2003). Categorial fluidity in Chinese and its implications for part-of-speech tagging. In A. Copestake, & J. Hajic (Eds.), *Proceedings of the 10th Conference of the European Chapter of the Association for Computational Linguistics* (pp. 115–118). Budapest, Hungary: Association for Computational Linguistics.

Levy, R. (2008). Expectation-based syntactic comprehension. *Cognition, 106*, 1126–1177.

Levy, R., & Andrew, G. (2006). Tregex and Tsurgeon: Tools for querying and manipulating tree data structures. In N. Calzolari, K. Choukri, A. Gangemi, B. Maegaard, J. Mariani, J. Odjik, & D. Tapias (Eds.), *Proceedings of the 5th International Conference on Language Resources and Evaluation* (pp. 2231–2234). Genoa, Italy: European Language Resources Association.

Levy, R., Fedorenko, E., & Gibson, E. (2013). The syntactic complexity of Russian relative clauses. *Journal of Memory and Language, 69*, 461–495.

Levy, R., & Keller, F. (2013). Expectation and locality effects in German verb-final structures. *Journal of Memory and Language, 68*, 199–222.

Lewis, R. L., & Vasishth, S. (2005). An activation-based model of sentence processing as skilled memory retrieval. *Cognitive Science, 29*, 1–45.

Lewis, R. L., Vasishth, S., & Van Dyke, J. (2006). Computational principles of working memory in sentence comprehension. *Trends in Cognitive Sciences, 10*(10), 447–454.

Lin, C.-J. C. (2008). The processing foundation of head-final relative clauses. *Language and Linguistics, 9*, 813–839.

Lin, C.-J. C. (2014). Effect of thematic order on the comprehension of Chinese relative clauses. *Lingua, 140*, 180–206.

Lin, C.-J. C., & Bever, T. G. (2006). Subject preference in the processing of relative clauses in Chinese. In D. Baumer, D. Montero, & M. Scanlon (Eds.), *Proceedings of the 25th West Coast Conference on Formal Linguistics* (pp. 254–260). Somerville, MA: Cascadilla Press.

Lin, C.-J. C., & Bever, T. G. (2011). Garden path and the comprehension of head-final relative clauses. In H. Yamashita, Y. Hirose, & J. Packard (Eds.), *Processing and producing head-final structures* (pp. 277–297). Dordrecht, Netherlands: Springer.

Lin, Y., & Garnsey, S. (2011). Animacy and the resolution of temporary ambiguity in relative clause comprehension in Mandarin. In H. Yamashita, Y. Hirose, & J. Packard (Eds.), *Processing and producing head-final structures* (pp. 241–275). Dordrecht, Netherlands: Springer.

Liu, T., Zhou, T., & Yang, Y. (2011). Universality of processing advantage for subject relative clause: An ERP study of Chinese relative clauses. *Linguistic Sciences (*语言科学*), 10*(1), 1–20.

MacWhinney, B. (2005). The emergence of grammar from perspective. In D. Pecher & R. A. Swann (Eds.), *Grounding cognition: The role of perception and action in memory, language, and thinking* (pp. 198–223). New York, NY: Cambridge University Press.

Miller, G., & Chomsky, N. (1963). Finitary models of language users. In R. D. Luce, R. R. Bush, & E. Galanter (Eds.). *Handbook of mathematical psychology* (Vol. 2, pp. 419–491). New York, NY: John Wiley and Sons.

Miyamoto, E. T., & Nakamura, M. (2003). Subject/object asymmetries in the processing of relative clauses in Japanese. In G. Garding, & M. Tsujimura (Eds.), *Proceedings of the 22nd West Coast Conference on Formal Linguistics* (pp. 342–355). Somerville, MA: Cascadilla Press.

O'Grady, W. (2007). *Syntactic development*. Chicago, IL: University of Chicago Press.

Pickering, M. J., & Barry, G. (1991). Sentence processing without empty categories. *Language and Cognitive Processes, 6*(3), 229–259.

Pickering, M. J., & Garrod, S. (2007). *Do people use language production to make predictions during comprehension? Trends in Cognitive Sciences, 11*(3), 105–110.

Pickering, M. J., & Van Gompel, R. P. (2006). Syntactic parsing. In M. Traxler & M. Gernsbacher (Eds.), *Handbook of psycholinguistics* (2nd ed., pp. 455–503). Amsterdam, Netherlands: Academic Press.

Qiao, X., Shen, L., & Forster, K. (2012). Relative clause processing in Mandarin: Evidence from the maze task. *Language and Cognitive Processes, 27*(4), 611–630.

Simmons, J. P., Nelson, L. D., & Simonsohn, U. (2011). False-positive psychology: Undisclosed flexibility in data collection and analysis allows presenting anything as significant. *Psychological Science, 22*(11), 1359–1366.

Staub, A. (2010). Eye movements and processing difficulty in object relative clauses. *Cognition, 116*(1), 71–86.

Traxler, M. J., & Pickering, M. J. (1996). Plausibility and the processing of unbounded dependencies: An eye-tracking study. *Journal of Memory and Language, 35*, 454–475.

Ueno, M., & Garnsey, S. M. (2008). An ERP study of the processing of subject and object relative clauses in Japanese. *Language and Cognitive Processes, 23*(5), 646–688.

Vasishth, S., Chen, Z., Li, Q., & Guo, G.-L. (2013). Processing Chinese relative clauses: Evidence for the subject-relative advantage. *PLoS One, 8*(10), e77006.

Vasishth, S., & Drenhaus, H. (2011). Locality in German. *Dialogue and Discourse, 1*, 59–82.

Vasishth, S., & Lewis, R. L. (2006). Argument-head distance and processing complexity: Explaining both locality and antilocality effects. *Language, 82*(4), 767–794.

Vasishth, S., Suckow, K., Lewis, R. L., & Kern, S. (2010). Short-term forgetting in sentence comprehension: Crosslinguistic evidence from head-final structures. *Language and Cognitive Processes, 25*(4), 533–567.

Wu, F. (2009). Factors affecting relative clause processing in Mandarin: Corpus and behavioral evidence (PhD thesis). University of Southern California, Los Angeles, CA.

Wu, F. (2011). Frequency issues of classifier configurations for processing Mandarin object-extracted relative clauses: A corpus study. *Corpus Linguistics and Linguistic Theory, 7*(2), 203–207.

Wu, F. (2013). Will more heads increase processing difficulties? Evidence from Chinese. *Journal of Foreign Languages (*外国语*), 36*(3), 60–67.

Wu, F., Kaiser, E., & Andersen, E. (2010). Subject preference, head animacy, and lexical cues: A corpus study of relative clauses in Chinese. In H. Yamashita, Y. Hirose, & J. Packard (Eds.), *Processing and producing head-final structures* (pp. 173–194). Dordrecht, NL: Springer.

Wu, F., Kaiser, E., & Andersen, E. (2012). Animacy effects in Chinese relative clause processing. *Language and Cognitive Processes, 27*(10), 1489–1524.

Wu, F., Luo, Y., & Zhou, X. (2014). Building Chinese relative clause structures with lexical and syntactic cues: Evidence from visual world eye-tracking and reading times. *Language, Cognition and Neuroscience, 29*(10), 1205–1226.

Xue, N., Xia, F., Chiou, F.-D., & Palmer, M. (2005). The Penn Chinese TreeBank: Phrase structure annotation of a large corpus. *Natural Language Engineering, 11*(2), 207–238.

Yoshida, M., Aoshima, S., & Phillips, C. (2004). Relative clause prediction in Japanese. In *17th Annual CUNY Conference on Human Sentence Processing*. College Park, MD.

Yoshida, M., & Yoon, S. (2014). When are clause-final verbs facilitated in Korean? In *The 2nd East Asian Psycholinguistics Colloquium*. Chicago, IL.

Yun, J., Whitman, J., & Hale, J.T. (2010). Subject-object asymmetries in Korean sentence comprehension. In S. Ohlsson and R. Catrambone (Eds.), *Proceedings of the 32nd Annual Meeting of the Cognitive Science Society* (pp. 2152–2157). Portland, OR: Cognitive Science Society.

Zhang, Q., & Jiang, H. (2010). Processing advantages in relative clauses and the local grammar complexity account. *Foreign Languages Research (*外语研究*), 124*(6), 19–26.