COGNITIVE SCIENCE
A Multidisciplinary Journal

Cognitive Science (2017) 1–33

ISSN: 0364-0213 print / 1551-6709 online
DOI: 10.1111/cogs.12551

Effects of Early Cues on the Processing of Chinese Relative Clauses: Evidence for Experience-Based Theories

Fuyun Wu,[a] Elsi Kaiser,[b] Shravan Vasishth[c]

[a]*School of Foreign Languages, Shanghai Jiao Tong University*
[b]*Department of Linguistics, University of Southern California*
[c]*Department of Linguistics, University of Potsdam*

Received 13 October 2015; received in revised form 12 August 2017; accepted 16 August 2017

Abstract

We used Chinese prenominal relative clauses (RCs) to test the predictions of two competing accounts of sentence comprehension difficulty: the experience-based account of Levy (2008) and the Dependency Locality Theory (DLT; Gibson, 2000). Given that in Chinese RCs, a classifier and/or a passive marker BEI can be added to the sentence-initial position, we manipulated the presence/absence of classifiers and the presence/absence of BEI, such that BEI sentences were passivized subject-extracted RCs, and no-BEI sentences were standard object-extracted RCs. We conducted two self-paced reading experiments, using the same critical stimuli but somewhat different filler items. Reading time patterns from both experiments showed facilitative effects of BEI within and beyond RC regions, and delayed facilitative effects of classifiers, suggesting that cues that occur before a clear signal of an upcoming RC can help Chinese comprehenders to anticipate RC structures. The data patterns are not predicted by the DLT, but they are consistent with the predictions of experience-based theories.

Keywords: Storage cost; Experience; Relative clause; Chinese; Classifiers; BEI

1. Introduction

In recent years the field of psycholinguistics has seen a growing interest in relative clause (RC) processing, because the complexity of RC structures provides a good test case to examine factors affecting processing difficulty within and across languages, and to evaluate theories of sentence comprehension. To process an RC structure (e.g., "the

Correspondence should be sent to Fuyun Wu, School of Foreign Languages, Shanghai Jiao Tong University, 800 Dong Chuan Road, Shanghai 200240, China. E-mail: fywu@sjtu.edu.cn

reporter$_i$ [that the stone hit t_i]"), a comprehender needs to recognize the beginning of the RC (i.e., the RC boundary, as marked by the relativizer "that") and establish a relationship between the head noun ("the reporter," also called the "filler") and the empty nominal within the RC (the gap) over a distance, thereby completing the filler-gap dependency (e.g., Clifton & Frazier, 1989; Frazier & d'Arcais, 1989). Researchers have not yet reached a consensus regarding how the human processor—subject to limited memory resources—arrives at the target RC structure quickly and efficiently (e.g., Gibson, 1998, 2000; Konieczny, 2000; Levy & Keller, 2013; MacDonald, 2013; Staub & Clifton, 2006; Vasishth & Lewis, 2006).

Recent work on RC processing has been guided by two general approaches: working memory–based theories and experience-based theories, both of which have a number of variants (for detailed discussions, see Gibson & Wu, 2013; Levy & Keller, 2013; Levy, Fedorenko, & Gibson, 2013). Directly related to our work on Chinese RC processing, we focus on one influential version of the working memory–based approach, namely Gibson's (1998, 2000) Dependency Locality Theory (DLT), because the predictions of this particular theory have been tested with Chinese RCs in the past decade (Gibson & Wu, 2013; Hsiao & Gibson, 2003).

According to the DLT, human parsing systems consume working memory resources in the process of keeping track of syntactic heads over a linear distance between the head and its dependents. Memory costs are calculated in terms of two kinds of processes[1]: (i) the *storage cost* (measured in Memory Units) of maintaining syntactic heads required to complete a phrase-structural dependency, and (ii) the *integration cost* (measured in Energy Units) of integrating a current word into an existing structure. Thus, the *storage cost* component of DLT predicts that the higher the number of predicted syntactic heads stored in working memory, the greater the computational difficulty; the *integration cost* component of DLT predicts that the longer the linear distance between a head (e.g., a filler) and its dependent (e.g., a gap) in terms of intervening discourse referents (nouns and verbs), the greater the integration cost.

In contrast, for experience-based theories in general, the key predictor for processing difficulties is comprehenders' experience (familiarity) with structures or comprehenders' ability to predict likely structures to occur in a particular context. There are several theories that can be grouped under the label of "experience-based" theories, including the word-order frequency theory (Bever, 1970; MacDonald & Christiansen, 2002), the Production-Distribution-Comprehension account (Gennari & MacDonald, 2008; MacDonald, 2013), surprisal/expectation (Hale, 2001; Levy, 2008), and entropy-reduction accounts (Hale, 2003). In word-order frequency theories, surface orderings of word categories in the input that occur more frequently or resemble the canonical word order in a language should be easier to process. Note, however, that word-order frequency theories do not make precise predictions about where in the sentence the processing difficulties occur. Theories based on surprisal (Hale, 2001; Levy, 2008) or entropy-reduction (Hale, 2003, 2006) propose that comprehenders formulate and update conditional probabilities of upcoming input based on various structural features of the words already seen. While surprisal can be quantified using corpora or sentence

completion data (e.g., Levy & Keller, 2013), it is not always easy to know the precise predictions they make due to the limited existence of corpora that faithfully represent the correct structural patterns within and across languages (see Levy & Manning 2003 for evaluations of Chinese Treebank corpus coding procedure). Furthermore, sometimes the target structure under discussion might not be easily found in certain genres of corpora.

Both the DLT and experience-based theories are supported by prior work. For the DLT, substantial evidence has been obtained for its *integration cost* metric in various structures in English: Adding more intervening discourse referents makes processing more difficult in regions where dependencies are integrated (see Gibson, 1998, for an overview)—a robust phenomenon commonly known as the locality effect. However, increasing evidence also indicates that adding a preverbal dependent facilitates, rather than hinders, processing at the clause-final verb in head-final German (Konieczny, 2000), Hindi (Vasishth & Lewis, 2006), and Japanese (Nakatani & Gibson, 2008). Such evidence against locality has been used to argue for the expectation-based account of Levy (2008), a variant of the experience-based theories. Recent work has also shown evidence for both locality effects and expectation effects in English (Demberg & Keller, 2008; Jaeger, Fedorenko, Hofmeister, & Gibson, 2008; Staub, 2010), German (Levy & Keller, 2013; Vasishth & Drenhaus, 2011), Hindi (Husain, Vasishth, & Srinivasan, 2014, 2015), Persian (Safavi, Husain, & Vasishth, 2016), and Russian (Levy et al., 2013).

Existing work on head-final RCs in Chinese has shown mixed results, with some favoring DLT's storage cost metric (Hsiao & Gibson, 2003) or integration cost metric (B. Gibson & Wu, 2013; Lin & Garnsey, 2010), and others arguing against DLT (C. Chen & Vasishth, 2010; Jäger, Chen, Li, Lin, & Vasishth, 2015; Lin & Bever, 2011; Vasishth, Chen, Li, & Kuo, 2013). It is worth noting that most studies on Chinese RC processing focused on testing DLT's integration cost metric at the head noun and did not directly test the storage-cost metric. One reason could be the word order difference, that is, verb-noun (VN) order in subject-extracted RCs and noun-verb (NV) order in object-extracted RCs. Due to this fact, the two-word RC region was usually treated as one single segment. Interestingly, in their earliest work on Chinese RCs, Hsiao and Gibson (2003, p. 14) only found evidence for DLT's storage-cost metric, not the integration cost metric, suggesting that it is empirically possible to distinguish these two cognitive processes involved in sentence parsing.

Because (i) few studies contrast DLT's storage-cost metric with experience-based theories, and (ii) evidence from Chinese RCs is still mixed, the present study aims to fill these gaps by exploring the predictions of these two theories for RCs in Chinese. Chinese provides an interesting test case because it has mixed word order (Huang, 1982). It unusually combines SVO word order with noun-final properties (Dryer, 1992), making RC-recognition a relatively delayed process compared with head-initial RCs in English where the head noun occurs prior to the RC (Hawkins, 2004, p. 145). Consider an object-extracted RC in Chinese in (1).

(1) Object-extracted RC (ORC)

[$_{RC}$ shikuai[2] zazhong t_i de] jizhe$_i$

stone hit t_i DE reporter$_i$

‘the reporter that the stone hit _’

Prior to the adnominal DE, the initial NV (“stone hit”) is temporarily ambiguous between a simple main clause parse and an RC parse. The next available adnominal DE, unlike the relativizer “that” in an English RC whose presence clearly marks a clausal boundary, might not be a reliable signal for an upcoming RC, because (i) it occurs at the right edge of the RC, immediately before the RC head (*jizhe* “reporter”), and (ii) it is homophonic, also occurring in possessive, attributive, and noun-complement structures, in addition to restrictive RCs (Li & Thompson, 1981; Simpson, 1999). Thus, it is not until the disambiguating head noun is seen that the ultimate gapped-RC structure can be built.

Taking advantage of the head-final nature of the Chinese RC construction, we can create a situation where constituents that are part of an RC structure are added to the leftmost edge of RC, which allows us to compare the predictions made by the DLT and experience-based theories. Specifically, we use a clause-initial demonstrative classifier and/or the passive marker BEI.

1.1. Demonstrative classifiers

In Chinese, a classifier is grammatically required for any noun that has a demonstrative (Dem) (Huang, 1982; Li, 1998; Li & Thompson, 1981), and it must be semantically congruent with the host noun. If we add a demonstrative and a human-denoting classifier *wei* to the left edge of the object-extracted RC (1), as in (2), it is possible that the comprehender may use the incongruence/mismatch between the classifier and the local inanimate noun (“stone”) to anticipate that a congruent/matching noun will come up later on, and thus start to expect an RC. Note that as marked in (2), an additional classifier-(head)noun dependency needs to be completed at the head noun, in addition to the filler-gap dependency that is commonly present in (1–2).

(2) object-extracted RC (ORC) with a clause-initial Dem-Cl

na-wei [$_{RC}$ shikuai zazhong t_i de] jizhe$_i$

that-CL$_{human/*stone}$ stone hit t_i DE reporter$_i$

‘the reporter who the stone hit _’

1.2. Passive marker BEI

If we add the passive marker BEI to the left edge of (1), as in (3), it is very likely that a comprehender will expect an RC. Due to Chinese grammar, the presence of BEI minimally requires a verb (e.g., *zazhong* "hit"), an optional agent or (in this case) instrument (e.g., *shikuai* "stone"), and—crucially—an obligatory patient (*jizhe* "reporter") in order to complete a passive event, forming the so-called "BEI construction".[3] Thus, the obligatory presence of a patient noun in a typical passive structure, together with the presence of the adnominal DE, would lead a Chinese comprehender to expect an upcoming patient noun, which must be the head of an RC structure.

According to syntactic analyses of Chinese passives (Huang, Li, & Li, 2009), the head noun "reporter" receives a patient role in its post-verbal base position, but it raises to the subject position of the BEI-marked RC clause (*the reporter was hit*). Therefore, we treat RCs with BEI (ex. 3) as involving subject-extraction. Compared with the standard RC structure, the *passivized* SRC is more complex due to more intervening heads between the gap-filler dependency, which, by the DLT, would entail more processing difficulties.

(3) Passive subject-extracted RC (SRC) with a clause-initial BEI

[$_{RC}$ t_i	**bei**	shikuai	zazhong	de]	jizhe$_i$
t_i	**BEI**	stone	hit	DE	reporter$_i$

'the reporter who _ was hit by the stone'

1.3. BEI and classifier

If we add *both* a mismatching classifier and the passive marker BEI to the left edge of (1),[4] as in (4), it is highly likely that a comprehender will start to expect an RC immediately upon encountering BEI. This is because in the Chinese BEI construction, the presence of a patient undergoing or experiencing the passive event is obligatory, and the *optional* noun following BEI, when present, has to be interpreted as the agent/instrument of an action. Thus, the Chinese comprehender may use the passive marker BEI, together with the preceding classifier *wei* whose presence entails a human (head) noun yet to come, to construct an RC structure, which is the only possible continuation to end the sentence.

(4) Passive subject-extracted RC (SRC) with clausal-initial Dem-Cl and BEI

na-wei	[$_{RC}$ t_i	**bei**	shikuai	zazhong	de]	jizhe$_i$
that-CL$_{human/*stone}$	t_i	**BEI**	stone	hit	DE	reporter$_i$

'the reporter who _ was hit by the stone'

From the perspective of experience-based theories, the presence of the demonstrative-classifier sequence (Dem-Cl) in (2) and/or the presence of BEI in (3–4) are likely

to help comprehenders not to entertain a main clause misparse prior to DE, but to expect an RC. In contrast, according to the DLT, these clause-initial cues potentially increase storage costs and integration costs over long-distance attachments between heads and their dependents, including (i) classifiers and their matching (head)nouns as in (2 & 4), and (ii) gaps and their fillers as in the BEI sentences (ex. (3 & 4)) where the filler-gap dependency is longer than that in no-BEI sentences (ex. (1–2)). In this paper, we assess the predictions made by the DLT and by experience-based theories in two self-paced reading experiments. Before going into the experimental details, we first present an overview of existing work on the role of the classifier and the passive marker BEI in Chinese RC processing.

2. Existing work on Chinese RC processing

Existing processing work investigating the role of classifiers in Chinese has mostly involved *object-extracted RCs* where classifiers are adjacent to local RC-subjects but distant from their hosting heads (Hsu, Hurewitz, & Phillips, 2006; Hsu, Phillips, & Yoshida, 2005; Wu, Haskell, & Andersen, 2006; Wu, Luo, & Zhou, 2014a). These studies compared a classifier-mismatch condition with a classifier-match condition, by manipulating the semantic (in)congruity between a preceding classifier and its local noun. The logic is that if classifiers can serve as a cue for the hosting RC-head, then parsing will be facilitated at the head noun in the classifier-mismatch condition relative to the classifier-match condition.

However, the results from existing studies in Chinese are rather mixed. When sentences were presented in isolation, no processing facilitation was found at the head noun; rather, a long-lasting slowdown at the embedded noun was incurred by mismatching classifiers compared with matching classifiers (Hsu et al., 2005). When preceded by RC-facilitative discourse contexts in which the presence of two referents rendered an RC necessary, some self-paced reading studies found facilitative effects at the head (Hsu, 2006; Hsu et al., 2006), while other visual world eye-tracking studies showed that comprehenders were able to use mismatching classifiers to prevent garden-pathing on a direct object parse, but needed time to decide upon the correct RC parse (Wu, Sheng, & Zhou, 2014b; Wu et al., 2006, 2014a).

Regarding the role of the passive marker BEI in RCs, the few studies that have explored its effects appear to support experience-based theories. In a self-paced reading study, Kuo and Vasishth (2006) showed that adding a sentence-initial passive marker BEI to object-extracted RCs yielded numerically faster reading times than ORCs without BEI. The results, although statistically marginal, are inconsistent with the predictions of the storage cost metric of the DLT.

Furthermore, in a visual world eye-tracking study, Wu et al. (2014a,b) found that the co-presence of a mismatching classifier and BEI in object-modifying RCs not only prevented comprehenders from garden-pathing but also successfully helped them to fixate at the target picture at the earliest possible time. This suggests that, contrary to

what the DLT would predict, the presence of two pre-RC cues greatly facilitated RC expectations.

In sum, while existing evidence suggests that the passive marker BEI in conjunction with a preceding classifier might pre-activate the RC structure, the question remains whether BEI or classifiers *alone* facilitate RC processing. It is worth noting that prior research on the role of pre-RC classifiers has almost exclusively used stimuli with classifiers, comparing classifier-match and classifier-mismatch conditions. Thus, more work is needed to investigate whether adding a classifier to an otherwise bare RC may impede or facilitate recognition of an RC head.

3. Goals and predictions

The main goal of our study is to test the predictions made by the DLT and experience-based theories of sentence processing. We manipulated (i) Classifier (absent or present) and (ii) BEI (absent or present), yielding four conditions: *no-Classifier, no-BEI* (5a), *Classifier, no-BEI* (5b), *no-Classifier, BEI* (5c), and *Classifier, BEI* (5d). In the Classifier conditions (5b & 5d), the classifier (*wei*) globally matches the head noun (*jizhe* "the reporter"), but locally mismatches the embedded noun (*shikuai* "the stone").[5] In our design, only person-denoting classifiers *wei* and *ming* were used, to keep the head nouns animate.

(5) a. **no-CL, no-BEI**
[shikuai zazhong t_i de] jizhe$_i$ aosangde huangu sizhou.
stone hit DE reporter distressfully look-about surroundings
'The reporter that the stone hit _ looked about his surroundings in distress.'

b. **CL, no-BEI**
na-wei [shikuai zazhong t_i de] jizhe$_i$ aosangde huangu sizhou.
that-CL_{human} stone hit DE reporter distressfully look-about surroundings
'The reporter that the stone hit _ cautiously looked about his surroundings in distress.'

c. **no-CL, BEI**
[t_i bei shikuai zazhong de] jizhe$_i$ aosangde huangu sizhou
BEI stone hit DE reporter distressfully look-about surroundings
'The reporter that _ was hit by the stone looked about his surroundings in distress.'

d. **CL, BEI**

na-wei [t_i bei shikuai zazhong de] jizhe$_i$ aosangde huangu sizhou
that-CL_{human} BEI stone hit DE reporter distressfully look-about surroundings
'The reporter that _ was hit by the stone looked about his surroundings in distress.'

A secondary goal of our study is to shed light on potential processing asymmetries between SRCs and ORCs. In our design, the basic word order is the same across conditions, but sentences without BEI contain ORCs, whereas sentences with BEI contain *(passivized)* SRCs.[6] Thus, comparing them can allow us to gain insights into the debate between the processing ease of SRCs and ORCs (e.g., Gibson & Wu, 2013; Hsiao & Gibson, 2003; Jäger et al., 2015; Kuo & Vasishth, 2006; Lin, 2014; Lin & Bever, 2006, 2011; Vasishth et al., 2013).

We first consider the predictions of experience-based theories for our design and then the predictions of the DLT.

3.1. Experience-based theories

To derive the predictions of experience-based theories, we estimated the *structural frequencies* of different conditions from a news corpus and a sentence-completion test. From the Chinese Treebank 5.0 corpus (Palmer, Chiou, Xue, & Lee, 2005), we manually coded the RC-internal verbs and extracted 331 transitive RCs (see Wu, Kaiser, & Andersen, 2011, for details). We found only five tokens of ORCs preceded by classifiers, none having the *mismatch-match configuration* (as in our *CL, no-BEI* condition, ex. (5b)) where a classifier is adjacent to an incongruent RC-subject and distant from a congruent RC-head (Wu, 2011). We found 51 tokens of *passivized* SRCs, accounting for 15% of the total transitive RCs, and 9 tokens of SRCs with both a classifier and a passive marker. Since the target structures are rare in the newswire corpus, it is difficult to estimate precise conditional probabilities at each word across conditions (cf. Jäger et al., 2015; Levy et al., 2013).

Interestingly, among the five tokens of ORCs identified in Wu (2011) that correspond to our example (5b), the local nouns are either dropped or have an additional word intervening between the mismatching classifier and the local noun, most probably to avoid possible incidence of grammatical disharmony[7] triggered by a mismatch between a classifier and a local noun appearing next to each other. This suggests that the local classifier-noun mismatch present in the *CL, no-BEI* condition (5b) might be rare in Chinese comprehenders' experience. Thus, encountering this kind of mismatching combination might induce a "mismatch penalty" or lexical disruption[8] during comprehension.

3.2. Norming study

Given the difficulty of estimating the relevant probabilities based on corpora, we conducted a sentence-completion norming study, to test whether the sentence-initial cues can help native speakers of Chinese to anticipate an RC.

Sentence fragments were truncated versions of the target RC sentences (5a–d), consisting of the words (presented in Chinese characters) prior to the adnominal DE, as in (6). Twenty-four sets in four versions were randomized with 24 filler items of comparable length. Ninety-three college students at Shanghai International Studies University (SISU) took this sentence-completion test for course credit. None of them participated in the self-paced reading experiments.

(6) a. no-CL, no-BEI
shikuai zazhong _____。 (stone hit _____.)

b. CL, no-BEI
nawei shikuai zazhong _____。 (that-CL_{human} stone hit _____.)

c. no-CL, BEI
bei shikuai zazhong _____。 (BEI stone hit _____.)

d. CL, BEI
nawei bei shikuai zazhong _____。(that-CL_{human} BEI stone hit __.)

We obtained a total of 2,206 sentences (26 missing data points: 18 in the *CL, no-BEI* condition, 5 in the *no-CL, BEI* condition, and 3 in the *CL, BEI* condition). Continuations were classified into four structural types:

1. (gapped) RCs as the target structure.
2. Main clauses, where the provided fragments served as a clausal subject (7a), part of a simple sentence with either a dropped subject (7b) or an elided NP (7c), or part of a subordinate clause in a main clause (7d).
3. NP-complement or adjunct RC, where continuations resembled the RC structure, but no gap can be identified within the clause, as in (8a–b).
4. Error continuations, where participants omitted an aspect marker or a comma[9], or overlooked the local classifier-noun incongruity by producing an ungrammatical main clause continuation—mostly in the *CL, no-BEI* condition (9a, cf. a similar error noted by Hsu, 2006), or neglected the compatibility of the classifier and RC head noun by producing an ungrammatical RC—mostly in the *CL, BEI* condition (9b).

(7) a. Clausal-subject continuation in the *CL, no-BEI* condition
naming shuifa chengjie yuan shizong le.
that-CL_{human} tax-law punish staff disappear ASP
'The staff member of tax-law punishment disappeared.

b. Simple sentence with a dropped subject in the *no-CL, BEI* condition
bei liuyan dihui le shengyu.
PASS rumor denigrate ASP reputation
'(someone)'s reputation was denigrated by rumors.

c. Simple sentence with NP-ellipsis in the *CL, BEI* condition
nawei bei dazibao gongji de hen can.
that-CL_{human} PASS big-character-poster attack Degree very badly
'That (person) was attacked by the big-character-poster very badly.'

d. Part of subordinate clause in the *no-CL, no-BEI* condition
shitou zazhong hou, dangchang siwang.
stone hit after, at-the-scene die
'(Someone) died right on the spot after the stone hit (him/her).'

(8) a. (gapless) NP-complement in the *no-CL, BEI* condition
bei bengdai baozha de yangzi hen kexiao
PASS bandage wrap DE look very funny
'The way of being wrapped up by bandage looks very funny.'

b. (gapless) Adjunct RC in the *no-CL, no-BEI* condition
wangzhan xuanchuan de fangshi shi duozhongduoyang de.
internet advocate DE method IS various SFP
'The methods with which the internet advocates are varied.

(9) a. Error continuation in the *CL, no-BEI* condition
* *naming midian gaofa le jianxi*.
that-CL_{human} cipher-telegraph inform ASP traitor

b. Error continuation in the *CL, BEI* condition
* *nawei bei shuicao chanrao de yu zhongyu baituo le*
that-CL_{human} PASS weed entwine REL fish finally break-away ASP
shufu, you huode le ziyou.
control, again gain ASP freedom

Table 1 shows the number and percentage of each continuation type for the four types of fragments. We focus on the completion rate of the target RC structure. Of the total 558 completions in the *no-CL, no-BEI* condition (5a), only 22 (3.94%) were RCs, whereas 552 (98.92%) were main clauses. In contrast, RCs were the most likely structure that participants produced in the other three conditions: Of the 540 total completions in the *CL, no-BEI* condition, 347 (64.26%) were RCs; of the 553 total completions in the *no-CL, BEI* condition, 321 (58.05%) were RCs; and of the 550 total completions in the *CL, BEI* condition, 526 (94.77%) were RCs. A generalized linear mixed model with a binomial link function and crossed varying intercepts for subjects and items shows main effects of CL ($z = 20$) and BEI ($z = 19$) and a CL × BEI interaction ($z = -4$); varying slopes were not fit as the model failed to converge. This suggests that prior to DE, two cues in the *CL, BEI* condition will unambiguously and strongly activate the RC structure; one cue will also highly activate the RC structure: the presence of a mismatching classifier alone (*CL, no-BEI*) triggered substantially more RC continuations than the bare condition (64.26% vs. 3.94%), and so did the presence of a passive marker BEI alone (*no-CL, BEI*: 58.05% vs. 3.94%). However, with only one cue, participants also produced 26.11% error structures in the *CL, no-BEI* condition, and 35.26% main clause structures in the *no-CL, BEI* condition. It is worth noting that the errors in the *CL, no-BEI* condition are overwhelmingly cases where comprehenders ignored the classifier mismatch and made an ungrammatical main clause continuation. This suggests that the comprehender might not be able to *fully* use the Classifier alone as the categorical cue to RC structure that a "competence grammar" of Mandarin would license.

Both the corpus data and sentence completion results presented above allow us to derive two predictions of experience-based accounts. The first relates to the *local disruption* that the classifier causes due to a mismatch with the RC-internal noun (e.g., "stone" in ex. (5b)). Given that a semantically congruent classifier-noun sequence forming a determiner phrase (DP) is very common in both written and spoken Chinese, and that our corpus data show the local classifier-noun mismatch configuration in ORCs is virtually non-existent (Wu, 2011), we suggest that Chinese comprehenders might experience rather high surprisal[10] upon encountering the RC-internal noun in (5b), and this effect might spill over to the following regions given the nature of self-paced reading paradigm. This

Table 1
Number and percentage of each continuation for the four types of fragments in Experiment 1

	Relative Clause		Main Clause		Gapless		Error		Sum
Condition	#	%	#	%	#	%	#	%	#
No-CL, no-BEI	22	3.94	522	93.55	8	1.43	6	1.08	558
CL, no-BEI	347	64.26	41	7.59	11	2.04	141	26.11	540
No-CL, BEI	321	58.05	195	35.26	11	1.99	26	4.70	553
CL, BEI	526	94.77	13	1.44	0	0	16	2.88	555

prediction is further supported by existing processing work showing that a mismatching classifier in a canonical ORC could induce rather severe lexical disruption, particularly when the RC modifies the sentential subject (Hsu, 2006). Given that our experimental stimuli all contain subject-modifying RCs, the experience-based theories would predict that the *CL, no-BEI condition* (5b) might incur processing difficulty at the RC-internal noun, compared with the bare *no-CL, no-BEI* condition (5a), which would yield little parsing difficulty with the initial character strings *NV*... conforming to the canonical SVO word order in Chinese.

The second prediction relates to the anticipatory structure built at the relativizer DE and possibly subsequent spillover regions as a result of encountering the classifier and/or the passive cue BEI. The sentence completion data suggest that the conditional probability for Chinese comprehenders to anticipate an RC in these regions is highest in the presence of two cues (mismatching classifier and BEI), lowest in the condition with bare nouns, and at an intermediate level when only one cue is present. Thus, experience-based theories would predict that reading times at DE and/or the disambiguating head noun (or possibly in the combined region of DE + HeadNoun) and beyond should be fastest in the *CL, BEI* condition (5d) and slowest in the *no-CL, no-BEI* condition (5a). The other two conditions (5b–c) are predicted to fall somewhere in between.

3.3. DLT

Table 2 shows the specific predictions of the DLT at each word, with words aligned by part of speech.

3.4. Storage-cost metric

Consider the no-BEI conditions (i.e., ORCs, 5a-b) first. Assuming that the ultimate structure to be constructed is an RC (given the mismatching classifier cue, which, syntactically speaking, *unambiguously* marks the clausal boundary), then the presence of a classifier (5b) necessarily initiates the projection of a semantically congruent noun which is the head of the RC, whereas in the bare ORC (5a), no projection of a head NP is necessary until after the presence of DE.[11] Thus, the storage cost metric predicts that at the RC-internal noun and verb regions, the classifier-present condition (5b) is more difficult than the classifier-absent condition (5a).

Turning now to the BEI conditions (i.e., passive SRCs, 5c–d), assuming that the ultimate structure to be constructed is an RC (given the sentence-initial BEI cue alone or in conjunction with the mismatching classifier cue), the storage cost metric predicts no processing differences, regardless of whether there is one or two cues. (See Introduction regarding the minimal syntactic projections required by the passive marker BEI.)

As mentioned earlier, in (5c) where no classifier is present, a simple passive matrix clause with a null patient is also possible at the instrument noun following BEI, as in *(pro) bei shikuai zazhong* "(pro) was hit by the stone."[12] But this parse has to be revised at the next available adnominal DE and/or the head noun, which indicates that the correct

Table 2
Illustration of storage-based DLT's word-by-word cost profile

No-CL, no-BEI (object-extracted RC)

		[shikuai stone	zazhong t_i] hit	de DE	jizhe$_i$ ... reportert$_i$
Storage cost	Syntactic heads needed	V	N	N,V	V
	Memory Unit	1	1	2	1
Integration cost/attachment	New discourse	1	1	0	1
	Filler-gap	0	0	0	0
	Classifier-noun	0	0	0	0
	Energy Unit	1	1	0	0

CL, no-BEI (object-extracted RC)

		na-wei that-CL	[shikuai stone	zazhong t_i] hit	de DE	jizhe$_i$ ... reporter$_i$
Storage cost	Syntactic heads needed	N,V	V,N,DE,V	V,DE,N	N,V	V
	Memory Unit	2	4	3	2	1
Integration cost/attachment	New discourse	0	1	1	0	1
	Filler-gap	0	0	0	0	0
	Classifier-noun	0	0	0	0	2
	Energy Unit	0	1	1	0	2

No-CL, BEI (subject-extracted RC)

		[t_i BEI PASS	shikuai stone	zazhong] hit	de DE	jizhe$_i$ ... reporter$_i$
Storage cost	Syntactic heads needed	N,V,N (DE,V)	V,N(DE,V)	N(V,DE)	N,V	V
	Memory Unit	3(5)	2(4)	1(3)	2	1
Integration cost/attachment	New discourse	0	1	1	0	1
	Filler-gap	0	0	0	0	2
	Classifier-noun	0	0	0	0	0
	Energy Unit	0	1	1	0	2

CL, BEI (subject-extracted RC)

		na-wei [t_i that-CL	BEI PASS	shikuai stone	zazhong] hit	de DE	jizhe$_i$ ... reporter$_i$
Storage cost	Syntactic heads needed	N,V	N,V,N,DE,V	V,N,DE,V	V,DE,N	N,V	V
	Memory Unit	2	**5**	4	3	2	1
Integration cost/attachment	New discourse	0	0	1	1	0	1
	Filler-gap	0	0	0	0	0	2
	Classifier-noun	0	0	0	0	0	2
	Energy Unit	0	0	1	1	0	**4**

Note. For the *no-CL, BEI* condition, we list in regions prior to DE the minimal number of heads required for a passive matrix clause and put *in parentheses* those heads projected for the RC analysis.

parse is an RC. Given that the use of null arguments is necessarily licensed by particular properties of preceding discourse (Huang, 1982; Li & Thompson, 1981), but no discourse contexts are provided in our experiments, the RC analysis is highly probable and is more plausible than a passive matrix clause, as confirmed by our norming data. Thus, it will be assumed here that the same number of heads is projected in (5c) as in (5d). In Table 2 for (5c), we list those heads projected for the RC analysis and put *in parentheses* the minimal (number of) heads required for a passive matrix clause prior to DE.

In sum, the storage cost metric predicts that at the RC-internal noun (*shikuai* "stone") and RC-verb regions (*zazhong* "hit"), the *no-classifier, no-BEI* condition (5a) should be easier than the other three conditions. Beginning from the adnominal DE and/or at the head noun where the ultimate parse is evidently an RC, there should be no parsing differences among the four conditions.

3.5. Integration cost metric

According to Gibson (1998, p. 8, p. 14; 2000, p. 102), integration costs are incremented by the number of new discourse referents (i.e., nouns and verb) that intervene between the head and its dependent. Applied to our experimental stimuli, we can identify two types of dependencies/attachments: (i) filler-gap dependencies; (ii) classifier-noun dependencies. In the case of filler-gap dependencies, two discourse referents (i.e., "hit" and "reporter") are intervening in the BEI sentences (5c–d), but not in the no-BEI sentences (5a–b), and the filler-gap dependency is linearly longer in the BEI sentences than the no-BEI sentences. In the case of classifier-noun dependencies, two discourse referents are intervening in the classifier-present conditions (5b, 5d), but not in the classifier-absent conditions (5a, 5c). Taken together, the integration cost at the head noun will be highest in the *CL, BEI* condition (5d), lowest in the *no-CL, no-BEI* condition (5a), and intermediate in conditions with either a classifier (5b) or a BEI (5c).

We summarize the predictions of the two theories in Table 3.

To assess the contrasting predictions of the DLT and experience-based theories, we report two self-paced reading experiments that explored the processing consequences of one and/or two sentence-initial cues in Chinese RCs. To anticipate our findings, the results show that, contrary to what the DLT would predict, the presence of additional cues actually helps to pre-build RC structures, thus supporting experience-based theories.

Table 3
Summary of predictions across conditions at different loci by DLT's storage-cost and integration cost metrics and by experience-based accounts ("<"means faster,">" means slower)

Theory		Prediction	Region
Experience-based accounts	Local disruption	*No-CL, no-BEI* < *CL, no-BEI*	Before DE
	Cue facilitation	*No-CL, no-BEI* > *CL, no-BEI*; *no-CL, BEI* > *CL, BEI*	DE and beyond
DLT	Storage cost	*No-CL, no-BEI* < *CL, no-BEI*; *no-CL, BEI*; *CL, BEI*	Within RC
	Integration cost	*No-CL, no-BEI* < *CL, no-BEI*; *no-CL, BEI* < *CL, BEI*	Head noun

4. Experiment 1

4.1. Methods

4.1.1. Participants

Sixty native speakers of Mandarin Chinese from Fudan University participated in this experiment in exchange for Chinese RMB 15. Their mean age was 23 years.

4.1.2. Materials and design

The experiment manipulated (i) Classifier (absent vs. present) and (ii) BEI (absent vs. present), yielding four conditions as in (5). Our critical stimuli were adapted from the third experiment of Wu et al. (2012), where word frequency, the plausibility of the event, and the likelihood of the event (for reversed animacy configurations) were all controlled. The experiment contained 24 target items.

In addition, 48 *filler items* were constructed. Half of those filler items superficially resembled object-extracted RCs up until the *NV*. . . part: 20 were *gapless* adjunct clauses expressing the reason, the manner, or the instrument of an action/event, and four were simple SVO sentences. In the other half of the filler items, eight were attributive clauses or prodropped possessives, both superficially resembling subject-extracted RCs up until the *VO*. . . part, four were subject-extracted RCs, and 12 were BA subject-extracted RCs (see ex. 10)[13] which were intended to counter-balance the structural prominence of the passive marker BEI in the critical stimuli. Furthermore, half of the fillers contained a demonstrative-classifier sequence, corresponding to the occurrence of such a sequence in half of the target items. All experimental stimuli are provided in Supplementary Material File I, and filler items in Supplementary Material File II.

(10) Example of a filler item: Subject-extracted RC with BA

t_i ba gouhuo shengqilai de na-ge nansheng$_i$ daizhe yi-ding bangqiu mao
BA bonfire raise-up DE that-CL boy-student wear-ASP one-CL baseball hat
'The boy who __ lit the bonfire was wearing a baseball hat.'

4.1.3. Procedure

A word-by-word, moving-window self-paced reading experiment was run on a PC laptop using Linger software developed by Doug Rohde. Participants read the sentences at their own speed, and then answered a yes/no comprehension question by pushing the F/J key. The questions asked about different parts of the sentences, half having "yes" answers and the other half "no." No feedback was provided unless the questions were incorrectly answered, in which case the computer flashed "You are wrong" in Chinese.

4.1.4. Data analysis

We aligned the data of the four conditions by regions after classifiers and BEI, focusing on seven positions (i.e., four critical RC regions and three post-head spillover regions). Separate linear mixed-effects models were fitted to each position after BEI, using the lme4 package in R (version 3.3; CRAN project; the R Foundation for Statistical Computing, 2011). The analyses of reading times were carried out on reciprocal-transformed values in order to stabilize variance and to achieve approximately normal residuals (Box & Cox, 1964). For each region, varying intercepts and varying slopes were fit (without intercept-slope correlation parameters) for the random effects of participants and items. When a variance component was zero, we removed this from the model. We used sum-contrasts coding to test for main effects of Classifier (classifier coded as 1 and no-classifier as -1) and BEI (BEI coded as 1 and no-BEI as -1). In cases where interactions were detected, we further defined two sets of sum contrasts nested within classifiers (coded as a main effect) and BEI (coded as a main effect) for pairwise comparisons. Residuals of linear mixed models were always checked to ensure that there were no serious deviations from the normality assumption. We took an absolute t-value equal to or above 2 to reach statistical significance at $\alpha = 0.05$.

4.2. Results

4.2.1. Comprehension question accuracy

The mean comprehension accuracy was high overall: 95% for target trials and 96% for fillers. On the target trials, a generalized linear mixed model with a binomial link function and crossed varying intercepts and slopes for subjects and items shows no main effect of CL ($z = 1.15$) and no interactions ($z = -0.53$), but only a main effect of BEI ($z = 3.63$): the accuracy rates of the BEI conditions (*no-CL, BEI*: 98.3%; *CL, BEI*: 97.5%) were higher than the no-BEI conditions (*no-CL, no-BEI*: 91.7%; *CL, no-BEI*: 90.6%).

4.2.2. Word-by-word reading times

Fig. 1 shows the means and 95% confidence intervals of the main effects and interaction in each region of interest, along with raw reading times by region. Table 4 reports the statistical results for each region.

At the RC-internal noun (“stone”), and the RC-verb (“hit”), we found no main effects or interactions.

To further explore the lexical-disruption effects predicted by experience-based theories, we ran additional statistical analyses testing RTs of the other three conditions against that of the *Cl, no-BEI* condition. We redefined contrasts with the *Cl, no-BEI* condition (5b) as the baseline, and we fitted non-correlation maximal model at the RC-internal NP and the RC-verb. No significant results were found in these two regions, although the *CL, no-BEI* condition was read numerically slower against the other three conditions.

At the relativizer DE, there was a main effect of Classifier, a main effect of BEI, and an interaction. Follow-up tests using nested contrasts showed that the *CL, no-BEI* condition (5b) was read slowest. Although the presence of a mismatching classifier in (5b) led

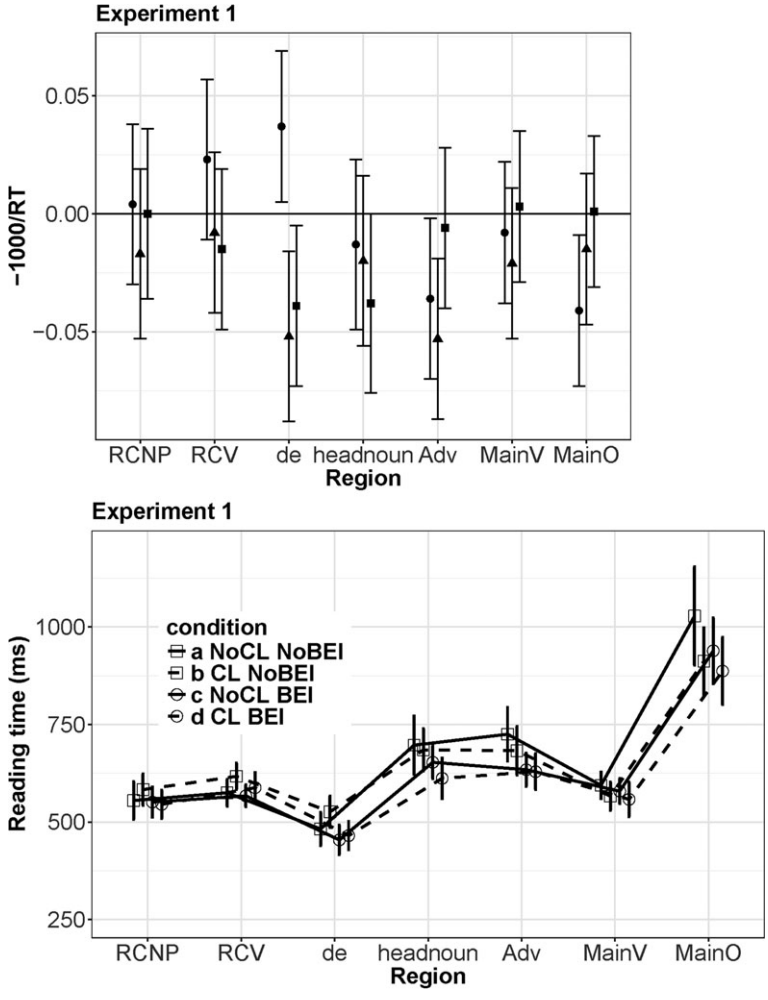

Fig. 1. Summary of Experiment 1 results: shown are the estimates of the main effects and interaction by region (upper plot), with 95% confidence intervals, and mean reading times by region, with 95% confidence intervals (lower plot).

to significantly slower RTs compared with the *no-CL, no-BEI* ($b = 0.075$, $SE = 0.023$, $t = 3.34$), this processing disadvantage of Classifier was blocked by the additional presence of BEI, as shown by the significant faster RTs in the *CL, BEI* condition than the *CL, no-BEI* conditions ($b = -0.045$, $SE = 0.013$, $t = -3.31$).

At the RC-head ("reporter"), there was a marginal interaction[14] ($t = 1.99$), but no main effect of Classifier or main effect of BEI.

Since Gibson and Wu (2013) analyzed the DE and head noun region by combining them, we also did so in order to compare our results with theirs. For the combined DE+head noun region analysis, we found a main effect of BEI and an interaction, but no main effect of Classifier. Follow-up tests show that without BEI, the presence of CL led

Table 4
Main effects of classifier and BEI and their interaction by region of interest in Experiment 1. The dependent variable is reciprocal-transformed reading time

Region	Contrast	Coef.	*SE*	*t*-value
RC-internal noun	CL	0.003	0.017	0.23
	BEI	−0.017	0.018	−0.96
	CL × BEI	0.0002	0.018	0.01
RC-internal V	CL	0.023	0.017	1.35
	BEI	−0.008	0.017	−0.45
	CL × BEI	−0.015	0.017	−0.88
DE	**CL**	0.0366	0.016	**2.29***
	BEI	−0.052	0.016	**−3.28***
	CL × BEI	−0.039	0.017	**−2.28***
Head noun	CL	−0.013	0.018	−0.73
	BEI	−0.020	0.018	−1.08
	CL × BEI	−0.038	0.019	−1.99
Adverb	**CL**	−0.036	0.017	**−2.13***
	BEI	−0.053	0.017	**−3.18***
	CL × BEI	−0.006	0.017	−0.34
Main verb	CL	−0.002	0.016	−0.11
	BEI	−0.023	0.016	−1.48
	CL × BEI	0.006	0.016	0.35
Main object	**CL**	−0.041	0.016	**−2.52***
	BEI	−0.015	0.016	−0.94
	CL × BEI	0.001	0.016	0.07

Bold *t*-values marked with * mean the results reach statistical significance.

to a processing disadvantage, with the *CL, no-BEI* condition being read slower than the *no-CL, no-BEI* condition ($b = 0.026$, $SE = 0.011$, $t = 2.32$), but with an additional BEI, the presence of CL is clearly advantageous, with the *CL, BEI* condition being read faster than the *CL, no-BEI* condition ($b = -0.04$, $SE = 0.012$, $t = -3.25$).

At the adverb ("distressfully"), there was a main effect of BEI: Conditions with BEI were read faster than conditions without BEI. There was a main effect of classifier, reflecting its anticipatory effect: Conditions with classifiers were read faster than conditions without classifiers. We found no interaction.

At the matrix verb ("looked about"), we found no effects.

At the matrix object ("surroundings"), there was a main effect of classifier: Conditions with classifiers were read faster than conditions without classifiers. No other effects were found.

4.3. Discussion

The results of Experiment 1 confirm that comprehenders can use pre-RC cues to anticipate an upcoming RC structure, and that the presence of two pre-RC cues activates the RC parse more strongly than one pre-RC cue. Our results also indicate that while either

BEI or a classifier alone clearly facilitates processing, the cueing effect of BEI occurs earlier and is stronger than that of a classifier. These findings are largely consistent with experience-based theories, but they are not predicted by the DLT.

Crucially within RCs, prior to the relativizer DE, the *no-CL, no-BEI* condition (i.e., with no pre-RC cues) was *not* processed fastest—in fact, we found no significant effects whatsoever at the RC-internal noun and the RC-verb. Nor was the *CL, BEI* condition (i.e., with two pre-RC cues) processed slowest—rather, it is the *CL, no-BEI* condition (i.e., with the classifier cue) that was processed slowest at the relativizer DE and in the combined region of DE+head noun. Furthermore, beginning from the relativizer DE and continuing to the end of the sentence, there were processing differences among the four conditions. All these results are contrary to what the *storage cost metric* would predict.

In addition, the *integration cost metric* predicts that at the head noun, the integration cost should be lowest in the *no-CL, no-BEI* condition, highest in the *CL, BEI* condition, and intermediate in the *CL, no-BEI,* and *no-CL, BEI* conditions. But our data do not show any evidence for this.

Our data are largely consistent with the two predictions of experience-based theories as summarized in Table 3. Regarding the prediction of lexical disruptions induced by the local classifier-noun incongruity, we found that the *Classifier, no-BEI* condition was read slowest within the RC region: This effect is numerically so at the RC-internal NP and RC-verb, and only becomes significant in the spillover region of DE. Note that unlike Hsu (2006) where the lexical-disruption effect occurred immediately at the local RC subject, in our study the effect is delayed. This might in part be due to the different experimental design in Hsu (2006), where comparing the classifier-match to classifier-mismatch conditions might have boosted lexical-disruption effects.

Regarding the second prediction, namely the cue-facilitation effects, our reading time data confirmed (i) the strong facilitative effect of two cues within the RC region and (ii) the facilitative effect of one pre-RC cue in the main clause. Specifically, at the relativizer DE and in the combined DE+head noun region, the *CL, BEI* condition (i.e., two cues) was read faster than the *CL, no-BEI* condition (i.e., one CL cue). This consistent effect suggests that two cues are more effective than a mismatching classifier cue in pre-building RCs unambiguously, replicating Wu et al. (2014a).

Our data beyond the RC region also show that Classifier and BEI can each alone serve as a predictor for the RC structure, although the timing for their *individual* contribution to facilitative processing of sentences—reflected by the main effect of Classifier and the main effect of BEI, respectively—occurs rather late, specifically at the adverb and the main object in the matrix clause. The delayed effect of mismatching classifier echoes the findings in Wu et al. (2014a), suggesting that despite an initial processing disadvantage, the classifier cue can lead to a facilitation.

Overall, Experiment 1 shows (i) local disruption effect at the relativizer DE, and (ii) facilitative effects of pre-RC cues, consistent with the predictions of experience-based theories.

While we find evidence against the storage- and integration-based metrics of DLT, one might argue that the facilitative effects of BEI in parsing RC sentences might be due to the relatively high frequency of the subject-extraction sentences that participants

encountered in Experiment 1. Specifically, there are both filler items (12 BA sentences) and target items (12 BEI sentences) involving subject extraction where the head nouns are base-generated within the RC. This means at least one third (24/72) of our experimental sentences are SRCs, in contrast to 12 ORCs. Thus, more occurrences of SRC in the stimuli might result in structural priming effects, which could facilitate processing of the BEI sentences (i.e., SRCs), but not the no-BEI sentences (i.e., ORCs). To eliminate this potential confound, we conducted a follow-up experiment.

5. Experiment 2

The purpose of Experiment 2 is to replicate the findings of Experiment 1, using revised filler items such that the frequency of SRCs is balanced.

5.1. Methods

5.1.1. Participants

Fifty-eight native speakers of Mandarin Chinese from Fudan University participated in Experiment 2 in exchange for RMB 20 yuan. Their mean age was 21. None of them had participated in Experiment 1 or the norming study.

5.1.2. Materials and design

The critical stimuli were exactly the same as in Experiment 1. The only difference lies in the filler items: The original 12 BA RCs were replaced by non-RC sentences beginning with *VN...*, specifically, eight complex sentences starting with a subordinate clause and four noun-complement structures. In addition, the original four SRC filler sentences were replaced by simple clauses with a passive marker BEI. Thus, out of 48 filler items, 20 were noun-complement structures or complex sentences beginning with *VN...*, 20 were adjunct RCs beginning with *NV...*, and 8 were simple clauses. Note that different functions of the adnominal DE were used in filler items. See Supplementary Material File III for the 12 new filler items.

5.1.3. Procedure

The same procedure was used as in Experiment 1.

5.1.4. Results

5.1.4.1. Comprehension question accuracy: The mean comprehension accuracy was high overall: 93% for target trials and 95% for fillers. On the target trials, a generalized linear mixed model with a binomial link function and crossed varying intercepts and slopes for subjects and items shows no main effect of CL ($z = -0.86$) and no interactions ($z = 1.21$), but only a main effect of BEI ($z = -2.495$): The accuracy rates of the BEI conditions (*no-CL, BEI*: 90.52%; *CL, BEI*: 92.82%) were lower than the no-BEI conditions (*no-CL, no-BEI*: 94.83%; *CL, no-BEI*: 94.25%).

5.1.4.2. Word-by-word reading times: We used reciprocal reading times for statistical analyses, as in Experiment 1. Eight reading-time values lower than 150 ms were removed, as they skewed the residuals even with the transformed data; this affected 0.07% of the data. Fig. 2 presents the main effects and interactions across conditions for each region as well as raw reading times by region. Table 5 reports the statistical results for each region.

At the RC-internal noun ("stone"), we found no main effects or interactions.

At the RC-verb ("hit"), we found a main effect of BEI, but no main effect of Classifier and no interactions. The BEI conditions were read faster than the no-BEI conditions.

To further explore the validity of the lexical-disruption prediction prior to the relativizer DE, we also conducted two additional analyses by redefining the *CL, no-BEI*

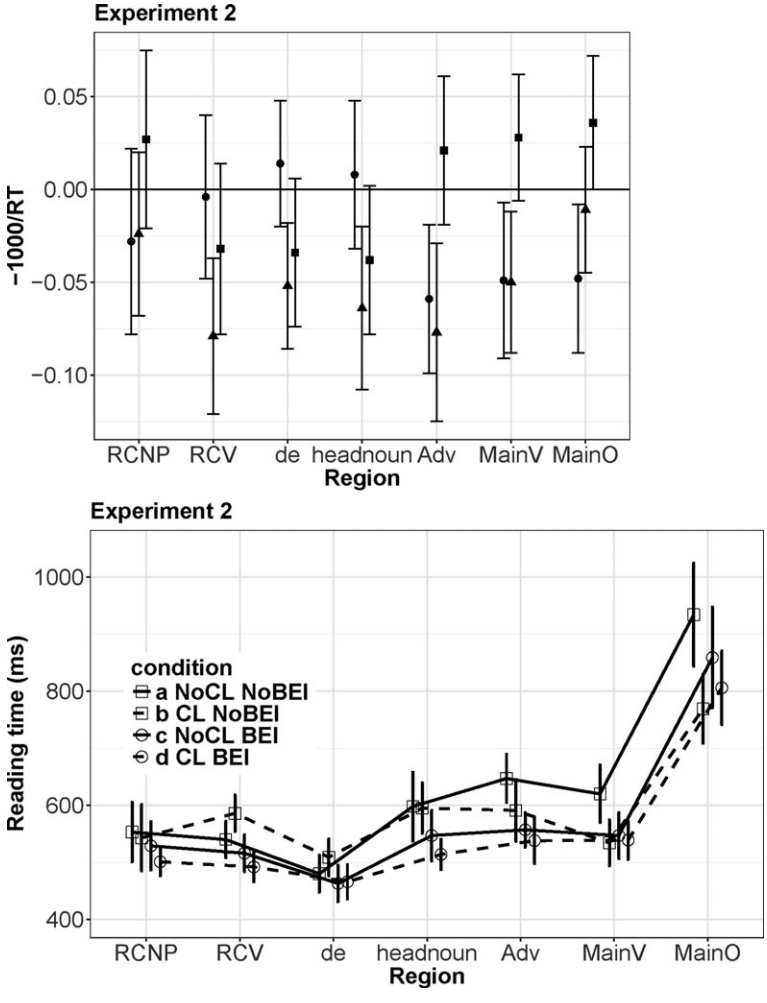

Fig. 2. Summary of Experiment 2 results: shown are the estimates of the main effects and interaction by region (upper plot), with 95% confidence intervals, and mean reading times by region, with 95% confidence intervals (lower plot).

Table 5
Main effects of classifier and BEI and their interaction by region of interest in Experiment 2. The dependent variable is reciprocal-transformed reading time

Region	Contrast	Coef.	*SE*	*t*-value
RC-internal noun	CL	−0.028	0.025	−1.10
	BEI	−0.024	0.022	−1.10
	CL × BEI	0.027	0.024	1.11
RC-internal V	CL	−0.004	0.022	−0.19
	BEI	−0.079	0.021	**−3.75***
	CL x BEI	−0.032	0.023	−1.37
DE	CL	0.0136	0.017	0.78
	BEI	−0.052	0.017	**−2.98***
	CL × BEI	−0.034	0.020	−1.70
Head noun	CL	0.008	0.020	0.39
	BEI	−0.064	0.022	**−2.95***
	CL × BEI	−0.038	0.020	−1.91
Adverb	**CL**	−0.059	0.020	**−2.92***
	BEI	−0.077	0.024	**−3.23***
	CL × BEI	0.021	0.020	1.06
Main verb	**CL**	−0.049	0.021	**−2.32***
	BEI	−0.050	0.019	**−2.63***
	CL × BEI	0.028	0.017	1.66
Main object	**CL**	−0.048	0.020	**−2.40***
	BEI	−0.011	0.017	−0.63
	CL × BEI	0.036	0.018	1.97

condition as the baseline. At the RC-internal noun, the *no-CL, no-BEI* condition was marginally slower than the *CL, no-BEI* condition ($b = 0.11$, $SE = 0.06$, $t = 1.90$). At the RC-verb, no difference was found between the bare *no-CL, no-BEI* condition and the *CL, no-BEI* condition ($t = -0.98$), but the *CL, no-BEI* condition was significantly slower than the *no-CL, BEI* condition ($b = -0.15$, $SE = 0.06$, $t = -2.52$) and the *CL, BEI* condition ($b = -0.22$, $SE = 0.06$, $t = -3.95$).

At the relativizer DE, we found a main effect of BEI: The BEI conditions were read faster than the no-BEI conditions. There were no main effect of Classifier and no interactions.

At the head noun ("reporter"), we found a main effect of BEI-facilitation, a marginal interaction[15] ($t = -1.91$), but no main effect of Classifier.

When combining DE and the head noun into one region, we found a main effect of BEI and an interaction, but no main effect of Classifier. Follow-up tests show that, similar to Experiment 1, the *CL, no-BEI* condition was slower than the *no-CL, no-BEI* condition ($b = 0.026$, $SE = 0.013$, $t = 2.02$), indicating lexical-disruption effects. However, with the additional BEI cue, the *CL, BEI* condition was faster than the *CL, no-BEI* condition ($b = -0.05$, $SE = 0.015$, $t = -3.47$).

At the adverb ("distressfully") and the matrix verb ("looked about"), we consistently found a main effect of classifier: Classifier-present conditions were read faster than classifier-absent conditions. We also found a main effect of BEI: Conditions with BEI were read faster than conditions without BEI. We found no interactions.

At the main object ("surroundings"), we found a main effect of classifier-facilitation, a marginal interaction[16] (t = 1.97), but no main effect of BEI.

5.1.5. Discussion

Experiment 2 replicates the essential findings of Experiment 1. Again, no lexical-disruption effects were found at the RC-internal noun but occurred in the spillover DE+head noun region. Facilitatory effects of BEI were now found much earlier within RCs and continued until the main clause. Although the *CL, no-BEI* condition showed an initial processing disadvantage at the RC-verb and in the combined region of DE+head noun, the pattern of reading times reversed beginning at the adverb and continuing to the end of the sentence, indicative of late facilitatory effects of classifier. Thus, even when we remove the possibility of priming of SRC structures, which was a potential concern in Experiment 1, we nevertheless find the same basic pattern of results.

In addition, Experiment 2 yields two additional patterns that further corroborate our claims:

First, instead of null effects in Experiment 1 at the RC-verb ("hit"), we now find that the earliest effect of BEI-facilitation already surfaced at the RC-verb and remained significant until the main verb. This suggests that—corresponding to our sentence-completion data—the presence of BEI can immediately exert its cueing effect and strongly activate RC expectations. We will come back to this early effect of BEI in the General Discussion.

Second, at the main verb we see facilitatory (main) effects of both classifier and BEI, whereas no effects reached significance in Experiment 1 in this region.

6. General discussion

In this paper, we set out to compare and test the predictions made by the Dependency Locality Theory (DLT) and experience-based theories, using prenominal relative clauses in Chinese. We conducted two self-paced reading experiments using identical critical stimuli, in which incomplete dependents that are yet-to-be integrated into their heads over a long distance (including filler-gap dependencies and classifier-noun attachment) were created at the left edge of RCs. We manipulated the presence/absence of pre-RC classifiers and the presence/absence of the passive marker BEI in object-extracted RCs. This allowed us to test both ORCs (without BEI) and the corresponding *passivized* SRCs (with BEI). We obtained converging evidence showing that Chinese comprehenders were sensitive to the availability of pre-RC cues to predict the head-final RC structure before the head noun was seen. The overall results were contrary to the predictions of the DLT (Gibson, 1998, 2000), but consistent with the predictions of experience-based theories (e.g., Hale, 2001; Levy, 2008; MacDonald, 2013).

The **storage cost metric** predicts that the *no-CL, no-BEI* condition should be easiest to process prior to the relativizer DE. However, this prediction is not borne out by our data. Similarly, the **integration cost metric** predicts that at the head noun, sentences with no cues will be easiest to process, those with two cues most difficult, and those with one cue

somewhere in between. While we indeed found interactions at DE (Experiment 1) and in the combined region of DE+head noun (Experiments 1 & 2), the direction of processing difficulty was different from what was predicted, because it is a processing advantage, instead of a disadvantage, that we found in the *CL, BEI* sentences. Taken together, our results from the two experiments pose challenges to both the storage-cost and integration cost metrics of the DLT.

As a whole, our results fit better with **experience-based theories**, which predict that comprehenders are sensitive to the statistical regularities in lexical and structural input, and can use their prior experience with linguistic signals to predict the upcoming structure. Specifically, given that (i) corpus analyses revealed virtually zero tokens of the locally mismatching, global-matching classifier configuration (Wu, 2011), together with (ii) Chinese speakers' preference for locally congruent classifier-noun constituents to form a determiner phrase (DP) licensed by Chinese grammar (Hsu, 2006; Wu et al., 2014a), experience-based theories predict *lexical-disruption effects* in the *CL, no-BEI* condition induced by the unfamiliar combination of a classifier that mismatches the adjacent RC-subject noun. Furthermore, the sentence completion norming data which we used to generate experience-based predictions showed that the rate of RC continuations is (i) highest in the sentences with two cues (*CL, BEI*), (ii) intermediate in the sentences with one cue (*CL, no-BEI* and *no-CL, BEI*), and (iii) extremely low in the sentences without any cues (*no-CL, no-BEI*). Thus, experience-based theories also predict that cue-facilitation effects vary depending on the availability of the BEI and/or Classifier cues.

The reading-time patterns of two experiments align well with these two predictions: Both experiments show an interaction between BEI and Classifier within the RC region, which modulates a main effect of BEI and/or a main effect of classifier. Specifically at the relativizer DE (Experiment 1) and in the combined region of DE+head noun (Experiments 1 and 2), (i) the *CL, no-BEI* condition was consistently slower than the *no-CL, no-BEI* condition, confirming the lexical-disruption prediction, and (ii) the *CL, BEI* condition was consistently faster than the *CL, no-BEI* condition, supporting the cue-facilitation prediction. In fact, the cue-facilitation effect of BEI even showed up as early as at the RC-verb in Experiment 2. These patterns suggest that despite the initial processing disadvantage, the mismatching classifier cue can nevertheless be facilitative in RC parsing in the presence of the BEI cue; in other words, two cues activate the RC structure more strongly than one Classifier cue.

Furthermore, consistent with the cue-facilitation prediction of experience-based accounts, an early occurrence of a cue, be it a lexical classifier or a syntactic BEI, can guide comprehenders to build the RC structure incrementally, and the syntactic predictability for RCs is proportional to the strength of cues (also see similar findings in Wu et al., 2014a). Although the facilitative effects of BEI were consistently found *early*—either within the RC (in Experiment 2) or modulated by an interaction (at DE in Experiment 1 and in the combined region of DE+head noun in Experiments 1 and 2)—the facilitative effects of classifiers were *delayed*, reaching significance only in the main clause regions. As predicted by experience-based accounts, this is presumably due to the lexical disruptions associated with the classifier cue that mismatches with the subsequent noun and the structural rarity of classifier-noun incongruity (Wu, 2011). The presence of

BEI, however, removes the lexical disruption of the *CL, no-BEI* condition, allowing the main effect of BEI to emerge early.

Together, these results support experience-based theories. An increased number of syntactic heads does not necessarily result in a filler-gap dependency or a classifier-noun attachment relationship becoming harder to processes. Instead, once an RC analysis is considered, it gets reinforced over time as more character strings unfold, resulting in subsequent processing ease.

6.1. Early facilitative effect of BEI at RC-internal verb

One related finding worth discussing is the early BEI-facilitative effect found in Experiment 2, starting at the verb inside the RC. The timing of this effect might strike one as odd, because presumably before the relativizer DE, participants (perhaps particularly those in the *no-CL, BEI* condition) might not yet know whether this verb was in an embedded clause or in a main clause. One possibility could be that the facilitatory effect of BEI at RC-verb simply reflects a penalty for inanimate subjects in the no-BEI conditions. However, Experiment 1—with exactly the same animacy configuration—does not show such a penalty effect at the RC-noun or RC-verb, which seems to rule out this explanation.

Another possible explanation is that the early BEI-facilitative effect reflects the processing ease of subject-extracted RCs (BEI conditions) relative to object-extracted RCs (no-BEI conditions). If this account were true, then presumably the BEI-facilitation effect would have been stronger in Experiment 1 due to potential structural priming occurring with SRCs in filler items. But no such effects at the RC-verb were found in Experiment 1. Furthermore, this SRC/ORC-based explanation would also not explain why, in both experiments, a CL×BEI interaction was found at the relativizer DE, where the two object-extracted RC types differed in their processing ease (the *CL, no-BEI* conditions were slower than the *no-CL, no-BEI* conditions). This is unexpected if the BEI-facilitative effect is driven by the prevailing superiority of SRCs to ORCs.

Instead, we would like to argue that the facilitative effect of BEI at the RC-verb is due to a few experience-based factors conspiring together.

First, in Chinese the sequence of BEI+verb is a high-frequency, low-surprisal part-of-speech-bigram.[17] Coupled with the fact that BEI+verb is very near to the left edge of the sentence, participants' reading times would be facilitated when a high-frequency function word or a sequence of words was presented at the beginning of the sentence.

Second, the early effect of BEI at RC-verb might also be boosted by (i) the lexical-disruption effect that spills over into the RC-verb in the *CL, no-BEI* condition, slowing down RTs in this condition, and (ii) the garden-path disambiguation effect that is strongest in the *CL, BEI* condition, speeding up RTs in this condition. In other words, it may be that the presence of BEI removed both the lexical disruption of the *CL, no-BEI* condition and the garden-path effect of the *no-CL, no-BEI* condition, leading to the main effect of BEI in the RC-verb region of Experiment 2.

Third, some idiosyncratic factor might also be at play in the early BEI-facilitative effect. In Experiment 2, out of the 60 items that participants saw, 12 target items

contained a sentence-initial BEI, whereas none of the filler sentences did. This might have reinforced the RC parse for participants reading the *no-CL, BEI* sentences, rendering the activation level of the main clause parse even lower. This suggests that even in the course of an experiment, Chinese comprehenders could learn statistical regularities of linguistic input to adjust their expectations for the ultimate structures.

In sum, the early effect of BEI in Experiment 2 can be largely accounted for by different factors in terms of experience-based theories.

6.2. Subject-extracted vs. object-extracted RCs: Evidence for a probabilistic approach to storage cost?

The presence/absence of BEI is connected to the debate regarding the ease of processing subject-extracted RCs (SRCs) vs. object-extracted RCs (ORCs) in Chinese. In our experiments, lexical variations due to different parts of speech were well controlled, yet we consistently find that BEI sentences (i.e., passive SRCs) were processed faster than no-BEI sentences (i.e., canonical ORCs) both within and beyond the RC. Thus, the main effect of BEI may in part be attributable to the universal subject preference (Keenan & Comrie, 1977). But we believe this cannot be the whole story, because it does not offer an explanation for the facilitative effects in the two (no-BEI) ORC conditions.

Crucially, our finding that BEI sentences are easier to process than no-BEI sentences is contrary to what the classic definitions for the storage and integration/retrieval metrics of the DLT predict, and it also diverges from the results of Hsiao and Gibson (2003) and Gibson and Wu (2013).

When Hsiao and Gibson (2003) derived the word-by-word predictions in their classic paper on the processing asymmetry in Chinese subject-extracted vs. object-extracted RCs, they assumed that the ultimate structure to be built is an RC, and therefore, more heads are projected in SRCs than ORCs prior to the relativizer DE. However, this assumption has recently been challenged by corpus-based investigations. Substantial evidence from structural frequency and the computation of conditional probabilities has shown that SRCs occur more frequently than ORCs in Chinese in both written and spoken corpora (Chen, Grove, & Hale, 2012; Hsiao & Gibson, 2003; Hsiao & MacDonald, 2013; Jäger et al., 2015; Kuo & Vasishth, 2006; Ming & Chen, 2010; Pu, 2007; Sheng & Wu, 2013; Wu et al., 2011 Wu, Kaiser, & Andersen, 2009; Wu et al., 2011), and the surprisal contrast indicates a preference for SRCs over ORCs in Chinese (Chen et al., 2012; Jäger et al., 2015). In addition, given the structural ambiguities inherent in Chinese RCs, it is not clear whether comprehenders construct a target structure in the very beginning. Thus, it is possible that the costs incurred by predicted heads are a function of a speaker's linguistic experience (rather than a fixed amount for a target RC structure yet to be built; see similar points in Hsiao & MacDonald, 2013; Levy & Keller, 2013), and that the structural probabilities constantly vary given the words already seen and incoming words yet to be seen. We, therefore, suggest that given existing empirical evidence and our findings, DLT's storage-based metric could perhaps be expanded to incorporate probabilistic parsing and comprehenders' linguistic experience.

6.3. Alternative accounts

Finally, we discuss alternative interpretations of the lack of support for the DLT. DLT has some degree of freedom as to exactly which syntactic heads count toward storage and integration costs. Although we count both filler-gap and classifier-noun dependencies for DLT's integration cost, it is not clear whether classifier-noun integrations should count. An alternative explanation could be that storage and/or filler-gap integration incurs costs, but these costs are outweighed by the expectation-derived benefits.[17] It is worth noting that recent research indicates that DLT's locality effects can override expectation effects under conditions of high memory load, a key example being relative clauses with nested structures (Demberg & Keller, 2008; Levy & Keller, 2013; Vasishth & Lewis, 2006). Boston, Hale, Vasishth, and Kliegal (2011) also argue that both surprisal and retrieval cost are independent predictors of reading difficulty (also see Demberg & Keller, 2008; Jaeger et al., 2008; Vasishth & Drenhaus, 2011). In our experiments, we only used simple RCs that modified sentential subjects. It is therefore possible that the memory load is not very high, and that the cost induced by a large number of heads is not high enough to be observable with self-paced reading. Thus, our study can be regarded as a case where locality effects cannot override anticipation effects, because memory load was not unusually high. An interesting direction for future work would be to see whether, under conditions of high memory load (e.g., in a dual-task paradigm), the processing of RCs in Chinese would in fact reveal effects of storage and/or integration costs.

6.4. Conclusion and future directions

Our results from two self-paced reading experiments on Chinese RCs reveal facilitative effects of BEI within and beyond RC regions and relatively delayed facilitative effects of classifiers. These facilitative effects suggest that incomplete heads that occur before a clear signal of an upcoming RC can help Chinese comprehenders to anticipate RC structures—an outcome that is not predicted by the DLT but is consistent with the predictions of experience-based theories.

An open question for future work has to do with animacy. The animacy configuration used in this study is uncommon because (i) the heads of our object-extracted RCs are animate, whereas corpus findings show object-extracted RCs typically have inanimate heads (Hsiao & MacDonald, 2013; Wu et al., 2012), and (ii) our passivized subject-extracted RCs contain an inanimate instrument following BEI, whereas Chinese passive construction typically requires an animate agent. Thus, future work might employ more common animacy configurations to see whether the conclusion still holds.

Acknowledgments

We thank audiences at the 2008 Western Conference on Linguistics, the 83rd Annual meeting of the Linguistics Society of America, the 8th Generative Linguistic to the Old

World in Asia, where earlier versions of some of this research were presented. We also thank Todd Haskell and Elaine Andersen for their suggestions regarding early stages of this research. We thank Yanan Sheng, Yao Chen, and Jun Lyu for assistance in running participants and annotating the norming data, and Lixin Jin for assistance in carrying out norming studies. This research was supported by grants from the New Century Excellent Talents in University (NCET-12-0907) and the National Social Science Foundation of China (13BYY152) to the first author.

Notes

1. Different versions of DLT differ somewhat in how they define the processes involved in sentence comprehension: One variant suggests (i) storage costs and (ii) retrieval/integration costs, without distinguishing retrieval and integration costs (e.g., Gibson & Wu, 2013); another variant suggests (i) storage and retrieval costs and (ii) integration costs (e.g., Levy et al., 2013). Overall, the cognitive resources involved in maintenance, retrieval, and integration of stored-element representations are limited. We focus on the version presented in Gibson (2000).
2. Here the RC-internal noun "stone" has the thematic role of "instrument" rather than "agent," and the underlying structure of this sentence is "[$_{RC}$ (pro) [with] stone hit] reporter," meaning "the reporter whom$_i$ [someone used a stone to hit t_i]." In somewhat simplified version, here we treat the instrument "stone" as an atypical agent taking the grammatical subject position.
3. Given that Chinese is a pro-drop language (Li & Thompson, 1981), it is also possible for a Chinese comprehender to posit a null subject immediately upon encountering the sentence-initial BEI, linking the obligatory patient noun of the BEI construction to this null subject (i.e., "(pro) was hit by a stone"). But without discourse context that licenses an omissible subject, this main clause reading would have to be discarded as more input comes in, specifically, as soon as the adnominal DE is seen.
4. It is possible to have a classifier after BEI, but this continuation is ruled out because it will ultimately result in a double-embedded RC structure, as in *BEI na-wei shikuai zazhong de jizhe dale de ren* "the person who is hit by the reporter who is hit by the stone," which is too complex to be naturally produced or comprehended by Chinese speakers.
5. Here, we used a special "reversed" animacy configuration that is non-canonical for all four conditions. The ORCs (5a–b) in our stimuli were always headed by human referents, against the findings from Chinese corpora showing that head nouns of object-extracted RCs prefer to be inanimate (Hsiao & MacDonald, 2013; Wu, Kaiser, & Andersen, 2012). While head nouns of subject-extracted RCs do not show particular preference for animacy (Wu et al., 2012), the passive structure typically requires a human agent (Li & Thompson, 1981), whereas the *passive* SRCs (5c–d) in our stimuli consistently contained a non-canonical "agent," or more precisely an

"instrument." Thus, we do not think the reversed animacy configuration would put the standard ORCs (5a–b) at a processing disadvantage, because any such processing disadvantage, if present, would be offset by the non-canonical usage of "instrument" in the passive SRCs (5c–d).

6. Note however, there is no consensus in Chinese syntax community about whether BEI in these structures truly creates an SRC. Thus, it is not obvious that the BEI/no-BEI contrast creates an ORC/SRC contrast.
7. This grammatical disharmony is more likely to be semantic than syntactic, because classifiers with *local-mismatch, global-match* configuration can be used by Chinese comprehenders to construct a syntactically sound structure, namely an ORC, as shown by behavioral (Hsu, 2006; Wu et al., 2014a) and neurophysiological studies (Hsu, Tsai, Yang, & Chen, 2014; Wu et al., 2014b).
8. We would like to note that experience-based theories might not be the only type of theory that makes the lexical-disruption prediction. However, what is crucial here is that experience-based theories clearly do make this prediction, whereas the DLT does not.
9. An example of error continuation due to omission of comma is given below. While this might count as a main clause continuation if a more lenient criterion were adopted, doing that would not affect the overall distribution/data patterns.

   *<u>*bei*</u> <u>*shoulei zhashang*</u> *zhanshi tuidao erxian*

   PASS grenade blow-hurt solider retreat-to secondary line
10. Strictly speaking, this prediction does not come directly from our off-line sentence completion data, because we presented both the classifier and RC-internal noun as part of the preamble, and most participants either managed to use the local classifier-noun mismatch to construct the correct RC or failed to detect the mismatch and produced a considerable rate of errors.
11. It is possible that the head noun can be dropped altogether in a headless RC. Given that this possibility applies to all conditions, this option would not change the prediction of the storage-cost metric.
12. As shown by our sentence-completion norming results, the percentage of such continuation is 35.26% (195 tokens), which is less than the RC continuations (321 tokens, 58.05%). While the production rate of matrix clauses appears non-trivial, the DLT—assuming a ranked or fully parallel parser—would predict that the *no-CL, BEI* Condition might be as difficult as, if not more difficult than, the *CL, BEI* Condition. But—preempting our results section—this prediction is again not borne out by our self-paced RT data
13. These 12 BA sentences contain the agent-markers *ba, jiang, and ling*, forming the so-called BA construction in which the noun following those markers is the patient and the noun preceding those markers is the agent.
14. Unpacking the marginal interaction at the head noun by a sliding contrast (Venables & Ripley, 2000) shows that the *CL, BEI* condition (5d) was read fastest.
15. Follow-up tests show that the facilitation of BEI was found only when a classifier cue was present (*CL, BEI* vs. *CL, no-BEI*: $b = -0.10$, $SE = 0.029$, $t = -3.58$),

but not when the classifier was absent (*no-CL, BEI* vs. *no-CL, no-BEI*: $t = -1.13$). While this finding should not to be interpreted without qualification, it is as predicted by experience-based theories, but is inconsistent with the DLT.

16. Follow-up tests showed that the facilitatory effects of classifiers were only found in the no-BEI conditions (the *CL, no-BEI* vs. *no-CL, no-BEI* conditions: $b = -0.084$, $SE = 0.024$, $t = -3.53$), but not in the BEI conditions (the *CL, BEI* vs. *no-CL, BEI* conditions: $t = -0.44$). While we refrain from discussing marginal interactions further, it is worth noting that this finding is not consistent with the DLT, but it is as predicted by experience-based theories.
17. We thank Roger Levy for suggesting this.

References

Bever, T. G. (1970). The cognitive basis for linguistic structures. In J. R. Hayes (Ed.), *Cognition and the development of language* (pp. 279–362). New York: Wiley.

Boston, M., Hale, J., Vasishth, S., & Kliegal, R. (2011). Parallelism and syntactic processes in reading difficulty. *Language and Cognitive Processes*, *26*(3), 301–349. https://doi.org/10.1080/01690965.2010.492228

Box, G., & Cox, D. (1964). An analysis of transformations. *Journal of the Royal Statistical Society. Series B (Methodological)*, *26*(2), 211–252. http://www.jstor.org/stable/2984418

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

Chen, Z., & Vasishth, S. (2010). Locality cost in sentence comprehension: Evidence from Chinese. Poster presented at the 84th Annual Meeting of the Linguistic Society of America (LSA). Baltimore, MD.

Clifton, C., & Frazier, L. (1989). Comprehending sentences with long-distance dependencies. In G. Carlson, & M. Tanenhaus (Eds.), *Linguistic structure in language processing* (pp. 213–217). Amsterdam: Springer.

Demberg, V., & Keller, F. (2008). Data from eye-tracking corpora as evidence for theories of syntactic processing complexity. *Cognition*, *101*, 193–210. https://doi.org/10.1016/j.cognition.2008.07.008

Dryer, M. S. (1992). The Greenbergian word order correlations. *Language*, *68*(1), 81–138. https://doi.org/10.1353/lan.1992.0028

Frazier, L., & d'Arcais, F. (1989). Filler-driven parsing: A study of gap filling in Dutch. *Journal of Memory and Language*, *28*, 331–344.

Gennari, S. P., & MacDonald, M. C. (2008). Semantic indeterminacy in object relative clauses. *Journal of Memory and Language*, *58*(2), 161–187. https://doi.org/10.1016/j.jml.2007.07.004

Gibson, E. (1998). Linguistic complexity: Locality of syntactic dependencies. *Cognition*, *68*(1), 1–76. https://doi.org/10.1016/S0010-0277(98)00034-1

Gibson, E. (2000). The dependency locality theory: A distance-based theory of linguistic complexity. In A. Marantz, Y. Miyashita, & W. O'Neil (Eds.), *Image, language, brain* (pp. 313–353). Cambridge, MA: MIT Press.

Gibson, E., & Wu, I. (2013). Processing Chinese relative clauses in context. *Language and Cognitive Processes*, *28*(1–2), 125–155. https://doi.org/10.1080/01690965.2010.536656

Hale, J. (2001). A probabilistic early parser as a psycholinguistic model. In *Proceedings of the 2nd Conference of the North American Chapter of the Association for Computational Linguistics* (pp. 159–

166). Pittsburgh, PA: Association for Computational Linguistics (Volume 2). doi: 10.3115/1073336.1073357

Hale, J. (2003). The information conveyed by words in sentences. *Journal of Psycholinguistic Research*, *32* (2), 101–123. https://doi.org/10.1023/A:1022492123056

Hale, J. (2006). Uncertainty about the rest of the sentence. *Cognitive science*, *30*(4), 609–642. https://doi.org/10.1207/s15516709cog0000_64

Hawkins, J. A. (2004). *Efficiency and complexity in grammars*. Oxford, UK: Oxford University Press. https://doi.org/10.1093/acprof:oso/9780199252695.001.0001

Hsiao, F., & Gibson, E. (2003). Processing relative clauses in Chinese. *Cognition*, *90*(1), 3–27. https://doi.org/10.1016/S0010-0277(03)00124-0

Hsiao, Y., & MacDonald, M. (2013). Experience and generalization in a connectionist model of Mandarin Chinese relative clause processing. *Frontiers in Psychology: Language Sciences*, *4*, 767. https://doi.org/10.3389.fpsyg.2013.00767

Hsu, C.-C. N. (2006). Issues in head-final relative clauses in Chinese—Derivation, processing and acquisition. Unpublished doctoral dissertation. University of Delaware.

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

Hsu, C.-C. N., Phillips, C., & Yoshida, S. (2005). Cues for head-final relative clauses in Chinese. Poster presented at the 18th Annual CUNY Sentence Processing Conference. Tucson, AZ.

Hsu, C.-C., Tsai, S.-H., Yang, C.-L., & Chen, J.-Y. (2014). Processing classifier-noun agreement in a long distance: An ERP study on Mandarin Chinese. *Brain and Language*, *137*, 14–28. https://doi.org/10.1016/j.bandl.2014.07.002

Huang, C. T. J. (1982). Logical relations in Chinese and the theory of grammar. Unpublished Ph.D. dissertation. MIT.

Huang, C. J., Li, Y.-H. A., & Li, Y. (2009). *The syntax of Chinese*. Cambridge, UK: Cambridge University Press. https://doi.org/10.1017/CBO9781139166935

Husain, S., Vasishth, S., & Srinivasan, N. (2014). Strong expectations cancel locality effects: Evidence from Hindi. *PLoS ONE*, *9*(7), 1–14. https://doi.org/10.1371/journal.pone.0100986

Husain, S., Vasishth, S., & Srinivasan, N. (2015). Integration and prediction difficulty in Hindi sentence comprehension: Evidence from an eye-tracking corpus. *Journal of Eye Movement Research*, *8*(2), 1–12. https://doi.org/10.16910/jemr.8.2.3

Jaeger, T. F., Fedorenko, E., Hofmeister, P., & Gibson, E. (2008). Expectation-based syntactic processing: Anti-locality outside of head-final languages. Oral presentation at the 21st CUNY Sentence Processing Conference. Chapel Hill, NC.

Jäger, L., Chen, Z., Li, Q., Lin, C., & Vasishth, S. (2015). The subject-relative advantage in Chinese: Evidence for expectation-based processing. *Journal of Memory and Language*, *79–80*, 97–120. https://doi.org/10.1016/j.jml.2014.10.005

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. https://doi.org/10.1023/A:1026528912821

Kuo, K., & Vasishth, S. (2006). Processing relative clauses: evidence from Chinese. Unpublished manuscript, University of Potsdam.

Levy, R. (2008). Expectation-based syntactic comprehension. *Cognition*, *106*(3), 1126–1177. https://doi.org/10.1016/j.cognition.2007.05.006

Levy, R., Fedorenko, E., & Gibson, E. (2013). The syntactic complexity of Russian relative clauses. *Journal of Memory and Language*, *69*(4), 461–495. https://doi.org/10.1016/j.jml.2012.10.005

Levy, R., & Keller, F. (2013). Expectation and locality effects in German verb-final structures. *Journal of Memory and Language*, *68*(2), 199–222. https://doi.org/10.1016/j.jml.2012.02.005

Levy, R., & Manning, C. (2003). Is it harder to parse Chinese, or the Chinese Treebank? *Proceedings of the 41st Annual Meeting on Association for Computational Linguistics*, vol. 1, 439–446. https://doi.org/10.3115/1075096.1075152

Li, A. Y.-H. (1998). Argument determiner phrases and number phrases. *Linguistic Inquiry*, *29*(4), 693–702. https://doi.org/10.1162/ling.1998.29.4.693

Li, C., & Thompson, S. (1981). Mandarin Chinese: A functional reference grammar. University of California Press, Berkeley.

Lin, C. (2014). Effect of thematic order on the comprehension of Chinese relative clauses. *Lingua*, *140*, 180–206. https://doi.org/10.1016/j.lingua.2013.12.003.

Lin, C., & Bever, T. (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 Proceedings Project.

Lin, C., & Bever, T. (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. 241–275). Amsterdam: Springer. https://doi.org/10.1007/978-90-481-9213-7_13

Lin, B., & Garnsey, S. M. (2010). 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). Amsterdam: Springer. https://doi.org/10.1007/978-90-481-9213-7_12

MacDonald, M. C. (2013). How language production shapes language form and comprehension. *Frontiers in Psychology*, *4*, 1–16. https://doi.org/10.3389/fpsyg.2013.00226

MacDonald, M. C., & Christiansen, M. H. (2002). Reassessing working memory: Comment on Just and Carpenter (1992) and Waters and Caplan (1996). *Psychological Review*, *109*(1), 35–54. https://doi.org/10.1037//0033-295X.109.1.35

Ming, T., & Chen, L. (2010). A discourse-pragmatic study of the word order variation in Chinese relative clauses. *Journal of pragmatics*, *42*(1), 168–189. https://doi.org/10.1016/j.pragma.2009.05.023

Nakatani, K., & Gibson, E. (2008). Distinguishing theories of syntactic expectation cost in sentence comprehension: Evidence from Japanese. *Linguistics*, *46*, 63–86. https://doi.org/10.1515/LING.2008.003

Palmer, M., Chiou, F.-D., Xue, N., & Lee, T. K. (2005). Chinese Treebank 5.0. Linguistic Data Consortium, Philadelphia.

Pu, M. (2007). The distribution of relative clauses in Chinese discourse. *Discourse Processes*, *43*(1), 25–53. https://doi.org/10.1207/s15326950dp4301_2

R Development Core Team (2011), *R: A Language and Environment for Statistical Computing.*. Vienna, Austria : the R Foundation for Statistical Computing ISBN: 3-900051-07-0. Available online at http://www.R-project.org/.

Safavi, M. S., Husain, S., & Vasishth, S. (2016). Dependency resolution difficulty increases with distance in persian separable complex predicates: Evidence for expectation and memory-based accounts. *Frontiers in psychology*, *7*(403), 1–15. https://doi.org/10.33389/psyg.2016.00403

Sheng, Y., & Wu, F. (2013). Audience design affects classifier positioning in Chinese relative clauses: Evidence from spoken corpus and sentence-production data. In N. Goto, K. Otaki, A. Sato, & K. Takita (Eds.), *Proceedings of GLOW in Asia IX 2012* (pp. 271–282). Mie: Mie University.

Simpson, A. (1999). On the status of modifying DE and the structure of the Chinese DP. In S.-W. Tang & L. L. Chen-Sheng (Eds.), *On the formal way to Chinese languages* (pp. 74–101). Stanford, CA: CSLI.

Staub, A., & Clifton, C. (2006). Syntactic prediction in language comprehension: Evidence from either. . .or. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *32*(2), 425–436. https://doi.org/10.1037/0278-7393.32.2.425

Vasishth, S., Chen, Z., Li, Q., & Kuo, G.-L. (2013). Processing Chinese relative clauses: Evidence for the universal subject preference. *PLoS ONE*, *8*(10), e77006. https://doi.org/10.1371/journal.pone.0077006

Vasishth, S., & Drenhaus, H. (2011). Locality effects in German. *Dialogue and Discourse*, *2*(1), 59–82. https://doi.org/10.5087/dad.2011.104

Vasishth, S., & Lewis, R. L. (2006). Argument-head distance and processing complexity: Explaining both locality and anti-locality effects. *Language*, *82*(4), 767–794. https://doi.org/10.1353/lan.2006.0236

Venables, W. N., & Ripley, B. D. (2000). *Modern applied statistics with S-PLUS*. New York: Springer Science & Business Media.

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), 203–227. https://doi.org/10.1515/cllt.2011.010

Wu, F., Haskell, T., & Andersen, E. (2006). The interaction of lexical, syntactic, and discourse factors in online Chinese parsing: Evidence from eye-tracking. Poster at the 19th Annual CUNY Sentence Processing Conference. New York, NY.

Wu, F., Kaiser, E., & Andersen, E. (2009). The effects of classifiers in predicting Chinese relative clauses. In M. Grosvald & D. Soares (Eds.), *Proceedings of the Western Conference on Linguistics (WECOL)* (pp. 330–339). Retrieved from www.fresnostate.edu/artshum/linguistics/wecolproceedings.html Davis: University of California.

Wu, F., Kaiser, E., & Andersen, E. (2011). 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. 173–193). Amsterdam: Springer. https://doi.org/10.1007/978-90-481-9213-7_9

Wu, F., Kaiser, E., & Andersen, E. (2012). Animacy effects in Chinese relative clause processing. *Language and Cognitive Processes*, *27*(10), 1489–1524. https://doi.org/10.1080/01690965.2011.614423

Wu, F., Luo, Y., & Zhou, X. (2014a). Building Chinese relative clause structures with lexical and syntactic cues: Evidence from visual-world eyetracking and reading times. *Language, Cognition and Neuroscience*, *29*(10), 1205–1226. https://doi.org/10.1080/01690965.2013.841969

Wu, F., Sheng, Y., & Zhou, X. (2014b). Animacy modulates cueing strength of mismatching classifiers in marking clausal boundary: ERP evidence from Chinese relatives. Paper presented at the 20th Annual Conference on Architectures and Mechanisms for Language Processing (AMLaP). September 4–6, 2014. University of Edinburgh, UK.

Supporting Information

Additional Supporting Information may be found online in the supporting information tab for this article:

Supplementary Material File I. The target items of Experiments 1 and 2. Please note that the sentences are glossed with the Chinese structure in mind/to help make the Chinese structure clear, and they are not fully translated into English.

Supplementary Material File II. The filler items of Experiment 1.

Supplementary Material File III. The twelve new filler items in Experiment 2 that replaced the original corresponding fillers in the Experiment 1.