Structural Expectations in Chinese Relative Clause Comprehension

Zhong Chen, Kyle Grove, and John Hale
Cornell University

1. Introduction

Relative clauses (RC) are among the most well-studied constructions in the field of psycholinguistics. A wide variety of work explores a robust processing asymmetry such that subject relatives (SRs) are easier to process than object relatives (ORs). For example, English shows a subject advantage, as demonstrated by a number of studies involving different measures including: self-paced reading (King & Just, 1991), eye-tracking (Traxler, Morris & Seely, 2002), ERP (King & Kutas, 1995), fMRI (Just, Carpenter, Keller, Eddy & Thulborn, 1996); and PET (Stromswold, Caplan, Alpert & Rauch, 1996).

A robust finding in the literature suggests that subject preference seems to be a universal processing phenomenon in RCs. Some pieces of evidence come from Dutch (Frazier, 1987), French (Frauenfelder, Segui & Mehler, 1980), German (Schriefers, Friederici & Kühn, 1995), Japanese (Miyamoto & Nakamura, 2003) and Korean (Kwon, Polinsky & Kluender, 2006).

In order to account for the universal processing pattern of RCs, several theories are proposed, such as: WORD ORDER (Bever, 1970; MacDonald & Christiansen, 2002), the ACCESSIBILITY HIERARCHY (Keenan & Comrie, 1977), EXPERIENCE/FREQUENCY-BASED ACCOUNTS (Mitchell, Cuetos, Corley & Brysbaert, 1995; Hale, 2001), STRUCTURE-BASED APPROACHES (O'Grady, 1997; Hawkins, 2004) and WORKING MEMORY (Gibson, 2000; Lewis & Vasishth, 2005).

Chinese RCs are valuable in testing those theories and have attracted much attention due to the controversial processing asymmetry in the literature. Inspired by Hale (2001), this work applies an information-theoretic linking theory, *surprisal*, as a complexity metric to derive the pattern of incremental processing difficulty in Chinese RCs. Under the principle that readers apply their knowledge of grammar in perceiving syntactic structures (Chomsky, 1965), this model investigates the role that frequency-drivened structural expectations play in parsing, especially when temporary ambiguities occur. It makes predictions which remain consistent with a universal SR preference in sentence comprehension.

1.1. Controversial Processing Asymmetry in Chinese RCs

In the face of the remarkably regular cross-linguistic preference for SRs, the head-final Chinese presents a puzzling irregularity. Hsiao & Gibson (2003) (hereafter HG) present a study in which Chinese ORs are found easier to comprehend than SRs. They take this finding as an important piece of evidence for a universally applicable working memory account of sentence processing, the Dependency Locality Theory or DLT (Gibson, 1998, 2000). In particular, the integration cost metric of the DLT explains the OR advantage well. It claims that the cognitive cost of integrating the head noun extracted out of the RC and the gap it left reflects on the number of discourse referents (e.g. noun phrases) in between. In Chinese RCs (1), the head noun *guanyuan* "official" is more distant from the gap (its co-indexed empty category e) in SRs (1a) than in ORs (1b). In (1a), there is one discourse referent *fuhao* "tycoon" situated between the gap and the head noun while no discourse referent exists on the same path in (1b). The integration cost at SR's head noun is thus higher, which indicates more comprehension efforts expended.

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

[	e_i	邀请	富豪	的]	官员$_i$	打了	记者
[	e_i	yaoqing	fuhao	de]	guanyuan$_i$	da-le	jizhe
[	e_i	invite	tycoon	DE]	official	hit	reporter

'The official who invited the tycoon hit the reporter.'

*The authors thank Chien-Jer Lin for providing us the experimental data. We are much indebted to discussions with Shravan Vasishth and Jiwon Yun. This paper is based upon work supported by the NSF Grant No. 0741666.

b. Subject-modifying OR (S-OR)

[	富豪	邀请	e_i	的	]	官员$_i$	打了	记者
[	fuhao	yaoqing	e_i	de	]	guanyuan$_i$	da-le	jizhe
[	tycoon	invite	e_i	DE	]	official	hit	reporter

'The official who the tycoon invited hit the reporter.'

If the OR advantage in Chinese RCs is reliable, it will cast doubt on other theoretical accounts. First, Keenan & Comrie's Accessibility Hierarchy (1977) will be ruled out because it claims that subjects are generally easier to extract and relativize than objects. Structure-based approaches (O'Grady, 1997; Hawkins, 2004) also face severe challenges. This class of theory predicts more comprehension difficulties in ORs since the relativized element is embedded deeper as an object. Finally, the experience-based accounts (Mitchell et al., 1995; Hale, 2001) are problematic because they predicts a universal SR preference due to the cross-linguistic fact that SRs tend to occur more frequently than ORs.

Since the initial investigation by HG, mixed results have been reported on the Chinese RC processing. Table 1 summarizes a number of studies that suggest either a subject preference or an object one. Crucially, some of them argue that SRs in Chinese are actually easier to comprehend than ORs.

SR Preference	OR Preference
Lin & Bever (2006, 2007, 2011)	Hsiao & Gibson (2003)
Lin (2008)	Hsu & Chen (2007)
Wu (2009)	Lin & Garnsey (2011)
Chen, Li, Kuo & Vasishth (submitted)	Gibson & Wu (2011)

Table 1: The processing asymmetry found in Chinese RC studies

Among those results that defend the universal subject preference, Lin & Bever present a series of self-paced reading studies which all provide counterevidence against HG. For example, in Lin & Bever (2006), they not only investigate RCs that modify matrix subjects (as examined by HG), but also test the comprehension of the other two types of RCs in (2) where the head noun is the matrix object. Both (2a) and (2b) begin with the matrix subject and the main verb. Figure 1 plots the reading time per word across four types of RCs listed in (1-2) and clearly shows that SRs are read faster than ORs in Chinese.

(2) a. Object-modifying SR (O-SR)

记者	打了	[	e_i	邀请	富豪	的	]	官员$_i$
jizhe	da-le	[	e_i	yaoqing	fuhao	de	]	guanyuan$_i$
reporter	hit	[	e_i	invite	tycoon	DE	]	official

'The reporter hit the official who invited the tycoon.'

b. Object-modifying OR (O-OR)

记者	打了	[	富豪	邀请	e_i	的	]	官员$_i$
jizhe	da-le	[	fuhao	yaoqing	e_i	de	]	guanyuan$_i$
reporter	hit	[	tycoon	invite	e_i	DE	]	official

'The reporter hit the official who the tycoon invited.'

1.2. Garden-path in Chinese RCs

In Figure 1, the subject preference also seems to be more prominent in object-modifying RCs. Lin & Bever (2011) discuss this issue and argue that the salient SR advantage in object-modifying conditions could be due to a garden-path effect (Bever, 1970) existing in object-modifying ORs (O-OR). For example in (2b), the sentence-initial words "reporter hit tycoon" consist of an illusory main clause sequence (N V N). Therefore, the structure reanalysis imposed at the next word, the relative verb *yaoqing* "invite", will slow down the reading. Lin & Bever (2011) conduct a control experiment in which participants are given specific instructions about the existence and the position of an RC in each experimental

sentence. In this way, participants will know where to expect an upcoming RC in the sentence, which presumably reduces the size effect of temporary ambiguities. A comparison between the control experiment and their 2006 study demonstrates that when readers are told about the location of the RC, there is no significant reading time difference in the ambiguous/disambiguation regions before the relativizer *de*. This finding suggests that O-ORs are indeed suspectible to garden-path readings by the main clause illusion (see also the strong effect in O-ORs reported in Packard et al. (2011) and Chen et al. (submitted)).

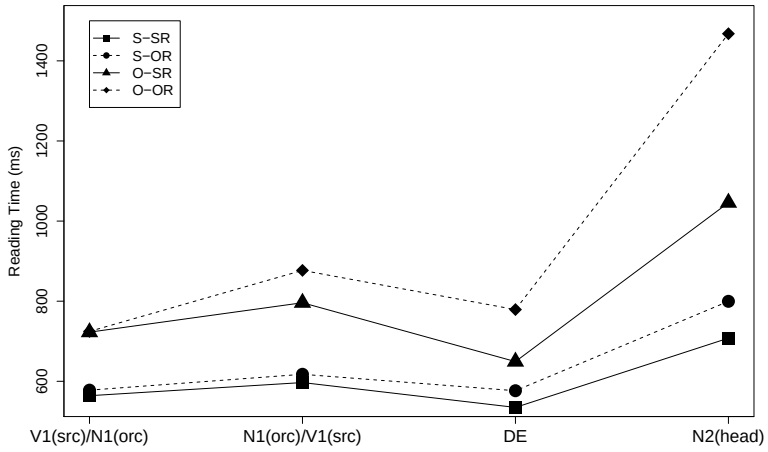


Figure 1: Reading times of each region in Chinese RCs (Lin & Bever 2006)

As an attempt at modeling the processing asymmetry in RCs, it is interesting to explore the relationship between the structure/experience-based accounts which suggest a universal SR advantage and the possible disambiguation in real time. In the rest of the paper, we describe modeling work that not only predicts the SR preference in Chinese, but also addresses the interaction between people's linguistic experience with relative clauses and their resolution of temporary ambiguities.

2. Background

2.1. Surprisal Theory

Attneave (1959) introduces the term "surprisal" to the field of cognitive science. This conception, based on a commonsense idea that low-probability events are surprising, has been revived by Hale (2001) in models of human sentence processing. Hale's proposal claims that the difficulty encountered during incremental comprehension could be modeled by word-to-word surprisals.

Under this proposal, reading times are predicted by comparing the total cognitive effort expended before and after each word. The surprisal at the word "w_n" situated between positions $n-1$ and n equals the log-ratio of two prefix probabilities, as denoted in Equation (3). The probability of a prefix string here is the *total probability* of all parses generated by the specified grammar.[1]

$$surprisal(w_n) = \log_2 \frac{P(w_0 \ \ldots \ w_{n-1})}{P(w_0 \ \ldots \ w_n)} \tag{3}$$

This information-theoretic notion of surprise can therefore be viewed as the disconfirmation of alternative continuations weighted by their probabilities and is in particular suitable to address the resolution of temporary ambiguities in comprehension. For example, using a probablistic Earley parsing algorithm (Earley, 1970) modified by Stolcke (1995), the computed surprisal in Hale (2001) correctly predicts the reading time abnormality at the sentence final word "fell" in the following renowned example of garden path effect in English.

[1]Under the assumption that parses with a extremely low probability will hardly be considered by the human reader, it is realistic to limit the calculation of total probability to a certain level of parallism (i.e. only computing on a few highly possible interpretations) if using a broad-coverage grammar (Boston, Hale, Vasishth & Kliegl, 2011).

(4) $_0$ The $_1$ horse $_2$ raced $_3$ past $_4$ the $_5$ barn $_6$ fell $_7$. (Bever, 1970)

Depending on the analysis of the verb "raced", the prefix string at position 6, "The horse raced past the barn", has both a highly possible main-clause reading and a less-preferred reduced relative reading. Integrating it with the upcoming word "fell", the parser will conduct a structural reanalysis. The reduced relative structure could be constructed however with much efforts. This process results in a sharply increased reading time. The surprisal model quantifies the amount of disambiguation efforts spent at the word "fell" by comparing the total probability of the first six-word string $P(w_0 \ \ldots \ w_6)$ and that of the first seven-word string $P(w_0 \ \ldots \ w_7)$. Since there is a big loss of probablity mass from the main clause structure, the ratio between the two reflects a high probability over a low one and thus a large surprisal.

2.2. Probabilistic Context-Free Grammar

The calculation of surprisal relies on *language models*. Both Hale (2001) and Levy (2008) build probabilistic context-free phrase structure grammars (PCFG) (Charniak, 1993) in their implementation of surprisal theories, e.g.

$$\begin{array}{llll} 1.0 & \mathrm{S} & \longrightarrow & \mathrm{NP\ VP} \\ 0.5 & \mathrm{NP} & \longrightarrow & \mathrm{DT\ N} \\ 0.5 & \mathrm{NP} & \longrightarrow & \mathrm{NP\ VP} \\ \vdots & \vdots & & \end{array} \quad \Longrightarrow \quad \begin{array}{ccc} & \mathrm{S}_{0.5} & \\ \mathrm{NP}_{0.5} & & \mathrm{VP} \\ \mathrm{DT} & \mathrm{N} & \end{array}$$

In this example, the nonterminal S could be freely substituted by symbols NP and VP no matter what context it has. In a PCFG, each derivation is associated with a probability. The probability of a parse is the product of the probabilities of all the derivations used. For instance, there is only half a chance of rewriting a noun phrase NP to a determiner DT followed by a noun N. The probability of the tree structure above is therefore the product of probabilities of two grammar rules involved: $1 \times 0.5 = 0.5$.

3. Procedure

In the current study, we prepared a small PCFG in Table 2 that covers four types of Chinese RCs discussed earlier where they either modify matrix subjects (1a–b) or matrix objects (2a–b). The grammar adopts a Generalised Phrase Structure Grammar (GPSG)-style analysis of RC (Gazdar, Klein, Pullum & Sag, 1985) and in addition recognizes Chinese as a *pro*-drop language that allows omitted subjects or objects.

0.9860	S	→	NPSBJ	VP				
0.0140	S	→	NPRC	VP	0.5	V	→	*yaoqing*
0.5945	NPRC	→	CPSR	NP	0.5	V	→	*dale*
0.4055	NPRC	→	CPOR	NP	0.3333	N	→	*fuhao*
0.9771	VP	→	V	NPOBJ	0.3333	N	→	*guanyuan*
0.0229	VP	→	V	NPRC	0.3334	N	→	*jizhe*
1.0	CPSR	→	VP	DEC	1.0	DEC	→	*de*
1.0	CPOR	→	S/NP	DEC	1.0	NP/NP	→	ϵ
1.0	S/NP	→	NPSBJ	VP/NP	1.0	pro	→	ϵ
1.0	VP/NP	→	V	NP/NP				
0.3562	NPSBJ	→	NP					
0.6438	NPSBJ	→	pro					
0.9927	NPOBJ	→	NP					
0.0073	NPOBJ	→	pro					
1.0	NP	→	N					

Table 2: A small PCFG for Chinese RCs with *pro*-drop rules

The probability on each grammar rule is the estimation based on the number of certain type of structures observed in the corpus. We employ the pattern matching tool Tregex (Levy & Andrew, 2006)

on the Chinese Treebank 7.0 (Xue et al., 2010) to obtain attestation counts for relevant structural types[2], as listed in Table 3. These counts faithfully reflect two key aspects of Chinese: SRs are more frequent than ORs (SRs 59.45% versus OR 40.55%); most *pro*s are dropped at the subject position and object *pro*-drops are rare.

Subject Type	**Count**	**Object Type**	**Count**	**RC Type**	**Count**
Nominal	15996	Nominal	20708	SR	1293
pro	28913	*pro*	152	OR	882
RC	636	RC	488		

Table 3: Attestation counts from the Chinese Treebank 7.0

Table 4 lists the Tregex queries we use to elicit RCs. The RC (the first "@CP") is the immediate left sister of the extracted subject/object (the "@NP" at last). There is always a "WH" operator in the RC that shares the same index with the extracted NP. Note that in both queries we define the regular verb "VV" as the only verb type in the search for RCs. This excludes instances involving the predicate adjective "VA" or the copula "VC" which are usually avoided in previous experimental stimuli designs.

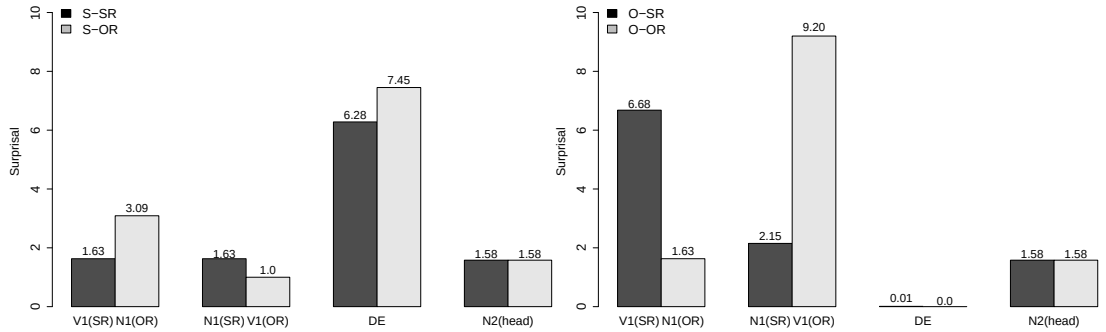
SR	(@**CP** <, /^WH.*-([0-9]+)$/#1%index <- (@CP <, (@IP < (NP-SBJ <<: /^*T*-([0-9]+)$/#1%index $+ (@VP < **VV** < NP-OBJ))) <2 DEC)) $+ @**NP**
OR	(@**CP** <, /^WH.*-([0-9]+)$/#1%index < (@CP <, (@IP < (NP-SBJ $+ (@VP < **VV** < (NP-OBJ <<: /^*T*-([0-9]+)$/#1%index)))) <2 DEC)) $+ @**NP**

Table 4: Tregex queries for Chinese RCs

We use a statistical prefix parsing system (Grove, 2010) to obtain surprisals for ambiguous prefixes in each of the four Chinese RC examples. It employs an exhaustive bottom-up chart parsing strategy (Goodman, 1999) in the style of Shieber, Schabes & Pereira (1995). The parser constructs a probabilistic model at each prefix and calculates the inside probability of all non-terminals. At each prefix, the worked out surprisal value reflects all of the syntactic alternatives that the grammar defines.

4. Results

The surprisal results derive the SR/OR asymmetry across four types of Chinese RCs in Examples (1–2). Figure 2 plots the surprisal values that compare the reading difficulties encountered at each word between the SR and the OR.



(a) Total: S-SR 11.12 vs. S-OR 13.12 (b) Total: O-SR 10.42 vs. O-OR 12.41

Figure 2: Surprisals predict the SR/OR asymmetry in Chinese RCs

We calculate the total surprisal of an RC by adding surprisal values of all four words. In both

[2]Chinese Treebank 7.0 is the latest release of the Chinese Treebank project by the Linguistic Data Consortium. It cosists 51,447 fully parsed sentences or 1,196,329 words.

subject-modifying and object-modifying conditions, the total surprisal of the SR is lower than that of the OR, indicating that SRs are easier to process as a whole. This prediction patterns the empirical observations reported in studies such as Lin & Bever (2006) and Chen et al. (submitted).

In addtion to the total surprisal, we also examine the critical disambiguation region at each condition. Table 5 schematically summarizes the resolution of temporary ambiguities at these positions in Chinese RCs. See Appendix for bare phrase structures spanning each prefix string before and after the critical disambiguation point.

	Before "DE"		**After "DE"**	**Surprisal**
a. S-SR	*pro* V N ... *e* V N ...	$\Longrightarrow$	~~*pro* V N ...~~ *e* V N DE ...	6.28
b. S-OR	N V (N ...) N V (*e* ...)	$\Longrightarrow$	~~N V (N ...)~~ N V *e* DE ...	7.45
	Before "N_1(SR)"		**After "N_1(SR)"**	
c. O-SR	N V *e* V ...	$\Longrightarrow$	N V *e* V N ...	2.15
	Before "V_1(OR)"		**After "V_1(OR)"**	
d. O-OR	N V N N V N (V *e* ...)	$\Longrightarrow$	~~N V N~~ N V N V *e* ...	9.20

Table 5: The resolution of temporary ambiguities in Chinese RCs

The relativizer "DE" is the critical disambiguation word in subject-modifying RCs (1a–b). The surprisal contrast (S-SR 6.28 vs S-OR 7.45) at this position suggests an SR preference, as illustrated in Figure 2(a). In S-SRs, upon encountering "DE" the ambiguous prefix string "invite tycoon ..." (V N ...) has two possible readings in which the "missing" subject could either be a dropped *pro* or an empty category *e*. If it is a *pro*, this prefix string will be a main clause, otherwise an SR. Integrating the relativizer "DE" will resolve this ambiguity due to the only avaliable SR reading. On the contrary, S-ORs have a prefix string of "tycoon invite ..." (N V ...) with both a main clause reading and an OR reading. Similarly, the main clause structure no longer exists in the face of "DE", which leaves the OR as the only option. In both conditions, the losing reading is the main clause. However, since Chinese ORs are less frequent than SRs, choosing the OR analysis indicates a relatively larger loss of probability mass and thus predicts more disambiguation efforts, as demonstrated by a bigger surprisal value.

The surprisal values in Figure 2(b) suggest that the SR preference in Chinese should be more pronounced in object-modifying RCs (2a–b) where the critical point of disambiguation is the second word. This critical point is the fourth word in the complete sentence, if taking into account the common matrix subject and the main verb "reporter hit ...". In O-SRs, the RC reading has been decided before "N_1(SR)" due to a sequence of "N V V ..." and cannot impose much additional disambiguation burden. On the contrary, in O-ORs, the illusory main-clause reading exists at "N_1(OR)" (N V N ...) and is disconfirmed upon reaching "V_1(OR)" (N V N V *e* ...). Choosing the less-preferred OR analysis triggers a garden-path effect and is evidenced by a sharper surprisal value contrast (O-SR 2.15 vs O-OR 9.20) than the subject advantage found in subject-modifying RCs, as also argued by Lin & Bever (2011).

5. Discussion

In this modeling, the expectation of the most possible structure given a prefix string predicts a subject preference in Chinese RCs. It also precisely illustrates the more prominent comprehension difficulty difference between SRs and ORs when they modify the matrix object.

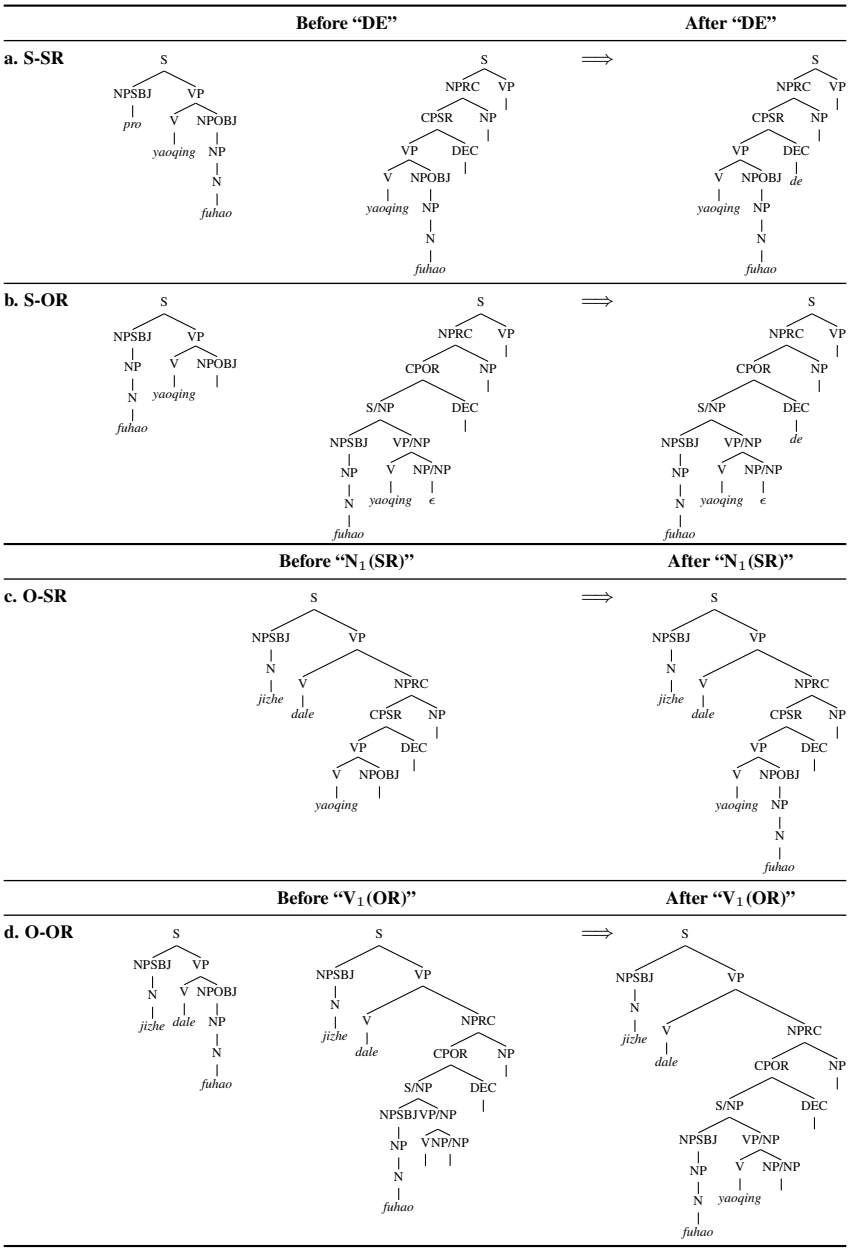
It occurs to us that the results of the modeling suggest an early effect, as compared to empirical results where SR are read faster at the head noun and/or the word after.[3] However, this result could in fact be congruent with a computational system which is more eager to make a parsing decision than sentence

[3] Staub (2010) discusses the experience-based accounts (Hale, 2001; Levy, 2008) and the working memory explanation (Gibson, 1998; Grodner & Gibson, 2005) on their different predictions about where the reading difficulty first appears in an English OR. Two eye-movement experiments suggest that both methods contribute to the difficulty.

comprehenders who are in unrushed experimental environments, e.g. in self-paced reading experiments. One could argue that the effect found in regions after the relativizer *de* is potentially due to the often observed spillover effect from the preceding regions (Mitchell, 1984; Vasishth & Lewis, 2006). Since the function word *de* could easily be ignored in reading, the efforts spent on resolving the temporary ambiguity might therefore be delayed until reading the following word(s).

To conclude, the modeling work presented in this paper provides an accurate account of processing difficulties across four RC constructions in Chinese. The technical notion *surprisal* therefore unifies difficulty caused by incremental ambiguity resolution and difficulty caused by readers' linguistic experience, here, the structural (in)frequency. This empirically robust theory remains consistent with a universal SR preference in sentence comprehension.

Appendix: Resolving temporary ambiguities in Chinese RCs



References

Attneave, Fred (1959). *Applications of Information Theory to Psychology: A summary of basic concepts, methods and results*. Holt, Rinehart and Winston.

Bever, Thomas G. (1970). The cognitive basis for linguistic structures. Hayes, J. R. (ed.), *Cognition and the Development of Language*, John Wiley, 279–360.

Boston, Marisa F., John T. Hale, Shravan Vasishth & Reinhold Kliegl (2011). Parallelism and syntactic processes in reading difficulty. *Language and Cognitive Processes* 26:3, 301–349.

Charniak, Eugene (1993). *Statistical Language Learning*. MIT Press, Cambridge, MA.

Chen, Zhong, Qiang Li, Kuei-Lan Kuo & Shravan Vasishth (submitted). Processing Chinese relative clauses: Evidence for the universal subject preference .

Chomsky, Noam (1965). *Aspects of the Theory of Syntax*. MIT Press, Cambridge, MA.

Earley, Jay (1970). An efficient context-free parsing algorithm. *Communications of the Association for Computing Machinery* 13:2, 451–455.

Frauenfelder, Ulrich, Juan Segui & Jacques Mehler (1980). Monitoring around the relative clause. *Journal of Verbal Learning and Verbal Behavior* 19, 328–337.

Frazier, Lyn (1987). Syntactic processing: Evidence from Dutch. *Natural Language & Linguistic Theory* 5:4, 519–559.

Gazdar, Gerald, Ewan Klein, Geoffrey K. Pullum & Ivan A. Sag (1985). *Generalized Phrase Structure Grammar*. Basil Blackwell, Oxford.

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

Gibson, Edward (2000). Dependency locality theory: A distance-based theory of linguistic complexity. Marantz, Alec, Yasushi Miyashita & Wayne O'Neil (eds.), *Image, Language, brain: Papers from the First Mind Articulation Project Symposium*, MIT Press, Cambridge, MA.

Gibson, Edward & Hsiao-Hung Iris Wu (2011). Processing Chinese relative clauses in context. *Language and Cognitive Processes* .

Goodman, Joshua (1999). Semiring parsing. *Computational Linguistics* 25, 573–605.

Grodner, Daniel & Edward Gibson (2005). Consequences of the serial nature of linguistic input. *Cognitive Science* 29, 261–290.

Grove, Kyle (2010). MCFGCKY v. 1.0. URL `http://conf.ling.cornell.edu/kgrove/mcfgcky/`.

Hale, John T. (2001). A probabilistic Earley parser as a psycholinguistic model. *Proceedings of the Second Meeting of the North American Chapter of the Association for Computational Linguistics*, Pittsburgh, PA.

Hawkins, John A. (2004). *Efficiency and Complexity in Grammars*. Oxford University Press.

Hsiao, Franny & Edward Gibson (2003). Processing relative clauses in Chinese. *Cognition* 90, 3–27.

Hsu, Chun-Chieh & Jenn-Yeu Chen (2007). A new look at the subject-object asymmetry: The effects of linear distance and structural distance on the processing of head-final relative clauses in Chinese. *Interdisciplinary Approaches to Relative Clauses (REL07)*.

Just, Marcel Adam, Patricia A. Carpenter, Timothy A. Keller, William F. Eddy & Keith R. Thulborn (1996). Brain Activation Modulated by Sentence Comprehension. *Science* 274:5284, p. 114.

Keenan, Edward & B. Comrie (1977). Noun phrase accessibility and universal grammar. *Linguistic Inquiry* .

King, Jonathan & Marcel Adam Just (1991). Individual differences in syntactic processing: The role of working memory. *Journal of Memory and Language* 30, 580–602.

King, Jonathan W. & Marta Kutas (1995). Who did what and when? Using word- and clause-level ERPS to monitor working memory usage in reading. *Journal of cognitive neuroscience* 7:3, 376–395.

Kwon, Nayoung, Maria Polinsky & Robert Kluender (2006). Subject Preference in Korean. Baumer, Donald, David Montero & Michael Scanlon (eds.), *Proceedings of the 25th West Coast Conference on Formal Linguistics*, Cascadilla Proceedings Project, Somerville, MA, 1–14.

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

Levy, Roger & Galen Andrew (2006). Tregex and Tsurgeon: tools for querying and manipulating tree data structures. *Proceedings of the 5th International Conference on Language Resources and Evaluation (LREC 06)*.

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

Lin, Chien-Jer Charles (2008). The processing foundation of head-final relative clauses. *Language and Linguistics* 9.4, 813–839.

Lin, Chien-Jer Charles & Thomas G. Bever (2006). Subject preference in the processing of relative clauses in Chinese. Baumer, Donald, David Montero & Michael Scanlon (eds.), *Proceedings of the 25th West Coast Conference on Formal Linguistics*, Cascadilla Proceedings Project, Sommerville, MA, 254–260.

Lin, Chien-Jer Charles & Thomas G. Bever (2007). Processing doubly-embedded head-final relative clauses. *Interdisciplinary Approaches to Relative Clauses*, Cambridge, UK.

Lin, Chien-Jer Charles & Thomas G. Bever (2011). Garden path and the comprehension of head-final relative clauses. Yamashita, Hiroko, Yuki Hirose & Jerome L. Packard (eds.), *Processing and Producing Head-final Structures,* Studies in Theoretical Psycholinguistics, Springer, 277–297.

Lin, Yowyu Brian & S. M. Garnsey (2011). Animacy and the resolution of temporary ambiguity in relative clause comprehension in mandarin. Yamashita, Hiroko, Yuki Hirose & Jerome L. Packard (eds.), *Processing and Producing Head-final Structures,* Studies in Theoretical Psycholinguistics, Springer, 241–276.

MacDonald, Maryellen C. & Morten H. Christiansen (2002). Reassessing working memory: A reply to Just and Carpenter and Waters and Caplan. *Psychological Review* 109:1, 35–54.

Mitchell, Don C. (1984). An evaluation of subject-paced reading tasks and other methods of investigating immediate processes in reading. Kieras, D. E. & M.A. Just (eds.), *New Methods in Reading Comprehension Research*, Erlbaum, Hillsdale, N.J.

Mitchell, Don C., Fernando Cuetos, Martin M. B. Corley & Marc Brysbaert (1995). Exposure-based models of human parsing: Evidence for the use of coarse-grained (nonlexical) statistical records. *Journal of Psycholinguistic Research* 24, 469–488.

Miyamoto, Edson T. & Michiko Nakamura (2003). Subject/object asymmetries in the processing of relative clauses in Japanese. Garding, G. & M. Tsujimura (eds.), *Proceedings of the 22nd West Coast Conference on Formal Linguistics*, 342–355.

O'Grady, William (1997). *Syntactic Development*. The University of Chicago Press.

Packard, Jerome L., Zheng Ye & Xiaolin Zhou (2011). Filler-gap processing in Mandarin relative clauses: Evidence from event-related potentials. Yamashita, Hiroko, Yuki Hirose & Jerry Packard (eds.), *Processing and Producing Head-final Structures,* Studies in Theoretical Psycholinguistics, Springer, vol. 38, 219–240.

Schriefers, Herbert, Angela Friederici & Katja Kühn (1995). The Processing of Locally Ambiguous Relative Clauses in German. *Journal of Memory and Language* 34, 499–520.

Shieber, Stuart M., Yves Schabes & Fernando C.N. Pereira (1995). Principles and implementation of deductive parsing. *The Journal of Logic Programming* 24:1–2, 3–36.

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

Stolcke (1995). An efficient probabilistic context-free parsing algorithm that computes prefix probabilities. *Computational Linguistics* 21, 165–202.

Stromswold, Karin, David Caplan, Nathaniel Alpert & Scott Rauch (1996). Localization of Syntactic Comprehension by Positron Emission Tomography. *Brain and Language* 52:3, 452–473.

Traxler, Matthew J., Robin K. Morris & Rachel E. Seely (2002). Processing subject and object relative clauses: Evidence from eye movements. *Journal of Memory and Language* .

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

Wu, Fuyun (2009). *Factors Affecting Relative Clause Processing in Mandarin: Corpus and Behavioral Evidence*. Ph.D. thesis, University of Southern California.

Xue, Nianwen, Zixin Jiang, Xiuhong Zhong, Martha Palmer, Fei Xia, Fu-Dong Chiou & Meiyu Chang (2010). *Chinese Treebank 7.0*. Linguistic Data Consortium, Philadelphia.