

汉语关系从句与普遍句子加工理论的建立

Processing Relative Clauses in Mandarin Chinese

张子涵

2025/03/06

报告结构

1 引言

2 相关研究

2.1 关系从句的加工不对称性

2.2 两类句子加工理论

2.2.1 基于期望的理论 (Expectation-based Theory)

2.2.2 基于记忆的理论 (Memory-based Theory)

2.3 双重嵌套结构的结构性遗忘

2.4 理论融合的尝试: Lossy-context Surprisal Model (LCS)

3 研究问题

4 讨论

Introduction

- 以Chomsky为代表的生成语法学家认为“迭代”（Recursion）是人类语言能力的核心，反应在语言中最典型的代表即为关系从句（Relative Clause, RC）这种理论上具有无限嵌套能力的结构。
 - 动物实验表明，学习以RC为代表的层级性结构是人类独有的语言能力 (e.g., Hauser et al., 2002)
- 多数语言中都存在着关系从句，但其加工机制目前尚不明晰。最初被关注到的现象是主语关系从句（Subject Relative Clauses, SRCs）和宾语关系从句（Object Relative Clauses, ORCs）的加工不对称性。
- 来自多数语言的研究均发现了SRCs的加工优势（e.g., 英语：Ford, 1983; Gibson, 1998；荷兰语：Frazier, 1987；法语：Holmes & O'Regan, 1981 etc.）-> 是否具有跨语言共性？

SRCs: The book [that is on the table] belongs to me.

[喜欢小红的]同学离开了。

ORCs: The book [that I read yesterday] was interesting.

[小红喜欢的]同学离开了。

Table 1. Preferences for Relative-Clause Processing

Language	RC position	Preference	Task	References
Brazilian Portuguese	Postnominal	SRC	RSVP	Gouvea (2003)
Dutch	Postnominal	SRC	Self-paced reading	Frazier (1987)
		SRC	Self-paced reading	Mak, Vonk, & Schriefers (2002)
			Eye-movement monitoring	
English	Postnominal	SRC	Continuous Lexical Decision Task	Ford (1983)
		SRC	Self-paced reading	King and Just (1991), Gibson, Desmet, Grodner, Watson, & Ko (2005)
		SRC	Eye-movement monitoring	Traxler, Morris, & Seely (2002)
French	Postnominal		ERP	King & Kutas, 1995
		SRC	phoneme-monitoring task	Frauenfelder, Segui, & Mehler (1980)
		SRC	click-monitoring	Cohen & Mehler (1996)
		SRC	Eye-movement monitoring	Holmes and O'Regan (1981)
German	Postnominal	SRC	Self-paced reading	Schriefers, Friederici, and Kuhn (1995)
		SRC	ERP	Mecklinger, Schriefers, Steinhauer, & Friederici's (1995)
Mandarin	Prenominal	ORC	Self-paced reading	Hsiao & Gibson (2003)
		SRC	Self-paced reading	Experiment 1 (Lin, 2006)
Japanese	Prenominal	SRC	Self-paced reading	Miyamoto & Nakamura (2003)
Korean	Prenominal	SRC	Self-paced reading	Kwon, Polinsky, & Kluender (2004)

- 中心语前置和中心语后置的语言均发现了SRCs的加工优势；

- 从类型学上看，汉语是唯一同时具备SVO语序和中心词后置的语言

-> 印欧语及日韩语的现象及理论解释可能无法推广至汉语，而汉语的结果对于建立普遍句子加工理论十分重要；

- 关于汉语关系从句的加工优势不对称性，目前的实验结果并不一致 (e.g., Hsiao & Gibson, 2003; Lin & Bever, 2006; Gibson & Wu, 2013)。

Hsiao & Gibson (2003)

- 实验范式：自定义步速阅读（Self-Paced Reading, SPR）

- 实验材料

[富豪邀请e的]官员心怀不轨但是善于隐藏。(ORC)

[e邀请富豪的]官员心怀不轨但是善于隐藏。(SRC)

[[富豪邀请e_i的]法官_i勾结e_k的]官员_k心怀不轨但是善于隐藏。(ORC+ORC)

[e_i邀请[e_k勾结法官的]富豪_i的]官员_k心怀不轨但是善于隐藏。(SRC+SRC)

- 实验结果

单层 RCs：在前两个词上ORCs的RT更短；在关系化标记‘的’以及之后区域没有显著差异；

双层 RCs：ORCs的RT总体显著更短；

-> 认为汉语与其他语言不同，具有ORCs的加工优势

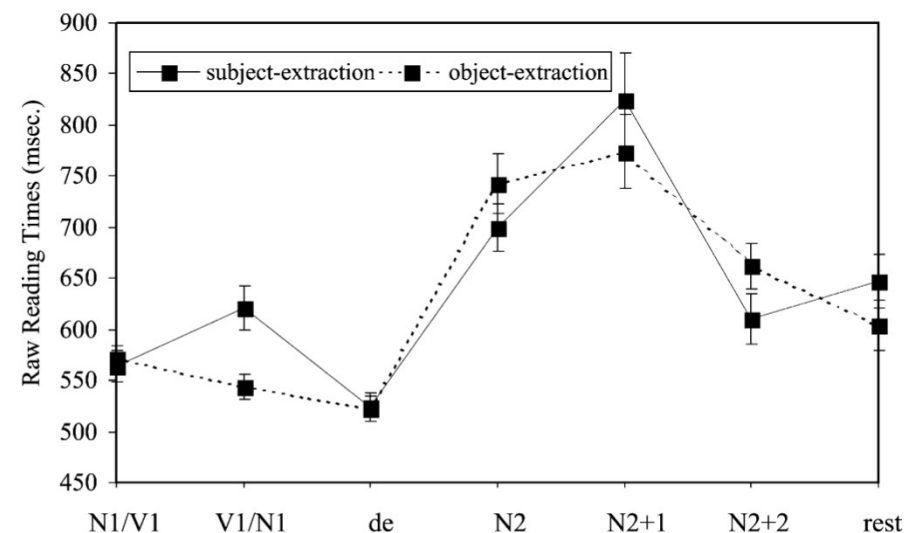


Fig. 1. Plot of mean (standard error) raw RTs per word for the singly-embedded conditions in (2).

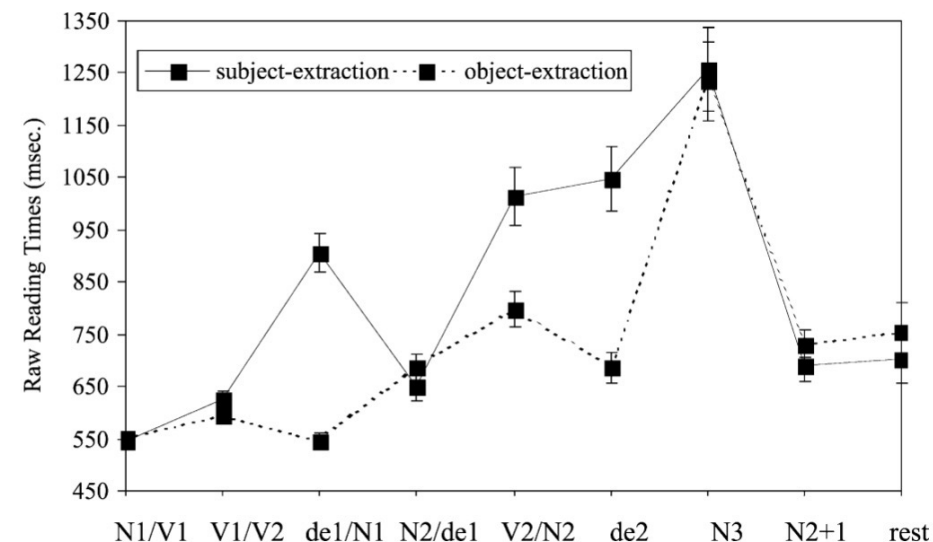
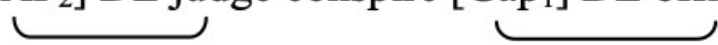


Fig. 2. Plot of mean (standard error) raw RTs per word for the doubly-embedded conditions in (3).

Lin & Bever (2006)

批判Hsiao & Gibson (2003) 的实验设计

- 单层 RCs: 在关系化标记“的”之前的部分的RTs更短不能证明ORCs具有加工优势;
[富豪邀请e的]官员 (NP+VP) vs. [e邀请富豪的]官员 (VP+NP)
- 双层 RCs: 双层嵌套ORCs的序列依赖 (serial dependency) 比SRCs的中心依赖 (central dependency) 更好处理;
 - a. Subject relative clauses embedded in subject relative clauses
[GAP₁] invite [GAP₂] conspire judge DE tycoon DE official have bad intentions

 - b. Object relative clauses embedded in object relative clauses
tycoon invite [GAP₂] DE judge conspire [Gap₁] DE official have bad intentions

- 实验材料中的动词存在句法歧义, 有的动词除了可以带名词性宾语还能带句子补语 (如, 邀请他 vs. 邀请他来) 以及动词性补语 (如, 开始工作、喜欢跑步);

Lin & Bever (2006)

- 实验设计：2（SRCs vs. ORCs）* 2（修饰主语 vs. 修饰宾语）
- 实验材料：使用不含句法歧义的动词，只能接名词性宾语；
- 实验结果：在关系化标记“的”和其后的中心词上 SRCs 的RTs显著更短；

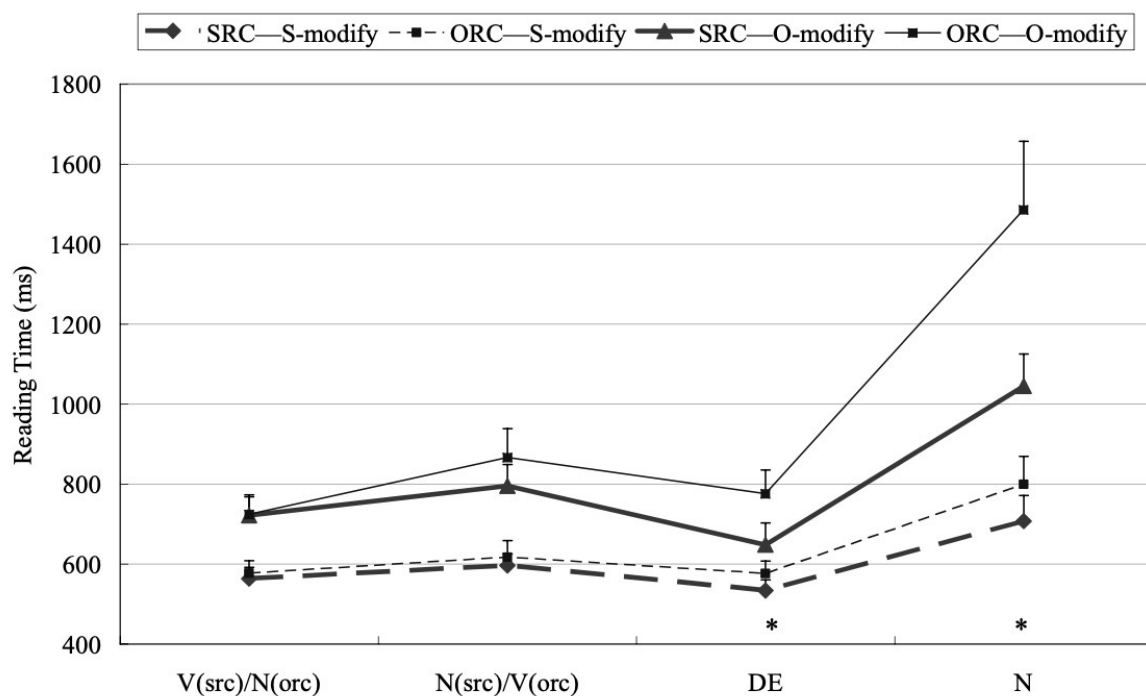


Figure 1. Reading time of each region within regular relative clauses (Lin, 2006: 131)
(SRC = subject-extracted relative clause, ORC = object-extracted relative clause)

使用这类动词是否合理？

多数为动作动词，如吃、喝、买、穿等；

富豪邀请的官员 (ORC) vs. 邀请富豪的官员 (SRC)

张三吃的苹果 (ORC) vs.
吃张三的苹果 (VP+NP)
吃苹果的张三 (SRC)

Gibson & Wu (2013)

- 回应Lin & Bever (2006) :
 - 其结果中的SRCs加工优势有可能是零语境条件下暂时的句子解析歧义导致的;
 - 图中 O-modify 内部差异明显, 而 S-modify 内部差异并不明显;

修饰主语的RCs:

[邀请富豪的]官员心怀不轨 (SRC) – VP + NP 开头, 可以被暂时解析为主句

[富豪邀请的]官员心怀不轨 (ORC) – NP + VP 开头, 容易被暂时解析为主句

修饰宾语的RCs:

议员撞倒了[勾引院长的]少女 (SRC) – ‘勾引’ 会直接被解析为从句, 无歧义

议员撞倒了[院长勾引的]少女 (ORC) – ‘院长’ 容易被解析为直接宾语

Gibson & Wu (2013)

- 实验设计：使用语境诱导句子的解析方式，尽量避免暂时的歧义解析；
- 实验材料：

在快速道路拥堵的车流中，有一台重型机车（X）追着一台轿车（y1）。另外一台轿车（y2）看到之后，就追着那台重型机车。

小明说：我听说一个高中生开着其中一台轿车，而另外一台是一个中年妇女开着。

小明说：那个高中生是开着哪一台车呢？

小梅说：[追重型机车的]车是高中生开的车。（SRC）

小梅说：[重型机车追的]车是高中生开的车。（ORC）

- 实验结果：

在支持性语境下，汉语 ORCs 的 RTs 显著更短；

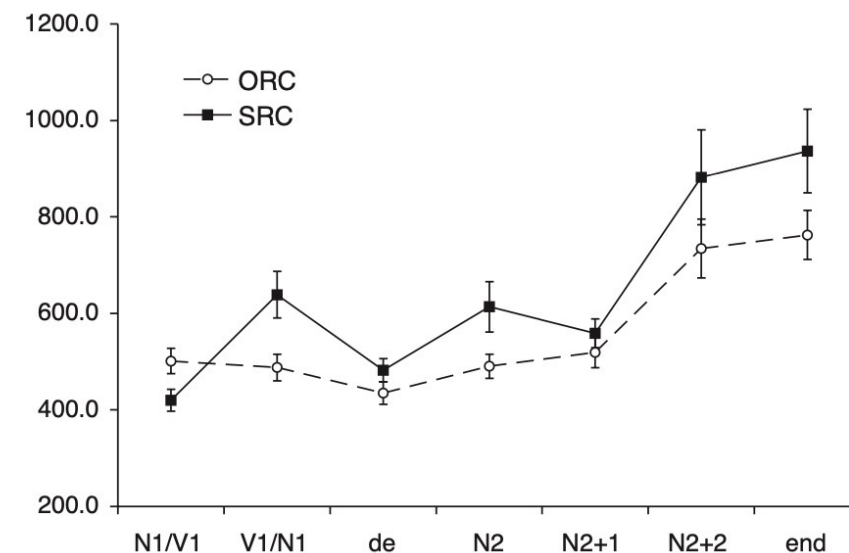


Figure 1. Reading times per word in the target sentence. Error bars indicate standard errors of the mean.

句子加工理论

如何从理论上解释关系从句的加工不对称性？

- 一个好的理论应该可以预测跨语言的不同表现；
- 目前存在的理论可大致被归为两类：
 - 基于记忆（认知资源）的理论 (Memory-based Theory)
 - 基于期望（经验统计）的理论 (Expectation-based Theory)

Memory-based Theory

- 基本理念：句子的加工难度取决于其所需的工作记忆资源。
- Dependency Locality Theory (DLT; Gibson, 1998)
 - 句子加工所需要的成本主要包括两部分两部分
 - Storage Costs: 用于存储未完成的依存结构（maintain incomplete dependencies）；
 - Retrieval (or Integration) Costs: 用于在将当前词嵌入到句法结构中时检索之前的表征，主要通过当前词到其依存结构的线性距离计算；
 - 预测英语存在 SRCs 加工优势，而汉语存在 ORCs 加工优势；
- 常规语序假说 (Macdonald & Christiansen, 2002) :
 - 汉语和英语的基本语序都是SVO，汉语ORCs符合该基本语序而SRCs不符合，因此存在ORCs加工优势；英语相反，SRCs符合SVO语序而ORCs不符合，故存在SRCs加工优势；

SRCs: The book [that e is on the table] belongs to me.

ORCs: The book [that I read e yesterday] was interesting.

[e喜欢小红的]同学离开了。

[小红喜欢e的]同学离开了。

Expectation-based Theory

- 基本理念：人们加工句子的难度是通过他们在过去遇到类似的单词和结构的经验来预测的，即结构越常见，加工它的难度就越小。
- Accessibility Theory (Keenan & Comrie, 1977)
 - 不同语言的RCs能够提取的句法位置的范围不同，由此形成了一个等级序列，例如，所有语言的RCs都可以提取主语，因此主语位置被认为是最容易访问的（accessible），提取成本最低。由于人们在容易访问的位置提取的频率更高，所以处理更容易。
- Surprisal Theory (Hale, 2001; Levy, 2008)
 - 在人们进行阅读时，每遇到一个新词会根据其加工难度动态地分配给其一定的RTs，加工难度与该词的可预测性 (Predictability) 相关，它可以用惊异度 ($\text{Surprisal} = -\log P(w|c)$) 进行量化表示；
- 由于在英汉语料库中，SRCs 的出现频率均高于 ORCs，因此预测英语存在 SRCs 的加工优势，汉语也存在 SRCs 的加工优势。

两种理论的解释力

- Memory-based Theory 解释了出现加工困难的位置
 - Expectation-based Theory 预测困难应出现在不常见处（如英语ORCs中的主语）；
 - Memory-based Theory 预测困难应出现在形成依赖关系处（如英语ORCs中的动词）；
 - 实验结果表明困难出现在动词处；

SRCs: The book [that e is on the table] belongs to me.

[e喜欢小红的]同学离开了。

ORCs: The book [that I read e yesterday] was interesting.

[小红喜欢e的]同学离开了。

- Expectation-based Theory 解释了ORCs 主语的生命体效应
 - Traxler et al. (2002) 发现，当ORCs的主语是有生命的，修饰无生客体（inanimate patient）的ORCs更容易处理（如，男孩扔的石头）；当ORCs的嵌入主体是无生命的时，修饰有生客体的ORCs更难处理（如，石头击中的登山员）-> 主客体都是有生的动词较少
- 但也存在二者均无法完全解释的现象，如结构性遗忘（Structural Forgetting）。

Structural Forgetting in Doubly Nested RCs

Missing VP effects in English (Gibson & Thomas, 1999)

- 实验材料 (ORCs + ORCs)

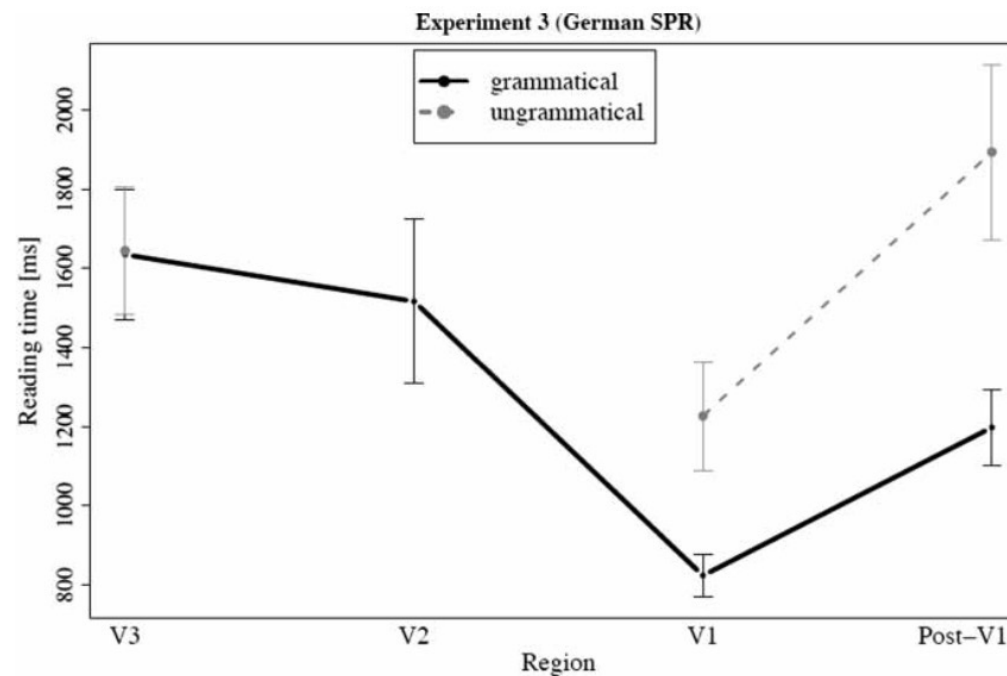
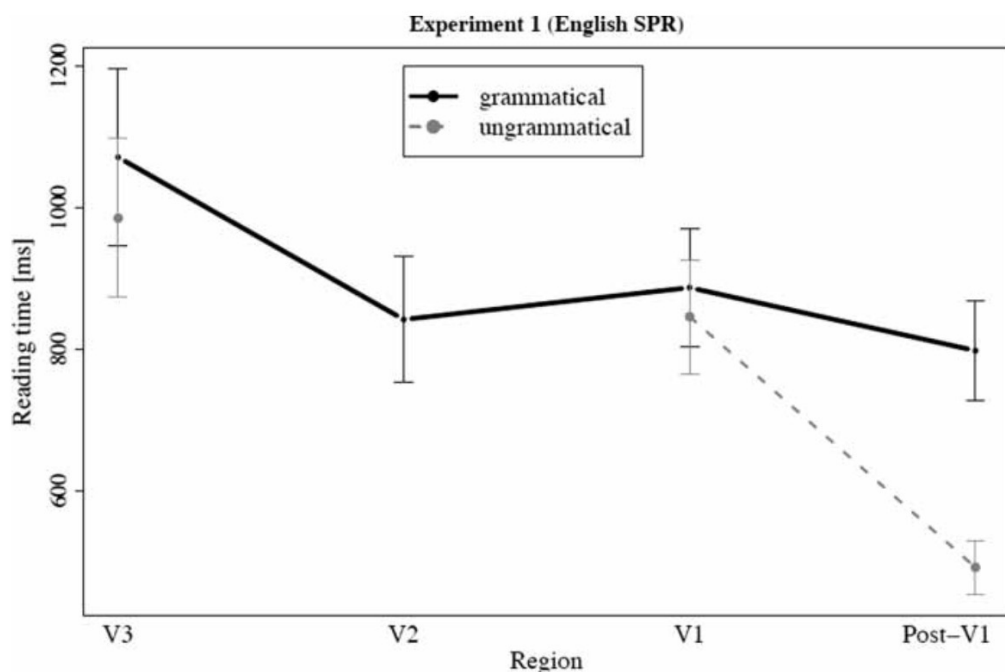
(a) * The apartment_{NP1} that [the maid_{NP2} who [the cleaning service_{NP3} had sent over_{VP3}]_{orc} (missing _{VP2})]_{orc} was well-decorated_{VP1}. (Ungrammatical)

(b) The apartment_{NP1} that [the maid_{NP2} who [the cleaning service_{NP3} had sent over_{VP3}]_{orc} was cleaning _{VP2} every week]_{orc} was well-decorated_{VP1}. (Grammatical)

- 实验结果：在英语上，缺省VP2的不合法句子有时比合法句子接受度更高，而省略VP1和VP3则会导致接受度降低；
- High Memory Cost Pruning Hypothesis（拓展 DLT，当工作记忆负担过大时如何？）
 - 在超出工作记忆负荷后，会遗忘造成工作记忆负荷最大的依赖项；
 - NP2的记忆负荷量最大，所以人们会遗忘NP2以缓解工作记忆压力，从而导致VP2的缺省符合人脑中的句子表征，导致其接受度高。

Vasishth et al. (2010): Verb-final Structure Evidence

- VP missing 并非跨语言的普遍现象 (Frank et al., 2016)
- 基于语言经验的推测：动词后置结构 (Verb-final Structure) 可能使德语解析器习惯于维持 VP 预测，从而减少 VP 遗忘的可能性。德语从句中的动词始终位于句末，受试者在处理过程中会始终期待 VP 出现，不会像英语一样遗忘 VP2 -> 对应 Expectation-based theory



Lossy-context Surprisal Model (LCS)

- Memory-based Theory : 无法解释为何在英语中存在的VP-missing在德语中不存在;
- Expectation-based Theory: 无法解释它为何不会分配给不合语法的句子零概率或小概率, 也无法解释缺省位置等细节的加工机制;

-> 将二者结合的尝试: Lossy-context Surprisal Model (Futrell et al., 2020; Futrell & Levy, 2017)

- 基本观点: 标准的Surprisal Theory 认为 Surprisal 来自对于上下文的完美的记忆表征, LCS承认了记忆表征的局限性, 认为对于当前词的推断来自于有噪声的上下文。

$$\textit{Surprisal} = -\log P(w|c)$$

$$\textit{Lossy-context Surprisal} = -\log P(w|c') = -\log \sum_c P(w|c) P(c|c')$$

LCS能够解释英语和德语之间跨语言的结构性遗忘现象

能否推广到汉语？

Rule	Probability
$S \rightarrow NP\ V$	1
$NP \rightarrow N$	$1 - m$
$NP \rightarrow N\ RC$	mr
$NP \rightarrow N\ PP$	$m(1 - r)$
$PP \rightarrow P\ NP$	1
$RC \rightarrow C\ V\ NP$	s
$RC \rightarrow C\ NP\ V$	$1 - s$

Table 1: Toy grammar used to demonstrate verb forgetting. Nouns are postmodified with probability m ; a postmodifier is a relative clause with probability r , and a relative clause is V-initial with probability s . For practical reasons we bound non-terminal rewrites of NP at 2.

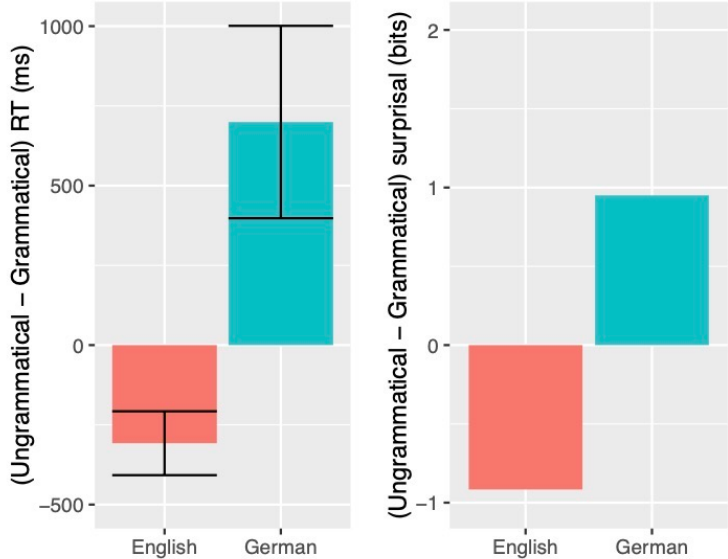


Figure 1: Differences in reaction times for ungrammatical continuations minus grammatical continuations, compared to noisy surprisal differences. RT data comes from self-paced reading experiments in Vasishth et al. (2010) in the post-VP region. The noisy surprisal predictions are produced with $d = .2$, $m = .5$, $r = .5$ fixed, and $s = .8$ for English and $s = 0$ for German.

NP-Missing in Mandarin Chinese (Hang et al., 2022)

- 汉语双重嵌套RCs中存在类似的结构性遗忘，但表现为 NP-missing;

- (a) * The apartment_{NP1} that [the maid_{NP2} who [the cleaning service_{NP3} had sent over_{VP3}]_{orc} (missing _{VP2})]_{orc} was well-decorated_{VP1}.
- (b) (b) The apartment_{NP1} that [the maid_{NP2} who [the cleaning service_{NP3} had sent over_{VP3}]_{orc} was cleaning _{VP2} every week]_{orc} was well-decorated_{VP1}.
- (c) 总理正在_{V1} 会见[曾_{V2} 责备过[刚_{V3} 审理_{NP1} 贪污案不久的]_{NP2} 法官好几次的]_{NP3} 部长。
- (d) *总理正在_{V1} 会见[曾_{V2} 责备过[刚_{V3} 审理_{NP1} 贪污案不久的]_{NP2} 法官]。

- LCS 不能解释汉语的NP-missing，使用上下文无关文法的概念模拟计算预测说明汉语与德语、荷兰语更相似;

Interference Theory

Haussler & Bader (2015)

- 解析器在构建句法结构时，会错误地将某个 VP 连接到主句，而不是相应的从句。这种错误连接导致VP省略，但解析器不会察觉到问题，从而形成结构性遗忘。

Bader (2016)

- 在英语中，由于主句和从句的句法结构相似，解析器更容易混淆两个 VP 的连接位置；
- 在德语中，由于主句和从句具有不同的动词位置，解析器更容易区分 VP 的连接位置，因此不会产生结构性遗忘。

Thematic Relations Perspectives (Hang et al., 2022)

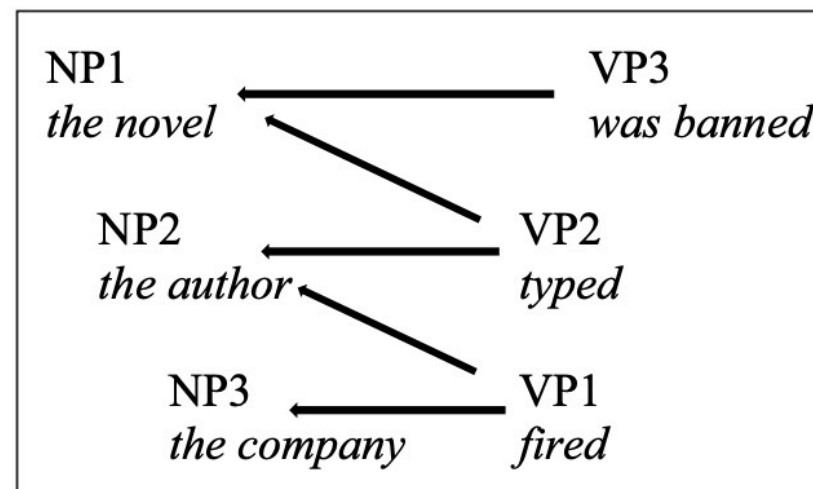
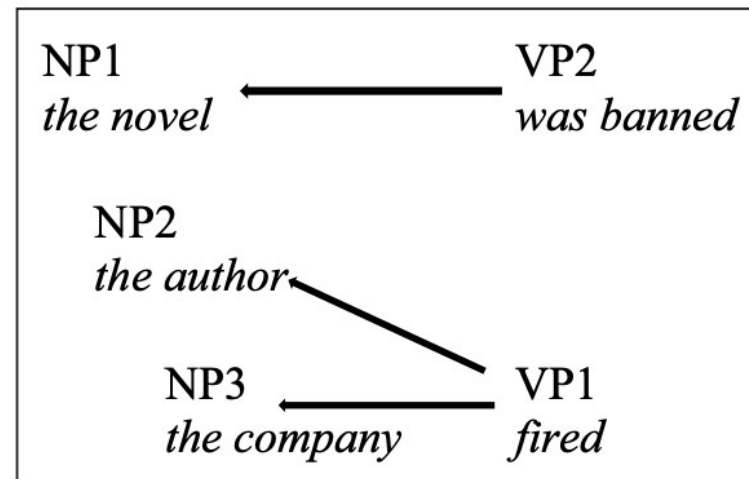
- 核心假设

- 解析器的核心目标之一是建立句中所有论元和谓词之间的语义关系，即谁做了什么。
- 如果所有的名词（NPs）和动词（VPs）都能成功形成合理的题元关系，那么即使句子在句法上不完整，它依然可能被误判为完整的。

[_{NP1} The novel] that [_{NP2} the horror author] who [_{NP3} the publishing company] [_{VP1} had recently fired] ---- [_{VP3} was banned by the local library].

- 其中的题元关系指向

- NP1 (the novel) → NP3 (was banned by the local library)
- NP2 (the horror author) → VP1 (had recently fired)
- NP3 (the publishing company) → VP1 (had recently fired)



Doubly Nested Relative Clauses in English

- ORC + ORC (n1 n2 n3 v3 v2 v1)

The manuscript_{n1} that [the graduate student_{n2} [who the new catalog_{n3} had confused_{v3}]_{orc} was studying_{v2}]_{orc} was missing_{v1} a page.

- SRC + SRC (n1 v2 n2 v3 n3 v1)

The scientist_{n1} [who admire_{v2} the professor_{n2} [who won_{v3} the award_{n3}]_{src}]_{src} wrote_{v1} a paper.

- ORC + SRC (n1 n2 v3 n3 v2 v1)

The manuscript_{n1} that [the graduate student_{n2} [who won_{v3} a prize_{n3}]_{src} was studying_{v2}]_{orc} was missing_{v1} a page.

- SRC + ORC (n1 v2 n2 n3 v3 v1)

The scientist_{n1} [who admire_{v2} the professor_{n2} [who the new catalog_{n3} had confused_{v3}]_{orc}]_{src} wrote_{v1} a paper.

Doubly Nested Relative Clauses in Mandarin Chinese

- ORC + ORC (v1 n2 v2 n3 v3 n1)

张三爱上了v1[[小红n2喜欢v2的]同学n3赶走v3了的]小明n1。

- SRC + SRC (v1 v2 v3 n3 n2 n1)

张三爱上了v1[赶走了v2[喜欢v3小红n3的]同学n2的]小明n1。

- ORC + SRC (v1 v2 n2 n3 v3 n1)

张三爱上了v1[[喜欢v2小红n2的]同学n3赶走了v3的]小明n1。

- SRC + ORC (v1 v2 n3 v3 n2 n1)

张三爱上了v1[赶走了v2[小红n3喜欢v3的]同学n2的]小明n1。

尚未解决的问题们

- 汉语究竟存在ORCs加工优势还是SRCs加工优势？
- 汉语的NP-Missing为何不能用LCS解释？
 - Huang et al. (2021) 的结果是否可靠？
 - 其仅使用基于上下文无关文法的概念模型进行计算；
 - 课进行自定义步速阅读实验，使用RR-LCS Model (Hahn et al., 2022) 估计lcs，拟合RTs；
- 汉语的NP-Missing该如何解释，又如何验证？
 - 如何更好地结合memory-based和expectation-based model？
- 汉语的RCs加工不对称性与NP-Missing之间的关系？其他嵌套组合是否会产生类似的结果？

.....

Reference

- Ford, M. (1983). A method for obtaining measures of local parsing complexity throughout sentences. *Journal of verbal learning and verbal behavior*, 22(2), 203-218.
- Frank, S. L., Trompenaars, T., & Vasishth, S. (2016). Cross-linguistic differences in processing double-embedded relative clauses: Working-memory constraints or language statistics?. *Cognitive science*, 40(3), 554-578.
- Frazier, L. (1987). Syntactic processing: evidence from Dutch. *Natural Language & Linguistic Theory*, 5(4), 519-559.
- Futrell, R., & Levy, R. (2017, April). Noisy-context surprisal as a human sentence processing cost model. In *Proceedings of the 15th conference of the european chapter of the association for computational linguistics: Volume 1, long papers* (pp. 688-698)
- Gibson, E., & Thomas, J. (1999). Memory limitations and structural forgetting: The perception of complex ungrammatical sentences as grammatical. *Language and Cognitive Processes*, 14(3), 225-248.
- Gibson, E. (1998). Linguistic complexity: Locality of syntactic dependencies. *Cognition*, 68(1), 1-76.
- Gibson, E., & Wu, H. H. I. (2013). Processing Chinese relative clauses in context. *Language and Cognitive Processes*, 28(1-2), 125-155.
- Hale, J. (2001). A probabilistic Earley parser as a psycholinguistic model. In *Second meeting of the north american chapter of the association for computational linguistics*.

- Hao, H., Yang, Y., & Hahn, M. (2024). Information Locality in the Processing of Classifier-Noun Dependencies in Mandarin Chinese. In *Proceedings of the Annual Meeting of the Cognitive Science Society* (Vol. 46).
- Hauser, M. D., Chomsky, N., & Fitch, W. T. (2002). The faculty of language: what is it, who has it, and how did it evolve?. *science*, 298(5598), 1569-1579.
- Holmes, V. M., & O'Regan, J. K. (1981). Eye fixation patterns during the reading of relative-clause sentences. *Journal of Verbal Learning and Verbal Behavior*, 20(4), 417-430.
- Huang, N., & Phillips, C. (2021). When missing NPs make double center-embedding sentences acceptable. *Glossa: a journal of general linguistics*, 6(1).
- Hsiao, F.P., & Gibson, E. (2003). Processing relative clauses in Chinese. *Cognition*, 90, 3-27.
- Lin, C. J. C., & Bever, T. G. (2006, September). Subject preference in the processing of relative clauses in Chinese. In *Proceedings of the 25th west coast conference on formal linguistics* (Vol. 25, pp. 254-260).
- McCurdy, K., & Hahn, M. (2024, November). Lossy context surprisal predicts task-dependent patterns in relative clause processing. In *Proceedings of the 28th Conference on Computational Natural Language Learning* (pp. 36-45).
- Vasishth, S., Suckow, K., Lewis, R. L., & Kern, S. (2010). Short-term forgetting in sentence comprehension: Crosslinguistic evidence from verb-final structures. *Language and Cognitive Processes*, 25(4), 533-567.
- Yang, Q., Yang, C., Liao, Y., & Zhuang, Z. (2024). Grammaticality illusion or ambiguous interpretation? Event-related potentials reveal the nature of the missing-NP effect in Mandarin centre-embedded structures. *arXiv preprint arXiv:2402.11282*.