

The role of discourse-level predictability in sentence comprehension

Edward Gibson¹, Ev Fedorenko¹ & Tomoko Ishizuka²

¹MIT Department of Brain and Cognitive Sciences, ²UCLA Department of Linguistics
egibson@mit.edu, evelina9@mit.edu, zukka@ucla.edu

BACKGROUND

Relative clause complexity effects have been studied extensively in **null contexts**:

Object-extraction: The man who the bouncer punched ...
Subject-extraction: The man who punched the bouncer ...

Null context observation: *Object-extractions are more complex than subject-extractions.* (e.g., King & Just, 1991; Gibson, 1998, 2000; Gordon et al., 2001, 2004; Grodner & Gibson, 2005; Lewis & Vasishth, 2006).

Theories: Working-memory-based.

Concern: The processing of some syntactic constructions may be **context-dependent**: the **interdependence of syntactic structure and function.** E.g., Chafe (1967, 1987); Givón (1979, 1984); Halliday (1970); Prince (1981); Gundel et al. (1988).

Previous work: Interactions between context and syntactic structural preferences (Crain & Steedman, 1985; Altmann & Steedman, 1988)

MV/RR ambiguity in reading: e.g., Spivey-Knowlton & Tanenhaus (1994); c.f. Ferreira & Clifton (1986).
PP attachment in spoken language: Tanenhaus et al. (1995).

Information structure and syntax:

Relative clauses in context: Gibson et al. (2005).

Main clause word order: Finnish (Kaiser & Trueswell, 2004); Russian (Fedorenko & Levy, 2006, 2007); English (Brown et al., 2007).

The goal of the current research: to investigate whether the processing of RCs is affected by (a) *contextual support*; and (b) *contextual predictability*. If the RC complexity effects disappear in contextually-supported environments, then the working memory accounts of the observed difference may be undermined.

EXPERIMENT 1

Design & Materials

Method: Self-paced reading (24 participants, 16 items).

Materials: Each **target item** consists of:

- Two context sentences (*sentence-by-sentence presentation*)
- A statement and question from "Bill" (*sentence-by-sentence presentation*)
- A statement by "Susan" in response to "Bill", including the **target RC**. (*word-by-word presentation*)

Two men visited a bar.

A bouncer punched one of them, and then the other man punched the bouncer.

Bill: I think that the bar owner knows one of the two men very well. Which one?

Susan: The man who **the bouncer punched / punched the bouncer** was the one who the bar owner knows very well.

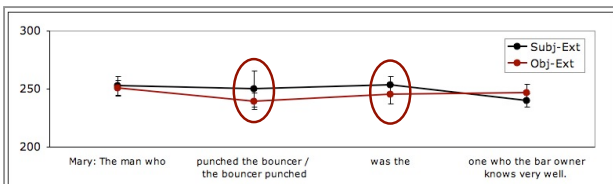
Contextual predictability: A modified NP is expected as the answer to Bill's question, likely a relative clause: *the man + modifier*

Filler items: The 32 filler items had the same general format (two context sentences, then a statement and question by "Bill", followed by a word-by-word response from "Susan"):

- 16 included different kinds of questions from Bill (e.g., "Is that right?" "WTF?");
- 16 were like the targets in distinguishing two relevant entities in the context with the same head noun *N*, but (a) none had the format of the targets (two *N*'s; *N*₁ acts on *X*; *X* acts on *N*₂); and (b) none of Mary's responses were initiated "The *N* ..."

Therefore, the context from the preceding sentences and the current sentence context causes a **100% expectation for an RC: 50% for each of the possible continuations.**

Results



No reliable RT differences in any region ($F_s < 1$)

Response accuracies: Object-extractions: 88.5%; Subject-extractions: 89.0%; $F_s < 1$

EXPERIMENT 2

Design & Materials

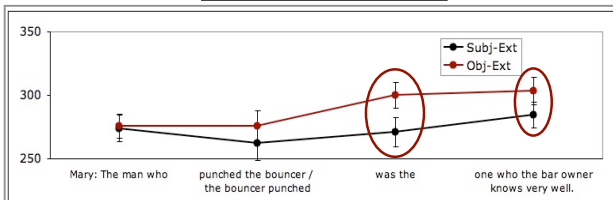
Method: Self-paced reading (30 participants, 16 items).

Materials: Same target items as Experiment 1, plus the same 32 filler items from Experiment 1.

Filler items: 16 additional filler items were included, all with the same sentential context as the targets (two *N*'s; *N*₁ acts on *A*; *B* acts on *N*₂); Eight of these had passive RC continuations (e.g., "The horse that was kicked by the mule ...")

The target RC was therefore **not easily predictable from the experimental context.** (See Ishizuka, Nakatani & Gibson, 2006, for a similar design, with the same results.)

Results



RTs: Obj-extractions slower than subject-extractions

RC region: $F(1,29)=3.12$; $p = .09$; $F(2,15)=2.04$; $p = .18$ (marginal / numerical)

Main verb region: $F(1,29)=10.3$; $p < .005$; $F(2,15)=16.4$; $p < .001$

Rest of sentence: $F(1,29)=15.4$; $p < .001$; $F(2,15)=15.5$; $p < .001$

Response accuracies: Object-extractions: 81.7%; Subject-extractions: 81.7%; $F_s < 1$.

EXPERIMENT 3

Design & Materials

Method: Self-paced reading (32 participants, 24 items, 48 fillers).

Materials:

Each **target item** consists of:

- Two context sentences (*sentence-by-sentence presentation*)
- A statement by "Mary" including the **first target RC** (*region-by-region presentation*)
- A statement by "John" in response to "Mary", including the **second target RC**. (*region-by-region presentation*)

After the meeting, a manager and two administrators examined the accounting books. One of the administrators questioned the manager and then the manager questioned the other administrator. Mary: I heard that the administrator that **the manager questioned / questioned the manager** had a problem with the company.

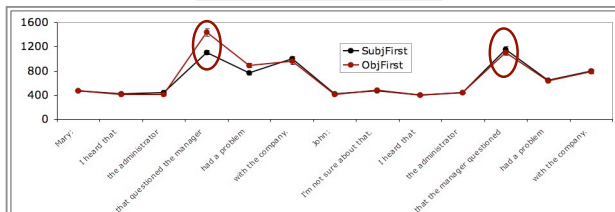
John: I'm not sure about that. I heard that the administrator that **questioned the manager / the manager questioned** had a problem with the company.

To control for plausibility, the order of the nouns acting on each other was included as a factor (e.g., Order 1: 1 manager & 2 administrators; Order 2: 2 managers & 1 administrator)

Contextual predictability:

- The first RC is **contextually supported, but unpredictable** in the preceding context.
- The second RC is **contextually supported, and highly predictable** in the preceding context.

Results



RC RTs:

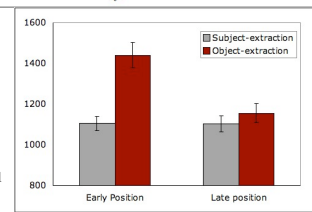
Early region: $F(1,31)=16.6$; $p < .001$;

$F(2,123)=31.3$; $p < .001$

Late region: $F_s < 1$

Response accuracies:

Obj-ext-first: 75.0%; Subj-ext-first: 76.3%; $F_s < 1$



SUMMARY AND CONCLUSIONS

Summary:

When the existence of an RC is highly predictable dependent on the previous context, we observed no differences between subject and object-extracted RCs.

However, when an RC was not highly predicted by the preceding context, there was a reliable difference between subject- and object-extracted RCs, even in contextually supported conditions.

Conclusions:

- As in syntax (Hale, 2001; Levy, 2005), contextual predictability plays an important role in sentence comprehension (c.f., Konieczny, 2000).
- Contextual support does not seem to affect the subject- vs. object-extraction complexity effect, when predictability is factored out. If anything, the effects appear to be more robust in supportive and not highly predictive contexts than in null contexts.