

Who next? Consequences of syntax-semantics mismatches for likelihood of mention

Elsi Kaiser, Edward Holsinger & David Li {emkaiser, holsinge, lidc} @ usc.edu
University of Southern California

CUNY 2011

Introduction

Background: Likelihood-of-mention in subsequent discourse

- A fundamental question underlying language production and comprehension: Referent-tracking, including *who speakers will talk about next*.
 - Likelihood-of-mention is connected to referents' cognitive accessibility, likelihood of pronominalization, which is influenced by multiple factors, including both syntax and semantics (e.g. Arnold 2008)
 - Likelihood-of-mention is primarily sensitive to semantic information; syntactic information matters for choice of referring expression (Kehler et al. 2008, see also Fukumura & van Gompel 2010)
- How much does syntactic information (grammatical role) matter when it comes to likelihood-of-mention? Is it 'swamped' by semantic information?
- Probe this question by looking at situations involving *syntax-semantics mismatches*

Research questions: To better understand how syntactic and semantic information modulate discourse flow => Do (mis)matches in syntactic and thematic prominence influence likelihood-of-mention?

- Syntactic prominence:** {subject > object}
- Thematic prominence:** (Belletti & Rizzi 1988, Grimshaw 1990, Jackendoff 1990, etc etc)
 - {agent > patient}
 - {experiencer > stimulus}

Exp 1: Agent-patient verbs: mismatch in passive

Mary_{agent} tickled Lisa_{patient}.
Lisa_{patient} was tickled by Mary_{agent}. ← Thematically prominent, but syntactically non-prominent position

Exp 2: Stimulus-experiencer verbs: no mismatch in passive

Mary_{stimulus} annoyed Lisa_{experiencer}.
Lisa_{experiencer} was annoyed by Mary_{stimulus}.

Exp 1: Agent-Patient verbs

- Is likelihood-of-mention influenced by (i) grammatical role, (ii) thematic role, or (iii) interplay of syntax and semantics? => specifically: *Mismatch* in syntactic prominence and thematic prominence
- Sentence continuation w/ **agent-patient verbs** (e.g. kicked, tickled, slapped). Voice was manipulated:
 - Active voice:** Syntactic and thematic prominence **match**
 - (1) Mary_{agent} slapped Lisa_{patient}.
 - Passive voice:** Syntactic and thematic prominence **mismatch**
 - (2) Lisa_{patient} was slapped by Mary_{agent}.
 - Prominent thematic role in syntactically non-prominent, non-canonical position (*by*-phrase) => mismatch
 - by*-phrase is optional, relatively infrequent in corpora
- Task:** See picture, hear auditory prompt, provide spoken continuation. Also eye-tracked (data not reported here).
- 24 targets, 36 fillers, 24 participants
- Coherence relations known to influence likelihood-of-mention -- control this by using causal relations. => With agent-patient verbs, **causal relation** focuses **patient** (object in active voice) (e.g. Kehler 2002)



Sample image

- (3a) Mary_{agent} slapped Lisa_{pat} at the zoo. {As a result/Then}... [Lisa?]
- (3b) Lisa_{pat} was slapped by Mary_{agent} at the zoo. {As a result/Then}... [Lisa?]
- 'Then' is ambiguous between causal and narrative/temporal use, included to see if need for decision-making / active processing strengthens causal focusing effects. (No)

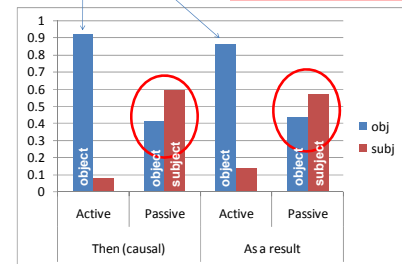
Terminology: **Subject** => grammatical subject. **Object** => direct object or object in *by*-phrase

Results

Which character do people start their continuations with, preceding subject or object?

With actives, significant object/patient preference (as expected) p's <.01

With passives, **competition** between subject/patient and object/agent. (Choices do not differ significantly from chance.)



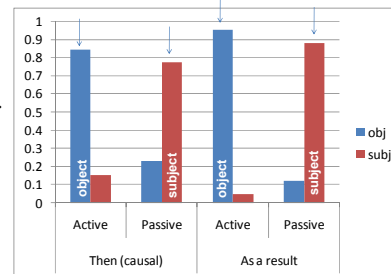
- Expectations influenced by **mismatch-triggered inferencing**:
- Encountering a **thematically-prominent argument** in a syntactically low-prominence, optional position => **signals that referent will be mentioned again**
- Marked configuration => triggers in an implicature?
- Seems to disrupt thematic-role related expectation

Exp 2: Stimulus-Exp verbs

- Q1** = Could effects in Exp 1 be due to word order?
 - e.g. preference for the patient modulated by linear recency
- Q2** = What happens when there is no syntax-semantics mismatch?
 - 24 new participants.
 - Same set-up, same task but with stimulus-experiencer verbs (e.g. annoyed, amused, frightened)
 - (4a) Mary_{stim} annoyed Lisa_{exp} at the zoo. {As a result/Then}...
 - (4b) Lisa_{exp} was annoyed by Mary_{stim} at the zoo. {As a result/Then}...
- Stim-Exp verbs followed by **causal relation** => focuses **experiencer** (e.g. Stevenson et al 1994, Rohde 2008, work on IC-1 verbs)
- Passivization of a stimulus-Exp verb does **not result in syntax-semantics mismatch**:
 - Object in *by*-phrase is thematically lower-ranked: {experiencer > stimulus}

Results

Which character do people start their continuations with, preceding subject or object?



Both actives and passives show **experiencer preference** (p's <.01)
=> Object with actives
=> Subject with passives
No clear effects of voice

Exp. 1 results not simply due to linear recency.

When referent in *by*-phrase does not involve syntax-semantics mismatch, no special future status.

Conclusions

What influences likelihood-of-mention in subsequent discourse?

- Likelihood-of-mention is influenced by the **interplay of syntax and semantics**.
 - Oversimplification to say that a particular syntactic or thematic role is consistently correlated with increased likelihood-of-mention.
- What's important: **Mapping between syntax and semantics**
- Consequences of 'demoting' a thematically-prominent entity by a non-canonical syntactic construction (passivization).
- Comprehenders draw inferences from unusual argument configurations.
- As a whole, these findings
 - (i) corroborate existing work regarding the **effects of thematic roles and coherence relations** (=>causal relations focus patient/experiencer role)
 - (ii) but show that syntax-semantics mismatches **can disrupt these effects** (Exp 1).
- Our results argue against a purely syntactic or a purely semantic/thematic approach to reference tracking => rather, interactive referent-tracking system, sensitive to markedness and mismatches between syntactic and semantic prominence.

Thanks to Jackie Kim and Monica Do for help with the experiments.
This research was supported by NIH grant 1R01HD061457.