Coreference and focus in human sentence processing

Wind Cowles
Language and Cognition Lab
Department of Linguistics

Over-arching research interest

• Impact of optionality in natural human language
  • Why? Ubiquitous and linguistically universal
  • Applies at two key levels:
    • Multiple syntactic structures possible for same information
    • Multiple referential forms possible for same referent

At the heart of structural optionality: Information structure

• Lambrecht (1999) - Independent component within sentence grammar

Syntax  Information Structure  Semantics

• Function that mediates between the form of utterances and the current mental states of interlocutors

• (Alternative, compatible approach: Jackendoff (2002) - Interface between components within the grammar)

Formal instantiations of information structure

• Information structure is formally realized via:
  • Prosody
  • Specialized (morpho-)syntactic markers
  • Ordering and positioning of syntactic constituents
  • Particular grammatical constructions
  • Certain choices among lexical options
### Major categories: Topic and Focus

**What did the player do?**

The player dropped the ball.

**Topic**

- Aboutness condition
- File card metaphor

**Focus**

- Informative part
- “pragmatically non-recoverable”

---

**What happened?**

The player dropped the ball.

**Topic**

- Aboutness condition
- File card metaphor

**Focus**

- Informative part
- “pragmatically non-recoverable”

---

### Formal instantiations (generalized)

<table>
<thead>
<tr>
<th></th>
<th>Topic</th>
<th>Focus</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Prosody</strong></td>
<td>reduced pitch accents</td>
<td>prominent pitch accents</td>
</tr>
<tr>
<td><strong>Specialized (morpho-)syntactic markers</strong></td>
<td>e.g. -wa in Japanese</td>
<td>e.g. only in English</td>
</tr>
<tr>
<td><strong>Ordering/positioning of syntactic constituents</strong></td>
<td>early</td>
<td>late</td>
</tr>
<tr>
<td><strong>Particular grammatical constructions</strong></td>
<td>fronting</td>
<td>clefts</td>
</tr>
<tr>
<td><strong>Certain choices among lexical options</strong></td>
<td>pronouns</td>
<td>fuller forms of reference</td>
</tr>
</tbody>
</table>
Questions for Human Coreference Processing

• How is the antecedent referent determined?
  • What are the influential factors?
    • Contextual, antecedent features, anaphor features
  • What is the time course of this process?

Different types of anaphoric expressions

<table>
<thead>
<tr>
<th>in focus</th>
<th>activated</th>
<th>familiar</th>
<th>uniquely identifiable</th>
<th>referential</th>
<th>type identifiable</th>
</tr>
</thead>
<tbody>
<tr>
<td>it</td>
<td>this/that</td>
<td>that N</td>
<td>the N</td>
<td>(indef) this N</td>
<td>a N</td>
</tr>
</tbody>
</table>

Interactions with prominent/focused antecedents

A boxer and a trainer entered the weight room.
The boxer talked to the trainer.

Interactions with prominent/focused antecedents

A boxer and a trainer entered the weight room.
The boxer talked to the trainer.

**Pronominal**  He ...
**Categorical-Def** The athlete ...
**Categorical-Dem** That athlete ...
Repeated The boxer ...

... was preparing for the upcoming competition.

**Pronominal**  He ...
**Categorical-Def** The athlete ...
**Categorical-Dem** That athlete ...
Repeated The boxer ...

**Prominence “preference”**
Pronominal co-reference to prominent antecedents leads to faster reading times - even if they are unmentioned.
(e.g. Cornish, Garnham, Cowles, Fossard & Andre, 2005)
Interactions with prominent/focused antecedents

A boxer and a trainer entered the weight room. The **boxer** talked to the trainer.

**Inverse Semantic Distance Effects**

*Categorical co-reference to atypical exemplars in prominent syntactic positions leads to faster reading times (e.g. Almor, 1999; Cowles & Garnham, 2005)*

<table>
<thead>
<tr>
<th>Pronominal</th>
<th>Categorical-Def</th>
<th>Categorical-Dem</th>
</tr>
</thead>
<tbody>
<tr>
<td>He</td>
<td>The athlete</td>
<td>That athlete</td>
</tr>
<tr>
<td>Repeated</td>
<td>The boxer</td>
<td></td>
</tr>
</tbody>
</table>

**Repeated Name Penalty**

*Repeated co-reference to prominent antecedents leads to slower reading times (e.g. Gordon, Grosz & Gilliom, 1993; Gordon et al. 1999; Almor, 1999) but not always (Cowles & Dawidzuk, in press)*

Typicality Effects

1. The professor and her student arranged the transportation for their field trip.
2. She rented a car/boat for the second stage of the trip.
3. The vehicle was necessary for getting to the exploration site.

*e.g. Garrod & Sanford (1977)*
Inverse Typicality Effects

1. The professor and her student arranged the transportation for their field trip.
2a. It was the student that rented the car/boat.
2b. What the student rented was the car/boat.
3. The vehicle was necessary for getting to the exploration site.

Beyond Typicality and Clefts

- Cowles & Garnham (2005)
- Two Experiments Tested Conceptual and Inverse Conceptual Distance Effects
  - Exp 1: Clefts
  - Exp 2: No Clefts
    - 28 native English participants, 24 items

Exp 1: Design & Sample Materials

- Most Specific
- Distant Antecedent
- Close Antecedent
- Least Specific
- Anaphor

Focus Antecedent Setup:
- What the mongoose stood up to was the cobra.

Non-focus Antecedent Setup:
- It was the mongoose that stood up to the cobra.

Target sentence:
- The reptile hissed and got ready to strike.

Methods - Procedure

- Self-paced reading
**Results - Reading time at the anaphor**

*The reptile hissed and got ready to strike.*

![Graph showing reading times for subject NP (msec)]

- Non-Focus: 584, 584
- Focus: 574, 542

- 28 native English participants, 24 items

---

**Results - Residual time at the predicate**

*The reptile hissed and got ready to strike.*

![Graph showing residual times for subject NP (msec)]

- Non-Focus: 145, -23
- Focus: 623, -139

- 28 native English participants, 24 items

---

**Experiment 2: Sample Materials**

*Most Specific* — Distant Antecedent

*Least Specific* — Close Antecedent

Focus Antecedent Setup:
The snake/cobra frightened the hunter.

Non-focus Antecedent Setup:
The hunter was frightened by the snake/cobra

Target sentence:
The reptile looked ready to strike at once if threatened.
Results - Residual time at the predicate

The reptile *hissed and got ready to strike.*

Beyond Typicality and Clefts

- Inverse semantic distance effects are not limited to clefts or specific to typicality

- What does this mean?
  - Janus (Garnham and Cowles, 2008)
    - Anaphoric processing has two aspects:
      - Antecedent identification
      - Discourse structuring