To what extent does the difficulty of processing relative clauses parallel typological complexity?

Claire Delle Luche
Frédérique Gayraud
Bruno Martinie
Fanny Meunier-Hoen

Laboratoire Dynamique du Langage, Université Lyon 2 Lumière, Lyon, France
Complexity and language

- **Quantity = Complexity?**
  
  (1) *Please pick up four tomatoes, a pound of apricots, prune juice, shallots, six apples and a bag of carrots on the way home.*

  (2) *The man that the woman that the child hugged kissed laughed.*

  Quantity of information is not the predominant factor of complexity

- **Difficulty = Complexity?**
  
  – (2) is more difficult to process than (1), the syntactic structure of (2) is more complex than (1)’s.

  Difficulty of processing may be a way to test the language complexity.
Introduction (1)

Relativization in general

- RCs are sentential modifiers of nouns.
- The operation of relativization can be marked by different ways, but in any case, the relativized constituent is moved from its canonical position.
- Main idiosyncratic properties in the world languages:
  - The relativizer can be compulsory (French), optional (English, Arabic) or never available (Japanese).
  - Possible morphological variation of the mark according to the relativized position (French vs. Chinese).
  - Possible marking by special verbal morphology (Turkish).
  - In the RC, presence of a resumptive pronoun coindexed with the antecedent (Arabic, Hebrew etc.) or gap (English, French…).
Introduction (2)

Relativization in standard French

Example: center-embedded O RC

(3) L’homme fume. Le chien chasse l’homme.

(4) L’homme que le chien chasse fume.

- The object NP is moved from its canonical postverbal position
- Presence of a relativizer
- The morphology of the relativizer (QUE) indicates the relativized position (Object)
- Main relativizable positions in French: Subject (QUI), Object or Attribute (QUE), Indirect Object (Prep+ Relative pronoun or DONT), Genitive (Prep+ Relative pronoun or DONT)
Introduction (3)

  
  (5) The reporter that the senator attacked admitted the error.
  (6) The reporter that attacked the senator admitted the error.

- Several theoretical attempts to account for the S O difference
  - Canonical word order
  - Functions parallelism
  - Depth of embedding
  - Working memory cost
Canonical word order

- French canonical word order = SVO
- S RC: SV(O)
  
  (7) Le sénateur qui attaque le reporter admet l’erreur.
  
  (The senator that attacked the reporter admitted the error.)

- O RC: OSV
  
  (8) Le sénateur que le reporter attaque admet l’erreur.
  
  (The senator that the reporter attacked admitted the error.)
In a language acquisition perspective, Sheldon (1974) claims that (9) is easier to process than (10) and (11), because only in (9), the relativized NP *The singer* has the same function both in the matrix and in the relative clause:

(9) *She despises the singer that you admire.*

(10) *The singer that you admire is on stage tonight.*

(11) *She despises the singer that is on stage.*
Hawkins (1994) : in a language such as French, a S RC is predicted to be less complex to process than a O RC. The reason is the following : in O RCs, the size of the portion of the tree involved in the coindexation between t and the head noun is larger.
Working memory accounts of RC processing:

- King, & Just (1991)
- Caplan, & Waters (1999)
- Gibson (1998, 2000): the DLT, a computational approach
The DLTheory

- Two processing costs: memory and integration costs

Predictions

(8) O RC: Le sénateur que le reporter attaque admet l’erreur.

(7) S RC: Le sénateur qui attaque le reporter admet l’erreur.

Extended predictions

(12) OG RC: Le sénateur dont le reporter attaque le parti admet l’erreur.

(The senator whose party the reporter attacked admitted the error.)
A typological approach

The Accessibility Hierarchy (Keenan, & Comrie, 1987):
S > O > IO > OBL > G > OComp

Principles:
- « A language must be able to relativize [S position].
- [Strategies apply] on a contiguous segment of the AH.
- Strategies can that apply at one point of the AH may cease to apply at any lower levels. » (Keenan, & Comrie, 1987, p. 6)

Examples:
- S only: Malagasy
- S-G: French
- S-OComp: English
The AH and comprehension


Materials: 36 sentences (9 different positions in the AH)

(13) S: *I know that the girl who got the answer right is clever.*
(14) O: *He remembered that the sweets which Dave gave Sally were a treat.*
(15) OG: *The fact that the sailor whose ship Jim took had one leg is important.*
(16) OComp: *He remembered that the girl who Jane is older than could read.*

Results: the AH is supported by adult and children’s data

Comments:
- S, IO, Obl: one animate noun; SG, OG, OComp…: two animate nouns
- The procedure mixes production and comprehension
Predictions

- Canonical word order
- Functions parallelism
- Depth of embedding
- Working memory cost

S > O, IO, OBL, G
S > O > IO > OBL > G
Our experiment

Method

- Participants: 30 native speakers of French (21.26, s.d. 2.49)
- Materials: 36 experimental sentences, 3 conditions (S, O, OG)
  17) S: La danseuse qui regarde le chanteur appelle le studio (The dancer that is looking at the singer is calling the studio)
  18) O: La danseuse que le chanteur regarde appelle le studio (The dancer that the singer is looking at is calling the studio)
  19) OG: La danseuse dont le chanteur regarde les jambes appelle le studio (The dancer whose legs the singer is looking at is calling the studio)
- Procedure: Self-paced word-by-word, moving window display
  Comprehension question after each trial, no feedback
Results

- Participants with more than 35% of extreme reading times (<100ms or >1500ms) were discarded, thus leaving 27 for statistical analysis.

- Comprehension errors:
  Significant effect of RC type

![Mean error rate graph]

Approaches to complexity in language,
Helsinki, Finland, August 26 2005
Regions of interest:
- Object NP of the S and OG RCs
- Embedded verb
- Main clause verb

Mean reading times for each region
Object NP of the S and OG RCs
– No difference
Reading times (3)

- RC verb
  - Effect of RC type

![Graph showing the mean reading times for different RC types.](chart)

Mean RT on the RC verb

- S RC: 400 ms
- O RC: 600 ms
- OG RC: 600 ms

* indicates a significant difference.
Main verb
- No significant difference

Mean RT on the main verb

<table>
<thead>
<tr>
<th></th>
<th>S RC</th>
<th>O RC</th>
<th>OG RC</th>
</tr>
</thead>
<tbody>
<tr>
<td>RT (ms)</td>
<td>400</td>
<td>500</td>
<td>500</td>
</tr>
</tbody>
</table>
Discussion (1)

- S RCs are faster to read and generate less errors than the other types (O, OG)
- Different from previous findings
- Different from what is predicted by the depth of embedding and DL theories
- Compatible with canonical word order and functions parallelism theories
Discussion (2)

On theories of language processing:

– Memory cost is not infirmed by canonical word order nor functions parallelism theories

– Necessity to disentangle canonical word order from functions parallelism
Further experiment

<table>
<thead>
<tr>
<th></th>
<th>Canonical +</th>
<th>Canonical -</th>
</tr>
</thead>
<tbody>
<tr>
<td>Parallel +</td>
<td><em>L’homme qui regarde la femme promène le chien.</em></td>
<td><em>La femme regarde le chien que l’homme promène.</em></td>
</tr>
<tr>
<td>Parallel-</td>
<td><em>L’hui regarde la femme qui promène le chien.</em></td>
<td><em>L’hui que la femme regarde promène le chien.</em></td>
</tr>
</tbody>
</table>
Discussion (3)

- Complexity = Difficulty?
- Structural complexity (depth of embedding, number of syntactic nodes, long distance dependencies) does not automatically trigger processing difficulty
- Nevertheless it does not mean that complex sentences are easy to process
  
  *The lion that the gorilla that the tiger bit chased died*

- Difficulty is not equivalent to complexity
### Why?

<table>
<thead>
<tr>
<th></th>
<th>Complex structure</th>
<th>Simple structure</th>
</tr>
</thead>
<tbody>
<tr>
<td>Processing difficulty</td>
<td><strong>YES</strong></td>
<td>?</td>
</tr>
<tr>
<td>S RC &lt; O RC</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Processing ease</td>
<td><strong>Possible</strong></td>
<td><strong>YES</strong></td>
</tr>
<tr>
<td>O RC = OG RC</td>
<td></td>
<td>S RC &lt; O RC</td>
</tr>
<tr>
<td>S RC &lt; O RC</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>