Processing Turkish canonical and non-canonical sentences in context

Yuki HIROSE
(The University of Tokyo, yuki_hirose@mac.com)

INTRODUCTION

Aim: To examine the impact of the information status of the subject and object nouns on the processing of canonical (SOV) and non-canonical (OSV) sentences in Turkish.

Major findings:
1. Unlike previous studies, new-old information order was easier to process than old-new information order in Turkish.
2. Non-canonical sentences were read faster than canonical sentences at the matrix verb.

Background: Canonical sentences are generally easier to process than non-canonical sentences (Sakurada, 2003).

Three major factors:
1. Structural complexity: Canonical sentences are structurally less complex.
2. Structural frequency: Canonical sentences occur more frequently.
3. Discourse function: Discourse-old referents tend to occur earlier than discourse-new referents.

Evidence for discourse function from Finnish (Kaiser and Trueswell, 2004)
- When the subject and object are mentioned in the previous context, non-canonical (OVS) sentences are not more difficult to process than canonical (SVO) sentences.
- Processing difficulty of non-canonical sentences is due to violation of discourse factors, rather than structural complexity or frequency.

Remaining issues & motivation:
1. Kaiser & Trueswell investigate the processing of Finnish canonical (SVO) and non-canonical (OVS) sentences.
   - Can the impact of discourse be generalized to typologically distinct languages?
2. Kaiser & Trueswell used same nouns in preceding context and test sentences.
   - Are their results due to repeated noun benefit?
   - Using pronouns as discourse-old referents, the present study attempted to answer these questions in Turkish.

Basic facts about Turkish (Erguveril, 1979; Konrill, 1997)
- SOV language
- Scrambling is OK
- Overt case marking
- Pro-drop language
- Pre-verbal position is default focus position.

CONCLUSIONS & FUTURE STUDY

1. Canonical sentences are not always easier to process than non-canonical sentences.
2. Complexity or structural frequency alone cannot explain the processing of canonical SOV and non-canonical OSV sentences in Turkish.
3. Processing difficulty of non-canonical sentences may be due to violations of discourse demands, which may vary among languages.
4. Discourse function may also have some language-specific properties.

Future study:
- The observed processing pattern in the present study may be due to the use of pronouns.
- In order to examine the impact of noun-type on the processing of canonical and non-canonical sentences, we will conduct a follow up experiment with proper nouns.
- In order confirm how non-canonical sentences are used in the discourse, we will also conduct a corpus study.

EXPERIMENT

Methodology & Participants: Self-paced reading experiment; N = 35.

Experimental materials: 24 sets of target + 48 filler items.

RESULTS & DISCUSSION

EMBEDDED VERB
- Unexpectedly, the embedded verb was read faster after new-old noun order than old-new noun order.
- This result also cannot be captured by the discourse function account because it predicts an interaction in which the difficulty of non-canonical sentences is reduced after the old-new noun order.
- This result is inconsistent with the information structure of Turkish because discourse-old referents do not tend to occur in the pre-verbal focus position.
- The use of pronouns in preverbal position might have speeded up the reading times of embedded verbs.
- Some aspects of discourse demand may be language-specific.

MATRIX VERB
- Surprisingly, the matrix verb was read faster after non-canonical sentences than canonical sentences.
- In the canonical sentences, sentence-initial subject nouns might have been interpreted as subject of matrix clause; clause boundary ambiguity might have caused longer reading times in canonical sentences.
- The difference within the canonical sentences might be due to the use of proper nouns and pronouns. Proper nouns might be easier to be interpreted as the matrix subject compared to pronouns.
- Matrix subject might have been interpreted as the antecedent of the object noun in the canonical new-old condition. This binding ambiguity might have caused a processing cost.