The empirical probe: voice mismatch under ellipsis

Verb phrase ellipsis (VPE) resolution is strongly conditioned by identity (syntactic or semantic) between the antecedent and ellipsis clauses [3].

Jack wrote the advertisement. David did too. (Voice match)

*Jack wrote the advertisement. The poster was too. (Voice mismatch)

VPE constructions that lack identity, including those involving voice-mismatch, have shown gradient instead of categorical acceptability judgments. Discourse coherence [4], information structure [5], and processing strategies [6, 7] have all been proposed to account for the gradient acceptability.

We explore the possibility that at least part of the gradient acceptability arises from the memory retrieval process itself, due to the influence of “what could have been said”.

Experiment 1: the production baseline for active vs. passive structures

Experiment 1 replicated previous results in [8] to establish two classes of verbs (Experiencer and Regular) that have different production preferences for active and passive structures.

Thirty-two word pairs of tripletts were adopted and modified from Ferreira (1994), e.g. for an experiencer verb “Jack, advertisement, anger” and for a regular verb “Jack, advertisement, write”.

Examples of experiencer verbs we used: lire, amaze, impress, frustrate, excite, etc.

Examples of regular verbs we used: label, complete, mute, finish, reject, etc.

Participants (N=50, Mechanical Turk) were given two NPs and one verb and were instructed to create a full sentence. The verb was always the last element in the string and the order of the two NPs was randomized across trials.

Experiment 2: Acceptability Judgments

Participants (N=50, Mechanical Turk) rated acceptability (1-7 scale) on sentences (item N=32) like the example below.

2×2×2 design: the voice feature at the ellipsis site (active vs passive); the structural match between the antecedent and the ellipsis site (voice match vs. mismatch); and the type of verb in the antecedent clause (experiencer vs. regular verbs, the verbs are identical as Experiment 1).

Regular verb antecedent

a. Jack wrote the advertisement. David did too. (active-active)

b. Jack wrote the advertisement. The poster was too. (active-passive)

c. The advertisement was written by Jack. The quote was too. (passive-active)

d. The advertisement was written by Jack. David did too. (passive-passive)

Experiencer verb antecedent

a. The advertisement was written by Jack. The quote was too. (active-active)

b. The advertisement was written by Jack. David did too. (active-passive)

c. Jack was angered by the advertisement. The quote was too. (passive-active)

d. Jack was angered by the advertisement. David did too. (passive-passive)

e. Jack was angered by the advertisement. David did too. (active-active)

Results for Experiment 1 and 2

Experiment 1 results: Frequency of producing active, passive and other structures:

- significant effect of verb type (p<.0001)
- Participants produced predominantly active structures for the regular verbs, but many more passive structures for the experiencer verbs.

Experiment 2 results: Acceptability judgments (z-transformed):

- There is an expected penalty for mismatch on all voice mismatch conditions (Match vs. Mismatch, p<.0001).
- There is also a significant Verb x Voice interaction (p<.001), with an experiencer verb antecedent, the ellipsis site prefers passive voice, but with a regular verb antecedent, ellipsis prefers active voice.

Discussion and conclusion

Establishing the antecedent-ellipsis relationship is affected by the alternative structures with which the antecedent could have been expressed:

- If the antecedent verb prefers passive over the active structure (e.g. with experiencer verbs), this facilitates the processing of ellipsis sites with passive voice; the opposite patterns were observed if the antecedent verb prefers active over the passive structure (e.g. with regular verbs).

- This effect cuts across the factor whether structural identity holds between the antecedent and the ellipsis site. The effect is also present in online measures (SPR). It therefore appears that the influence of production alternatives from the antecedent is not a secondary strategy after the parser fails to establish identity.

- By incorporating production into the memory retrieval process, our findings provide a mechanistic link between production and comprehension. The proposal provides mechanistic instantiation to the "noisy channel" approach [9] to comprehension in general, and the "repair-based" analysis to voice mismatch under ellipsis in particular [7]. We also extend previous work that discussed the memory retrieval mechanism in ellipsis resolution [10].