Word-order uncertainty induces alternative, non-veridical structures in online comprehension

Gabriel Doyle and Roger Levy – UC San Diego

1. BACKGROUND
Sentence processing occurs in noisy conditions. But traditional approach does not account for noise:
- word identities & order set veridically before parsing starts
Previous work has found evidence for noisy representations of word identity:
- disambiguation by later context [1,2]
- regressive eye movements [3,4]
- non-veridical garden-pathing [5]

Is there noise in the word order representation too?
- potential sources: speaker error, perceptual noise, poor memory

2. NOISY WORD ORDER REPRESENTATION
We look for noise in the word order representation using sentences where an incorrect word order can temporarily create an alternative syntactic structure, e.g.:

```
The journalist that the fact surprised came to the press conference late.
```

2A. AT RC VERB
At RC verb (surprised), with noisy word order, two potential parses:

```
The journalist [that the fact surprised]RC (verb) (Veridical word order; relative clause)
The fact [that the journalist surprised (someone)]CC (Swapped word order; complement clause)
```

Swapped order CC becomes more likely at RC verb because:
- fact unlikely subject for RC verb
- journalist likely subject for RC verb

2B. AT MAIN VERB
Veridical-RC and swap-CC parses make different predictions about upcoming word:
- RC in veridical reading has its arguments, main verb expected
- CC in swap reading is missing an object, noun phrase expected

Swapped order CC becomes unlikely at main verb (wrote) because:
- no object for CC verb (surprised)
- CC verb unlikely to be intransitive
- fact implausible subject for main verb

2C. HYPOTHESIS
Changes in structural probability distribution at RCV & MV lead to processing slowdowns in both positions

3. STIMULI
2 ways to block the swap-CC parse:
- use who instead of that
- use RCNP strongly biased against CC

Each of these manipulations introduces confounds:
- run 2x2 to control for confounds
- look for superadditive slowdown on that−yesCC condition

2x2 design: \{ that \} \times \{ yesCC (NP likely to take a CC) \}

\[that-yesCC\] The journalist that the fact surprised came to the press conference late.
\[that-noCC\] The journalist that the article surprised came to the press conference late.
\[who-yesCC\] The journalist who the fact surprised came to the press conference late.
\[who-noCC\] The journalist who the article surprised came to the press conference late.

4. READING TIMES
Moving window self-paced reading time study, with 24 items, 68 participants. 90% accuracy on comprehension questions.

```
Predicted superadditive difficulty on that−yesCC appears at both RCV and MV
```

Not predicted by existing frameworks or word-identity uncertainty:
- inanimate RCNP should have similar interferences
- word-identity uncertainty wouldn’t differ across conditions

5. CONCLUSIONS
- Evidence that readers’ representations of the order of previously-read words are noisy
- These noisy representations lead to pursuit of parses inconsistent with true word order
- Consistent with rational maintenance of uncertainty in response to noisy input [6]

REFERENCES