When Speech Becomes Writing: The Case of Disfluencies

Aida Tarighat, Martin Corley

s2184935@ed.ac.uk | martin.corley@ed.ac.uk

For human readers, written disfluencies (e.g., *um*) render the text more speech-like which in turn opens it up to less literal interpretation.

literal fluent — I'm sure negative peer pressure leads to mostly **idiotic** decisions.

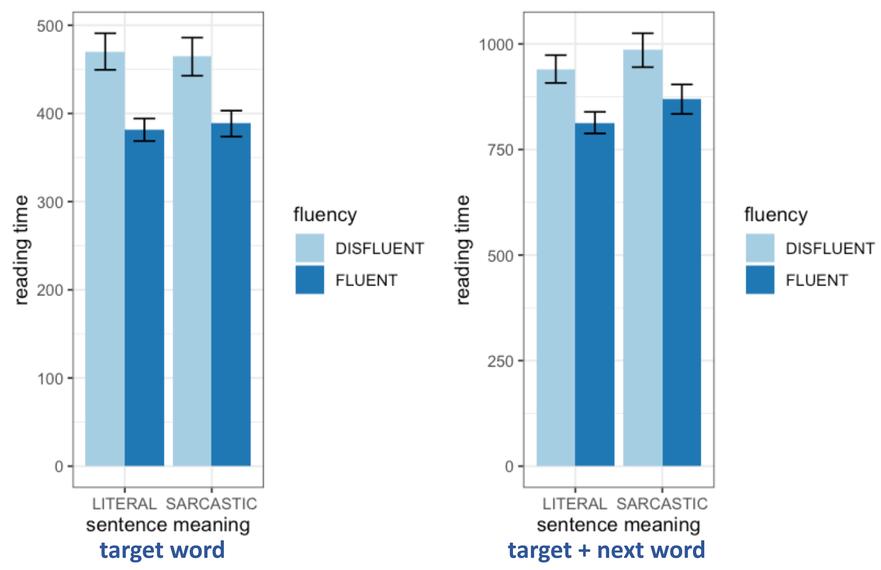
literal disfluent — I'm sure negative peer pressure leads to mostly ***um*** **idiotic** decisions.

non-literal fluent — I'm sure negative peer pressure leads to mostly **clever** decisions.

non-literal disfluent — I'm sure negative peer pressure leads to mostly ***um*** **clever** decisions.

- **When preceded by *um*, sentences compatible with non-literal meaning are not faster to read.**
- **Literal sentences are faster to read than non-literal sentences.**
- **Disfluent sentences are slower to read than fluent sentences.**

self-paced reading



99 participants, word-by-word self-paced reading of 24 items

eye-tracking experiment

natural reading prosody not disrupted

48 counterbalanced pairs, same character count:

`merry(literal)/feral (non-literal)`
`feral(literal)/merry (non-literal)`

commas before and after *um*:

`… be, um, merry when …`

UK Research and Innovation

THE UNIVERSITY of EDINBURGH informatics

THE UNIVERSITY of EDINBURGH School of Philosophy, Psychology and Language Sciences

NLP UKRI CENTRE FOR DOCTORAL TRAINING

Informatics Research Ethics Process no. 789617