



A Corpus-Based Analysis of Temporal Connectives Used in Language

Martha Valmana Crocker

Department Psychology, Western Institute for Neuroscience, University of Western Ontario

Background

- We are able to understand and communicate time through the use of human language.
- The relationship between time and language concerns the order in which we describe past, present and future events.
- Before* and *after* are temporal connectives that link multiple events in a sentence.
- Sentences tested in Munte et al. (1998) are in the form of *before* and *after* sentence-initial clauses.

Before sentence-initial clause example:

Before the scientist submitted the article, the journal changed its policy

After sentence-initial clause example:

After the scientist submitted the article, the journal changed its policy

- Results of the studies found that *before* sentence-initial clauses are harder to comprehend than *after* sentence-initial clauses (Munte et al., 1998).

Previous results suggest that comprehension difficulties occur when there is a mismatch between the real-world temporal order of events and the linguistically presented order of events (Politzer-Ahles et al., 2017). We have performed a corpus analysis to investigate additional factors that may influence the comprehension of *before/after* in natural language.

How is *before/after* used in natural language?

Methods

424 million sentences taken from Wikipedia (2018) and OpenSubtitles

9 million sentences contain *before* and/or *after*

2 million sentences are *before/after* - initial

40,000 *before/after* - initial sentences randomly selected

Corpus Coding

- First Word/ Phrase Category

Category	Example from Corpus
Event Noun	Before breakfast , he eats pomegranate.
Verbing	After joining the marine corps, he participated in World War II.
Date	Before 1866 , the city was patrolled by town constables.
Time Reference	After six days , the boats reached Cartagena.
Time + Event	After five years of playing football , I got a college degree.
Pronouns	Before I return to London, let's have a meal together.
Proper Noun	After Peter leaves, the lights go out.
People	Before the police find him, we have to catch him.
Things	Before the fruit enters the factory, it is washed and sanitized
Other	Anything that does not fit into the categories listed above

- Number of Words
- Tense
- Presence of a Verb in First Phrase
- Number of Clauses/ Ideas

Example Sentences

After holding the Olympics, Utah became home to two national governing bodies of sport.

- First Word/ Phrase Category – Verbing
- Past tense
- Verb present in first phrase
- 2 clauses/ ideas

Before I go, there is something I have to tell you.

- First Word/ Phrase Category – Pronoun
- Present Tense
- Verb present in first phrase
- 2 clauses/ ideas

Results and Discussion

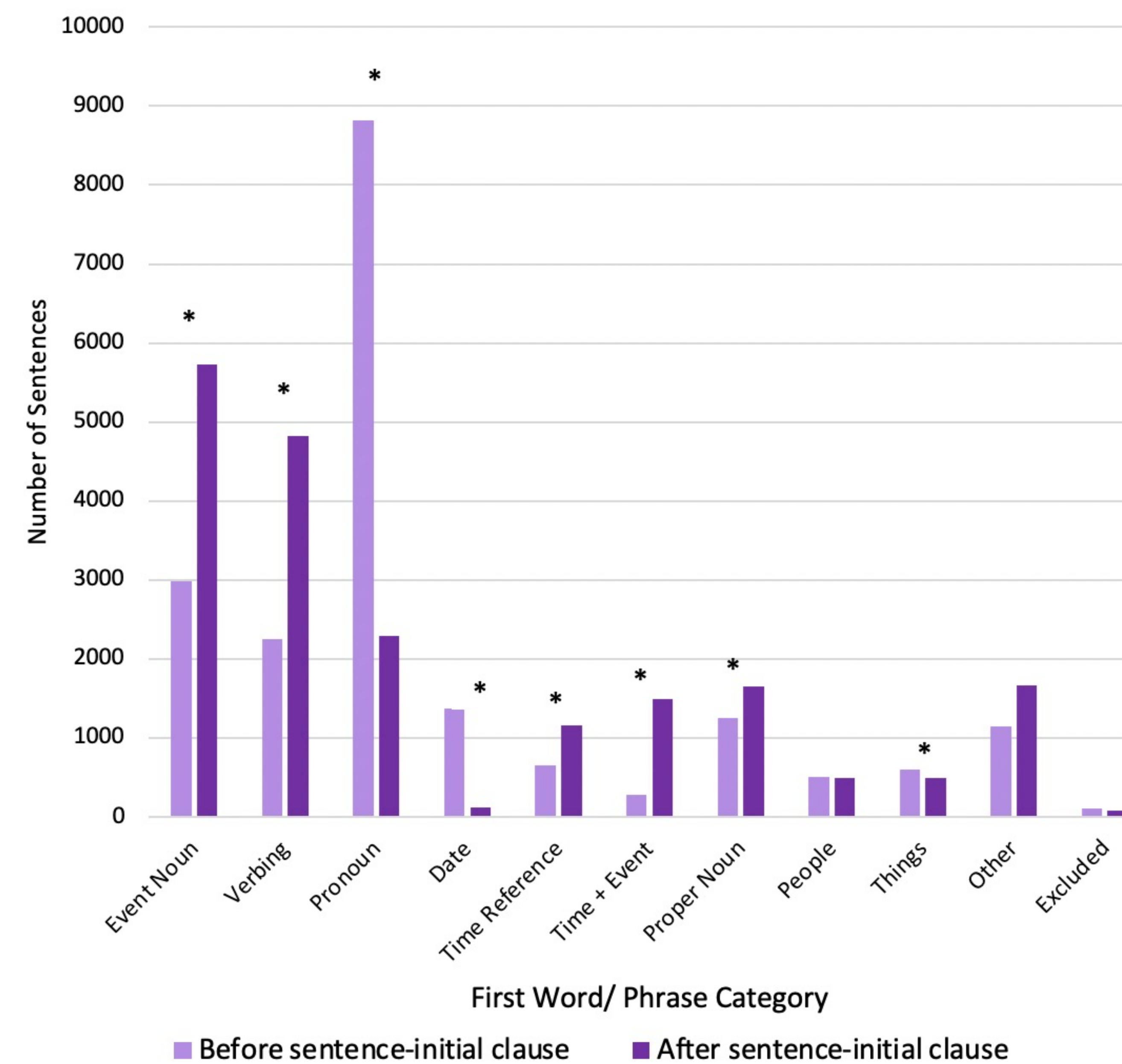


Fig. 1. Number of before vs after sentence-initial clauses in each first word/phrase category. * indicates statistical significance ($p < 0.05$) calculated according to the Chi-Squared Goodness of Fit Test.

Before sentence-initial clauses	After sentence-initial clauses
Significantly more contain a pronoun in the first word/phrase	Significantly more contain event nouns in the first word/phrase
Significantly more contain a date in the first word/phrase	Significantly more contain a verb ending in -ing in the first word/phrase
Significantly more contain things in the first word/phrase	Significantly more contain a time and event in the first word/phrase
Significantly more are in the present tense	Significantly more contain a proper noun in the first word/phrase
Significantly more are in the future tense	Significantly more are in the past tense
Significantly more contain a verb in the first clause	Significantly more are made up of 2 clauses

- From the 2 million *before/after* – initial sentences, there were 5.5X more *after* than *before*

Conclusion

- There are significant differences in the structure of *before/after* – initial sentences used in natural language.
- Most notably, *before/after* sentences differ in their first word/phrase category.
- Our results are important as previous studies that investigated the comprehension of *before/after* sentences only used a small proportion of the sentence types and structures found in the corpus.
- While temporal mismatch may cause difficulties in the comprehension of *before* sentences, the content and structure of these sentences may also play a role.

Next Steps....

- Investigate if the context of a sentence leads to anticipation of the use of *before* or *after*.
- Create sentence stimuli using the proportions found from this study to ensure they reflect the natural use of language.
- Conduct an eye tracking and EEG study using newly created stimuli.

Acknowledgements

I would like to thank Kimberley Henderson and Riya Kaur, which whom I collaborated on this corpus analysis. I also thank my supervisor, Dr. Ken McRae for his continuous guidance and encouragement. I would also like to thank the McRae Cognitive Science Lab for their support.