

Exploring Cultural Differences in Language Usage: The Case of Negation

Svetlozara Stoytcheva, Dov Cohen, and Catherine Blake
University of Illinois at Urbana-Champaign

BACKGROUND

- Prior research suggests that speakers of Asian languages are more likely to use negation than English-speakers
- Our goal is to explore this theory using empirical data from news stories
- We used natural language processing to compare negation usage in two newspapers: the *New York Times* and *Xinhua News (English Edition)*

Based on prior work, our hypothesis was that negation would be more common in articles from *Xinhua News* than in articles from the *New York Times*.

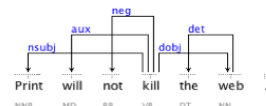
DATA

The data for this project comes from the AQUAINT Corpus of English Language News Texts, which covers the period 1998-2000.

Source	Number of sentences	Number of articles
<i>New York Times</i>	16,923,893	313,291
<i>Xinhua News (English)</i>	3,431,663	290,243

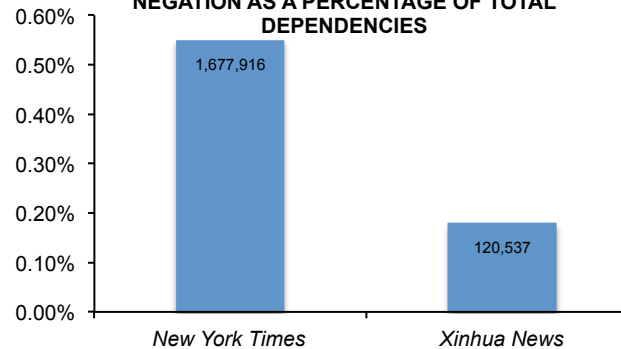
METHODS

- Sentences were preprocessed and parsed using the Stanford Natural Language Parser (v.3.2)
- The parser generates typed dependencies: grammatical features of a sentence and their relationships
- Data was loaded into an Oracle SQL Developer (v. 3.2.20.09) database
- Normalized word tables were generated
- Identifiers were assigned to link the normalized word tables, dependency output, and the original sentence

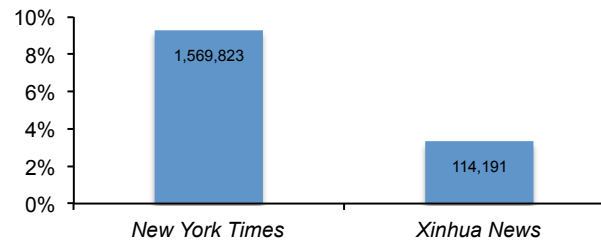


Example of a typed dependency structure.

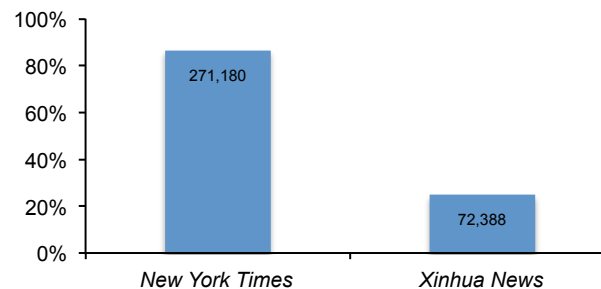
NEGATION AS A PERCENTAGE OF TOTAL DEPENDENCIES



PERCENTAGE OF SENTENCES CONTAINING NEGATION



PERCENTAGE OF ARTICLES CONTAINING NEGATION



RESULTS

- Negation represents 0.55% of typed dependencies in the *New York Times* (versus 0.18% in *Xinhua News*)
- 9.28% of sentences and 86.56% of articles in the *New York Times* contain one or more instances of negation
- In *Xinhua News*, 3.33% of sentences and 24.94% of articles in *Xinhua News* contain one or more instances of negation
- Negation is roughly 3 times more common in the *New York Times* than in *Xinhua News*

CONCLUSIONS AND NEXT STEPS

The prevalent theory suggests that Chinese authors would use negation more frequently than American authors. Our empirical results show that this theory does not hold in news texts.

Directions for future research:

- Examine how negation functions in these texts and in what contexts it appears
- Apply these methods to additional texts and compare uses of negation in other genres

SELECTED REFERENCES

- De Marneffe, M. C., MacCartney, B., & Manning, C. 2006. Generating typed dependency parses from phrase structure parses. Proceedings of the LREC, 6, 449-454. http://nlp.stanford.edu/pubs/LREC06_dependencies.pdf
- LDC. (2013). The AQUAINT Corpus of English News Texts. Linguistic Data Consortium Catalog. Retrieved from: <http://catalog.ldc.upenn.edu/LDC2002T31>
- Leung, A. K.-Y., & Cohen, D. (2011). Within- and between-culture variation: Individual differences and the cultural logics of honor, face, and dignity cultures. *Journal of personality and social psychology*, 100, 507-526. doi: 10.1037/a0022151
- Triandis, H.C. (1994). *Culture and social behavior*. New York: McGraw-Hill.

ACKNOWLEDGEMENTS

We would like to thank Craig Evans, Senior Research Programmer at the Graduate School of Library & Information Science, University of Illinois at Urbana-Champaign for his assistance with data processing and analysis.

This research is made possible in part by a grant from the U.S. Institute of Museum and Library Services (IMLS), Laura Bush 21st Century Librarian Program Grant Number RE-05-12-0054-12 Socio-technical Data Analytics (SODA).