

## Premise

### Text as a proxy for its author

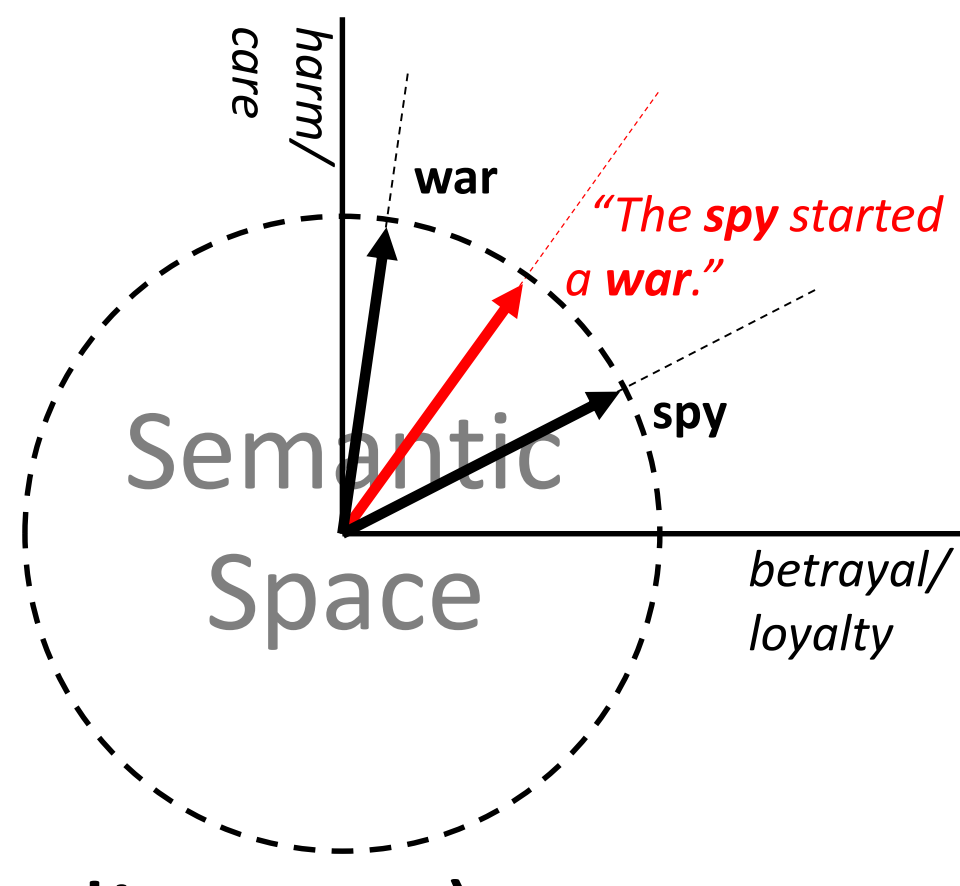
The psychological makeup of a person influences their use of language.

It is possible to identify these influences and use them to test psychological theories.

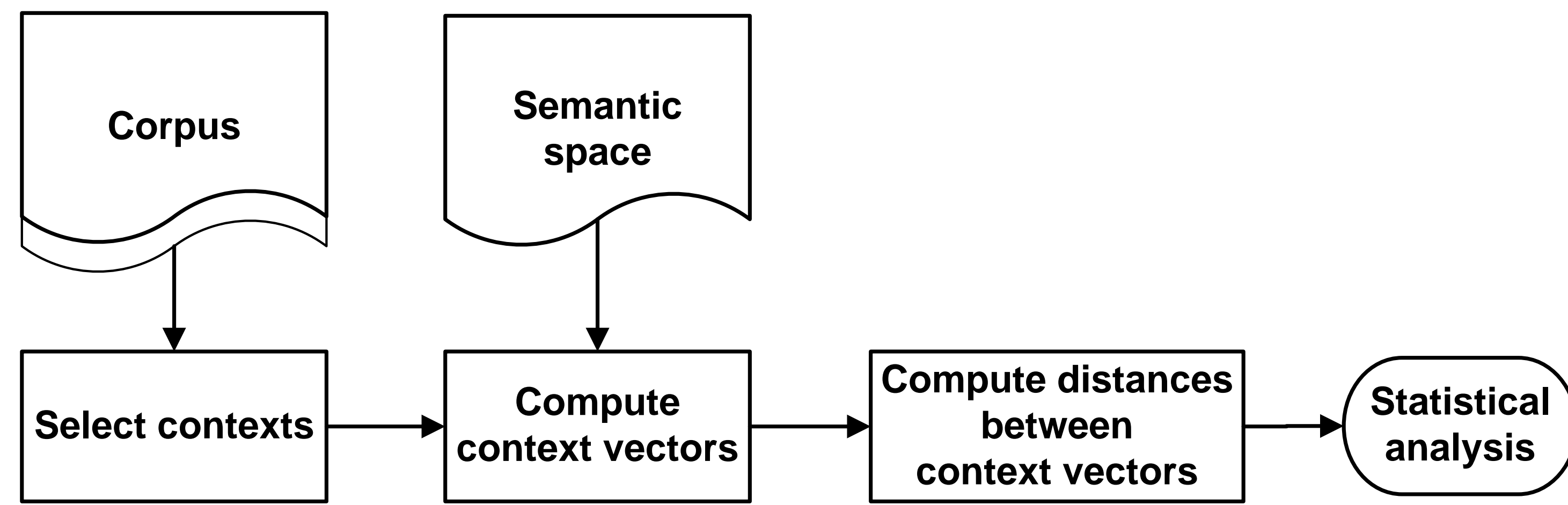
## Overview

### Words represented as vectors in a semantic space

- ❖ **Data points:** Word spans
  - ❖ e.g. Contexts of use, Tweets, etc.
  - ❖ Identified and labeled based on *theory*
  - ❖ Represented by sum of word vectors



- ❖ **Basic measure:** Angle between vectors (i.e. distance)

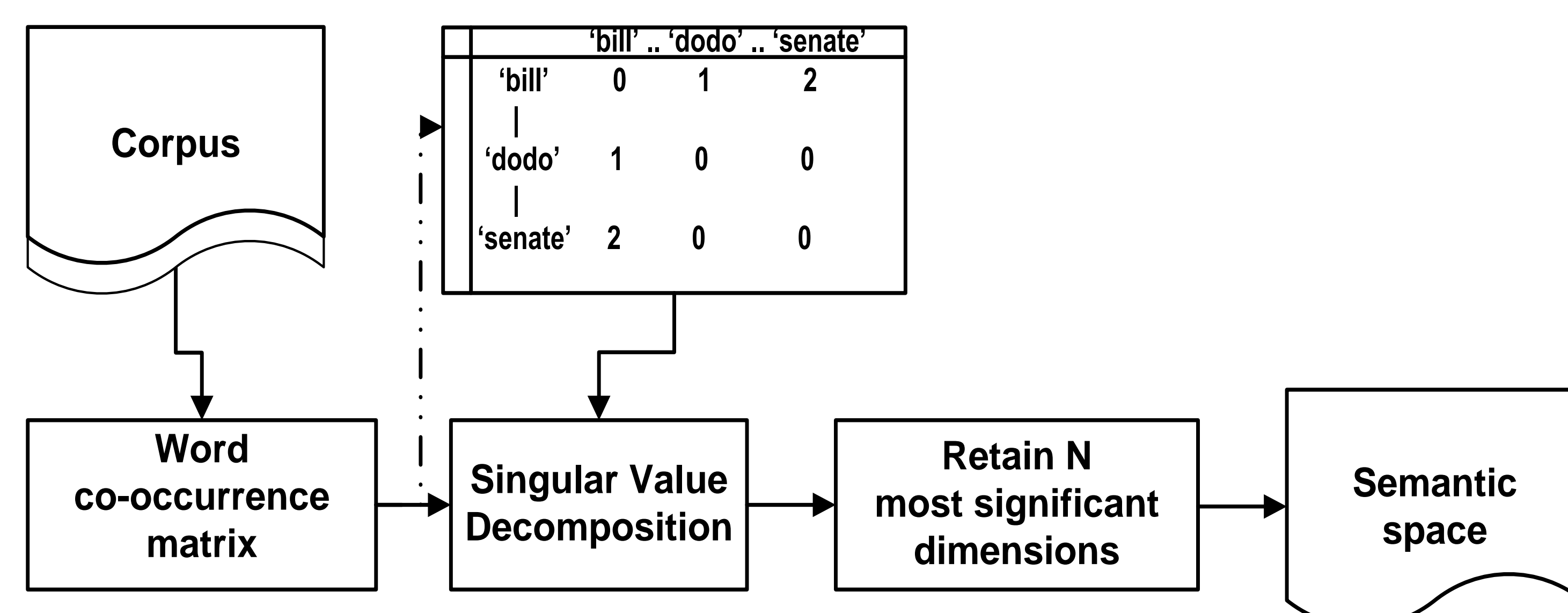


## Semantic Spaces

**Topicality:** Words that occur together are likely to relate to the same topic

- ❖ Based on *Infomap* (Takayama et al. 1999)
  - ❖ Similar to *Latent Semantic Analysis* (Deerwester et al., 1990)

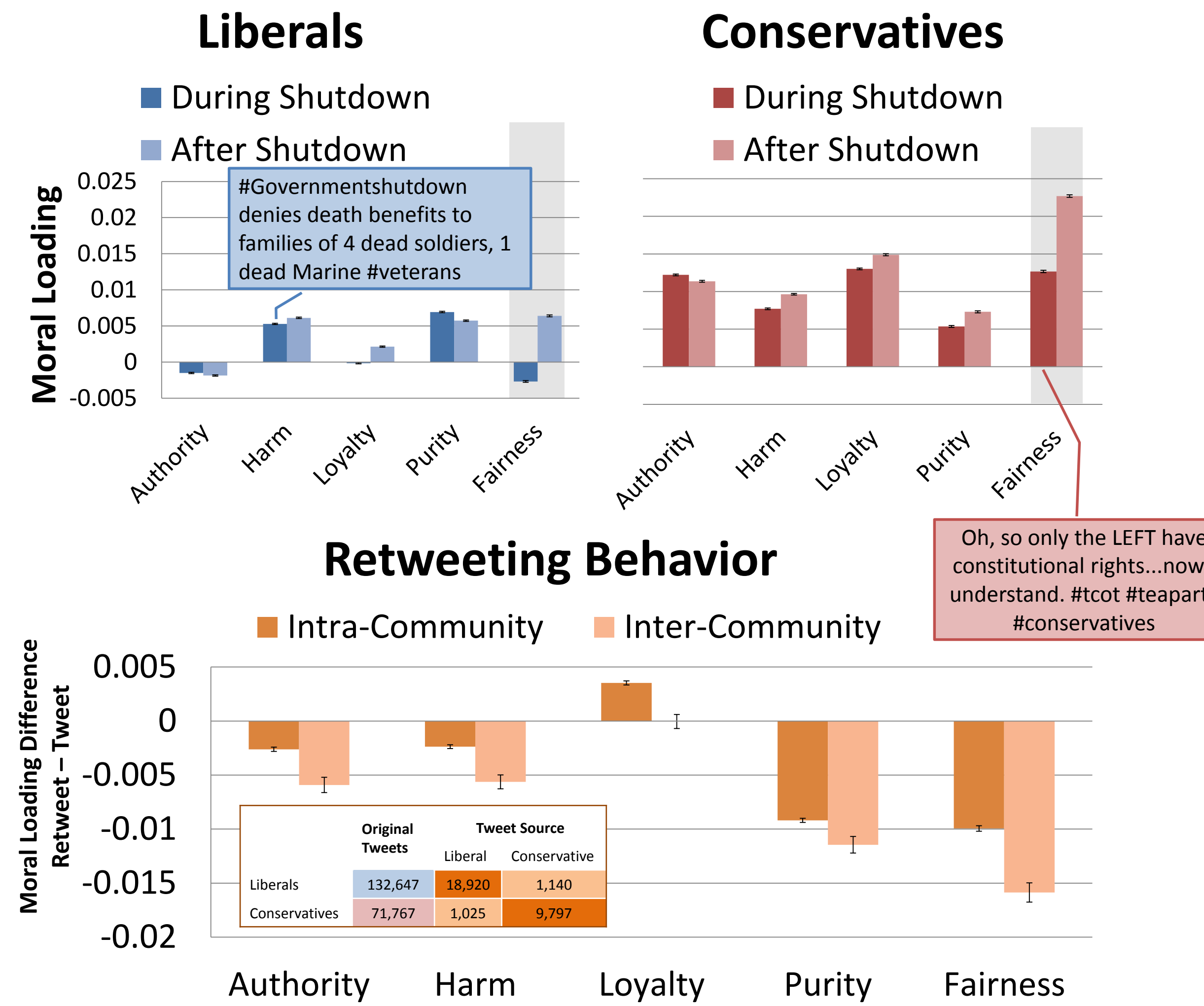
- ❖ Semantic space is based on patterns of *word co-occurrence*
  - ❖ Linear space generated using Singular Value Decomposition (generalized factoring)



## Moral Rhetoric

### Federal Shutdown (Tweeter; ~9.5M words)

Analysis based on Moral Foundations Theory (Haidt & Joseph 2004)



Source: Sagi & Dehghani (2014)

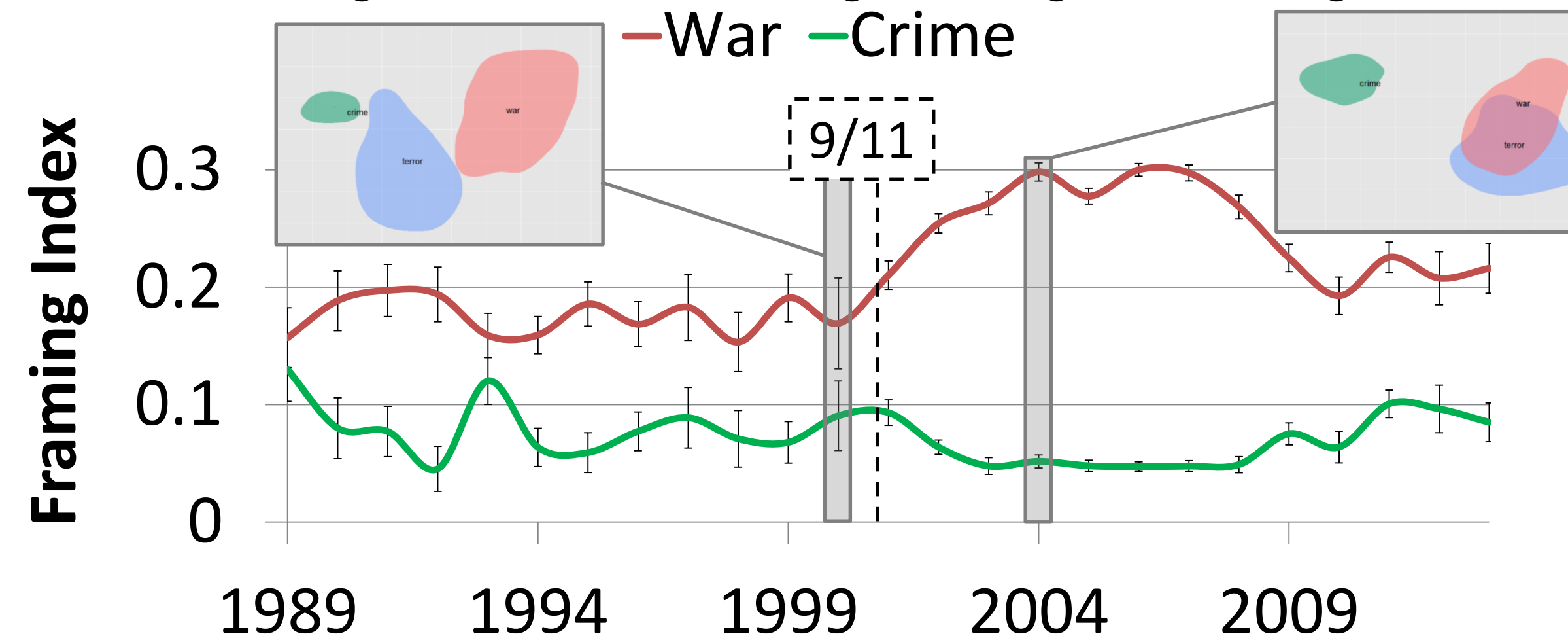
Error bars in graphs represent standard error of the mean

## Framing effects

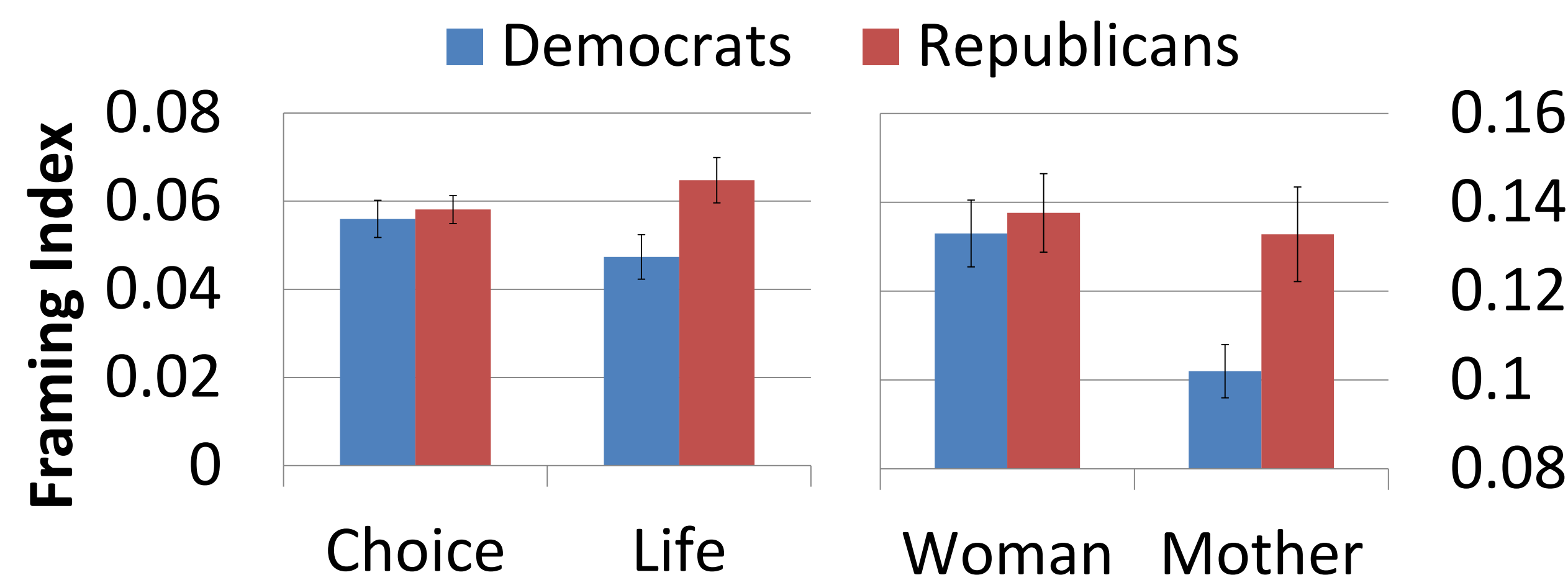
### Framing of 'terror' over time

(Senate speeches; 180M words)

Higher numbers indicate a greater degree of framing



### Framing of 'abortion' by party



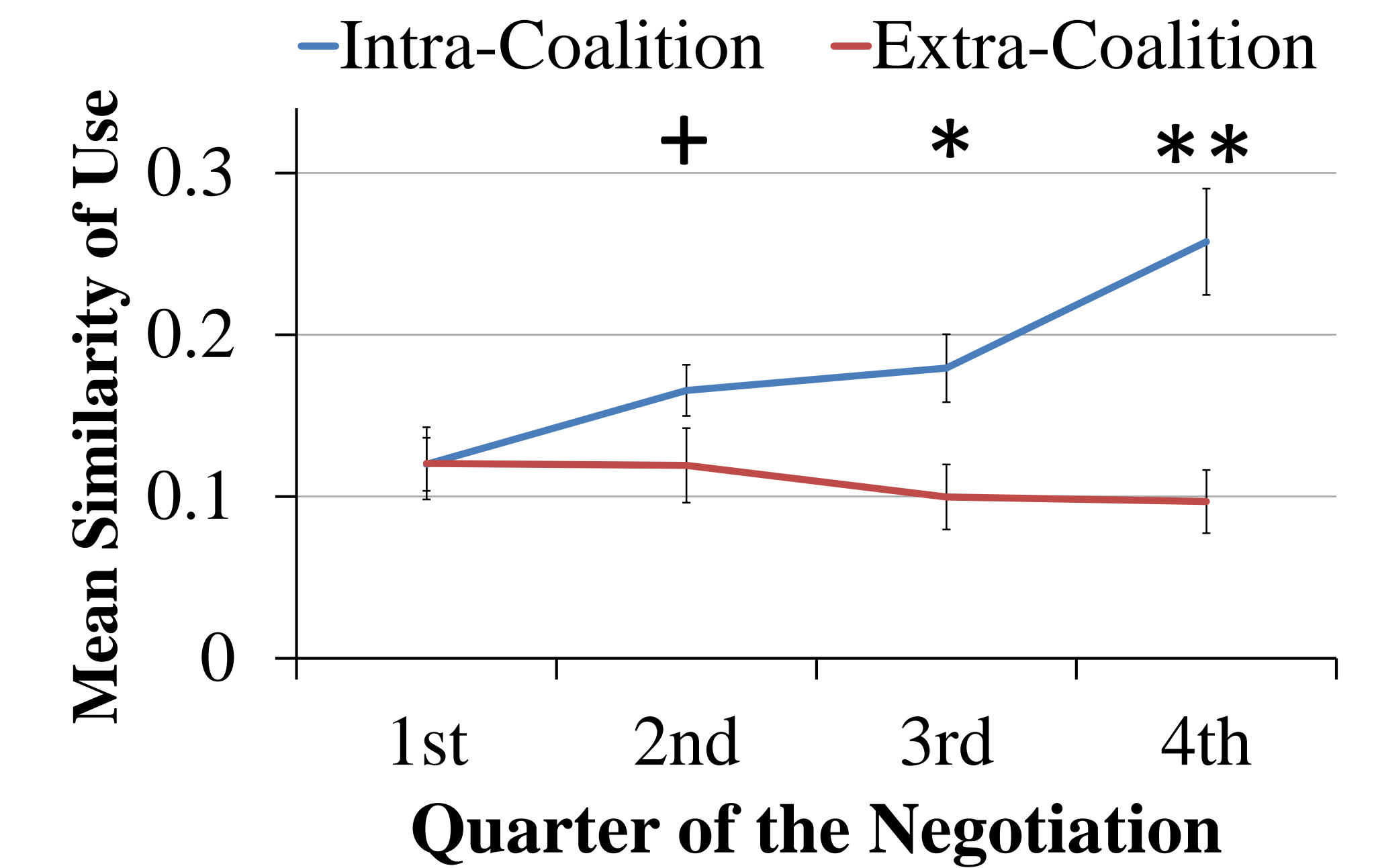
Source: Sagi, Diermeier, & Kaufmann (2013)

Error bars in graphs represent standard error of the mean

## Convergence in language use

### Reaching agreement

(3-party negotiations; ~0.5M words)



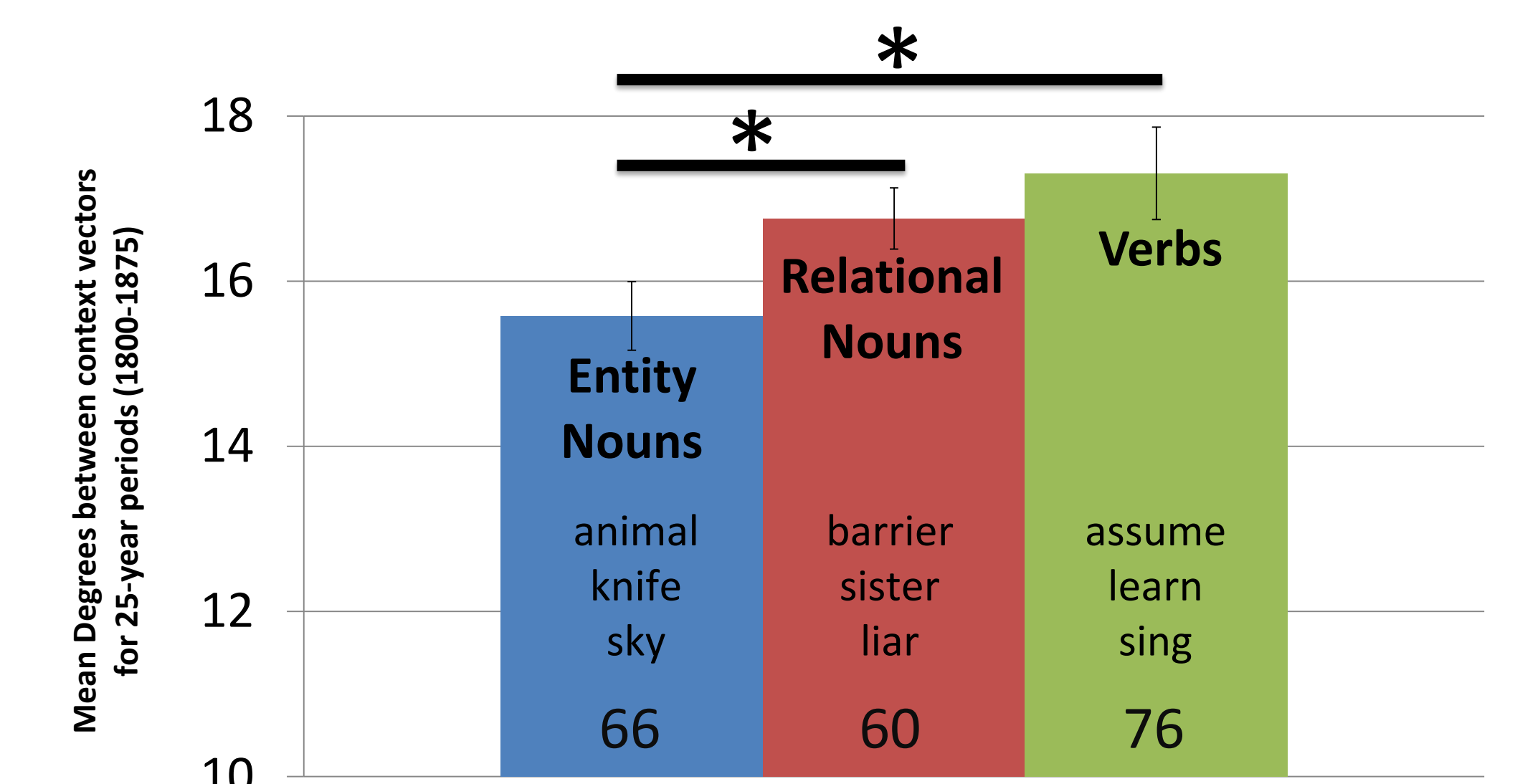
Source: Sagi & Diermeier (2013; submitted)

Error bars in graphs represent standard error of the mean

## Relationality and Stability of Meaning

### Nouns vs. Verbs (Project Gutenberg; 240M words)

Average angle for word context vectors by grammatical category.



Stability of meaning is related to **Relationality**

Similarly to *semantic mutability* (Gentner & France, 1988)

Source: Sagi (2010)

Error bars in graphs represent standard error of the mean, number of words at bottom of bars.

## References

- Deerwester, S., Dumais, S. T., Furnas, G. W., Landauer, T. K., & Harshman, R. (1990) Indexing by Latent Semantic Analysis. *Journal of the American Society for Information Science*, 41, 391-407.
- Gentner, D., & France, I. M. (1988) The verb mutability effect: Studies of the combinatorial semantics of nouns and verbs. In S. L. Small, G. W. Cottrell, & M. K. Tanenhaus (Eds.), *Lexical ambiguity resolution: Perspectives from psycholinguistics, neuropsychology, and artificial intelligence* (pp. 343-382). San Mateo, CA: Kaufmann.
- Infomap [Computer Software]. (2007) <http://infomap-nlp.sourceforge.net/> Stanford, CA.
- Sagi, E. (2010). Nouns are more stable than Verbs: Patterns of semantic change in 19th century English. In S. Ohlsson & Catrambone, R., *Proceedings of the 32th Annual Conference of the Cognitive Science Society* (p. 640)
- Sagi, E., & Dehghani, M. (2014) Moral Rhetoric in Twitter: A Case Study of the U.S. Federal Shutdown of 2013. In P. Bello, Guarini, M., McShane, M., & Scassellati, B., *Proceedings of the 36th Annual Conference of the Cognitive Science Society* (pp. 1347-1352).
- Sagi, E., & Diermeier, D. (2013) The Time Course of Language Use in Multiparty Negotiations. In M. Knauft, Pauen, M., Sebanz, N., & Wachsmuth, I., *Proceedings of the 35th Annual Conference of the Cognitive Science Society* (pp. 3343-3347).
- Sagi, E., Diermeier, D., & Kaufmann, S. (2013) Identifying Issue Frames in Text. *PLoS ONE*, 8(7), e69185.
- Takayama, Y., Flournoy, R., Kaufmann, S. & Peters, S. (1999) Information retrieval based on domain-specific word associations. In Cercone, N. and Naruedomkul K. (eds.), *Proceedings of the Pacific Association for Computational Linguistics (PAACLING'99)*, Waterloo, Canada. 155-161.