Neural correlates of semantic coherence in English and Chinese speakers during natural language comprehension

Jixing Li
Cornell University

Christophe Pallier
INSERM-CEA Cognitive Neuroimaging Unit

Yiming Yang
Jiangsu Normal University

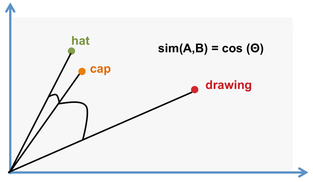
John Hale
Cornell University

Introduction

Semantically incoherent words induce comprehension difficulty:

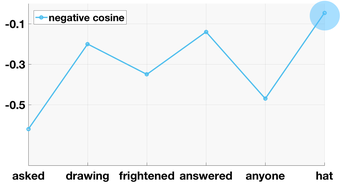
"I asked them if my drawing frightened them. They answered me why would anyone be frightened by a hat."
— *The Little Prince*

Quantify semantic similarities: the meaning of a word is represented as a vector computed from the distribution of words around it.



Words often co-occur have similar word embeddings and the cosine of their angle is larger.

Negative cosine between word embeddings of the current content word and the mean of the previous ten content words indicates semantic coherence.



Research question: What part of the brain works hard when encountering a semantically incoherent word? Is there a **cross-linguistic difference** in the neural correlates of semantic coherence?

Methods

English (n=31) and Chinese participants (n=33) listened to the whole book *The Little Prince* (~100 mins) in English and Chinese respectively in the fMRI scanner.

Word embeddings were generated by Mikolov et al.'s (2013) skipgram model, trained on English (n=4275675) and Chinese (n=744150) wikipedia articles.

Story annotation:

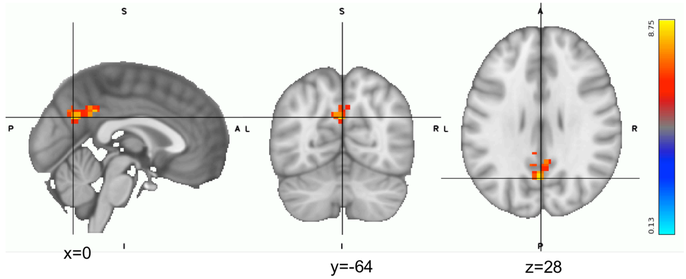
word	onset	offset	word frequency (log)	semantic coherence
once	0.1132	0.7283	2.54	-0.16454
when	0.7282	0.91941	3.31	
I	0.91941	1.0249	4.60	
was	1.0249	1.1582	3.75	-0.16454
six	1.1582	1.4641	2.30	
years	1.4641	1.6712	2.75	-0.15463
old	1.6712	1.9729	2.78	-0.049209
…	…	…	…	…

GLM analysis:
4 nuisance regressors:
word rate: offset of each word in time;
word frequency: log-frequency of each word in movie subtitles (Brysbaert & New, 2009; Cai & Brysbaert, 2010)
RMS intensity at every 10 ms of the audio;
f0 at every 10 ms of the audio.

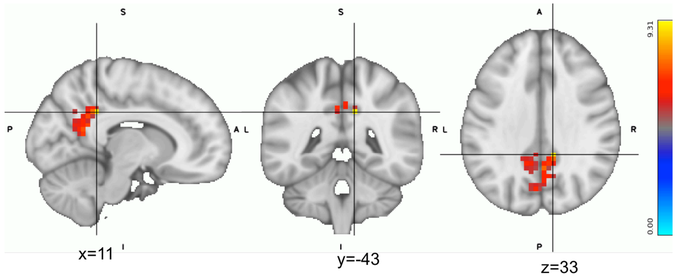
Regressors are convolved with the canonical HRF using SPM12.

Results

English:



Chinese:



English and Chinese comparison ($p < 0.05$ *FWE*, $k >= 20$)

	Region	MNI coordinates	Cluster size	Max. t-value
English	Precuneus cortex	0 -64 28	370	8.77
Chinese	Precuneus cortex	11 -43 33	149	9.22
	left aSTG	-56 -10 -11	50	7.96
	right aSTG	56 2 -15	41	7.84
English > Chinese	–	–	–	–
Chinese > English	left aSTG	-52 2 -8	16	6.48

Conclusion

The **precuneus cortex** seems to be associated with semantic coherence for both English and Chinese speakers.

Consistent with a number of studies reporting the precuneus for semantic processing (e.g., Binder et al. 2009). Different from the semantic combination region LATL reported in Bemis & Pylkkänen (2011), suggesting that semantic coherence might be a cognitively different process from semantic combination.

Precuneus has also been associated with different story characters in a narrative (Wehbe et al., 2014). This suggests that semantic coherence and discourse coherence might invoke similar cognitive mechanisms.

References

Bemis, D., & Pylkkänen, L. (2011). Simple composition: An meg investigation into the comprehension of minimal linguistic phrases. *Journal of Neuroscience, 31*, 2801-2814.

Binder, J., Desai, R., Graves,W., & Conant, L. (2009). Where is the semantic system? a critical review and meta-analysis of 120 functional neuroimaging studies. Cerebral Cortex, 19, 2767-2796.

Brysbaert, M., & New, B. (2009). Moving beyond Kücera and Francis: A critical evaluation of current word frequency norms and the introduction of a new and improved word frequency measure for american english. *Behavior Research Methods, 41*, 977-990.

Cai, Q., & Brysbaert, M. (2010). Subtlex-ch: Chinese word and character frequencies based on film subtitles. *PLoS ONE, 5*, e10729.

Mikolov, T., Chen, K., Corrado, G., & Dean, J. (2013). Efficient estimation of word representations in vector space. *arXiv* preprint arXiv:1301.3781.

Wehbe, L., Murphy, B., Talukdar, P., Fyshe, A., Ramdas, A., & Mitchell, T. (2014). Simultaneously uncovering the patterns of brain regions involved in different story reading subprocesses. *PloS ONE, 9*, e112575.

Acknowledgment

This work is supported by NSF award # 1607441