Semantic shift in social networks

Bill Noble[*], Asad Sayeed[*], Raquel Fernández[†], Staffan Larsson[*]

August 5–6, 2021 (StarSem)

University of Gothenburg,[*] University of Amsterdam[†]

CLASP centre for linguistic theory and studies in probability

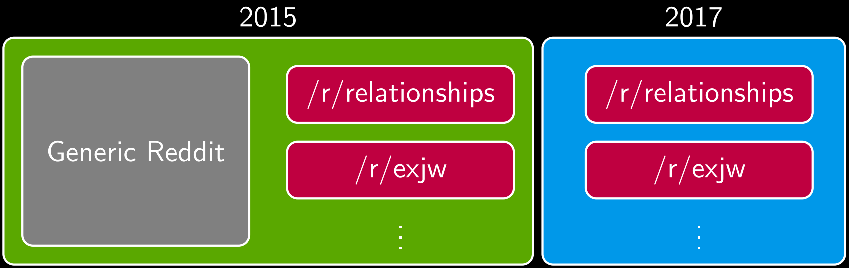
European Research Council
Established by the European Commission

Semantic shift in social networks

- Words change in meaning over time
- The meaning of words is with respect to the community in which they are used (e.g., Clark, 1996; Stalnaker, 2002)
- Simulations and laboratory experiments have suggested that the structure of a community can affect how quickly or in what way words change in meaning (Lev-Ari, 2018; Raviv et al., 2019).

Data: Reddit comments

- Social media comments
  - threaded replies
  - authorship identified by username
- Two time periods: 2015 and 2017 (one year gap)
- 46 randomly selected communities (avg. 282K comments per community)
- A larger "generic" 2015 corpus of comments randomly selected from all of Reddit (55M comments)



Diachronic skip-gram (Kim et al., 2014)

- Skip-gram with Negative Sampling (SGNS) tries to guess, for a given word, whether another word was drawn from its context window or not (i.e. if it is a negative sample)
- The diachronic skipgram procedure we followed is as follows (adapted from Del Tredici et al. (2019)):
  1. Train a base model, M_{15}, on the generic Reddit 2015 corpus.
  2. For each subreddit c:
     - 2.1 Initialize with the generic 2015 model and train a community-specific 2015 model: $M_{15} \rightarrow M^c_{15}$.
     - 2.2 Initialize with the 2015 community-specific model and train a community-specific 2017 model: $M^c_{15} \rightarrow M^c_{17}$.
  3. Finally, train a 2017 generic model by initalizing with the 2015 generic corpus and training on a (smaller) corpus of generic Reddit comments from 2017 (this is used to measure *generic change*: $M_{15} \rightarrow M_{17}$.

Naïve cosine change

With the aligned 2015/2017 word vectors, the most straight-forward way to measure change is as the cosine distance (we use angular distance) between the two vectors:

$$\Delta_c^{\cos}(w) = \frac{\cos^{-1}(\cos\,\mathrm{sim}(\vec{w}_{c,15}, \vec{w}_{c,17}))}{\pi}$$

where

$$\cos\,\mathrm{sim}(v_1, v_2) = \frac{v_1 \cdot v_2}{\|v_1\|\|v_2\|}$$

Naïve cosine change

With the aligned 2015/2017 word vectors, the most straight-forward way to measure change is as the cosine distance (we use angular distance) between the two vectors:

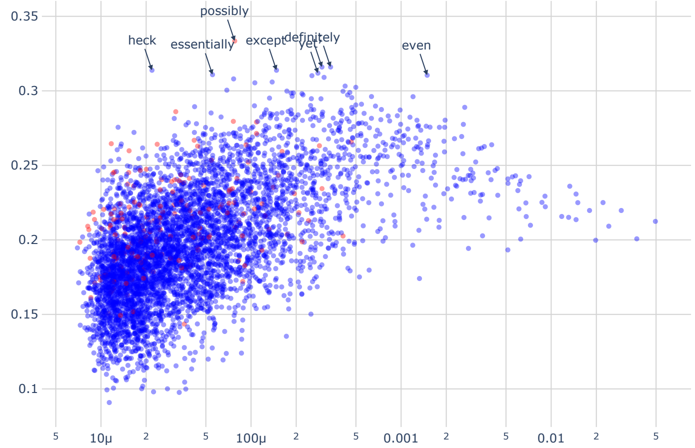
$$\Delta_c^{\cos}(w) = \frac{\cos^{-1}(\cos\ \mathrm{sim}(\vec{w}_{c,15}, \vec{w}_{c,17}))}{\pi}$$

where

$$\cos\ \mathrm{sim}(v_1, v_2) = \frac{v_1 \cdot v_2}{\|v_1\|\|v_2\|}$$

Problem: Naïve cosine change is inherently biased towards words that appear in more variable contexts—which has a strong correlation to frequency.

Naïve cosine change: /r/Toronto



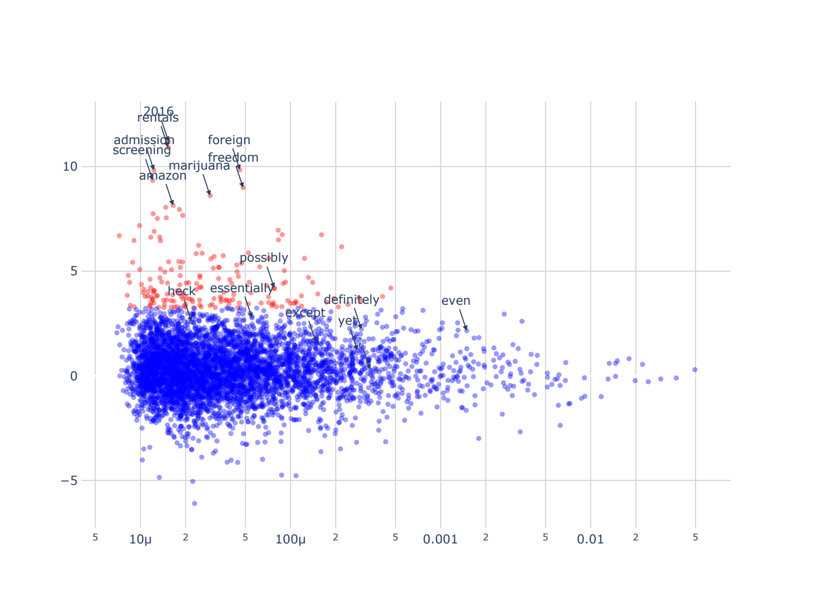
Rectified semantic change

To fix this, we use a method modified from Dubossarsky and Weinshall (2017). For each community:

1. Shuffle the 2015 and 2017 community-specific corpora and split them randomly to create a pseudo diachronic corpus with two "time periods".
2. Train diachronic SGNS models just as before
3. Repeat the first two steps 10 times.
4. For each community c and word w, compute the cosine "change" over the 10 pseudo-diachronic models. Take the average, $\bar{x}_{c,w}$, and standard deviation, $s_{c,w}$, and compute rectified change:

$$\Delta_c^*(w) = \frac{\Delta_c^{\cos}(w) - \bar{x}_{c,w}}{s_{c,w}\sqrt{1 + 1/n}}$$

Rectified semantic change: /r/Toronto

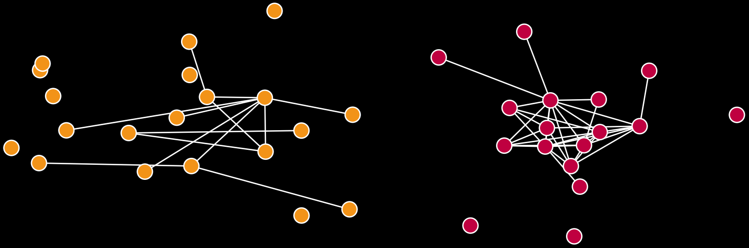


Predictive features

This gives us the following features:

Effect		Varies by
Size (2015)	S_{2015}	community
Stability	T	community
Clustering	C	community
Frequency (2015)	f_{2015}	token, community
Change in Frequency	f_{Δ}	token, community
Generic rectified change	Δ^*_G	token
Rectified change	Δ^*	token, community

We model the social network as a simple graph structure, where the vertices are community members and an edge is drawn between members with more than one interaction in a given time period.



(a) /r/relationships **(b)** /r/exjw

Figure 1: Sub-graphs of two communities with different clustering patterns. Left: $C = 0.04$; Right: $C = 0.42$.

Predictive model

- We use a linear mixed-effects model (with Δ^* as the dependent variable) to:
  - Assess the relationship between semantic shift and community-level features (including mixed effects)
  - Account for mediating effects among the word-level features
- The mixed effects model:

$$\Delta^*_{c,w} \sim (1|\text{community}) + S_{2015} * T * C + \Delta^*_{G,w} * f_{2015} * \Delta_f$$

Results: Word-level features[3]

Predictor	Coefficient	Standard Error
f_{2015} (frequency)[1]	-0.014	0.007
f_{Δ} (change in frequency)[2]	0.462	0.005
Δ^*_G (generic change)	0.055	0.003
$f_{2015} \cdot f_{\Delta}$	-0.026	0.001
$f_{2015} \cdot \Delta^*_G$	-0.012	0.006
$f_{\Delta} \cdot \Delta^*_G$	0.251	0.004
$f_{2015} \cdot f_{\Delta} \cdot \Delta^*_G$	-0.014	0.000

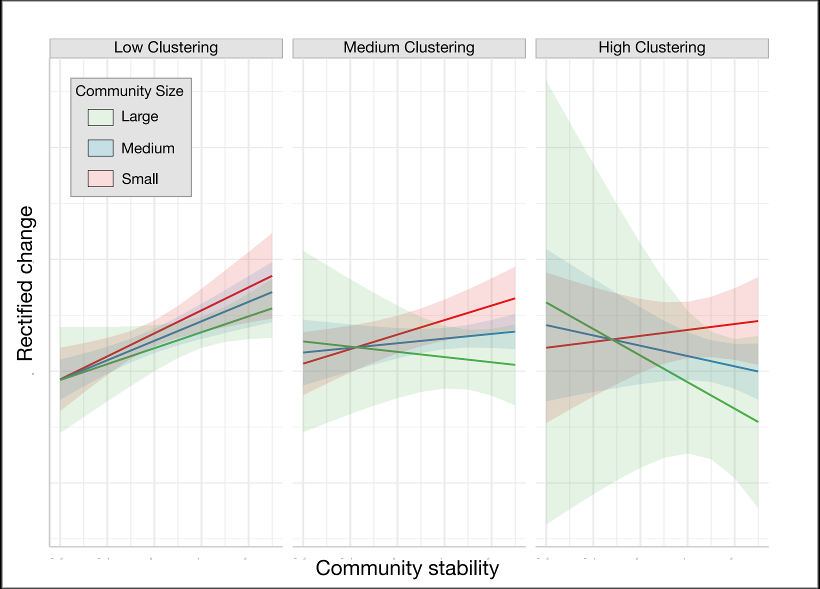
[1] Replicates Hamilton et al. (2016); Dubossarsky and Weinshall (2017)
[2] Replicates Del Tredici et al. (2019); Shoemark et al. (2019)
[3] All word-level features and interactions were found to be significant.

Results: Community-level features[4]

Predictor	Coefficient	Standard Error
Community intercept	0.250	0.069
S_{2015} (size)	-0.076	0.146
T (stability)	0.041	0.046
C (clustering)	-0.022	0.107
$S_{2015} \cdot T$	-0.088	0.076
$S_{2015} \cdot C$	-0.017	0.192
$T \cdot C$	-0.132	0.056
$S_{2015} \cdot T \cdot C$	-0.056	0.112

[4] Significant effects are marked in red.

Results: Community-level features



Future work

- How do these results generalize to different communicative settings and time frames?
- What kinds of change are taking place?
  - Broadening/narrowing of meaning
  - Metaphor/metanomy
- How are community-specific changes introduced and propagated?

Thanks for watching!

[4] Code for downloading data and running experiments is available at: https://github.com/GU-CLASP/semantic-shift-in-social-networks/

References

Herbert H. Clark. 1996. *Using Language*. Cambridge University Press.

Marco Del Tredici, Raquel Fernández, and Gemma Boleda. 2019. Short-Term Meaning Shift: A Distributional Exploration. In *Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics*, volume 1 (Long and Short Papers), pages 2069–2075, Minneapolis, Minnesota. Association for Computational Linguistics.

Haim Dubossarsky and Daphna Weinshall. 2017. Outta Control: Laws of Semantic Change and Inherent Biases in Word Representation Models. In *Proceedings of the 2017 Conference on Empirical Methods in Natural Language Processing*, pages 1136–1145, Copenhagen, Denmark. Association for Computational Linguistics.

William L. Hamilton, Jure Leskovec, and Dan Jurafsky. 2016. Diachronic Word Embeddings Reveal Statistical Laws of Semantic Change. In *Proceedings of the 54th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pages 1489–1501, Berlin, Germany. Association for Computational Linguistics.

Yoon Kim, Yi-I Chiu, Kentaro Hanaki, Darshan Hegde, and Slav Petrov. 2014. Temporal Analysis of Language through Neural Language Models. In *Proceedings of the ACL 2014 Workshop on Language Technologies and Computational Social Science*, pages 61–65, Baltimore, MD, USA. Association for Computational Linguistics.

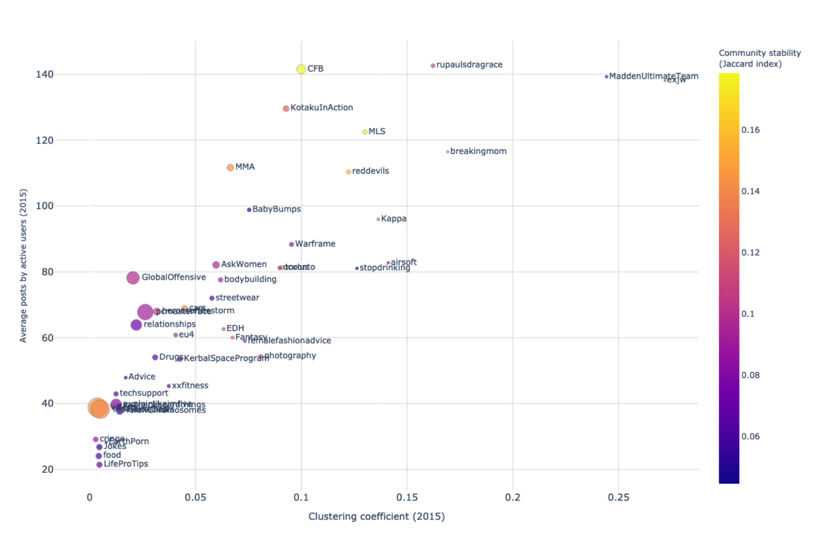
Shiri Lev-Ari. 2018. Social network size can influence linguistic malleability and the propagation of linguistic change. *Cognition*, 176:31–39.

Limor Raviv, Antje Meyer, and Shiri Lev-Ari. 2019. Larger communities create more systematic languages. *Proceedings of the Royal Society B: Biological Sciences*, 286(1907):20191262.

Philippa Shoemark, Farhana Ferdousi Liza, Dong Nguyen, Scott Hale, and Barbara McGillivray. 2019. Room to Glo: A Systematic Comparison of Semantic Change Detection Approaches with Word Embeddings. In *Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing and the 9th International Joint Conference on Natural Language Processing (EMNLP-IJCNLP)*, pages 66–76, Hong Kong, China. Association for Computational Linguistics.

Robert Stalnaker. 2002. Common Ground. *Linguistics and Philosophy*, 25(5-6):701–721.

Community features



Clustering coefficient

For an individual, i, the clustering coefficient C^i is defined as the proportion of possible connections that exist between individuals connected to i:

$$C_G^i = \frac{|\{\{j,k\} \in G \mid j,k \in N(i)|\}}{|N(i)|(|N(i)| - 1)} \tag{1}$$

where $N(i) = \{j \in U \mid \{i,j\} \in G\}$ is the *neighborhood* of i. The clustering coefficient for the community as a whole is the mean clustering coefficient of its members:

$$C_G = \frac{\sum_{i \in U} C_G^i}{|U|} \tag{2}$$