Semantic shift in social networks

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

[*]Dept. of Philosophy, Linguistics and Theory of Science, University of Gothenburg
{bill.noble,asad.sayeed,staffan.larsson}@gu.se

[†]University of Amsterdam
raquel.fernandez@uva.nl

Abstract

Just as the meaning of words is tied to the communities in which they are used, so too is semantic change. But how does lexical semantic change manifest differently across different communities? In this work, we investigate the relationship between community structure and semantic change in 45 communities from the social media website Reddit. We use distributional methods to quantify lexical semantic change and induce a social network on communities, based on interactions between members. We explore the relationship between semantic change and the *clustering coefficient* of a community's social network graph, as well as community *size* and *stability*. While none of these factors are found to be significant on their own, we report a significant effect of their three-way interaction. We also report on significant word-level effects of frequency and change in frequency, which replicate previous findings.

1 Introduction

The mechanisms and patterns of semantic change have a long history of study in linguistics (e.g., Paul, 1886; Bloomfield, 1933; Blank, 1999). However, historical accounts of semantic change typically consider meaning at the language level and, as Clark (1996) points out, referring to Lewis's (1969) account of convention, the meaning of a word "does not hold for a word *simpliciter*, but for a word *in a particular community*". This gives rise questions of how semantic change manifests differently in different communities. In this work, we explore relationship between semantic change and several community characteristics, including social network structure.

Social network analysis has long been a tool of sociolinguists studying variation and change (e.g., Bloomfield, 1933; Milroy and Milroy, 1985; Eckert, 1988), but our work differs somewhat from that tradition in both methodology and focus. Sociolinguists typically work with the social networks *of individuals*—their *ego networks*—how many people each speaker is connected to, what kind of relationships they have and, sometimes, how people in their immediate network are connected to each other. The ego network is convenient for sociolinguists using ethnographic methods; it is usually infeasible to recreate the entire social network of a large community (Sharma and Dodsworth, 2020). By studying online communities, we are able to define and compute several community-level structural characteristics including *size*, *stability*, and *social network clustering* (Section 5).

Another way that our work differs from the variationist approach is that we consider change on the level of *meaning*. With a few exceptions (e.g., Hasan, 2009), sociolinguistic research studies variation in linguistic *form* (phonology, morphology and syntax). Indeed, mainstream sociolinguists have expressed skepticism that semantics can be a proper subject of variational analysis at all (Lavandera, 1978; Weiner and Labov, 1983), since the received definition of linguistic variation concerns multiple forms expressing the same content—i.e., *different ways of saying the same thing*. With semantics at the top of the traditional linguistic hierarchy, there is no higher-order constant to which two *meanings* can refer. In this work, we instead consider *semantic shift*, which refers to changes in the meaning of a given lexical form (Newman, 2015).

For more traditional sociolinguistic variables, *social indexicality*—the association of a variant with social identities and ideology—is the main factor that mediates diffusion (Eckert, 2019). Since semantic variation can itself carry social and idealogical meaning (Hasan, 2009), there is good reason to think that it may be sensitive to some of the same aspects of community structure.

The focus on semantic shift is also made possible

Proceedings of the 10th Conference on Lexical and Computational Semantics, pages 26–37
August 5–6, 2021, Bangkok, Thailand (online)

by computational methodology—we model word meaning with distributional semantics (Section 4), which allows us to quantify short-term lexical semantic shifts at the community level.

In this study, we model the social networks of 45 English-language communities from the social media website Reddit,[1] and use diachronic word vectors to measure semantic change between two time periods one year apart. Then, we use a multistage linear mixed effects statistical model to test the effect of various community features on word-level semantic change.

2 Related work

In this section, we review work that uses computational methods to study linguistic variation and change in social context.

Distributional semantics Distributional methods, which model the meaning of a word with the contexts in which it appears, are a popular way to detect and quantify semantic change.[2] Several recent studies use distributional semantics to examine short-term semantic shift at the community level. Azarbonyad et al. (2017) use diachronic word vectors to study semantic change in political and media discourse, including in UK parliamentary debates, finding that word meaning changes differently depending on the political viewpoint of the speaker. Stewart et al. (2017) use diachronic word vectors to measure semantic change in the VKontakte social network during the Russia-Ukriaine crisis and find that changes in word frequency are predictive of semantic shift. Del Tredici et al. (2019) studied short-term semantic shift in the /r/LiverpoolFC community on Reddit, empirically validating the diachronic word vector model proposed by Kim et al. (2014) by correlating cosine distance between vectors from two different time periods with semantic change judgments collected from members of the community. In another study Del Tredici and Fernández (2017) find variations in word meaning across different Reddit communities, including communities organized around the same topic.

Social network analysis In an early example of using social network analysis to study the language online communities, Paolillo (1999) categorizes the relationships of users of an IRC channel as strong or weak ties, based on interaction frequency. They find that tie strength predicts the use of some online and community-specific forms but not others and conjecture that this difference is related the social meaning of those forms. Kooti et al. (2012) examined early Twitter conventions for attributing the source tweet to someone else (i.e., indicating that it is a *retweet*). They examined social network features, such as the size of a user's ego network, but did not find such features to be very predictive of convention adoption compared to global trends.

Communication games in a laboratory setting have also been used to examine the effect of social network structure on linguistic change. Raviv et al. (2019) quantified the communicative success, systematicity and stability of languages developed by "communities" of participants, but did not find a significant effect across the three different network structures that were tested. Lev-Ari (2018) found that individuals with larger real-world ego networks had less malleable semantic representations in the lab, and use computer simulations to argue that individuals with smaller ego networks therefore play an important role in the community-level propagation of linguistic change.

3 Data

To investigate semantic change in different communities, we use comments collected from the social media website Reddit.[3] On Reddit, users create *posts*, which consist of a link, image, or user-generated text, along with a *comment* section. Comments are threaded: users can comment on the post or reply to another user's comment.

Reddit is divided into forums called *subreddits*, which are typically organized around a topic of interest. While some forums—especially those organized around relatively niche topics—have a small tightly-knit community of users, others have a much looser community structure, with any given user posting and commenting infrequently.

Our dataset consists of comments from 45 randomly selected subreddits that were active in the years 2015–2017. In addition to the subreddit corpora, we created a generic Reddit corpus, consisting of comments sampled from every subreddit, including communities not in our sample. For both the generic corpus and the community-specific corpora, we constructed separate datasets for 2015 and 2017, leaving a one-year gap between them.

[1] https://www.reddit.com

[2] See Tahmasebi et al. (2018), Tang (2018), and Kutuzov et al. (2018) for recent surveys.

[3] Obtained from pushshift.io (Baumgartner et al., 2020).

The generic corpus consists of 55M comments for 2015 and 54M for 2017. For each of the selected subreddits, we sampled comments from 2015 and 2017 to construct two datasets of 5.4M tokens each (averaging 158K comments).[4]

4 Semantic change model

In this section, we describe how we quantify semantic change. We adopt a modeling procedure similar to that of Del Tredici et al. (2019), which is adapted from Kim et al. (2014)'s diachronic skip-gram with negative sampling (SGNS) model (Section 4.1). We define *naïve cosine change* for the community-specific and "generic" lexicons (Section 4.2). In Section 4.3, we use a control procedure adapted from Dubossarsky and Weinshall (2017) to account for noise in the naïve metric.

4.1 Diachronic SGNS

The strategy laid out by Kim et al. (2014) is to train a standard skip-gram language model on a corpus from some time period t_0, and then for each subsequent time period t_{n+1}, initialize a model with the same architecture with word vectors from time period t_n.[5] Del Tredici et al. (2019) adapts this procedure for a low-data setting by first training a *base* model on some large corpus, and initializing the t_0 model with vectors from that model. We follow the same framework. We train a base model, $M_{G,2015}$, on the generic 2015 corpus. Then, for each community, c, $M_{c,2015}$ is initialized with word vectors from $M_{G,2015}$ and trained on the community-specific 2015 corpus. Then, $M_{c,2017}$ is initialized from $M_{c,2015}$ and trained on the community-specific 2017 corpus. Additionally, we train a generic 2017 model, $M_{G,2017}$, which is initialized from $M_{G,2015}$ and trained on the generic 2017 corpus. See the supplementary materials for details on vocabulary and skip-gram model hyperparameters.

In the following, will write $\vec{w}_{c,t}$ for the word vector from $M_{c,t}$, corresponding word w.

[4]See Appendix A and B for details on community selection and data preprocessing. Code for downloading the data and the running experiments can be found at https://github.com/GU-CLASP/semantic-shift-in-social-networks.

[5]It is not clear in the original paper if the t_{n+1} model is initialized with only the word vectors from the previous time period, or if internal weights and context vectors are included as well. It seems that most subsequent implementations only carry over the word vector weights, though, which allows for more flexibility with the vocabulary. We follow this approach.

4.2 Naïve cosine change

We define *naïve cosine change* as the angular distance between corresponding word vectors from the two different time periods.[6]

For a community c, naïve cosine change is defined for all words in the vocabulary as follows:

$$\Delta_c^{\cos}(w) = \frac{\cos^{-1}(\cos\text{ sim}(\vec{w}_{c,2015}, \vec{w}_{c,2017}))}{\pi} \quad (1)$$

where

$$\cos\text{ sim}(v_1, v_2) = \frac{v_1 \cdot v_2}{\|v_1\|\|v_2\|} \quad (2)$$

Generic naïve cosine change, $\Delta_G^{\cos}$, is defined analogously.

Generally speaking, naïve cosine change has a strong track record as a semantic change metric, performing well in both human-annotated and synthetic evaluations (Hamilton et al., 2016b; Shoemark et al., 2019; Schlechtweg et al., 2020). Especially relevant to this work, Del Tredici et al. (2019) found cosine change to correlate with aggregated semantic change judgments collected from members of the /r/LiverpoolFC community on Reddit.

Model drift can distort cosine change, although this is mainly a problem with many serially-trained time periods (Shoemark et al., 2019). In a pilot study, we experimented with post-hoc aligned vector spaces and a neighborhood-based change metric (Hamilton et al., 2016a), but found minimal differences from the naïve metric.

A more serious concern for our purposes is the fact that naïve cosine change is inherently biased towards words that appear in more variable contexts. In the following section, we examine this issue more closely and define a *rectified change* metric that controls for noise. We discuss other limitations of the model in the final discussion section.

4.3 Rectified change score

Consider Figure 1 (left). Although naïve cosine change ranges *a priori* from 0 to 1, very few words score below 0.1. Even some of the most common function words have naïve cosine change above 0.2. Dubossarsky and Weinshall (2017) demonstrate that this bias is due to differences in the variance of different words' context distributions—if

[6]Some authors use $1 - \cos\text{ sim}$ as the cosine change metric, but angular distance is easier to interpret since it is a distance metric and ranges from 0 to 1.

a word appears in highly variable contexts, the SGNS model is more likely to pick up on differences between time periods, even if those differences are mere happenstance and not reflective of actual change. This is especially a problem in our case where the amount of data is relatively small.

We adapt the *shuffle control condition* described by Dubossarsky and Weinshall (2017) to address this problem. For each subreddit, we shuffle the 2015 and 2017 corpora together and split them randomly to create *pseudo-diachronic* corpora with two "time periods". Then, we train diachronic SGNS models just as before, including initializing the "first" model with word vectors from $M_{G,2015}$. We do this $n = 10$ times for each community, giving us, for each sample i, and each vocabulary item w, a *pseudo*-naïve cosine change, $\Delta^{\cos}_{c,i}(w)$. Since no *genuine* change can possibly have taken place between the shuffled corpora, $\Delta^{\cos}_{c,i}(w)$ is a sample from the noise distribution that contributes to w's naïve cosine change, based purely on the nosiness of its context distribution in c.

Next, we take the mean, $\bar{x}_{c,w}$ and sample standard deviation (using Bessel's correction of $n-1$ degrees of freedom), $s_{c,w}$, of the samples and compute *rectified change*, which we define as the t-statistic of the genuine naïve cosine change, given the estimated noise distribution:The resulting metric, although it is still more variable for less frequent words, is unbiased by the variance of the underlying context distribution (Figure 1, right).

$$\Delta^{*}_{c}(w) = \frac{\Delta^{\cos}_{c}(w) - \bar{x}_{c,w}}{s_{c,w}\sqrt{1+1/n}} \qquad (3)$$

We perform this same procedure with the generic change models (shuffling together the generic 2015 and 2017 corpora) and define *generic rectified change*, Δ^{*}_{G}, analogously.

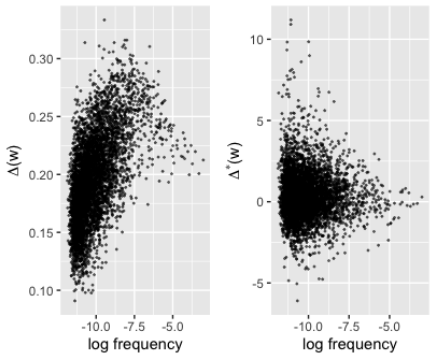

Figure 1: Naïve cosine change versus rectified change for words in the /r/toronto subreddit.

	$\Delta^{\cos}$	rank	Δ^{*}	rank	freq.
possibly	0.333	1	4.19	81	7.78
;	0.316	2	0.30	2519	33.89
definitely	0.316	3	2.23	450	29.68
heck	0.314	4	2.58	311	2.19
except	0.314	5	1.60	860	14.78
2016	0.260	303	11.19	1	1.54
rentals	0.245	576	10.91	2	1.53
foreign	0.218	1414	9.84	3	4.60
admission	0.221	1330	9.83	4	1.23
screening	0.245	582	9.34	5	1.21

Table 1: Top five tokens from /r/toronto, according to naïve cosine change and rectified change. Frequency is per 100k tokens.

Rectified change is a measure of how much higher (or lower) the measured naïve cosine change is than would be expected if the word's underlying context distribution hadn't changed at all. In other words, it quantifies the strength of the evidence that the word has changed. In our setup with 10 samples from the noise distribution, rectified change scores above 4.781 correspond to a 99.95% confidence that the change detected by the diachronic SGNS model was genuine. In addition to the analytical reasons for preferring rectified change and previous empirical work on historical change, we note that the highest scoring words for each community in our data are intuitively more varied and community-specific for rectified change. The naïve cosine change frequently ranks words with some kind of rhetorical or discourse connective function as the having changed the most (see Table 1 for examples).

5 Community features

In this section we characterize the structural features of the online communities in our dataset. Many of the features we define use the notion of *active members*. For a community c and time period t, the active members, $U_{c,t}$, is the set of members who made at least 10 posts in that period.

Size The size of a community may have an effect on semantic change. In communication game experiments, Raviv et al. (2019) found that larger communities of participants developed linguistic structure faster and more consistently than when they were grouped in smaller communities.

We define community size, $S_{2015} = |U_{c,2015}|$, as the number of active members in 2015.

Stability Community stability may also have an effect on semantic change. For example, communi-

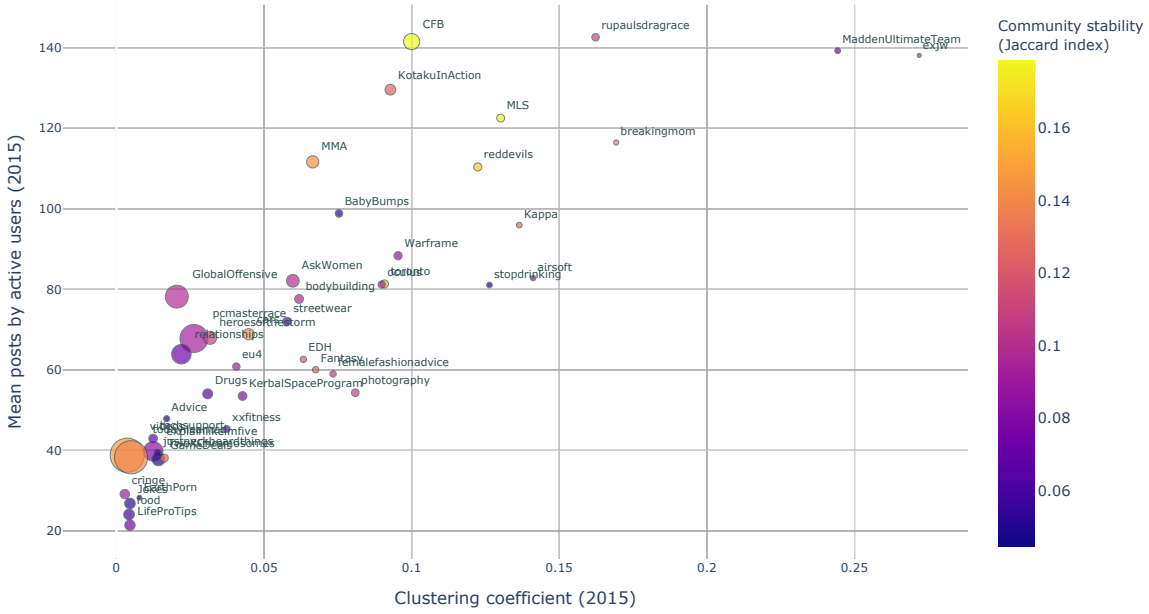

Figure 2: Community-level features for each of the 45 subreddits in our experiments. Dot size represents the community's active membership in 2015 (smallest = 1 679; largest = 118 625).

ties with stable membership have a better chance of building up community-specific common ground. On the other hand, stable communities may experience less change if such change tends to come from new community members, as some studies have suggested (Danescu-Niculescu-Mizil et al., 2013).

We define community stability as the Jaccard index between the sets of active members in 2015 and 2017. This metric, ranging from 0 to 1, captures how similar the community membership is between the two time periods.

$$T = \frac{|U_{c,t_0} \cap U_{c,t_1}|}{|U_{c,t_0} \cup U_{c,t_1}|} \tag{4}$$

Mean posts P_{2015} is the average number of posts per active members over the course of 2015.

5.1 Social network model

In this section, we define our model of social network structure and a measure of network connectivity, which we consider along with the other community features. First, we give some background and motivation for including this feature.

Social network connectivity can have seemingly-contradictory influences on linguistic change. Bloomfield (1933) observed that densely connected networks and strong social ties have a conservative influence on an individual's speech.

It is not clear whether this pattern will hold for semantic change since, as discussed by Sharma and Dodsworth (2020), different variables respond differently to different social network structures. We must also consider the evidence that an encounter with a novel or subtly unfamiliar word usage gives a speaker about the community's lexical *common ground* (Stalnaker, 2002; Clark, 1996). In more densely connected communities, such an exposure is better evidence that other speakers have been exposed to similar uses of the same word, either by the same speaker or, especially in the case of communities on social media, to the very same occurrence. For this reason, it could be that semantic change occurs faster in communities with dense clusters of strong social ties.

Clustering coefficient For each community, we define a graph model of its social network. For $a, b \in U_{c,2015}$, let $I(a, b)$ be the number of interactions between a and b in that community in 2015. Interactions are considered undirected (regardless of who is replying to whom) and we don't consider self-replies, meaning that $I(a, b) = I(b, a)$ and $I(a, a) = 0$. The two networks are thus defined:

$$G = \{\{a, b\} \mid I(a, b) > 1\} \tag{5}$$

Note that we do not consider a top-level com-

ment to be an interaction between the commenter and the creator of the post for two reasons: First, posts frequently do not contain any text written by the author—they are often just a link or photo. Second, the author of the post is not always the addressee of top-level comments, whereas in replies to comments, the author of the parent comment is always salient (though replies may of course be made with a wider audience in mind).

The *clustering coefficient* (Watts and Strogatz, 1998), measures the graph's tendency to form dense, interconnected clusters of nodes. For an individual, i, the clustering coefficient C^i is defined as the proportion of possible connections that exist between individuals connected to i in G:

$$C^i = \frac{|\{\{j,k\} \in G \mid j,k \in N(i)|\}}{|N(i)|(|N(i)|-1)} \quad (6)$$

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:

$$C_G = \frac{\sum_{i \in U} C^i_G}{|U|} \quad (7)$$

Note that C^i is precisely the measure of ego network density used in many sociolinguistic studies (Milroy, 1987), meaning that we would expect communities with higher clustering coefficients to exhibit less sociolinguistic change. We don't know whether the same effect holds for semantic change.

6 Predictive model

We perform an exploratory analysis of the data using multi-stage regressions and model selection by backwards elimination with semantic change, as measured by Δ^*, as the dependent variable.[7]

Since we fit the mixed effects model at the word level, in addition to the community-level independent variables described in Section 5, we consider two word-level features as fixed effects. See Table 2 for the full list of fixed effects.

Word frequency Since word frequency known to interact with semantic change (Hamilton et al., 2016b). we include the frequency of the token in the 2015 community corpus (f_{2015}) as a feature.

Change in frequency Additionally, we include the change in frequency between 2015 and 2017 ($f_\Delta = f_{2017} - f_{2015}$) as a feature since previous work suggests that increases in the frequency of a word often accompany semantic change (Wijaya and Yeniterzi, 2011; Kulkarni et al., 2015; Del Tredici et al., 2019).

Effect		Varies by
Mean posts (2015)	P_{2015}	community
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

Table 2: Fixed effect inputs to the statistical model. **Rectified change** is the dependent variable.

Community intercepts In addition to fixed effects, we use community-level random intercepts under the hypothesis that community topics have idiosyncratic reasons or lexical reasons for differences in semantic change rates to do with the community topics themselves, which we do not model.

6.1 Detecting multicollinearity

Before fitting the full model with interactions, we checked for multicollinearity via linear regressions with the standard `lm` function in R as well as the variance inflation factor (VIF) calculation provided by the `car` package in R. All the predictors were scaled and centered (n = 201 240 word-community combinations). We found that the distribution of Δ^* is fat-tailed (it is likely t-distributed). Nevertheless, it is bell-shaped and large enough that this should not be a problem. We ran a regression under the hypothesis $\Delta^* \sim S_{2015} + T + C + P_{2015} + \Delta^*_G + f_{2015} + f_\Delta$ (see Table 2) and calculated the VIF on this model. We found that P_{2015} had VIF higher than 2, the cutoff from Zuur et al. (2010). Removing it produced VIFs below the cutoff for the other predictors.[8]

We fit a linear mixed effects model (using the `lmer` command from the `lme4` package in R; Bates et al., 2015) with the remaining predictors

[7] The use of stepwise regression has been criticized for being a fallacious method for one-shot hypothesis testing but is a legitimate way to investigate the explanatory capacity of predictors. See https://dynamicecology.wordpress.com/2013/10/16/in-praise-of-exploratory-statistics/ for a discussion of the issue.

[8] Initially we defined a separate clustering metric C_{weak} for the *weak ties* network, analogous to the network defined in Section 5.1 but with edges between community members with exactly one interaction. However, this features was highly colinear with C and had a very high VIF when we tested it at this stage, so it was also excluded from further analysis.

in order to take into account the individual semantic change characteristics of community and word. (Model code and output will be placed on the web upon publication.)

We performed a regression on the model equation $\Delta_c^* \sim (1|\text{community}) + S_{2015} * T * C + \Delta_G^* * f_{2015} * \Delta_f$; that is, we included interactions among the community-level and word-level predictors.

6.2 Results

For the regression results (table 3), we do not report statistical significance directly from `lmer`. Instead, using R's `anova` function, we performed backwards elimination model selection (by stepwise removal of interactions and factors), and we report statistical significance based on p-values derived from the χ^2 log-likelihood ratio between models.

We found that all word-level fixed effects and their three-way interaction were significant at $p < 0.05$ in the model in terms of a χ^2 likelihood ratio test. The three-way word-level interaction $\Delta_G^* \cdot f_{2015} \cdot f_\Delta$ had a p-value too small to represent ($\chi^2(4) = 6380.751$) relative to a model with all predictors without the interaction (so terms $\Delta_G^* + f_{2015} + f_\Delta$) along with all the other predictors and interactions. Relative to the model without the three-way word level interaction, removing each word-level predictor individually yielded $p_{\Delta_G^*} = 7.059 \times 10^{-81}$ ($\chi^2(1) = 362.759$), $p_{f_{2015}} = 1.605 \times 10^{-26}$ ($\chi^2(1) = 113.587$), and p_{f_Δ} was too small to measure ($\chi^2(2) = 2070.095$).

The three-way interaction for the community-level features was significant at $p = 0.014$ ($\chi^2(4) = 12.530$), but none of the two-way interactions or the individual predictors were significant.[9]

We plotted the three-way interaction in Figure 3. Clustering coefficient and size are held at the mean and plus or minus one standard deviation from the mean. At low levels of clustering, all levels of size have a positive linear relationship on rectified change with respect to increasing stability.

At mean levels of clustering, the lower and mean levels of size retain the positive relationship but flatten out, and the high size level becomes negative. At one standard deviation above the mean for clustering, only the lowest size level remains positively sloped relative to stability. Confidence intervals increase dramatically as clustering increases (as there are fewer examples with higher coefficients).

[9]This means that all the individual predictors and two-way interactions must be part of the model, but their significant effect is conditioned on one another.

Predictor	Coefficient	SE
(intercept)	0.250	0.069
S_{2015}	-0.076	0.146
T	0.041	0.046
C	-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
f_{2015}	-0.014	0.007
f_Δ	0.462	0.005
Δ_G^*	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

Table 3: Fixed effect coefficients of the mixed effects model with standard errors. p-values for some predictors are reported in the text.

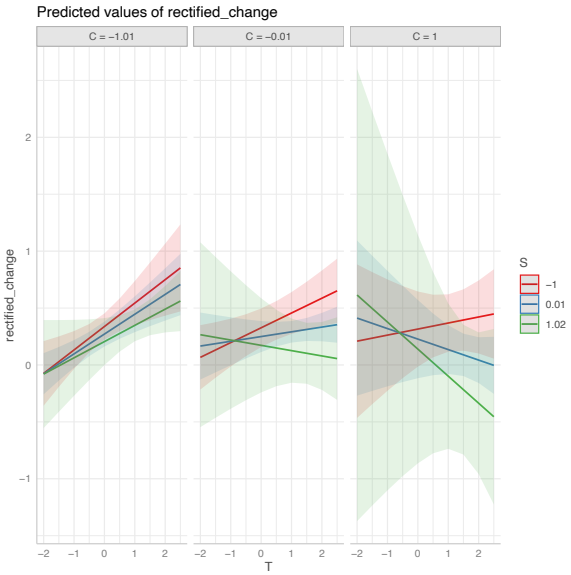

Figure 3: Plot of three-way interaction between community-level predictors vs. the rectified change using the `ggeffects` package. Each panel represents a fixed value for the clustering coefficient, specifically -1 st.dev. from the mean, the mean, and 1 st.dev from the mean. Similarly, each line represents the same three values for the size. The x-axis in each panel represents the group stability.

The effect of the random intercept is small (σ^2 = 0.019, SD = 0.138). This is the extent to which the type of community causes the intercept of rectified change to vary.

7 Discussion and conclusions

We conducted an exploratory statistical analysis of the relationship between semantic change and several word- and community-level predictive features. Rectified semantic change, our independent

variable, protects the results from certain systematic biases inherent in the traditional cosine change metric. By looking at online communities, we were able to compute a clustering coefficient on the social network graph of each community, as well as several other community-level structural features.

Community features and semantic change We found all three word-level features to be significant. Together with the intercept, f_Δ dominates the mixed-effects model, with greater changes in frequency associated with higher semantic change. This is in line with previous findings (Wijaya and Yeniterzi, 2011; Kulkarni et al., 2015; Del Tredici et al., 2019), but our study is the first to demonstrate this effect while controlling for noise effects.

Although the effect is much smaller, there is a negative relationship between semantic change and baseline frequency, f. This agrees with previous results about historical change (Hamilton et al., 2016a; Dubossarsky and Weinshall, 2017), but we note that, while we cannot compare the regression coefficients directly, it appears that frequency may have a much smaller effect on semantic change in the short-term setting; however, testing this hypothesis would require further research.

Semantic change in the generic lexicon also predicts community-level change, though it has a smaller effect than f_Δ. The interaction between f_Δ and Δ_G suggests that changes in frequency can predict whether generic lexicon changes in meaning will be picked up by a particular community.

We found that the three-way interaction between size, stability, and clustering, was significant: For communities with low clustering, there is a positive linear relationship between stability and semantic change (regardless of community size). For communities with average or high clustering, however, the positive relationship between stability and change only appears to hold for smaller communities. Note, however that the confidence intervals increase dramatically as clustering increases, since our sample of communities found fewer examples with high clustering.

We did not find significant correlations for any of the community-level features on their own. It is possible that a larger study with more communities or a more diverse set of communities would reveal some more universal effect, but we cannot make any conclusions from these results. The fact that the three-way interaction has a significant effect while none of the individual features did on their own demonstrates the complexity of relationship between structural community characteristics and semantic change.

Assumptions and limitations of the semantic change model In spite of our efforts to control for biases of cosine change, there are still some caveats when interpreting the results.

Like most distributional models of semantics, the diachronic SGNS model associates each word form with a single vector, meaning it is not sensitive to polysemy or homonymy. If a word with multiple senses undergoes changes in the relative frequency with which those senses are used, this would be reflected in the vector representation of the token that both senses are associated with, even if the meaning of either sense hasn't changed on its own.[10] However, many theories of semantic change emphasize the role of changing sense distributions as a mechanism for lexical semantic change, so it is not necessarily contrary to our aims of quantifying semantic change over the lexicon.

A related weakness of distributional semantics has to do with the distinction between meaning-in-use and lexical meaning. Even if we assume that distributional context is a faithful (if noisy) representation of the situated meaning of a word (cf. Lücking et al., 2019; Bisk et al., 2020; Bender and Koller, 2020), it might not capture the word's full *meaning potential* (Norén and Linell, 2007)—in the extreme, a word may have common ground semantic content that *could* be activated, but that happens not to appear in the corpus.

Moreover, changes in the topics discussed by the community may cause changes in the context distribution of words that don't reflect actual change in meaning. Consider the words at the top of the list for /r/toronto (Table 1). It's possible that some of those words appear due to changes in the sociopolitical topics people were discussing on the forum between 2015 and 2017. Similarly, the top word, 2016, presumably still refers to the same year, though the year itself went from being in the future to being in the past. Whether or not such a change counts as a *change in meaning* is naturally beyond the scope of this paper.

[10] Contextualized word representations (Peters et al., 2018; Devlin et al., 2019) don't have this shortcoming and have recently been used to investigate semantic change (Giulianelli et al., 2020; Vani et al., 2020), but extracting one vector per occurrence is computationally expensive and has therefore only been applied to small sets of target words.

Future work This work offers some insight into how semantic change and community structure interact, but there are still many open questions, including how these results generalize to communities in different communicative settings and over different time frames. Future work should take a closer look at the *kinds* of change (e.g., Blank, 1999) taking place. For example, are the meanings of words broadening or narrowing? How are existing community-level communicative resources used to create new word uses? Given that we can identify statistically significant changes in meaning over a relatively short period of time, it would also be interesting to investigate the circumstances of individual changes. For example, do community members with more central social network position tend to innovate more? How are early innovative uses received by the community? Is there a correlation between semantic change in a given time period and the frequency of explicit *word meaning negotiation* (Myrendal, 2019) in the same period?

Acknowledgements

This work was supported by grant 2014-39 from the Swedish Research Council (VR) for the establishment of the Centre for Linguistic Theory and Studies in Probability (CLASP) at the University of Gothenburg. This work was also supported by the Marianne and Marcus Wallenberg Foundation grant 2019.0214 for the Gothenburg Research Initiative for Politically Emergent Systems (GRIPES). This project has received funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (grant agreement No. 819455 Awarded to RF).

References

Hosein Azarbonyad, Mostafa Dehghani, Kaspar Beelen, Alexandra Arkut, Maarten Marx, and Jaap Kamps. 2017. Words are Malleable: Computing Semantic Shifts in Political and Media Discourse. In *Proceedings of the 2017 ACM on Conference on Information and Knowledge Management*, CIKM '17, pages 1509–1518, New York, NY, USA. Association for Computing Machinery.

Douglas Bates, Martin Mächler, Ben Bolker, and Steve Walker. 2015. Fitting Linear Mixed-Effects Models Using lme4. *Journal of Statistical Software*, 67(1):1–48.

Jason Baumgartner, Savvas Zannettou, Brian Keegan, Megan Squire, and Jeremy Blackburn. 2020. The Pushshift Reddit Dataset. *arXiv:2001.08435 [cs]*.

Emily M. Bender and Alexander Koller. 2020. Climbing towards NLU: On Meaning, Form, and Understanding in the Age of Data. In *ACL 2020*.

Yonatan Bisk, Ari Holtzman, Jesse Thomason, Jacob Andreas, Yoshua Bengio, Joyce Chai, Mirella Lapata, Angeliki Lazaridou, Jonathan May, Aleksandr Nisnevich, Nicolas Pinto, and Joseph Turian. 2020. Experience Grounds Language. In *Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing (EMNLP)*, pages 8718–8735, Online.

A. Christian Blank. 1999. Why do new meanings occur? A cognitive typology of the motivations for lexical semantic change. In Andreas Blank and Peter Koch, editors, *Historical Semantics and Cognition*. De Gruyter Mouton.

Leonard Bloomfield. 1933. *Language*, 14. impr edition. Allen & Unwin, London.

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

Cristian Danescu-Niculescu-Mizil, Robert West, Dan Jurafsky, Jure Leskovec, and Christopher Potts. 2013. No country for old members: User lifecycle and linguistic change in online communities. In *Proceedings of the 22nd International Conference on World Wide Web*, WWW '13, pages 307–318, Rio de Janeiro, Brazil. Association for Computing Machinery.

Marco Del Tredici and Raquel Fernández. 2017. Semantic Variation in Online Communities of Practice. In *IWCS 2017 - 12th International Conference on Computational Semantics - Long Papers*.

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.

Jacob Devlin, Ming-Wei Chang, Kenton Lee, and Kristina Toutanova. 2019. BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding. In *Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 1 (Long and Short Papers)*, pages 4171–4186, 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.

Penelope Eckert. 1988. Adolescent social structure and the spread of linguistic change. *Language in Society*, 17(2):183–207.

Penelope Eckert. 2019. The limits of meaning: Social indexicality, variation, and the cline of interiority. *Language*, 95(4):751–776.

Mario Giulianelli, Marco Del Tredici, and Raquel Fernández. 2020. Analysing Lexical Semantic Change with Contextualised Word Representations. In *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics*, pages 3960–3973, Online. Association for Computational Linguistics.

William L. Hamilton, Jure Leskovec, and Dan Jurafsky. 2016a. Cultural Shift or Linguistic Drift? Comparing Two Computational Measures of Semantic Change. In *Proceedings of the 2016 Conference on Empirical Methods in Natural Language Processing*, Austin, Texas. Association for Computational Linguistics.

William L. Hamilton, Jure Leskovec, and Dan Jurafsky. 2016b. 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.

Ruqaiya Hasan. 2009. *Collected Works of Ruqaiya Hasan. Vol. 2, Semantic Variation: Meaning in Society and in Sociolinguistics*. Equinox, London ; Oakville.

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.

Farshad Kooti, Winter A. Mason, Krishna P. Gummadi, and Meeyoung Cha. 2012. Predicting emerging social conventions in online social networks. In *Proceedings of the 21st ACM International Conference on Information and Knowledge Management - CIKM '12*, page 445, Maui, Hawaii, USA. ACM Press.

Vivek Kulkarni, Rami Al-Rfou, Bryan Perozzi, and Steven Skiena. 2015. Statistically Significant Detection of Linguistic Change. In *Proceedings of the 24th International Conference on World Wide Web - WWW '15*, pages 625–635, Florence, Italy. ACM Press.

Andrey Kutuzov, Lilja Øvrelid, Terrence Szymanski, and Erik Velldal. 2018. Diachronic word embeddings and semantic shifts: A survey. In *Proceedings of the 27th International Conference on Computational Linguistics*, pages 1384–1397, Santa Fe, New Mexico, USA. Association for Computational Linguistics.

Beatriz R. Lavandera. 1978. Where Does the Sociolinguistic Variable Stop? *Language in Society*, 7(2):171–182.

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

David K. Lewis. 1969. *Convention: A Philosophical Study*. Wiley-Blackwell.

Andy Lücking, Robin Cooper, Staffan Larsson, and Jonathan Ginzburg. 2019. Distribution is not enough: Going Firther. In *Proceedings of the Sixth Workshop on Natural Language and Computer Science*, pages 1–10, Gothenburg, Sweden. Association for Computational Linguistics.

James Milroy and Lesley Milroy. 1985. Linguistic change, social network and speaker innovation. *Journal of Linguistics*, 21(02):339.

Lesley Milroy. 1987. *Language and Social Networks*. Wiley.

Jenny Myrendal. 2019. Negotiating meanings online: Disagreements about word meaning in discussion forum communication - Jenny Myrendal, 2019. *Discourse Studies*, 21(3):317–339.

John Newman. 2015. Semantic shift. In Nick Riemer, editor, *The Routledge Handbook of Semantics*. Routledge, London.

Kerstin Norén and Per Linell. 2007. Meaning potentials and the interaction between lexis and contexts: An empirical substantiation. *Pragmatics*, 17(3):387–416.

J.C. Paolillo. 1999. The virtual speech community: Social network and language variation on IRC. In *Proceedings of the 32nd Annual Hawaii International Conference on Systems Sciences. 1999. HICSS-32. Abstracts and CD-ROM of Full Papers*, volume Track2, pages 10 pp.–, Torun, Italy.

Hermann Paul. 1886. *Prinzipien der Sprachgeschichte*. Max Niemeyer.

Matthew Peters, Mark Neumann, Mohit Iyyer, Matt Gardner, Christopher Clark, Kenton Lee, and Luke Zettlemoyer. 2018. Deep Contextualized Word Representations. In *Proceedings of the 2018 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 1 (Long Papers)*, pages 2227–2237, New Orleans, Louisiana. Association for Computational Linguistics.

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.

Dominik Schlechtweg, Barbara McGillivray, Simon Hengchen, Haim Dubossarsky, and Nina Tahmasebi. 2020. SemEval-2020 Task 1: Unsupervised Lexical Semantic Change Detection. In *Proceedings of the Fourteenth Workshop on Semantic Evaluation*, pages 1–23, Barcelona (online). International Committee for Computational Linguistics.

Devyani Sharma and Robin Dodsworth. 2020. Language Variation and Social Networks. *Annual Review of Linguistics*, 6(1):341–361.

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.

Ian Stewart, Dustin Arendt, Eric Bell, and Svitlana Volkova. 2017. Measuring, Predicting and Visualizing Short-Term Change in Word Representation and Usage in VKontakte Social Network. In *Proceedings of the Eleventh International AAAI Conference on Web and Social Media (ICWSM 2017)*, page 4, Montréal, Québec, Canada.

Nina Tahmasebi, Lars Borin, and Adam Jatowt. 2018. Survey of Computational Approaches to Diachronic Conceptual Change. *arXiv:1811.06278 [cs]*, 1(1).

Xuri Tang. 2018. A state-of-the-art of semantic change computation. *Natural Language Engineering*, 24(5):649–676.

K. Vani, Sandra Mitrovic, Alessandro Antonucci, and Fabio Rinaldi. 2020. SST-BERT at SemEval-2020 Task 1: Semantic Shift Tracing by Clustering in BERT-based Embedding Spaces. *arXiv:2010.00857 [cs]*.

Duncan J. Watts and Steven H. Strogatz. 1998. Collective dynamics of 'small-world' networks. *Nature*, 393(6684):440–442.

E. Judith Weiner and William Labov. 1983. Constraints on the agentless passive. *Journal of Linguistics*, 19(1):29–58.

Derry Tanti Wijaya and Reyyan Yeniterzi. 2011. Understanding semantic change of words over centuries. In *Proceedings of the 2011 International Workshop on DETecting and Exploiting Cultural diversiTy on the Social Web - DETECT '11*, page 35, Glasgow, Scotland, UK. ACM Press.

Alain F. Zuur, Elena N. Ieno, and Chris S. Elphick. 2010. A protocol for data exploration to avoid common statistical problems. *Methods in Ecology and Evolution*, 1(1):3–14.

A Subreddit selection

We randomly selected 50 subreddits from the set of all forums with at least 15,000 comments per month for each of the 36 months in the 2015–2017 period. We initially selected 50 subreddits but excluded five from further analysis: two which were primarily non-English, two with particularly short average comment lengths, and one where our procedure for identifying template-generated posts failed (see Section B for details).

B Data preprocessing

Below we describe the preprocessing procedure we used to prepare training data for our diachronic SGNS models.

Duplicate comments Before any text normalization steps (described below), we sought to remove duplicate template-generated posts by bots and moderating tools. Since this automated content frequently appears in only one of the two time periods, it can have an outsized effect on the cosine change score of words included in the template.

We identified these posts by comparing the tail (after the first 50 characters) any two posts of more than 50 characters in length. Posts marked as duplicate under this criteria were discarded (keeping one such post in each category). This preserves "natural" human-written duplicates, which tend to be short, while catching most template-generated content, where form-filled deviations tend to be relegated to the beginning of the post. Unfortunately, this criteria missed posts by a bot in the /r/jailbreak subreddit, resulting rectified semantic change score outliers for certain words in the bot's template. As a result, we excluded this community from analysis in the mixed-effects model.

Normalization and tokenization The text of comments was normalized as follows. We removed markdown formatting, extracting only rendered text. We exclude the content of block quotes, code blocks, and tables. We tokenized comments using the SpaCy tokenizer with the default English model (version 2.2.3). We lower-cased all tokens and removed whitespace, including linebreaks. Additionally, we removed tokens containing certain characters present in the 2015 data but absent in 2017, apparently due to text encoding changes made by Reddit. The removed characters were mostly emojis and certain Hangul graphemes and none were particularly common in our data (see [link] for a list of excluded characters).

C Vocabulary and SGNS training proceedure

For each community c we maintain a separate vocabulary, V_c. Words with at least 50 occurances in *both* the 2015 and 2017 time periods are included in the vocabulary. Likewise, the generic Reddit models have vocabulary V_G, which includes words with at least 500 occurances in both time periods.

All models were trained with the Gensim (v. 3.8.1) SGNS implementation, with 200 dimensional vectors for 50 epochs (for both the generic and community-spcefic models). For all other hyperparameters, we maintain the default hyperparameters (length 5 context window, 5 negative samples per word, inital learning rate of 0.025, subsampling threshold of 1×10^{-5}, and negative sampling distribution exponent of 0.75).

For $M_{c,2015}$, we randomly initialize vectors for words in $V_c \setminus V_G$. Words in $V_G \setminus V_c$ have no vector representation in $M_{c,2015}$ or $M_{c,2017}$.