Short-term meaning shift: a distributional exploration

Anonymous NAACL submission

Abstract

We present the first exploration of meaning shift over short periods of time in online communities using distributional representations. We create a small annotated dataset and use it to assess the performance of a standard model for meaning shift detection on short-term meaning shift. We find that the model has problems distinguishing meaning shift from referential phenomena, and propose a measure of contextual variability to remedy this.

1 Introduction

Semantic change has received increasing attention in empirical Computational Linguistics / NLP in the last few years (Tang, 2018; Kutuzov et al., 2018). Almost all studies so far have focused on meaning shift in long periods of time—decades to centuries. However, the genesis of meaning shift and the mechanisms that produce it operate at much shorter time spans, ranging from the online agreement on words' meaning in dyadic interactions (Brennan and Clark, 1996) to the rapid spread of new meanings in relatively small communities of people (Wenger, 1998; Eckert and McConnell-Ginet, 1992). This paper is, to the best of our knowledge, the first exploration of the latter phenomenon—which we call *short-term meaning shift*—using distributional representations.

Short-term shift is usually hard to observe in standard language, such as the language of books or news, which has been the focus of long-term shift studies (e.g., Hamilton et al., 2016; Kulkarni et al., 2015), since it takes a long time for a new meaning to be widely accepted in the standard language. We therefore focus on the language produced in an online community of speakers, in which the adoption of new meanings happens at a much faster pace (Clark, 1996; Hasan, 2009).

Our contribution is twofold. First, we create a small dataset of short-term shift for analysis and evaluation, and qualitatively analyze the types of meaning shift we find.[1] This is necessary because, unlike studies of long-term shift, we cannot rely on material previously gathered by linguists or lexicographers. Second, we test the behavior of a standard distributional model of semantic change when applied to short-term shift. Our results show that this model successfully detects most shifts in our data, but it overgeneralizes. Specifically, the model gets confused with contextual changes due to speakers in the community often talking about particular people and events, which are frequent on short time spans. We propose to use a measure of contextual variability to remedy this and showcase its potential to spot false positives of referential nature like these. We thus make progress in understanding the nature of semantic shift and towards improving computational models thereof.

2 Related Work

Distributional models of semantic change are based on the hypothesis that a change in context of use mirrors a change in meaning. This in turn stems from the Distributional Hypothesis, that states that similarity in meaning results in similarity in context of use (Harris, 1954). Therefore, all models (including ours) spot semantic shift as a change in the word representation in different time periods. This includes classic distributional representations obtained with count-based models, topic modelling or Latent Semantic Analysis (Sagi et al., 2011; Jatowt and Duh, 2014; Wijaya and Yeniterzi, 2011; Gulordava and Baroni, 2011) as well as more recent, neural-network based word embeddings (Kim et al., 2014; Kulkarni et al., 2015; Hamilton et al., 2016; Azarbonyad et al., 2017; Szymanski, 2017; Yao et al., 2018). Note, however, that distributional representations from

[1] Data and code will be made available upon publication.

different time periods are not easily comparable. We adapt the method by Kim et al. (2014) to enable comparison (see Section 3).

Unlike previous work, we focus on the language of online communities. Recent studies of this type of language have investigated the spread of new forms and meanings (Del Tredici and Fernández, 2017, 2018; Stewart and Eisenstein, 2018), competing lexical variants (Rotabi et al., 2017), and the relation between conventions in a community and the social standing of its members (Danescu-Niculescu-Mizil et al., 2013). None of these works has analyzed the ability of a distributional model to capture these phenomena, which is what we do in this paper for short-term meaning shift.

Evaluation of semantic shift is difficult, due to the lack of annotated datasets (Fermann and Lapata, 2016). For this reason, even for long-term shift, evaluation is usually performed by manually inspecting the n words whose representation changes the most according to the model under investigation (Hamilton et al., 2016; Kim et al., 2014). Our dataset allows for a more systematic evaluation and analysis, and enables comparison in future studies.

3 Experimental Setup

Data. We exploit user-generated language from an online forum of football fans, namely, the r/LiverpoolFC subreddit, one of the many communities hosted by the Reddit platform.[2] We focus on a short period of eight years, between 2011 and 2017. In order to enable a clearer observation of short-term meaning shift, we define two non-consecutive time bins: the first one (t_1) contains data from 2011–2013 and the second one (t_2) from 2017.[3] We also use a large sample of community-independent language for the initialization of the word vectors, namely, a random crawl from Reddit in 2013. Table 1 shows the size of each sample.

Model. In the method proposed by Kim et al. (2014), word embeddings for the first time bin t_1 are initialized randomly; then, given a sequence of time-related samples, embeddings for t_i are initialized using the embeddings of t_{i-1} and further updated. If at t_i the word is used in the same contexts as in t_{i-1}, its embedding will only be marginally updated, whereas a major change in the context of use will lead to a stronger update of the embedding. The model makes embeddings across time bins directly comparable.

sample	time bin	million tokens
Reddit$_{13}$	2013	~900
LiverpoolFC$_{13}$	2011–13	8.5
LiverpoolFC$_{17}$	2017	11.9

Table 1: Time bin and size of the datasets.

We implement the following steps: First, we create randomly initialized word embeddings with the large sample Reddit$_{13}$ to obtain meaning representations that are community-independent. We then use these embeddings to initialize those in LiverpoolFC$_{13}$, update the vectors on this sample, and thus obtain embeddings for time t_1. This step adapts the general embeddings to the LiverpoolFC community. Finally, we initialize the word embeddings for LiverpoolFC$_{17}$ with those of t_1, train on this sample, and get embeddings for t_2.

The vocabulary is defined as the intersection of the vocabularies of the three samples (Reddit$_{13}$, LiverpoolFC$_{13}$, LiverpoolFC$_{17}$), and includes 157k words. For Reddit$_{13}$, we include only words that occur at least 20 times in the sample, so as to ensure meaningful representations for each word, while for the other two samples we do not use any frequency threshold: Since the embeddings used for the initialization of LiverpoolFC$_{13}$ encode community-independent meanings, if a word doesn't occur in LiverpoolFC$_{13}$ its representation will simply be as in Reddit$_{13}$, which reflects the idea that if a word is not used in a community, then its meaning is not altered within that community. We train with standard skip-gram parameters (Levy et al., 2015): window 5, learning rate 0.01, embedding dimension 200, hierarchical softmax.

Evaluation dataset. Our dataset consists of 97 words from the r/LiverpoolFC subreddit with annotations by members of the subreddit —that is, community members with domain knowledge (needed for this task) but no linguistic background.

To ensure that we would get enough cases of semantic shift to enable a meaningful analysis, we started out from content words that increase their relative frequency between t_1 and t_2.[4] A threshold

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

[3]These choices ensure that the samples in these two time bins are approximately of the same size – see Table 1. The r/LiverpoolFC subreddit exists since 2009, but very little content was produced in 2009–2010.

[4]Frequency increase has been shown to positively correlate with meaning change (Wijaya and Yeniterzi, 2011; Kulkarni et al., 2015); although it is not a necessary condition, it is a reasonable starting point, as a random selection of

of 2 standard deviations above the mean yielded ∼200 words. The first author manually identified 34 semantic shift candidates among these words by analyzing their contexts of use in the r/LiverpoolFC data. Semantic shift is defined here as a change in the ontological type that a word denotes (cf. examples in Sec. 4). We added two types of confounders: 33 words with a significant frequency increase but not marked as meaning shift candidates, and 33 words with constant frequency between t_1 and t_2, included as a sanity check.[5]

The participants were shown the 100 words (in randomized order) together with randomly chosen contexts of usage from each time period (μ=4.7 contexts per word) and, for simplicity, were asked to make a binary decision about whether there was a change in meaning. As semantic shift is arguably a graded notion, we aggregate the annotations into a graded *semantic shift index*, ranging from 0 (no shift) to 1 (shift) depending on how many subjects spotted semantic change.[6] Overall, 26 members of r/LiverpoolFC participated in the survey, and each word received on average 8.8 judgements. Three words were discarded after analysis of the redditor data.[7] Further details, including the instructions to participants, are in the supplementary material.

4 Types of Meaning Shift

We identify three main types of shift in our data via qualitative analysis of examples with a shift index > 0.5: metonymy, metaphor, and meme.

In metonymic shifts, a highly salient characteristic of an entity is used to refer to it. Among these cases are, for example, *'highlighter'*, which in t_2 occurs in sentences like *'we are playing with the highlighter today'*, used to talk about a kit in a colour similar to that of a highlighter pen; or *'lean'*, in *'I hope a lean comes soon!'*, used to talk about hiring players due to new hires typically leaning on a Liverpool symbol when posing for a photo right after signing for the club. Particularly illustrative is the 'F5' example shown in Table 2. While 'F5' is initially used with its common usage of shortcut for refreshing a page (1), it then starts to denote the act of refreshing in order to get the latest news about the possible transfer of a new player to LiverpoolFC (2). This use catches on and many redditors use it to express their tension while waiting for good news (3-5),[8] though not all members are aware of the new meaning of the word (6). After the player finally signed for the team, someone leaves the *'F5 squad'* (7), and after a while, another member recalls the period in which the word was used (8).

(1)	*after losing the F5 key on my keyboard...*	18 Jun
(2)	*F5 tapping is so intense now. I want him*	28 Jun
(3)	*Don't think about it too much, man. Just F5*	1 Jul
(4)	*just woke up and thought it was f5 time*	3 Jul
(5)	*this was a happy f5*	13 Jul
(6)	*what is an F5?*	13 Jul
(7)	*I'm leaving the f5 squad for now*	5 Aug
(8)	*I made this during the f5 madness in the sub*	6 Sep

Table 2: Examples of use of 'F5' with time stamps, which illustrate the speed of the meaning shift process. All the examples are from LiverpoolFC$_{17}$.

Metaphorical shifts lead to a broadening of the original meaning of a word through analogy. For example, in t_2 *'shovel'* occurs in sentences such as *'welcome aboard, here is your shovel'* or *'you boys know how to shovel coal'*: the team is seen as a train that is running through the season, and every supporter is asked to figuratively contribute by shoving coal into the train boiler.

Finally, memes—where fans use a word to make jokes and be sarcastic—are another prominent source of meaning shift. For instance, as the club was about to sign a new player named Van Dijk, redditors exploited the homography of the surname with the common noun *'van'* and the shoe brand 'Vans': '*Rumour has it Van Djik was wearing these vans in the van*'. Jokes of this kind are positively received by the community ('*Hahah I love it. Anything with vans is instant karma!*') and quickly become frequent in it.

5 Modeling Results and Analysis

The positive correlation between cosine distance and semantic shift index (Pearson's r= 0.49, $p < 0.001$, see Figure 1) confirms the hypothesis that meaning shift is mirrored by a change in context of use. However, we also find systematic deviations.

False negatives. A small, but consistent group is that of words that undergo semantic shift but are

words would contain very few positive examples. Our dataset is thus biased towards precision over recall.

[5] All words have absolute frequency in range [50–500].

[6] The shift index is exclusively based on the judgements by the redditors, and does not consider the preliminary candidates selection done by us.

[7] The words are: 'discord' and 'owls' due to the homonymy with proper names not detected during survey's implementation; 'tracking' because the chosen examples clearly mislead the judgements of the redditors.

[8] Here the semantic change is accompanied by a change in the part of speech, and 'F5' becomes a denominal verb.

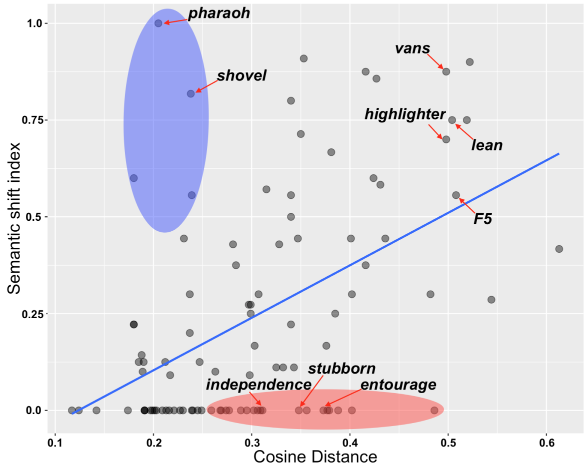

Figure 1: Semantic shift index vs. cosine distance in the evaluation dataset (Pearson's $r = 0.49$, $p < 0.001$). Red ellipsis: false positives; blue ellipsis: false negatives.

not captured by the model (blue ellipsis Figure 1; shift index>0.5, cosine distance<0.25). These are all metaphorical shifts; in particular, cases of extended metaphor (Werth, 1994), where the metaphor is developed throughout the whole text. For instance, besides the *'shovel'* example mentioned in Section 4, we find *'pharaoh'*, the nickname of an Egyptian player who joined Liverpool in 2017, used in sentences like *'approved by our new Pharaoh Tutankhamun'*, or *'our dear Egyptian Pharaoh, let's hope he becomes a God'*. Despite the metaphoric usage, the local context of these words is similar to the literal one, and so the model does not spot the meaning shift. We expect this to happen in long-term shift models, too, but we are not aware of results confirming this.

False positives. A larger group of problematic cases is that of words that do *not* undergo semantic shift despite showing relatively large differences in context between t_1 and t_2 (red ellipsis in Figure 1; shift index=0, cosine distance>0.25). Manual inspection reveals that most of these "errors" are due to a referential effect: words are used almost exclusively to refer to a specific person or event, and so the context of use is different from the contexts in t_1. For instance, *'stubborn'* is almost always used to talk about a coach who was not there in 2013 but only in 2017; *'entourage'*, for the entourage of a particular star of the team; *'independence'* for the political events in Catalonia (Spain). In all these cases, the meaning of the word stays the same, despite the change in context. In line with the Distributional Hypothesis, the model spots the context change, but it is not sensitive to its nature. We expect long-term shift to not be as susceptible to referential effects like these because embeddings are aggregated over a larger and more varied number of occurrences.

We expect that in referential cases the contexts of use will be *narrower* than for words with actual semantic shift, as they are specific to one person or event. Hence, a measure of *contextual variability* should help spot false positives. To test this hypothesis, we define contextual variability as follows: for a target word, we create a vector for each of its contexts (5 words on both sides of the target) in t_2 by averaging the embeddings of the words occurring in it, and define variability as the average pairwise cosine distance between context vectors. We find that contextual variability is indeed significantly correlated with semantic shift in our dataset (Pearson's $r = 0.55$, $p < 0.001$), while it is independent from cosine distance (Pearson's r= 0.18, $p > 0.05$). These two aspects are thus complementary. While both shift words and referential cases change context of use in t_2, context variability captures the fact that only in referential cases words occur in a restricted set of contexts. Figure 2 shows this effect visually. This result can inform future models of short-term meaning shift.

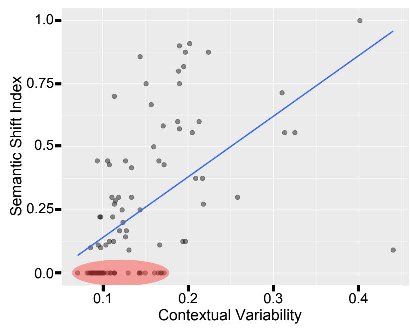

Figure 2: Semantic shift index vs. context variability. Red ellipsis: referential cases which are assigned high cosine distance values by the model (false positives).

6 Conclusion

The goal of this preliminary study was to bring to the attention of the NLP community short-term meaning shift, an under-studied problem in the field. We hope that it will spark further research into a phenomenon which, besides being of theoretical interest, has potential practical implications for NLP downstream tasks concerned with user-generated language, as modeling how word meanings rapidly change in communities would allow a better understanding of what their members say.

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*, pages 1509–1518. ACM.

Susan E. Brennan and Herbert H. Clark. 1996. Conceptual pacts and lexical choice in conversation. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 22(6):1482.

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*, pages 307–318. ACM.

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 and Raquel Fernández. 2018. The road to success: Assessing the fate of linguistic innovations in online communities. In *COLING 2018-27th International Conference on Computational Linguistics-Long papers*.

Penelope Eckert and Sally McConnell-Ginet. 1992. Communities of practice: Where language, gender, and power all live. In *Locating Power, Proceedings of the 1992 Berkeley Women and Language Conference*, pages 89–99.

Lea Frermann and Mirella Lapata. 2016. A bayesian model of diachronic meaning change. *Transactions of the Association for Computational Linguistics*, 4:31–45.

Kristina Gulordava and Marco Baroni. 2011. A distributional similarity approach to the detection of semantic change in the google books ngram corpus. In *Proceedings of the GEMS 2011 Workshop on GEometrical Models of Natural Language Semantics*, pages 67–71. 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)*, volume 1, pages 1489–1501.

Zellig S Harris. 1954. Distributional structure. *Word*, 10(2-3):146–162.

Ruqaiya Hasan. 2009. *Semantic variation: Meaning in society and in sociolinguistics*. Equinox London.

Adam Jatowt and Kevin Duh. 2014. A framework for analyzing semantic change of words across time. In *Proceedings of the 14th ACM/IEEE-CS Joint Conference on Digital Libraries*, pages 229–238. IEEE Press.

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. Association for Computational Linguistics.

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*, pages 625–635. ACM.

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 (COLING)*, pages 1384–1397. Association for Computational Linguistics.

Omer Levy, Yoav Goldberg, and Ido Dagan. 2015. Improving distributional similarity with lessons learned from word embeddings. *Transactions of the Association for Computational Linguistics*, 3:211–225.

Rahmtin Rotabi, Cristian Danescu-Niculescu-Mizil, and Jon Kleinberg. 2017. Competition and selection among conventions. In *Proceedings of the 26th International Conference on World Wide Web*, pages 1361–1370. International World Wide Web Conferences Steering Committee.

Eyal Sagi, Stefan Kaufmann, and Brady Clark. 2011. Tracing semantic change with latent semantic analysis. *Current methods in historical semantics*, pages 161–183.

Ian Stewart and Jacob Eisenstein. 2018. Making" fetch" happen: The influence of social and linguistic context on the success of lexical innovations. *Proceedings of Conference on Empirical Methods in Natural Language Processing (EMNLP)*.

Terrence Szymanski. 2017. Temporal word analogies: Identifying lexical replacement with diachronic word embeddings. In *Proceedings of the 55th Annual Meeting of the Association for Computational Linguistics (Volume 2: Short Papers)*, volume 2, pages 448–453.

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

Etienne Wenger. 1998. *Communities of practice: Learning, meaning, and identity*. Cambridge University Press.

Paul Werth. 1994. Extended metaphor—a text-world account. *Language and literature*, 3(2):79–103.

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*, pages 35–40. ACM.

Zijun Yao, Yifan Sun, Weicong Ding, Nikhil Rao, and Hui Xiong. 2018. Dynamic word embeddings for evolving semantic discovery. In *Proceedings of the Eleventh ACM International Conference on Web Search and Data Mining*, pages 673–681. ACM.