

A rich tradition of cross-linguistic, detailed studies in lexical semantics has generated hypotheses that are ripe for further quantitative exploration made possible by the outcomes of this project. For example, synchronically: there are cognitively different kinds of temperature sensations, which form overt or covert sub-senses (Koptjevskaja-Tamm, 2015), that may be handled by different grammatical constructions (Pustet, 2015). Diachronically: word senses often go from describing external reality to internal perception (Traugott and Dasher, 2002); change often proceeds from a higher sensory modality to a lower (Viberg, 1980); perception verbs develop into cognitive verbs (Sweetser, 1991); or the exact hues denoted by Swedish color terms may change in a predictable fashion over time (Vejdemo, 2017).

The first methods for automatic detection of word sense change were based on context vectors; they investigated semantic density (Sagi et al., 2009) and utilized mutual information (Gulordava and Baroni, 2011) to identify semantic change over time; both methods detect signals of change but neither aligns senses over time or determines what changed. Topic-based models (where topics are interpreted as senses) were used to detect novel senses by identifying new topics in a later corpus compared to an earlier one (Lau et al. (2012); Cook et al. (2014)), or by clustering topics over time (Wijaya and Yeniterzi, 2011). A dynamic topic model where topics for time t are learned with respect to topics from time $t-1$ is proposed by Frermann and Lapata (2016). With one exception, no alignment is made between topics to allow following the diachronic progression of a sense.

Graph-based models are utilized by Mitra et al. (2015, 2014) and Tahmasebi and Risse (2017) and aim to revealing complex relations between a word's senses by (a) modeling senses per se using WSI; and (b) aligning senses over time. The models allow differentiation of individual senses at different periods in time and the latter also group senses into linguistically related concepts (Cooper, 2005).

The largest body of work is done using word embeddings of different kinds in the last years (Basile et al., 2016; Kim et al., 2014; Zhang et al., 2016). Embeddings are trained on different time-sliced corpora and compared over time. Kulkarni et al. (2015) project words onto their frequency, POS and semantic vectors and propose a model for detecting statistically significant changes between time periods. Hamilton et al. (2016) view both similarity between a priori known pairs of words, and between a word's vectors over time to detect change and have released HistWords, a set of pre-trained historical word embeddings. Basile et al. (2016); Hamilton et al. (2016); Kulkarni et al. (2015) all propose different methods for projecting vectors from different time periods onto the same space to allow comparison. Bamler and Mandt (2017) propose using dynamic embeddings to completely avoid projection, which results in smoother diachronic embeddings.

Existing methods for detecting change based on word embeddings do not allow us to recover individual senses; they model words with one vector per time unit and detect meaning change as change in the direction of those vectors. Once a change point is found, the most similar words around that time are used to illustrate the change. However, the most similar terms will only represent the dominant sense; they will not reflect the other senses or capture stable parts of a word. If multi-sense embeddings are used, to allow senses to be modeled individually, we again face multiple vectors for each time point, over hundreds of years, leading to an overwhelming result space. Solving this reduction and investigating the tradeoff between a large results space and information loss during reduction, is at the core of this field and largely ignored in previous work.

The plan

Existing approaches to change detection derive sense representations first and compare these over time. The most recent approaches make use of vector representations, providing a methodology that can answer when something has changed, but not how. In addition, vector representations have the target word as one unit without differentiating between a word's senses and work poorly on smaller corpora (which our historical corpora are, but also our modern corpora might be if we split them up in e.g. yearly subsets). Therefore, we need to go beyond vector representations in the following ways, (i) model word senses separately e.g. using multi-sense embeddings; (ii) allow all senses to change individually; (iii) differentiate change types (novel /outdated senses, broadening/narrowing, related/un-related senses as well as stable senses); (iv) model the relation between the senses; and (v) handle smaller amounts of (noisy) text.

Our experiments, e.g., Tahmasebi et al (2017), show that a word like the *telephone* have senses in actual usage indicating that the word goes from being a *property of a building* to the belonging to each *household* to being a *tool for communication* much like a radio, television or a newspaper. The word

travel has always had the *moving from point A to point B* sense but was mostly used for describing literature; people read about traveling. First in the early 20th century, the word usage added the sense of actual travel by means of train, boat and eventually airplane, the literature sense still being valid but less frequent. This usage change is extremely interesting, reflects cultural change rather than explicit sense change, and is found only if we differentiate and align a word's senses over time.

We will investigate count-based methods like SVD_{ppmi} (Hamilton et al., 2016) that do not suffer from randomness, and the dynamic embeddings, which increase robustness by automatically aligning embeddings at different time points, will be extended to account for multiple word senses. In addition, we will go beyond analyzing individual words to capture changes in semantic fields.

Using sense-differentiated embeddings targeted for small amounts of data, and our reduction techniques, we will answer what happens to *all* senses of a word over time, allowing senses to change individually while others stay stable for the same word. We will focus on precision, which is particularly important for uptake, i.e., presenting good quality results without too much noise. We will find (and rank) good text passages that exemplify each meaning and change. As proof of concept, we will revisit and enlarge several previous manually executed case studies on synchronic word sense variation and diachronic word sense change in Swedish, and look deeper into active lexical semantics research fields such as temperature, color, and smell. We will approach the under-investigated topic of how word senses change with different speed and processes in different linguistic sources (the inertia of change differs in e.g. fiction and newspapers) and different time periods (some, but not all, published work have noted an increased level of SC after the world wars, see e.g. Juola (2003)).

- R. Bamler and S. Mandt. Dynamic word embeddings. In *ICML*, 2017.
- P. Basile, A. Caputo, R. Luisi, and G. Semeraro. Diachronic analysis of the italian language exploiting google ngram. In *Italian Conf. on Comp. Linguistics*, 2016.
- P. Cook, J. H. Lau, D. McCarthy, and T. Baldwin. Novel word-sense identification. In *COLING*, 2014.
- M. C. Cooper. A Mathematical Model of Historical Semantics and the Grouping of Word Meanings into Concepts. *Computational Linguistics*, 32(2):227–248, 2005.
- L. Frermann and M. Lapata. A bayesian model of diachronic meaning change. *TACL*, 4:31–45, 2016.
- K. Gulordava and M. Baroni. A distributional similarity approach to the detection of semantic change in the google books ngram corpus. In *GEMS workshop*, 2011.
- W. L. Hamilton, J. Leskovec, and D. Jurafsky. Diachronic word embeddings reveal statistical laws of semantic change. In *ACL*, 2016.
- P. Juola. The Time Course of Language Change. *Computers and the Humanities*, 37(1):77–96, 2003.
- Y. Kim, Y.-I. Chiu, K. Hanaki, D. Hegde, and S. Petrov. Temporal analysis of language through neural language models. In *LACSS*, 2014.
- M. Koptjevskaja-Tamm. *Introducing the linguistics of temperature*, volume 107 of *Typological Studies in Language*, page 1–40. John Benjamins Publishing Company, 2015.
- V. Kulkarni, R. Al-Rfou, B. Perozzi, and S. Skiena. Statistically significant detection of linguistic change. In *WWW*, 2015.
- J. H. Lau, P. Cook, D. McCarthy, D. Newman, and T. Baldwin. Word Sense Induction for Novel Sense Detection. In *EACL*, 2012.
- S. Mitra, R. Mitra, M. Riedl, C. Biemann, A. Mukherjee, and P. Goyal. That's sick dude!: Automatic identification of word sense change across different timescales. In *ACL*, 2014.
- S. Mitra, R. Mitra, S. K. Maity, M. Riedl, C. Biemann, P. Goyal, and A. Mukherjee. An automatic approach to identify word sense changes in text media across timescales. *NLE*, 21(05):773–798, 2015.
- R. Pustet. *The syntax of temperature predications*, page 889–916. John Benjamins Publishing Company, 2015.
- E. Sagi, S. Kaufmann, and B. Clark. Semantic density analysis: comparing word meaning across time and phonetic space. In *GEMS workshop*, 2009.
- E. Sweetser. *From Etymology to Pragmatics: Metaphorical and Cultural Aspects of Semantic Structure*. Cambridge University Press, July 1991.
- N. Tahmasebi and T. Risse. Finding individual word sense changes and their delay in appearance. In *RANLP*, 2017.
- E. Traugott and R. B. Dasher. *Regularity in semantic change*. Cambridge University Press, Cambridge ; New York, 2002.
- S. Vejdemo. *Triangulating Perspectives on Lexical Replacement: From Predictive Statistical Models to Descriptive Color Linguistics*. Stockholm University, 2017.
- A. Viberg. *Studier i kontrastiv lexikologi: perceptionsverb*. Stockholms Universitet, Inst. för lingvistik, 1980.
- D. T. Wijaya and R. Yeniterzi. Understanding semantic change of words over centuries. In *DETECT*, 2011.
- Y. Zhang, A. Jatowt, and K. Tanaka. Detecting evolution of concepts based on cause-effect relationships in online reviews. In *WWW*, 2016.