

# DismemBERT: Detecting Diachronic Lexical Semantic Change Using BERT Embeddings in an Unsupervised Knowledge-Free Setting

David Rother

TU Darmstadt

Thomas Haider

thomas.haider@ae.mpg.de

Steffen Eger

eger@ai.phys.tu-darmstadt.de

david.rother@stud.tu-darmstadt.de

## Abstract

This document contains the instructions for preparing a paper submitted to COLING-2020 or accepted for publication in its proceedings. The document itself conforms to its own specifications, and is therefore an example of what your manuscript should look like. These instructions should be used for both papers submitted for review and for final versions of accepted papers. Authors are asked to conform to all the directions reported in this document.

## 1 Introduction

Here is something regarding semantic change (Schlechtweg et al., 2018)

## 2 Related Work

### 2.1 Diachronic Lexical Semantic Change

With an increasing interest in Diachronic Lexical Semantic Change (LSC) there is a multitude of approaches. Three different word representations are commonly used. First are semantic vector representations such as word2vec (Mikolov et al., 2013), which represents each word with two different vectors for each time period respectively (Hamilton et al., 2016a; Hamilton et al., 2016b). The vectors themselves represent the co-occurrence statistics of the word in the given time period. (Frermann and Lapata, 2016) use bayesian learning

### 2.2 Diachronic Word Embeddings

### 2.3 Unsupervised Knowledge-free Sense Modelling

## 3 Corpora

## 4 Framework

## 5 Experiments

## 6 Evaluation

Hi

## 7 Conclusion

## References

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

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 Conference on Empirical Methods in Natural Language Processing. Conference on Empirical Methods in Natural Language Processing*, volume 2016, page 2116. NIH Public Access.

William L Hamilton, Jure Leskovec, and Dan Jurafsky. 2016b. Diachronic word embeddings reveal statistical laws of semantic change. *arXiv preprint arXiv:1605.09096*.

Tomas Mikolov, Kai Chen, Greg Corrado, and Jeffrey Dean. 2013. Efficient estimation of word representations in vector space. *arXiv preprint arXiv:1301.3781*.

Dominik Schlechtweg, Sabine Schulte im Walde, and Stefanie Eckmann. 2018. Diachronic usage relatedness (durel): A framework for the annotation of lexical semantic change. *arXiv preprint arXiv:1804.06517*.