BOOK REVIEW

Nina Tahmasebi, Lars Borin, Adam Jatowt, Yang Xu & Simon Hengchen, eds. *Computational Approaches to Semantic Change* (= *Language Variation*, 6). Berlin: Language Science Press, 2021. doi

Reviewed by Christin Beck (University of Konstanz)

While research on diachronic semantic change has a long-standing tradition in historical linguistics, computational linguists only recently began to take an interest in the topic. This interest is growing presently, with a surge of computational work on lexical semantic change over the past few years. One factor behind this growing interest is that investigations of lexical semantic change present an interesting testbed for state-of-the-art Natural Language Processing (NLP) technologies such as neural word embeddings. These embeddings are distributional representations of a word, generated by a neural language model (e.g., BERT (Devlin et al. 2019), word2vec (Mikolov et al. 2013) and fasttext (Joulin et al. 2016)) in the form of vectors that are able to represent the lexical meaning of a word to some extent (cf. Wiedemann et al. 2019).

The volume *Computational Approaches to Semantic Change* presents current computational applications of semantic change research, discussing and highlighting the challenges which remain for future work. The book comprises eleven chapters, covering an interesting variety of topics and a broad range of computational methodologies. The volume begins with a survey of the computational technologies developed for the detection of lexical semantic change (Chapter 1), and it is concluded by a chapter surveying visualization systems for lexical semantic change (Chapter 10) and a chapter that discusses future challenges (Chapter 11). The remaining chapters (Chapters 2–9) present individual research efforts, with different methodologies proposed for the investigation of lexical semantic change across different languages. In part, the book results from work presented at the 1st International Workshop on Historical Language Change,[1] which aimed at bringing historical and computational linguists closer together to initiate collaborative research efforts. This gives the impression that the book is meant for a computational as well as historical linguistic readership. Yet, despite its methodological extensiveness, the chapters are mostly technical in nature, focusing on describing the underlying algorithms and computational models rather than the linguistic use

1. https://languagechange.org/events/2019-acl-lcworkshop/

https://doi.org/10.1075/jhl.22063.bec | Published online: 30 October 2023
Journal of Historical Linguistics 14:2 (2024), pp. 376–384. ISSN 2210-2116 | E-ISSN 2210-2124

cases. Therefore, the volume will be more useful to computational linguists than to the historical linguistic community. Still, this book provides a great, hands-on overview of relevant state-of-the-art methodologies.

In the first chapter, "Survey of Computational Approaches to Lexical Semantic Change Detection", Nina Tahmasebi, Lars Borin and Adam Jatowt provide a detailed and extensive overview of the various computational techniques proposed in recent years for investigating lexical semantic change. Methods for modeling changes in word meaning are, for example, co-occurrence-based methods (e.g., Positive Pointwise Mutual Information (PPMI) scores, Turney & Pantel (2010)), neural word embeddings and topic models (e.g., Frermann & Lapata 2016). Overall, the authors discuss and compare a variety of different algorithms, methodologies and techniques, showing the chronological development of computational approaches to lexical semantic change detection. The survey ends with a sound discussion of methodological issues and current evaluation techniques.

In Chapter 2, "Semantic Changes in Harm-Related Concepts in English", Ekaterina Vylomova and Nick Haslam develop a computational model for the investigation of semantic change in psychological concepts. The authors investigate Haslam's (2016) 'concept creep' hypothesis in two corpora: a corpus consisting of psychology journal abstracts and a general domain corpus based on COCA (Davies 2008) and COHA (Davies 2012), using English data from the period 1970–2017. The hypothesis states that many harm-related concepts, e.g., *addiction, bullying, harassment, prejudice* and *trauma*, have undergone semantic broadening over the past half-century in Western societies. The authors' model is based on the calculation of word2vec embeddings of the concepts. Cosine similarities between the embeddings are computed to assess the semantic breadth of a word as well as changes over time (following Hamilton, Leskovec & Jurafsky's (2016) work). While the methodology is not by itself innovative, the chapter features an interesting application case and, in addition to presenting quantitative results, contains a detailed qualitative evaluation of the resulting data. This in turn sheds light on the 'concept creep' hypothesis, showing that *addiction, bullying*, and *harassment* have broadened over time, while *prejudice* and *trauma* remained relatively stable.

Chapter 3, "Computation of Semantic Change in Scientific Concepts: Case Study of 'Circular Economy'" by Sampriti Mahanty, Frank Boons, Julia Handl and Riza Batista-Navarro, investigates the semantic change of 'circular economy', a popular concept in sustainability studies, with the goal of detecting conceptual evolution in the process of scientific knowledge production. Their study is based on abstracts of English journal articles extracted from the Scopus database, dating from 2005 to 2019.[2] The authors propose a methodology which combines topic

2. https://www.scopus.com/home.uri

modeling, co-occurrence networks and word embeddings, creating a multi-angle analysis pipeline, which is nicely tailored to the purpose of studying the diachronic evolution of a concept and understanding the evolution of a concept with respect to other concepts in the semantic field. That is, topic modeling is used to detect change points in the corpus, co-occurrence networks based on specific keywords are built to detect the nature of changes in the concept, and finally, word2vec embeddings are computed for 'circular economy' to obtain global semantics, with cosine similarity used as a measure for assessing changes over time. The authors nicely elaborate on their method illustrating the analysis process and their methodology via hands-on visualizations. However, it seems that what the authors describe as semantic change is not a change in meaning per se. Rather, the concept 'circular economy' undergoes a change in topics with which it is associated, moving from *ecology, industrial economics* and *environmental management* to *business models, supply chain* and *product design*.

For a change, Stellan Petersson and Emma Sköldberg's chapter on "Semantic Change in Swedish – From a Lexicographic Perspective" (Chapter 4) does not present a computational application, but it discusses challenges with respect to the lexicographic description of semantic change in Swedish. The authors themselves are part of the editorial team of the dictionary *Svensk ordbok utgiven av Svenska Akademien* ('The Contemporary Dictionary of the Swedish Academy') and highlight several lexicographic issues related to semantic change on the lexical level associated with different types of change, e.g., changes in concepts and their reference, changes in constructional behavior and grammaticalization. Eventually, the authors discuss the implications that using (semi-)automatic computational linguistic methods would have for lexicographic research. The chapter differs from the remaining papers in the volume, but it presents interesting and relevant discussions for future research.

Chapter 5, "Historical Changes in Semantic Weights of Sub-Word Units" by Yang Xu and Zheng-sheng Zhang, examines meaning changes of 'sub-word units' (e.g., morphemes) over time and whether there are any changes in the role which these units play in determining overall word meanings, i.e., changes with respect to their 'semantic weight'. The study focuses on Chinese characters as sub-word units, whose semantic weights are modeled via neural network architectures based on the Character-incorporated Word Embedding model (CWE; Chen et al. 2015) and fasttext. In addition, the authors draw a comparison between the Chinese data and some Indo-European languages, i.e., English, French, German, Italian and Spanish, using data from Wikimedia,[3] as well as the Google Books

3. https://wikimediafoundation.org

Ngram corpus[4] dating from 1500 to 2000. However, the comparison is elusive: while the Chinese characters under consideration do indeed carry meaning, n-grams are considered for the other languages. This results in an arbitrary combination of letters which might not carry any linguistic content. Still, the results for Chinese are relevant and interesting. The authors find that in older Chinese words, characters tend to carry more semantic weight than in newer Chinese words. Moreover, they claim that this is evidence for a historical change in word formation in Chinese; in modern Chinese, multiple characters form a single semantic unit, while in older Chinese, one character corresponds to a semantic unit. Overall, this feeds into claims from qualitative studies that Chinese has moved from mono- to multisyllabicity. The Indo-European languages also tend to show increasing semantic weights on sub-word units over time. These results are taken to be evidence for the Indo-European synthetic to analytic pattern shift at the level of phrase composition. Although the claims for Indo-European need to be taken with a grain of salt, this seems to be an interesting starting point for further investigations into historical changes at the morpheme level – an area which is still understudied from the perspective of computational linguistics.

The next chapter, Chapter 6, "Chaining Algorithms and Historical Adjective Extension" by Karan Grewal and Yang Xu, investigates the emergence of novel adjective-noun pairings, where adjectives extend their meaning, claiming that the emergence adheres to a process of chaining which is sensitive to semantic neighborhood. More precisely, they claim that the 'new' and 'old' referents of a changing adjective are semantically close, with the corresponding historical word meaning extension process following an incremental pattern, e.g., *cold food* → *cold person* → *cold war*. This claim is tested via a set of probabilistic models that predict adjective noun-pairs from historical data stemming from the Google Books corpus (Michel et al. 2011), dating from 1850 to 2000. The different models represent different processes of chaining: (i) the exemplar model is based on the assumption that a query noun should be linked to the adjective category with the highest local semantic density around that noun; (ii) the prototype model, where the query noun is linked to the adjective that has the closest noun prototype in semantic space to the query noun; and (iii) the *k*-nearest neighbors model, selecting the *k* nouns closest to the query noun. The semantic space underlying these models is constructed using word2vec. They find that semantic neighborhood density plays an important role in the emergence of novel adjective-noun pairings, with all three models performing well and better than a frequency-based baseline, confirming insights from theoretical studies on chaining in the extension of grammat-

4. http://storage.googleapis.com/books/ngrams/books/datasetsv2.html

ical categories. In addition, the authors discuss the relevance of their work with respect to a general theory of word meaning extension.

In Chapter 7, “Cross-Lingual Laws of Semantic Change”, Ana-Sabina Uban, Alina Maria Ciobanu and Liviu P. Dinu propose a novel method for the investigation of semantic change by measuring ‘cognate divergence’. Cognate divergence is assessed between cross-lingual synchronic word embeddings with the corresponding cognates extracted from etymology dictionaries. The embeddings are trained on Wikipedia data using fasttext. Their approach is novel in that instead of examining texts from one language across different time intervals, the present meanings of cognates are compared across different languages, taking the cognates to be “snapshots in time of each of the word’s different histories of evolution” (p. 224). Five Romance sister languages are examined: Romanian, French, Italian, Spanish and Portuguese. Additionally, Latin is used for a historical comparison, and English is considered to gain an idea about Latin borrowings in a more remote language. The cross-linguistic approach is particularly interesting since previous computational studies focused on measuring change within a language, with multiple languages generally examined in parallel, not taking into account multilingual settings with language contact as a potential source of change. In addition, their methodology entails a fully automated framework for the detection and correction of deceptive cognate pairs, i.e., ‘false friends’: words which have similar surface forms that originally shared a meaning, but the meaning diverged over time. Moreover, laws of semantic change are put into relation with their approach by exploring the interrelation between frequency and polysemy with the degree of semantic change of words, showing that there is a positive correlation between semantic divergence and frequency as well as polysemy.

Next, in Chapter 8, “Structured Representation of Temporal Document Collections by Diachronic Linguistic Periodization”, Yijun Duan, Adam Jatowt and Masatoshi Yoshikawa present a methodology for learning time-aware word semantics, which is used for the periodization of diachronic document collections. As data basis, the New York Times dataset (Yao et al. 2018) is used, with the examined time span ranging from 1990 to 2016. The goal is to achieve a periodization scheme where within the same period, most words are semantically stable with respect to their senses, while there are recognizable semantic shifts between periods. The temporal-aware word embeddings are learned via an anchor-based joint matrix factorization framework based on time-stamped PPMI matrices. Shared frequent terms are used as anchors to align the embeddings to the same latent space across periods. The periodization task is treated as an optimization problem, with the aim of maximizing the overall differences in word semantics between all pairs of periods. For evaluation, standard evaluation metrics for text segmentation are used (Pk and WinDiff, see p. 264). Although the approach is mathematically solid, the linguistic relevance of the semantic periodization might be delusive. By looking at different words at once, significant developments and changes of individual words might be obscured: it is rather unlikely that each word changes at the same time and at the same rate. Nevertheless, this is a great starting point for future work on semantic periodization.

Chapter 9 presents a method for the investigation of "Lexical Semantic Change for Ancient Greek and Latin", which, in contrast to the previous chapters, is not based on a word embedding model. Instead, Valerio Perrone, Simon Hengchen, Marco Palma, Alessandro Vatri, Jim Q. Smith and Barbara McGillivray make use of a dynamic Bayesian mixture model, namely GASC (genre-aware semantic change), which allows for the direct embedding of expert judgements into the probabilistic framework. As the name suggests, GASC takes into account text genre as a relevant factor for the diachronic development of word senses in addition to distributional information. As such, GASC is able to differentiate between language change and genre prevalence. The data stems from two large diachronic corpora which are annotated for genre: LatinISE (McGillivray & Kilgarriff 2013) for Latin, with texts from the 2nd century BCE to the 21st century CE; and the Diorisis Annotated Ancient Greek Corpus (Vatri & McGillivray 2018) for Ancient Greek, with texts from the 8th century BCE to the 5th century CE. Their study focuses on polysemous words such as the Ancient Greek *mus*, which can mean 'mouse', 'muscle', or 'mussel'. Here, text genre plays an important role: the 'mouse' meaning is more likely to occur in narrative texts, while the 'muscle' meaning is dominant in technical texts. This genre-specific meaning difference is added as information to the model. By conditioning the model on the observed genre, the authors gain genre-specific distributions of words over their different senses. Finally, they show via a systematic comparison that Bayesian mixture models are competitive to state-of-the-art embedding based models in detecting binary lexical semantic change in Latin and Ancient Greek.[5]

In Chapter 10, "Computational Approaches to Lexical Semantic Change: Visualization Systems and Novel Applications", Adam Jatowt, Nina Tahmasebi and Lars Borin survey visualization tools and other interactive user interfaces designed for understanding lexical semantic change, as well as downstream computational applications of lexical semantic change methodologies. Without further ado, the chapter begins with presenting visualization systems that support a manual analysis of lexical semantic change. This first part is particularly interesting for readers with a historical linguistic background, since it provides a broad

5. It is noteworthy that SCAN (Frermann & Lapata 2016), the Bayesian model which GASC is based on and which does not include genre as factor, outperforms GASC with respect to the F1 score for Latin – a relevant point which is not discussed further.

overview of visualization systems that are ready to use for exploring datasets with respect to diachronic lexical semantic change. The remainder of the chapter surveys a number of other applications of approaches designed for the computational analysis of semantic change and lexical replacement, highlighting their usability for research outside of historical linguistics, e.g., culturomics (Michel et al. 2011) or Information Retrieval.

Chapter 11, "Challenges for Computational Lexical Semantic Change" by Simon Hengchen, Nina Tahmasebi, Dominik Schlechtweg and Haim Dubossarsky, concludes the book by outlining the several open challenges which remain for future computational research on lexical semantic change, as well as sketching potential research directions. For one, there are challenges with respect to the available data. For example, there is often not enough historical data available for many languages to allow for a robust use of neural embeddings and other machine learning technologies. In addition, it is unclear how well the methodologies that have been developed mainly for English are suited for other languages, and also, how well methods, which have been developed on the basis of modern languages, such as neural embeddings, can be transferred to historical languages. Furthermore, albeit the strong theoretical tradition of work on semantic change, terminology is an issue, with different taxonomies of change postulated. Moreover, what is modeled computationally in terms of meaning is still very blunt and primarily captures contextual similarity between lexical items, not necessarily reflecting different nuances of word meaning. Often, lexical meaning becomes conflated with cultural and/or topical information, and the resulting computational models do not necessarily reflect an actual change in lexical meaning. Additionally, the computational representations are often not interpretable, which make a qualitative assessment of the data extremely difficult, and qualitative feedback on the data by domain experts, i.e., historical linguists, is generally rare. Finally, there exists no devised, large-scale evaluation framework for lexical semantic change, with only little ground truth data available, which renders an overarching comparison of the effectiveness of the developed methodologies extremely difficult.

Despite these challenges, the computational methodologies developed for the investigation of lexical semantic change have great potential for historical linguistics, with the possibility of providing quantitative support for theoretical work. Overall, the present volume provides a great overview of computational approaches to semantic change and presents a great starting point for future collaborative research, which will be necessary to overcome many of the current challenges.

Funding

Open Access publication of this article was funded through a Transformative Agreement with University of Konstanz.

References

Chen, X., L. Xu, Z. Liu, M. Sun & H.-B. Luan. 2015. Joint Learning of Character and Word Embeddings. *Proceedings of IJCAI 2015*, 1236–1242. Cambridge, MA: AAAI Press.

Davies, M. 2008. *The Corpus of Contemporary American English (COCA): 560 Million Words, 1990–Present.*

Davies, M. 2012. Expanding Horizons in Historical Linguistics with the 400-Million Word Corpus of Historical American English. *Corpora* 7:2.121–157.

Devlin, J., M.-W. Chang, K. Lee & K. Toutanova. 2019. BERT: Pre-Training of Deep Bidirectional Transformers for Language Understanding. *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)* ed. by Jill Burstein, Christy Doran & Thamar Solorio, 4171–4186. Minneapolis, MN: Association for Computational Linguistics.

Frermann, L. & M. Lapata. 2016. A Bayesian Model of Diachronic Meaning Change. *Transactions of the ACL* 4:31–45.

Hamilton, W.L., J. Leskovec & D. Jurafsky. 2016. Diachronic Word Embeddings Reveal Statistical Laws of Semantic Change. *Proceedings of the 54th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)* ed. by Katrin Erk & Noah A. Smith, 1489–1501. Berlin: Association for Computational Linguistics.

Haslam, N. 2016. Concept Creep: Psychology's Expanding Concepts of Harm and Pathology. *Psychological Inquiry* 27:1.1–17.

Joulin, A., E. Grave, P. Bojanowski, M. Douze, H. Jégou & T. Mikolov, T. 2016. Fasttext.zip: Compressing Text Classification Models. *arXiv preprint arXiv:1612.03651*.

McGillivray, B. & A. Kilgarriff. 2013. Tools for Historical Corpus Research and a Corpus of Latin. *New Methods in Historical Corpus Linguistics* ed. by P. Bennett, M. Durrell, S. Scheible & R.J. Whitt, 247–257. Tübingen: Narr.

Michel, J.-B., Y.K. Shen, A.P. Aiden, A. Veres, M.K. Gray, The Google Books Team, J.P. Pickett, D. Hoiberg, D. Clancy, P. Norvig, J. Orwant, S. Pinker, M.A. Nowak & E.L. Aiden. 2011. Quantitative Analysis of Culture Using Millions of Digitized Books. *Science* 331:6014.176–182.

Mikolov, T., I. Sutskever, K. Chen, G. Corrado & J. Dean. 2013. Distributed Representations of Words and Phrases and their Compositionality. *Advances in Neural Information Processing Systems* 26:3111–3119.

Turney, P.D. & P. Pantel. 2010. From Frequency to Meaning: Vector Space Models of Semantics. *Journal of Artificial Intelligence Research* 37:1.141–188.

Vatri, A. & B. McGillivray. 2018. The Diorisis Ancient Greek Corpus. *Research Data Journal for the Humanities and Social Sciences* 3:1.55–65.

Wiedemann, G., S. Remus, A. Chawla & C. Biemann. 2019. Does BERT Make any Sense? Interpretable Word Sense Disambiguation with Contextualized Embeddings. In *Proceedings of KONVENS 2019*. Erlangen, Germany, October 9–11.

Yao, Z., Y. Sun, W. Ding, N. Rao & H. Xiong. 2018. Dynamic Word Embeddings for Evolving Semantic Discovery. In *Proceedings of the Eleventh ACM International Conference on Web Search and Data Mining*, 673–681. Marina Del Rey, CA, February 5–9.

Address for correspondence

Christin Beck
Department of Linguistics
University of Konstanz
Universitätsstraße 10
78464 Konstanz
Germany

christin.beck@uni-konstanz.de
https://orcid.org/0000-0002-9110-9727

Publication history

Date received: 14 November 2022
Date accepted: 24 February 2023
Published online: 30 October 2023