

- Liu, Yinhan, Myle Ott, Naman Goyal, Jingfei Du, Mandar Joshi, Danqi Chen, Omer Levy, Mike Lewis, Luke Zettlemoyer, and Veselin Stoyanov. 2019b. RoBERTa: A robustly optimized BERT pretraining approach. *CoRR*, abs/1907.11692.
- Lucas, Margery. 2000. Semantic priming without association: A meta-analytic review. *Psychonomic Bulletin & Review*, 7(4):618–630.
- Luong, Thang, Richard Socher, and Christopher Manning. 2013. Better word representations with recursive neural networks for morphology. In *Proceedings of CoNLL*, pages 104–113.
- Lyons, John. 1977. *Semantics*, volume 2. Cambridge University Press.
- Majid, Asifa, Melissa Bowerman, Miriam van Staden, and James S Boster. 2007. The semantic categories of cutting and breaking events: A crosslinguistic perspective. *Cognitive Linguistics*, 18(2):133–152.
- Malaviya, Chaitanya, Graham Neubig, and Patrick Littell. 2017. Learning language representations for typology prediction. In *Proceedings of EMNLP*, pages 2529–2535.
- Mantel, Nathan. 1967. The detection of disease clustering and a generalized regression approach. *Cancer Research*, 27(2 Part 1):209–220.
- McKeown, Kathleen R., Regina Barzilay, David Evans, Vasileios Hatzivassiloglou, Judith L. Klavans, Ani Nenkova, Carl Sable, Barry Schiffman, and Sergey Sigelman. 2002. Tracking and summarizing news on a daily basis with Columbia’s newsblaster. In *Proceedings of HLT*, page 280–285.
- Melamud, Oren, David McClosky, Siddharth Patwardhan, and Mohit Bansal. 2016. The role of context types and dimensionality in learning word embeddings. In *Proceedings of NAACL-HLT*, pages 1030–1040.
- Mikolov, Tomas, Edouard Grave, Piotr Bojanowski, Christian Puhrsch, and Armand Joulin. 2018. Advances in pre-training distributed word representations. In *Proceedings of LREC*.
- Mikolov, Tomas, Quoc V. Le, and Ilya Sutskever. 2013. Exploiting similarities among languages for machine translation. *arXiv preprint, CoRR*, abs/1309.4168.
- Mikolov, Tomas, Ilya Sutskever, Kai Chen, Gregory S. Corrado, and Jeffrey Dean. 2013. Distributed representations of words and phrases and their compositionality. In *Proceedings of NeurIPS*, pages 3111–3119.
- Miller, George A. 1995. WordNet: A lexical database for English. *Communications of the ACM*, pages 39–41.
- Mimno, David and Laure Thompson. 2017. The strange geometry of skip-gram with negative sampling. In *Proceedings of EMNLP*, pages 2873–2878.
- Mohiuddin, Tasnim and Shafiq Joty. 2019. Revisiting adversarial autoencoder for unsupervised word translation with cycle consistency and improved training. In *Proceedings of NAACL-HLT*, pages 3857–3867.
- Mrkšić, Nikola, Diarmuid Ó Séaghdha, Blaise Thomson, Milica Gašić, Lina Maria Rojas-Barahona, Pei-Hao Su, David Vandyke, Tsung-Hsien Wen, and Steve Young. 2016. Counter-fitting word vectors to linguistic constraints. In *Proceedings of NAACL-HLT*, pages 142–148.
- Mrkšić, Nikola, Ivan Vulić, Diarmuid Ó Séaghdha, Ira Leviant, Roi Reichart, Milica Gašić, Anna Korhonen, and Steve Young. 2017. Semantic specialisation of distributional word vector spaces using monolingual and cross-lingual constraints. *Transactions of the ACL*, 5:309–324.
- Mu, Jiaqi, Suma Bhat, and Pramod Viswanath. 2018. All-but-the-top: Simple and effective postprocessing for word representations. In *Proceedings of ICLR*.
- Mykowiecka, Agnieszka, Małgorzata Marciniak, and Piotr Rychlik. 2018. SimLex-999 for Polish. In *Proceedings of LREC*.
- Nelson, Douglas L., Cathy L. McEvoy, and Thomas A. Schreiber. 2004. The University of South Florida free association, rhyme, and word fragment norms. *Behavior Research Methods*, 36(3):402–407.
- Netisopakul, Ponrudee, Gerhard Wohlgemann, and Aleksei Pulich. 2019. Word similarity datasets for Thai: Construction and evaluation. *CoRR*, abs/1904.04307.
- Nivre, Joakim, Mitchell Abrams, Željko Agić, Lars Ahrenberg, Gabrielé Aleksandravičiūtė, Lene Antonsen, Katya Aplonova, Maria Jesus Aranzabe, Gashaw Arutie, Masayuki Asahara, et al. 2019. Universal Dependencies 2.4. LINDAT/CLARIN digital library at the Institute of Formal and Applied Linguistics (ÚFAL), Faculty of Mathematics and Physics, Charles University.
- Pearson, Karl. 1901. On lines and planes of closest fit to systems of points in space. *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science*, 2(11):559–572.
- Peters, Matthew, Mark Neumann, Mohit Iyyer, Matt Gardner, Christopher Clark,