

Peter LoBue and Alexander Yates. Types of common-sense knowledge needed for recognizing textual entailment. In *Proceedings of the 49th Annual Meeting of the Association for Computational Linguistics: Human Language Technologies: Short Papers - Volume 2*, HLT '11, pages 329–334, Stroudsburg, PA, USA, 2011. Association for Computational Linguistics. ISBN 978-1-932432-88-6. URL <http://dl.acm.org/citation.cfm?id=2002736.2002805>.

Annie Louis and Ani Nenkova. A corpus of general and specific sentences from news. In *LREC*, pages 1818–1821, 2012.

Nitin Madnani and Bonnie J Dorr. Generating phrasal and sentential paraphrases: A survey of data-driven methods. *Computational Linguistics*, 36(3):341–387, 2010.

Nitin Madnani, Joel Tetreault, and Martin Chodorow. Re-examining machine translation metrics for paraphrase identification. In *Proceedings of the 2012 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*, NAACL HLT '12, pages 182–190, Stroudsburg, PA, USA, 2012. Association for Computational Linguistics. ISBN 978-1-937284-20-6. URL <http://dl.acm.org/citation.cfm?id=2382029.2382055>.

H. B. Mann and D. R. Whitney. On a test of whether one of two random variables is stochastically larger than the other. *Ann. Math. Statist.*, 18(1):50–60, 03 1947. doi: 10.1214/aoms/1177730491. URL <https://doi.org/10.1214/aoms/1177730491>.

Michał Marcińczuk, Marcin Oleksy, and Jan Kocoń. Inforex - a collaborative system for text corpora annotation and analysis. In *Proceedings of RANLP-2017*, September 2017. URL https://doi.org/10.26615/978-954-452-049-6_063.

Mitchell P Marcus, Mary Ann Marcinkiewicz, and Beatrice Santorini. Building a large annotated corpus of english: The penn treebank. *Computational Linguistics*, 19(2):313–30, 1993.

Marco Marelli, Stefano Menini, Marco Baroni, Luisa Bentivogli, Raffaella Bernardi, Roberto Zamparelli, et al. A SICK cure for the evaluation of compositional distributional semantic models. In *LREC*, pages 216–223, 2014.

Maria Antònia Martí, Mariona Taulé, Venelin Kovatchev, and Maria Salamó. Discover: Distributional approach based on syntactic dependencies for discovering constructions. *Corpus Linguistics and Linguistic Theory*, 2019.

- Rada Mihalcea, Courtney Corley, and Carlo Strapparava. Corpus-based and knowledge-based measures of text semantic similarity. In *Proceedings of the 21st National Conference on Artificial Intelligence - Volume 1*, AAAI, 2006, page 775–780. AAAI Press, 2006. ISBN 9781577352815.
- Tomas Mikolov, Kai Chen, Greg Corrado, and Jeffrey Dean. Efficient estimation of word representations in vector space. *CoRR*, abs/1301.3781, 2013a.
- Tomas Mikolov, Ilya Sutskever, Kai Chen, Greg Corrado, and Jeffrey Dean. Distributed representations of words and phrases and their compositionality. In *Proceedings of the 26th International Conference on Neural Information Processing Systems - Volume 2*, NIPS’13, pages 3111–3119, USA, 2013b. Curran Associates Inc. URL <http://dl.acm.org/citation.cfm?id=2999792.2999959>.
- Tomas Mikolov, Wen-tau Yih, and Geoffrey Zweig. Linguistic regularities in continuous space word representations. In *HLT-NAACL*, pages 746–51, 2013c.
- George Miller. *WordNet: An electronic lexical database*. MIT press, 1998.
- George A. Miller. Wordnet: A lexical database for english. *Commun. ACM*, 38(11):39–41, November 1995. ISSN 0001-0782.
- Jeff Mitchell and Mirella Lapata. Composition in distributional models of semantics. *Cognitive Science*, 34(8):1388–1439, 2010.
- Hermann Moisl. *Cluster Analisys for Corpus Linguistics*. De Gruyter Mouton, 2015.
- K. Muischnek and H. Sajkan. Using collocation-finding methods to extract constructions and estimate their productivity. In *Workshop on Extracting and Using Constructions in Natural Language Processing*, NODALIDA 2009, 2009.
- Brian Murphy, Partha Pratim Talukdar, and Tom M Mitchell. Learning effective and interpretable semantic models using non-negative sparse embedding. In *COLING*, pages 1933–50, 2012.
- Aakanksha Naik, Abhilasha Ravichander, Norman Sadeh, Carolyn Rose, and Graham Neubig. Stress test evaluation for natural language inference. In *Proceedings of the 27th International Conference on Computational Linguistics*, pages 2340–2353, Santa Fe, New Mexico, USA, August 2018. Association for Computational Linguistics. URL <https://www.aclweb.org/anthology/C18-1198>.

- Vivi Nastase, Devon Fritz, and Anette Frank. Demodify: A dataset for analyzing contextual constraints on modifier deletion. In *Proceedings of LREC-2018*, 2018.
- Roberto Navigli and Simone Paolo Ponzetto. Babelnet: The automatic construction, evaluation and application of a wide-coverage multilingual semantic network. *Artif. Intell.*, 193:217–250, December 2012. ISSN 0004-3702.
- P.B. Nemenyi. *Distribution-free Multiple Comparisons*. PhD thesis, Princeton University, 1963.
- Yoshiki Niwa and Yoshihiko Nitta. Co-occurrence vectors from corpora vs. distance vectors from dictionaries. In *Proceedings of the 15th Conference on Computational Linguistics*, volume 1 of *COLING '94*, pages 304–309, Stroudsburg, PA, USA, 1994. Association for Computational Linguistics.
- Geoffrey Nunberg, Ivan A Sag, and Thomas Wasow. Idioms. *Language*, pages 491–538, 1994.
- Matthew Brook O’Donnell and Nick Ellis. Towards an inventory of english verb argument constructions. In *Proceedings of the NAACL HLT Workshop on Extracting and Using Constructions in Computational Linguistics*, EUCL ’10, pages 9–16, Stroudsburg, PA, USA, 2010. Association for Computational Linguistics.
- Sebastian Padó, Michel Galley, Dan Jurafsky, and Christopher D Manning. Textual entailment features for machine translation evaluation. In *Proceedings of the Fourth Workshop on Statistical Machine Translation*, pages 37–41. Association for Computational Linguistics, 2009.
- Lluís Padró and Evgeny Stenetsovsky. Freeling 3.0: Towards wider multilinguality. In Nicoletta Calzolari, Khalid Choukri, Thierry Declerck, Mehmet Ugur Dogan, Bente Maegaard, Joseph Mariani, Jan Odijk, and Stelios Piperidis, editors, *LREC*, pages 2473–9. European Language Resources Association (ELRA), 2012. ISBN 978-2-9517408-7-7.
- Ellie Pavlick, Pushpendre Rastogi, Juri Ganitkevitch, Benjamin Van Durme, and Chris Callison-Burch. PPDB 2.0: Better paraphrase ranking, fine-grained entailment relations, word embeddings, and style classification. In *Proceedings of the 53rd Annual Meeting of the Association for Computational Linguistics and the 7th International Joint Conference on Natural Language Processing (Volume 2: Short Papers)*, pages 425–430, Beijing, China, July 2015. Association for Computational Linguistics. doi: 10.3115/v1/P15-2070. URL <https://www.aclweb.org/anthology/P15-2070>.