Denis Paperno
denis.paperno@gmail.com

# Distributional Models for Lexical Semantics

## Data

Human and Machine Judgements about Russian Semantic Relatedness
http://russe.nlpub.org/downloads/

> Panchenko, A., Ustalov, D., Arefyev, N., Paperno, D., Konstantinova, N., Loukachevitch, N., & Biemann, C. (2016, April). Human and machine judgements for russian semantic relatedness. In *International Conference on Analysis of Images, Social Networks and Texts* (pp. 221-235). Springer

MEN word relatedness dataset: https://staff.fnwi.uva.nl/e.bruni/MEN

> E. Bruni, N. K. Tran and M. Baroni. Multimodal Distributional Semantics. Journal of Artificial Intelligence Research 49: 1-47.

WordSim353 (word similarity and relatedness): http://alfonseca.org/eng/research/wordsim353.html

> Eneko Agirre, Enrique Alfonseca, Keith Hall, Jana Kravalova, Marius Pasca, Aitor Soroa. A Study on Similarity and Relatedness Using Distributional and WordNet-based Approaches, In *Proceedings of NAACL-HLT 2009*.

## A short bibliography

0. Origins and foundations

Firth, John R. (1951/1957): Modes of meaning. In: Papers in Linguistics, 1934-1951. Oxford: Oxford University Press.

Harris Zellig S. 1951. Methods in Structural Linguistics. Chicago: University of Chicago Press.

Miller, George A. & Walter G. Charles, 1991. Contextual correlates of semantic similarity. Language and Cognitive Processes VI. 1-28.

1. Nature of distributional models and parameters of their creation.

Alessandro Lenci. Distributional semantics in linguistic and cognitive research. Rivista di Linguistica 20.1 (2008), pp. 1-31 http://linguistica.sns.it/RdL/20.1/ALenci.pdf

Turney, P. D., & Pantel, P. (2010). From frequency to meaning: Vector space models of semantics. *Journal of artificial intelligence research*, *37*, 141-188. http://www.jair.org/media/2934/live-2934-4846-jair.pdf

2. Simple (psycho)linguistic applications of distributional models (modeling semantic similarity and relatedness judgements, priming effects, etc.).

Landauer T. K., Dumais S. T. A solution to Plato's problem: The latent semantic analysis theory of acquisition, induction, and representation of knowledge //Psychological review. – 1997. – T. 104. –

№. 2. – С. 211.
http://www.indiana.edu/~clcl/Q550_WWW/Papers/Landauer_Dumais_1997.pdf
Levy, O., & Goldberg, Y. (2014). Linguistic regularities in sparse and explicit word representations. In *Proceedings of the eighteenth conference on computational natural language learning* (pp. 171-180).
https://levyomer.files.wordpress.com/2014/04/linguistic-regularities-in-sparse-and-explicit-word-representations-conll-2014.pdf
Griffiths, T. L., Steyvers, M., & Tenenbaum, J. B. (2007). Topics in semantic representation. *Psychological review*, *114*(2), 211. http://psiexp.ss.uci.edu/research/papers/Griffiths_Steyvers_Tenenbaum_2007.pdf

3. Neural Word Embedings

Mikolov T. et al. Efficient estimation of word representations in vector space //arXiv preprint arXiv:1301.3781. – 2013. https://arxiv.org/abs/1301.3781
Pennington J., Socher R., Manning C. D. Glove: Global vectors for word representation //EMNLP. – 2014. – Т. 14. – С. 1532-1543. http://www.aclweb.org/anthology/D14-1162
Baroni, M., Dinu, G., & Kruszewski, G. (2014, June). Don't count, predict! A systematic comparison of context-counting vs. context-predicting semantic vectors. In *ACL (1)* (pp. 238-247). http://anthology.aclweb.org/P/P14/P14-1023.pdf
Levy O., Goldberg Y. Neural word embedding as implicit matrix factorization //Advances in neural information processing systems. – 2014. – С. 2177-2185. http://papers.nips.cc/paper/5477-neural-word-embedding-as-implicit-matrix-factorization.pdf
Krebs A., Paperno D. When hyperparameters help: beneficial parameter combinations in distributional semantic models. Proceedings of the Fifth Joint Conference on Lexical and Computational Semantics. Association for Computational Linguistics – 2016. – С. 97-101. http://www.aclweb.org/anthology/S/S16/S16-2011v2.pdf
E. Vecchi, M. Marelli, R. Zamparelli and M. Baroni. 2017. Spicy adjectives and nominal donkeys: Capturing semantic deviance using compositionality in distributional spaces. Cognitive Science 41(1): 102-136.

4. Compositionality in distributional semantics ;applications to morphology ; applications to syntactic preferences.

M. Baroni, R. Bernardi and R. Zamparelli. 2014. Frege in space: A program for compositional distributional semantics. Linguistic Issues in Language Technologies 9(6): 5-110.Mitchell J., Lapata M. Vector-based Models of Semantic Composition //ACL. – 2008. – С. 236-244. http://anthology.aclweb.org/P/P08/P08-1028.pdf
Baroni M., Zamparelli R. Nouns are vectors, adjectives are matrices: Representing adjective-noun constructions in semantic space //Proceedings of the 2010 Conference on Empirical Methods in Natural Language Processing. – Association for Computational Linguistics, 2010. – С. 1183-1193.
http://clic.cimec.unitn.it/marco/publications/bz-adj-com-emnlp10.pdf
Paperno, D., & Baroni, M. (2016). When the Whole Is Less Than the Sum of Its Parts: How Composition Affects PMI Values in Distributional Semantic Vectors. *Computational Linguistics*, *42*(2). https://www.aclweb.org/anthology/J/J16/J16-2006.pdf
Marelli, M., & Baroni, M. (2015). Affixation in semantic space: Modeling morpheme meanings with compositional distributional semantics. *Psychological review*, *122*(3), 485. http://clic.cimec.unitn.it/composes/FRACSS/marelli-baroni-fracss-preprint.pdf
Lazaridou A. et al. Compositional-ly Derived Representations of Morphologically Complex Words in Distributional Semantics //ACL (1). – 2013. – С. 1517-1526.
http://www.aclweb.org/website/old_anthology/P/P13/P13-1149.pdf

5. Extracting semantic properties and relations from distributional representations.

Louwerse, M. M., & Zwaan, R. A. (2009). Language encodes geographical information. *Cognitive Science*, *33*(1), 51-73. https://pdfs.semanticscholar.org/950a/8a41a8dd8a8607a0ea59102b33bae96b58a7.pdf
Louwerse, M. M., & Benesh, N. (2012). Representing spatial structure through maps and language: Lord of the Rings encodes the spatial structure of Middle Earth. *Cognitive science*, *36*(8), 1556-1569. https://pdfs.semanticscholar.org/ac56/4fb7e53cc3627743c2178e59d6cbd05b28e9.pdf
Roller, S., Erk, K., & Boleda, G. (2014). Inclusive yet Selective: Supervised Distributional Hypernymy Detection. In *COLING* (pp. 1025-1036). http://www.cs.utexas.edu/~ml/papers/roller.coling14.pdf
Santus, E., Lenci, A., Lu, Q., & Im Walde, S. S. (2014, April). Chasing Hypernyms in Vector Spaces with Entropy. In *EACL* (pp. 38-42). http://www.aclweb.org/old_anthology/E/E14/E14-4008.pdf
Santus E. et al. ROOT13: spotting hypernyms, co-hyponyms and randoms //arXiv preprint arXiv:1603.08705. – 2016. https://arxiv.org/pdf/1603.08705.pdf
Kruszewski, G., Paperno, D., & Baroni, M. (2015). Deriving boolean structures from distributional vectors. *Transactions of the Association for Computational Linguistics*, *3*, 375-388. https://www.transacl.org/ojs/index.php/tacl/article/viewFile/616/138
Herbelot, A., & Vecchi, E. M. (2015). Building a shared world: mapping distributional to model-theoretic semantic spaces. In *EMNLP* (pp. 22-32). http://www.anthology.aclweb.org/D/D15/D15-1003.pdf

6. Distributional models and pragmatic alternatives.

Kruszewski, G., Paperno, D., Bernardi, R., & Baroni, M. (2016) There Is No Logical Negation Here, But There Are Alternatives: Modeling Conversational Negation with Distributional Semantics. *Computational Linguistics*, *42*(4). http://www.aclweb.org/anthology/J/J16/J16-4003.pdf
Van Tiel, B., Van Miltenburg, E., Zevakhina, N., & Geurts, B. (2016). Scalar diversity. *Journal of Semantics*, *33*(1), 137-175.

7. Applications to language typology and to historical linguistics.

William L. Hamilton, Jure Leskovec, Dan Jurafsky. Diachronic Word Embeddings Reveal Statistical Laws of Semantic Change. Proceedings of the 54th Annual Meeting of the Association for Computational Linguistics, pages 1489–1501,Berlin, Germany, August 7-12, 2016. http://www.aclweb.org/anthology/P/P16/P16-1141.pdf
(to appear) Haim Dubossarsky, Daphna Weinshall and Eitan Grossman (2017). Outta Control: Laws of Semantic Change and Inherent Biases in Word Representation Models. EMNLP 2017.
Daria Ryzhova, Maria Kyuseva, Denis Paperno. Typology of Adjectives Benchmark for Compositional Distributional Models. Proceedings of Language Resources and Evaluation Conference, 2016, pp. 1953-1957. http://www.lrec-conf.org/proceedings/lrec2016/pdf/665_Paper.pdf

8. The problem of grounding : distributional semantics and computer vision.

Lazaridou, A., Bruni, E., & Baroni, M. (2014). Is this a wampimuk? Cross-modal mapping between distributional semantics and the visual world. In *ACL (1)* (pp. 1403-1414). http://www.aclweb.org/old_anthology/P/P14/P14-1132.pdf
Kiela and Clark. 2015. Multi- and Cross-Modal Semantics Beyond Vision: Grounding in Auditory Perception. Proceedings of the 2015 Conference on Empirical Methods in Natural Language Processing, pages 2461–2470, Lisbon, Portugal, 17-21 September 2015.

https://www.cs.cmu.edu/~ark/EMNLP-2015/proceedings/EMNLP/pdf/EMNLP293.pdf