













- (almost) no bilingual data. In *Proceedings of the ACL2017 (Volume 1: Long Papers)*, pages 451–462.
- Artetxe, M. and H. Schwenk. 2019. Margin-based parallel corpus mining with multilingual sentence embeddings. In *Proceedings of the ACL2019 – Volume 1*, pages 3197–3203.
- Bahdanau, D., K. Cho, and Y. Bengio. 2015. Neural machine translation by jointly learning to align and translate. In *Proceedings of ICLR 2015*.
- Chan, W., N. Jaitly, Q. Le, and O. Vinyals. 2016. Listen, attend and spell: A neural network for large vocabulary conversational speech recognition. In *Proceedings of IEEE ICASSP 2016*, pages 4960–4964.
- Devlin, J., M. Chang, K. Lee, and K. Toutanova. 2019. BERT: Pre-training of deep bidirectional transformers for language understanding. In *Proceedings of the NAACL-HLT 2019*, pages 4171–4186.
- Dušek, O. and F. Jurčiček. 2016. Sequence-to-sequence generation for spoken dialogue via deep syntax trees and strings. In *Proceedings of the ACL2016 (Volume 2: Short Papers)*, pages 45–51.
- Freitag, M. and S. Roy. 2018. Unsupervised natural language generation with denoising autoencoders. In *Proceedings of the EMNLP 2018*, pages 3922–3929, October–November.
- Gagniuc, P. A. 2017. *Markov chains: from theory to implementation and experimentation*. John Wiley & Sons.
- Ghaddar, A. and P. Langlais. 2017. WiNER: A Wikipedia annotated corpus for named entity recognition. In *Proceedings of the IJCNLP 2017 (Volume 1: Long Papers)*, pages 413–422.
- Inaba, M. and K. Takahashi. 2016. Neural utterance ranking model for conversational dialogue systems. In *Proceedings of the SIGDIAL 2016*, pages 393–403.
- Inoue, H. 2018. Data augmentation by pairing samples for images classification. *arXiv preprint arXiv:1801.02929*.
- Junczys-Dowmunt, M. 2019. Microsoft translator at wmt 2019: Towards large-scale document-level neural machine translation. In *Proceedings of the WMT19 Translation (Volume 2: Shared Task Papers, Day 1)*, pages 225–233.
- Junczys-Dowmunt, M., R. Grundkiewicz, T. Dwojak, H. Hoang, K. Heafield, T. Neckermann, F. Seide, U. Germann, A. Fikri Aji, N. Bogoychev, A. F. T. Martins, and A. Birch. 2018. Marian: Fast neural machine translation in C++. In *Proceedings of ACL 2018, System Demonstrations*, pages 116–121.
- Jurafsky, D. and J. H. Martin. 2008. *Speech and Language Processing: An Introduction to Natural Language Processing, Computational Linguistics, and Speech Recognition*. Nueva Jersey: Prentice Hall.
- Lample, G., A. Conneau, M. Ranzato, L. Denoyer, and H. Jégou. 2018. Word translation without parallel data. In *Proceedings of the ICLR 2018*.
- Le, Q. and T. Mikolov. 2014. Distributed representations of sentences and documents. In *International conference on machine learning*, pages 1188–1196.
- Liu, T., K. Wang, L. Sha, B. Chang, and Z. Sui. 2018. Table-to-text generation by structure-aware seq2seq learning. In *32nd AAAI Conference on Artificial Intelligence*.
- Mikolov, T., K. Chen, G. S. Corrado, and J. Dean. 2013. Efficient estimation of word representations in vector space. *CoRR*, abs/1301.3781.
- Moritz, N., T. Hori, and J. L. Roux. 2019. Unidirectional Neural Network Architectures for End-to-End Automatic Speech Recognition. In *Proc. Interspeech 2019*, pages 76–80.
- Peters, M. E., M. Neumann, M. Iyyer, M. Gardner, C. Clark, K. Lee, and L. Zettlemoyer. 2018. Deep contextualized word representations. In *Proceedings NAACL-HLT 2018 – Volume 1*, pages 2227–2237.
- Pham, N.-Q., T.-S. Nguyen, J. Niehues, M. Müller, and A. Waibel. 2019. Very Deep Self-Attention Networks for End-to-End Speech Recognition. In *Proc. Interspeech 2019*, pages 66–70.

- Puzikov, Y. and I. Gurevych. 2018. E2E NLG challenge: Neural models vs. templates. In *Proceedings of the INLG 2018*, pages 463–471.
- Reimers, N. and I. Gurevych. 2019. Sentence-bert: Sentence embeddings using siamese bert-networks. In *Proceedings of the EMNLP-IJCNLP 2019*, pages 3982–3992.
- Ruiter, D., C. España-Bonet, and J. van Genabith. 2019. Self-Supervised Neural Machine Translation. In *Proceedings of the ACL 2019, Volume 2: Short Papers.*, pages 1828–1834.
- Sankar, C., S. Subramanian, C. Pal, S. Chandar, and Y. Bengio. 2019. Do neural dialog systems use the conversation history effectively? an empirical study. In *Proceedings of the ACL2019*, pages 32–37.
- Sennrich, R., B. Haddow, and A. Birch. 2016. Improving neural machine translation models with monolingual data. In *Proceedings of the ACL2016 (Volume 1: Long Papers)*, pages 86–96.
- Serban, I. V., A. Sordoni, Y. Bengio, A. Courville, and J. Pineau. 2016. Building end-to-end dialogue systems using generative hierarchical neural network models. In *30th AAAI Conference on Artificial Intelligence*.
- Sordoni, A., M. Galley, M. Auli, C. Brockett, Y. Ji, M. Mitchell, J.-Y. Nie, J. Gao, and B. Dolan. 2015. A neural network approach to context-sensitive generation of conversational responses. In *Proceedings of the NACCL-HLT 2015*, pages 196–205.
- Tanner, M. A. and W. H. Wong. 1987. The calculation of posterior distributions by data augmentation. *Journal of the American statistical Association*, 82(398):528–540.
- Taulé, M., M. A. Martí, and M. Recasens. 2008. AnCora: Multilevel annotated corpora for Catalan and Spanish. In *Proceedings of the LREC'08*.
- Tiedemann, J. 2012. Parallel data, tools and interfaces in OPUS. In *Proceedings of the LREC2012*, pages 2214–2218.
- Vaswani, A., N. Shazeer, N. Parmar, J. Uszkoreit, L. Jones, A. N. Gomez, L. Kaiser, and I. Polosukhin. 2017. Attention is all you need. In *Proceedings of the NIPS2017*, pages 6000–6010.
- Vinyals, O. and Q. V. Le. 2015. A neural conversational model. In *Proceedings of the 31st International Conference on Machine Learning*, volume 37.
- Wen, T.-H., M. Gašić, D. Kim, N. Mrkšić, P.-H. Su, D. Vandyke, and S. Young. 2015. Stochastic language generation in dialogue using recurrent neural networks with convolutional sentence reranking. In *Proceedings of the SIGDIAL2015*, pages 275–284.
- Yang, Z., Z. Dai, Y. Yang, J. G. Carbonell, R. Salakhutdinov, and Q. V. Le. 2019. Xlnet: Generalized autoregressive pretraining for language understanding. *ArXiv*, abs/1906.08237.