External References

The following are external references for my research application in a more readable format. The same references can be found as hyperlinks in the application form each time a work is referenced.

- Philip Arthur, Graham Neubig, and Satoshi Nakamura. 2016. Incorporating discrete translation lexicons into neural machine translation. In *Proceedings of the 2016 Conference on Empirical Methods in Natural Language Processing*, pages 1557–1567, Austin, Texas.

- Yonatan Belinkov, Nadir Durrani, Fahim Dalvi, Hassan Sajjad and James Glass. 2020. On the Linguistic Representational Power of Neural Machine Translation Models. *Journal of Computational Linguistics*. To appear.

- Luisa Bentivogli, Arianna Bisazza, Mauro Cettolo and Marcello Federico. 2016. Neural versus Phrase-Based Machine Translation Quality: a Case Study. In *Proceedings of the 2016 Conference on Empirical Methods in Natural Language Processing*, pages 257–267, Austin, Texas.

- David M. Blei, Andrew Y. Ng and Michael I. Jordan. 2003. Latent Dirichlet allocation. Journal of Machine Learning Research 3, pages 993-1022.

- Jane Bromley, J.W. Bentz, Leon Bottou, I. Guyon, Yann Lecun, C. Moore, Eduard Sackinger and R. Shah. 1993. Signature verification using a "SIAMESE" time delay neural network. *International Journal of Pattern Recognition and Artificial Intelligence*, 7(4), pages 669–688.

- Peter F. Brown, Stephen A. Della Pietra, Vincent J. Della Pietra and Robert L. Mercer. 1991. Word-Sense Disambiguation Using Statistical Methods, In *Proceedings of the 29th Annual Meeting of the Association for Computational Linguistics*, pages 264-270, Berkeley, California, USA

- Alexis Conneau, Kartikay Khandelwal, Naman Goyal, Vishrav Chaudhary, Guillaume Wenzek, Francisco Guzmán, Edouard Grave, Myle Ott, Luke Zettlemoyer and Veselin Stoyanov. 2019. Unsupervised Cross-lingual Representation Learning at Scale. arXiv:1911.02116

- Éric de La Clergerie, Benoît Sagot, Djamé Seddah. 2017. The ParisNLP entry at the ConLL UD Shared Task 2017: A Tale of a #ParsingTragedy. In *Proceedings of the CoNLL 2017 Shared Task: Multilingual Parsing from Raw Text to Universal Dependencies*, pages 243-253, Vancouver, Canada.

- Jacob Devlin, Ming-Wei Chang, Kenton Lee, Kristina Toutanova. 2019. BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding. In *Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*, pages 4171-4186, Minneapolis, Minnesota, USA.

- Georgiana Dinu, Prashant Mathur, Marcello Federico and Yaser Al-Onaizan. 2019. Training Neural Machine Translation To Apply Terminology Constraints. In *Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics*, pages 3063–3068, Florence, Italy.

- Christian Federmann and Will Lewis. 2016. Microsoft Speech Language Translation (MSLT) Corpus: The IWSLT 2016 release for English, French and German. Microsoft Research.

- Yang Feng, Shiyue Zhang, Andi Zhang, DongWang, and Andrew Abel. 2017. Memory-augmented neural machine translation. In *Proceedings of the 2017 Conference on Empirical Methods in Natural Language Processing*, pages 1390–1399, Copenhagen, Denmark.

- Markus Freitag and Yaser Al-Onaizan. 2016. Fast Domain Adaptation for Neural Machine Translation. arXiv:1612.06897

- Mercedes García-Martínez, Loïc Barrault and Fethi Bougares. 2016. Factored Neural Machine Translation Architectures. In *Proceedings of the International Workshop on Spoken Language Translation*, Seattle, USA.

- Jiatao Gu, Hany Hassan, Jacob Devlin and Victor O.K. Li. 2018. Universal Neural Machine Translation for Extremely Low Resource Languages. In *Proceedings of the 2018 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*, pages 344-354, New Orleans, Louisiana.

- Rogers P. Hall. 1989. Computational Approaches to Analogical Reasoning: A Comparative Analysis. *Artificial Intelligence 39*, pages 39-120.

- Xuanli He, Gholamreza Haffari and Mohammad Norouzi. 2018. Sequence to Sequence Mixture Model for Diverse Machine Translation. In *Proceedings of the 22nd Conference on Computational Natural Language Learning*, pages 583–592, Brussels, Belgium.

- Robert R. Hoffman. 1995. Monster Analogies. *AI Magazine 16*, pages 11-35.

- Pierre Isabelle, Colin Cherry and George Foster. 2017. A Challenge Set Approach to Evaluating Machine Translation. In *Proceedings of the 2017 Conference on Empirical Methods in Natural Language Processing*, pages 2486-2496, Copenhagen, Denmark.

- Ganesh Jawahar, Benoît Sagot, Djamé Seddah. 2019. What Does BERT Learn about the Structure of Language?. In *Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics*, pages 3651-3657, Florence, Italy.

- Josef Jon. 2018. Exploring Contextual Information in Neural Machine Translation, PhD Thesis, Brno University of Technology.

- Melvin Johnson, Mike Schuster, Quoc V. Le, Maxim Krikun, Yonghui Wu, Zhifeng Chen, Nikhil Thorat, Fernanda Viégas, Martin Wattenberg, Greg Corrado, Macduff Hughes and Jeffrey Dean. 2017. Google's Multilingual Neural Machine Translation System: Enabling Zero-Shot Translation. *Transactions of the Association for Computational Linguistics*, Volume 5.

- Karthikeyan K, Zihan Wang, Stephen Mayhew, Dan Roth. 2019. Cross-Lingual Ability of Multilingual BERT: An Empirical Study. In *Proceedings of The International Conference on Learning Representations*.

- Eliyahu Kiperwasser and Miguel Ballesteros. 2018. Scheduled Multi-Task Learning: From Syntax to Translation. *Transactions of the Association for Computational Linguistics*, Volume 6, pages 225–240.

- Sneha Kudugunta, Ankur Bapna, Isaac Caswell and Orhan Firat. 2019. Investigating Multilingual NMT Representations at Scale. In *Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing and the 9th International Joint Conference on Natural Language Processing*, pages 1565–1575, Hong Kong, China.

- Guillaume Lample and Alexis Conneau. 2019. Cross-lingual Language Model Pretraining. In *Advances in Neural Information Processing Systems*.

- Anne Lacheret-Dujour, Sylvain Kahane and Paola Pietrandrea. 2019. A prosodic and syntactic treebank for spoken French. *Studies in Corpus Linguistics*, 89. John Benjamins.

- Philippe Langlais and Alexandre Patry. 2007. Translating Unknown Words by Analogical Learning. In *Proceedings of the 2007 Joint Conference on Empirical Methods in Natural Language Processing and Computational Natural Language Learning*, pages 877-886, Prague, Czech Republic.

- Yves Lepage. 2005. Translation of Sentences by Analogy Principle. In *Proceedings of the Language and Technology Conference*, pages 21-23 Poznań, Poland.

- Ivan Lerner, Nicolas Paris, Xavier Tannier. 2020. Terminologies augmented recurrent neural network model for clinical named entity recognition, *Journal of Biomedical Informatics*, Volume 102.

- Pierre Lison, Jörg Tiedemann and Milen Kouylekov. 2018. OpenSubtitles2018: Statistical Rescoring of Sentence Alignments in Large, Noisy Parallel Corpora. In *Proceedings of the Eleventh International Conference on Language Resources and Evaluation*, pages 1742-1748, Miyazaki, Japan.

- Yinhan Liu, Jiatao Gu, Naman Goyal, Xian Li, Sergey Edunov, Marjan Ghazvininejad, Mike Lewis and Luke Zettlemoyer. 2020. Multilingual Denoising Pre-training for Neural Machine Translation. arXiv:2001.08210.

- Annie Louis and Bonnie Webber. 2014. Structured and unstructured cache models for SMT domain adaptation. In *Proceedings of the 14th Conference of the European Chapter of the Association for Computational Linguistics*, pages 155–163, Gothenburg, Sweden.

- Jonathan Mallinson, Rico Sennrich and Mirella Lapata. 2017. Paraphrasing Revisited with Neural Machine Translation In *Proceedings of the 15th Conference of the European Chapter of the Association for Computational Linguistics*, pages 881–893, Valencia, Spain.

- Louis Martin, Benjamin Muller, Pedro Javier Ortiz Suárez, Yoann Dupont, Laurent Romary, Éric Villemonte de la Clergerie, Djamé Seddah, Benoît Sagot. 2019. CamemBERT: a Tasty French Language Model. arXiv:1911.03894.

- Louis Martin, Benoît Sagot, Éric de la Clergerie and Antoine Bordes. 2020. Controllable Sentence Simplification. In *Proceedings of the 12th Language Resources and Evaluation Conference*, Marseille, France.

- Tsuyoshi Morimoto, Noriyoshi Uratani, Toshiyuki Takezawa, Osamu Furuse, Yasuhiro Sobashima, Hitoshi Iida, Atsushi Nakamura, Yoshinori Sagisaka, Norio Higuchi, and Yasuhiro Yamazaki. 1994. A Speech and Language Database for Speech Translation Research. In *Proceedings of the 3rd International Conference on Spoken Language Processing*, pages 1791–1794, Yokohama, Japan.

- Maria Nădejde, Siva Reddy, Rico Sennrich, Tomasz Dwojak, Marcin Junczys-Dowmunt, Philipp Koehn and Alexandra Birch. 2017. Predicting Target Language CCG Supertags Improves Neural Machine Translation. In *Proceedings of the Second Conference on Machine Translation*, pages 68-79, Copenhagen, Denmark.

- Makoto Nagao. 1984. A Framework of a Mechanical Translation between Japanese and English by Analogy Principle. In A. Elithorn and R. Barnerji, ed., *Artificial and Human Intelligence*, pages 173-180, North-Holland.

- Ramesh Nallapati, Bowen Zhou, Cicero dos Santos, Çağlar Gülçehre and, Bing Xiang. 2016. Abstractive Text Summarization using Sequence-to-sequence RNNs and Beyond. In *Proceedings of The 20th SIGNLL Conference on Computational Natural Language Learning*, pages 280–290, Berlin, Germany.

- Paul Neculoiu, Maarten Versteegh, Mihai Rotaru. 2016. Learning Text Similarity with Siamese Recurrent Networks. In *Proceedings of the 1st Workshop on Representation Learning for NLP*, pages 148-157, Berlin, Germany.

- Maximilian Nickel and Douwe Kiela. 2017. Poincaré Embeddings for Learning Hierarchical Representations. arXiv:1705.08039.

- Timothy Osborne and Sylvain Kahane. 2015. Elements of structural syntax. Amsterdam, John Benjamins.

- Kishore Papineni, Salim Roukos, Todd Ward and Wei-Jing Zhu. 2002. Bleu: a Method for Automatic Evaluation of Machine Translation. In *Proceedings of the 40th Annual Meeting of the Association for Computational Linguistics*, pages 311–318, Philadelphia, Pennsylvania, USA.

- Matthew Peters, Mark Neumann, Mohit Iyyer, Matt Gardner, Christopher Clark, Kenton Lee and Luke Zettlemoyer. 2018. Deep contextualized word representations. In *Proceedings of the 2018 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*, pages 2227-2237, New Orleans, Louisiana, USA.

- Telmo Pires, Eva Schlinger and Dan Garrette. 2019. How multilingual is Multilingual BERT? arXiv:1906.01502.

- Alexander M. Rush, Sumit Chopra and Jason Weston. 2015. A Neural Attention Model for Abstractive Sentence Summarization. In *Proceedings of the 2015 Conference on Empirical Methods in Natural Language Processing*, pages 379–389, Lisbon, Portugal.

- Danielle Saunders, Felix Stahlberg, Adrià de Gispert and Bill Byrne. 2019. Domain Adaptive Inference for Neural Machine Translation. In *Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics*, pages 222-228, Florence, Italy.

- Rico Sennrich, Barry Haddow and Alexandra Birch 2016a. Controlling Politeness in Neural Machine Translation via Side Constraints. In *Proceedings of the 2016 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*, pages 35–40, San Diego, California.

- Rico Sennrich, Barry Haddow, Alexandra Birch. 2016b. Improving Neural Machine Translation Models with Monolingual Data. In *Proceedings of the 54th Annual Meeting of the Association for Computational Linguistics*, pages 86-96, Berlin, Germany.

- Tianxiao Shen, Myle Ott, Michael Auli and Marc'Aurelio Ranzato. 2019. Mixture Models for Diverse Machine Translation: Tricks of the Trade. In *Proceedings of the 36th International Conference on Machine Learning*, pages 5719-5728, Long Beach, California, USA.

- Kaitao Song, Xu Tan, Tao Qin, Jianfeng Lu, Tie-Yan Liu. 2019. MASS: Masked Sequence to Sequence Pre-training for Language. Generation. In *Proceedings of the 36th International Conference on Machine Learning*.

- Toshiyuki Takezawa, Eiichiro Sumita, Fumiaki Sugaya, Hirofumi Yamamoto, and Seiichi Yamamoto. 2002. Toward a Broad-coverage Bilingual Corpus for Speech Translation of Travel Conversations in the Real World. In *Proceedings of the 3rd International Conference on Language Resources and Evaluation*, pages 147–152, Las Palmas, Canary Islands, Spain.

- Ashish Vaswani, Noam Shazeer, Niki Parmar, Jakob Uszkoreit, Llion Jones, Aidan N. Gomez, Łukasz Kaiser, and Illia Polosukhin. 2017. Attention is all you need. In *Advances in Neural Information Processing Systems 30*, pages 5998-6008. Curran Associates, Inc.

- Jesse Vig and Yonatan Belinkov. 2019. Analyzing the Structure of Attention in a Transformer Language Model. In *Proceedings of the 2019 ACL Workshop BlackboxNLP: Analyzing and Interpreting Neural Networks for NLP*, pages 63-76, Florence, Italy.

- Oriol Vinyals, Quoc Le. 2015. A Neural Conversational Model. In *Proceedings of the 32nd International Conference on Machine Learning*, Lille, France.

- Elena Voita, David Talbot, Fedor Moiseev, Rico Sennrich, Ivan Titov. 2019. Analyzing Multi-Head Self-Attention: Specialized Heads Do the Heavy Lifting, the Rest Can Be Pruned. In *Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics*, pages 56-62, Florence, Italy.

- David Yarowsky, Grace Ngai and Richard Wicentowski. 2001. Inducing Multilingual Text Analysis Tools via Robust Projection across Aligned Corpora. In *Proceedings of the First International Conference on Human Language Technology Research.*

- Xingxing Zhang and Mirella Lapata. 2017. Sentence Simplification with Deep Reinforcement Learning. In *Proceedings of the 2017 Conference on Empirical Methods in Natural Language Processing*, pages 584–594, Copenhagen, Denmark.