ALEXANDER M. FRASER - Publications List

Refereed Journal Publications

- Irina Bigoulaeva, Viktor Hangya, Iryna Gurevych, **Alexander Fraser** (2023). Label modification and bootstrapping for zero-shot cross-lingual hate speech detection. *Language Resources and Evaluation*, 57(4), pages 1515 to 1546

- Hassan Sajjad, Helmut Schmid, **Alexander Fraser**, Hinrich Schütze (2017). Statistical Models for Unsupervised, Semi-supervised and Supervised Transliteration Mining. *Computational Linguistics*, 43(2), pages 349 to 375

- Valentin Deyringer, **Alexander Fraser**, Helmut Schmid, Tsuyoshi Okita (2017). Parallelization of Neural Network Training for NLP with Hogwild! *The Prague Bulletin of Mathematical Linguistics*, No. 109, pages 29 to 38

- Anita Ramm, Riccardo Superbo, Dimitar Shterionov, Tony O'Dowd, **Alexander Fraser** (2017). Integration of a Multilingual Preordering Component into a Commercial SMT Platform. *The Prague Bulletin of Mathematical Linguistics*, No. 108, pages 61 to 72

- Nadir Durrani, Helmut Schmid, **Alexander Fraser**, Philipp Koehn, Hinrich Schütze (2015). The Operation Sequence Model – Combining N-Gram-based and Phrase-based Statistical Machine Translation. *Computational Linguistics*, 41(2), pages 157 to 186

- Aleš Tamchyna, Fabienne Braune, **Alexander Fraser**, Marine Carpuat, Hal Daumé III, Chris Quirk (2014). Integrating a Discriminative Classifier into Phrase-based and Hierarchical Decoding. *The Prague Bulletin of Mathematical Linguistics*, No. 101, pages 29 to 41

- **Alexander Fraser**, Helmut Schmid, Richárd Farkas, Renjing Wang, Hinrich Schütze (2013). Knowledge Sources for Constituent Parsing of German, a Morphologically Rich and Less-Configurational Language. *Computational Linguistics*, 39(1), pages 57 to 85

- **Alexander Fraser**, Daniel Marcu (2007). Measuring Word Alignment Quality for Statistical Machine Translation. *Computational Linguistics*, 33(3), pages 293 to 303

Book Chapters

- **Alexander Fraser**, William Wong (2008). The Language Weaver Statistical Machine Translation Software System. In A. Farghaly, editor, *Arabic Computational Linguistics*, CSLI Publications, Chapter 9, pages 257 to 288

Refereed Conference Publications

- Faeze Ghorbanpour, Viktor Hangya, **Alexander Fraser** (2025). Fine-Grained Transfer Learning for Harmful Content Detection through Label-Specific Soft Prompt Tuning. Accepted to *Proceedings of NAACL, to appear.*

- Katharina Hämmerl, Tomasz Limisiewicz, Jindřich Libovický, **Alexander Fraser** (2025). Beyond Literal Token Overlap: Token Alignability for Multilinguality. Accepted to *Proceedings of NAACL, to appear.*

- Yaqi Zhang, Viktor Hangya, **Alexander Fraser** (2025). LLM Sensitivity Challenges in Abusive Language Detection: Instruction-Tuned vs. Human Feedback. *Proceedings of the 31st International Conference on Computational Linguistics (COLING)*, pages 2765-2780, Abu Dhabi, UAE, January.

- Marion Di Marco, **Alexander Fraser** (2024). Subword Segmentation in LLMs: Looking at Inflection and Consistency. In *Proceedings of the 2024 Conference on Empirical Methods in Natural Language Processing (EMNLP)*, pages 12050-12060, Miami, FL, USA, November.

- Lukas Edman, Helmut Schmid, **Alexander Fraser** (2024). CUTE: Measuring LLMs' Understanding of Their Tokens. In *Proceedings of the 2024 Conference on Empirical Methods in Natural Language Processing (EMNLP)*, pages 3017-3026, Miami, FL, USA, November.

- Wen Lai, Viktor Hangya, **Alexander Fraser** (2024). Style-Specific Neurons for Steering LLMs in Text Style Transfer. In *Proceedings of the 2024 Conference on Empirical Methods in Natural Language Processing (EMNLP)*, pages 13427-13443, Miami, FL, USA, November.

- Hate Personified: Investigating the role of LLMs in content moderation (2024). Sarah Masud, Sahajpreet Singh, Viktor Hangya, **Alexander Fraser**, Tanmoy Chakraborty. In *Proceedings of the 2024 Conference on Empirical Methods in Natural Language Processing (EMNLP)*, pages 15847-15863, Miami, FL, USA, November.

- Wen Lai, Mohsen Mesgar, **Alexander Fraser** (2024). LLMs Beyond English: Scaling the Multilingual Capability of LLMs with Cross-Lingual Feedback. In *Findings of the Association for Computational Linguistics: ACL 2024*, pages 8186-8213, Bangkok, Thailand, August.

- Katharina Hämmerl, Jindřich Libovický, **Alexander Fraser** (2024). Understanding Cross-Lingual Alignment – A Survey. In *Findings of the Association for Computational Linguistics: ACL 2024*, pages 10922-10943, Bangkok, Thailand, August.

- Prisca Piccirilli, **Alexander Fraser**, Sabine Schulte im Walde (2024). VOLIMET: A Parallel Corpus of Literal and Metaphorical Verb-Object Pairs for English–German and English–French. *Proceedings of *SEM 2024*, pages 222-237, Mexico City, Mexico, June.

- Marion Weller-Di Marco, **Alexander Fraser** (2024). Analyzing the Understanding of Morphologically Complex Words in Large Language Models. *Proceedings of the 2024 Joint International Conference on Computational Linguistics, Language Resources and Evaluation (LREC-COLING 2024)*, pages 1009-1020, Torino, Italy, May.

- Viktor Hangya, **Alexander Fraser** (2024). How to Solve Few-Shot Abusive Content Detection Using the Data We Actually Have. *Proceedings of the 2024 Joint International Conference on Computational Linguistics, Language Resources and Evaluation (LREC-COLING 2024)*, pages 8307-8322, Torino, Italy, May.

- Fabian Simonjetz, Jussi Laasonen, Yunus Cobanoglu, **Alexander Fraser**, Enrique Jiménez (2024). Reconstruction of Cuneiform Literary Texts as Text Matching. *Proceedings of the 2024 Joint International Conference on Computational Linguistics, Language Resources and Evaluation (LREC-COLING 2024)*, pages 13712-13721, Torino, Italy, May.

- Marion Di Marco, Katharina Hämmerl, **Alexander Fraser** (2023). A Study on Accessing Linguistic Information in Pre-Trained Language Models by Using Prompts. In *Proceedings of the 2023 Conference on Empirical Methods in Natural Language Processing (EMNLP)*, pages 7328–7336, Singapore, December.

- Wen Lai, Alexandra Chronopoulou, **Alexander Fraser** (2023). Mitigating Data Imbalance and Representation Degeneration in Multilingual Machine Translation. In *Findings of the Association for Computational Linguistics: EMNLP 2023*, pages 14279-14294, Singapore, December.

- Katharina Hämmerl, Alina Fastowski, Jindřich Libovický, **Alexander Fraser** (2023). Exploring anisotropy and outliers in multilingual language models for cross-lingual semantic sentence similarity. In *Findings of the Association for Computational Linguistics: ACL 2023*, pages 7023-7037, Toronto, Canada, July.

- Katharina Hämmerl, Bjoern Deiseroth, Patrick Schramowski, Jindřich Libovický, Constantin Rothkopf, **Alexander Fraser**, Kristian Kersting (2023). Speaking Multiple Languages Affects the Moral Bias of Language Models. In *Findings of the Association for Computational Linguistics: ACL 2023*, pages 2137-2156, Toronto, Canada, July.

- Yihong Liu, Alexandra Chronopoulou, Hinrich Schütze, **Alexander Fraser** (2023). On the Copying Problem of Unsupervised NMT: A Training Schedule with a Language Discriminator Loss. In *The International Conference on Spoken Language Translation (IWSLT)*, pages 491-502, Toronto, Canada, July.

- Sophie Henning, William Beluch, **Alexander Fraser**, Annemarie Friedrich (2023). A Survey of Methods for Addressing Class Imbalance in Deep-Learning Based Natural Language Processing. In *Proceedings of the 17th Conference of the European Chapter of the Association for Computational Linguistics (EACL)*, pages 523–540, Dubrovnik, Croatia, May.

- Alexandra Chronopoulou, Matthew Peters, **Alexander Fraser**, Jesse Dodge (2023). AdapterSoup: Weight Averaging to Improve Generalization of Pretrained Language Models. In *Findings of the Association for Computational Linguistics: EACL 2023*, pages 2054–2063, Dubrovnik, Croatia, May.

- Viktor Hangya, Hossain Shaikh Saadi, **Alexander Fraser** (2022). Improving Low-Resource Languages in Pre-Trained Language Models. In *Proceedings of EMNLP*, pages 11993-12006, Abu Dhabi, UAE, December.

- Wen Lai, Alexandra Chronopoulou, **Alexander Fraser** (2022). m^4 Adapter: Multilingual Multi-Domain Adaptation for Machine Translation with a Meta-Adapter. In *Findings of EMNLP*, pages 4282-4296, Abu Dhabi, UAE, December.

- Wen Lai, Jindřich Libovický, **Alexander Fraser** (2022). Improving Both Domain Robustness and Domain Adaptability in Machine Translation. In *Proceedings of Coling*, pages 5191-5204, Gyeongju, Republic of Korea, October.

- Marius Gassen, Benjamin Hättasch, Benjamin Hilprecht, Nadja Geisler, **Alexander Fraser**, Carsten Binnig (2022). Demonstrating CAT: Synthesizing Data-Aware Conversational Agents for Transactional Databases. In *Proceedings of the 48th International Conference on Very Large Databases (VLDB) 2022 Demos*, pages 3586-3589, Sydney, Australia, September.

- Katharina Hämmerl, Jindřich Libovický, **Alexander Fraser** (2022). Combining static and contextualised multilingual embeddings. In *Findings of ACL*, pages 2316-2329, Dublin, Ireland, May.

- Jindřich Libovický, **Alexander Fraser** (2022). Why don't people use character-level machine translation? In *Findings of ACL*, pages 2470-2485, Dublin, Ireland, May.

- Viktor Hangya, Qianchu Liu, Dario Stojanovski, **Alexander Fraser**, Anna Korhonen (2021). Improving machine translation of rare and unseen word senses. In *Proceedings of the Sixth Conference on Machine Translation (WMT)*, pages 614-624, Punta Cana, Dominican Republic, November.

- Denis Peskov, Viktor Hangya, Jordan Boyd-Graber, **Alexander Fraser** (2021). Adapting Entities across Languages and Cultures. In *Findings of EMNLP 2021*, pages 3725-3750, November.

- Simon Riess, Matthias Huck, **Alexander Fraser** (2021). A Comparison of Sentence-Weighting Techniques for NMT. In *Proceedings of Machine Translation Summit (MT Summit)*, pages 176-187, August.

- Tobias Eder, Viktor Hangya, **Alexander Fraser** (2021). Anchor-based Bilingual Word Embeddings for Low-Resource Languages. In *Proceedings of the 2021 Annual Conference of the Association for Computational Linguistics (ACL)*, pages 227-232, August.

- Alexandra Chronopoulou, Dario Stojanovski, **Alexander Fraser** (2021). Improving the Lexical Ability of Pretrained Language Models for Unsupervised Neural Machine Translation. In *Proceedings of the 2021 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies (NAACL)*, pages 173-180, June.

- Silvia Severini, Viktor Hangya, **Alexander Fraser**, Hinrich Schütze (2020). Combining Word Embeddings with Bilingual Orthography Embeddings for Bilingual Dictionary Induction. In *Proceedings of the 28th International Conference on Computational Linguistics (COLING)*, pages 6044-6055, December.

- Dario Stojanovski, Benno Krojer, Denis Peskov, **Alexander Fraser** (2020). ContraCAT: Contrastive Coreference Analytical Templates for Machine Translation. In *Proceedings of the 28th International Conference on Computational Linguistics (COLING)*, pages 4732-4749, December.

- Alexandra Chronopoulou, Dario Stojanovski, **Alexander Fraser** (2020). Reusing a Pretrained Language Model on Languages with Limited Corpora for Unsupervised NMT. In *Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing (EMNLP)*, pages 2703-2711, November

- Jindřich Libovický, **Alexander Fraser** (2020). Towards Reasonably-Sized Character-Level Transformer NMT by Finetuning Subword Systems. In *Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing (EMNLP)*, pages 2572-2579, November

- Jindřich Libovický, Rudolf Rosa, **Alexander Fraser** (2020). On the Language Neutrality of Pre-trained Multilingual Representations. In *Findings of EMNLP 2020*, pages 1663-1674, November

- Viktor Hangya, **Alexander Fraser** (2020). Towards Handling Compositionality in Low-Resource Bilingual Word Induction. In *Proceedings of the 14th Conference of the Association for Machine Translation in the Americas (AMTA)*, pages 89-101, October

- Marion Weller-Di Marco, **Alexander Fraser** (2020). Modeling Word Formation in English-German Neural Machine Translation. In *Proceedings of the 2020 Annual Conference of the Association for Computational Linguistics (ACL)*, pages 4227-4232, July

- Leah Michel, Viktor Hangya, **Alexander Fraser** (2020). Exploring Bilingual Word Embeddings for Hiligaynon, a Low-Resource Language. In *Proceedings of the 12th Edition of the Language Resources and Evaluation Conference (LREC)*, pages 2573-2580, May

- Veronika Hintzen, **Alexander Fraser** (2019). To Act Or Not To Act - Annotating and Classifying Email Regarding Necessary Action. In *Proceedings of the Conference on Natural Language Processing (KONVENS)*, pages 50-58, Erlangen, Germany, October

- Dario Stojanovski, **Alexander Fraser** (2019). Improving Anaphora Resolution in Neural Machine Translation Using Curriculum Learning. In *Proceedings of the Machine Translation Summit 2019*, pages 140-150, Dublin, Ireland, August

- Viktor Hangya, **Alexander Fraser** (2019). Unsupervised Parallel Sentence Extraction with Parallel Segment Detection Helps Machine Translation. In *Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics (ACL)*, pages 1224-1234, Florence, Italy, July

- Matthias Huck, Viktor Hangya, **Alexander Fraser** (2019). Better OOV Translation with Bilingual Terminology Mining. In *Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics (ACL)*, pages 5809-5815, Florence, Italy, July

- Dario Stojanovski, **Alexander Fraser** (2018). Coreference and Coherence in Neural Machine Translation: A Study Using Oracle Experiments. In *Proceedings of the Third Conference on Machine Translation: Research Papers*, pages 49-60, Brussels, Belgium, October

- Viktor Hangya, Fabienne Braune, **Alexander Fraser**, Hinrich Schütze (2018). Two Methods for Domain Adaptation of Bilingual Tasks: Delightfully Simple and Broadly Applicable. In *Proceedings of the 54th Annual Conference of the Association for Computational Linguistics (ACL)*, pages 810-820, Melbourne, Australia, July

- Philipp Dufter, Mengjie Zhao, Martin Schmitt, **Alexander Fraser**, Hinrich Schütze (2018). Embedding Learning Through Multilingual Concept Induction. In *Proceedings of the 54th Annual Conference of the Association for Computational Linguistics (ACL)*, pages 1520-1530, Melbourne, Australia, July

- Fabienne Braune, Viktor Hangya, Tobias Eder, **Alexander Fraser** (2018). Evaluating Bilingual Word Embeddings on the Long Tail. In *Proceedings of the 16th Annual Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies (NAACL)*, pages 188-193, New Orleans, Louisiana, USA, June

- Costanza Conforti, Matthias Huck, **Alexander Fraser** (2018). Neural Morphological Tagging of Lemma Sequences for Machine Translation. In *Proceedings of the 13th Conference of the Association for Machine Translation in the Americas (AMTA), vol. 1: MT Research Track*, pages 39-53, Boston, Massachusetts, USA, March

- Matthias Huck, Simon Riess, **Alexander Fraser** (2017). Target-side Word Segmentation Strategies for Neural Machine Translation. In *Proceedings of the Second ACL Conference on Machine Translation (WMT)*, pages 56-67, Copenhagen, Denmark, September

- Aleš Tamchyna, Marion Weller-Di Marco, **Alexander Fraser** (2017). Modeling Target-Side Inflection in Neural Machine Translation. In *Proceedings of the Second ACL Conference on Machine Translation (WMT)*, pages 32-42, Copenhagen, Denmark, September

- Leonie Weißweiler, **Alexander Fraser** (2017). Developing a Stemmer for German Based on a Comparative Analysis of Publicly Available Stemmers. In *Proceedings of the German Society for Computational Linguistics and Language Technology (GSCL)*, Berlin, Germany, September

- Costanza Conforti, **Alexander Fraser** (2017). Supervised Word Sense Disambiguation for Venetan: a Proof-of-Concept Experiment. In *Proceedings of the 30th International Florida Artificial Intelligence Research Society Conference (FLAIRS), Special Track on Natural Language Processing of Ancient and Other Low-Resource Languages*, May, Florida, USA

- Matthias Huck, Aleš Tamchyna, Ondřej Bojar, **Alexander Fraser** (2017). Producing Unseen Morphological Variants in Statistical Machine Translation. In *Proceedings of the 15th Conference of the European Chapter of the Association for Computational Linguistics (EACL)*, pages 369 to 375, Valencia, Spain, April

- Marion Weller-Di Marco, **Alexander Fraser**, Sabine Schulte im Walde (2017). Addressing Problems across Linguistic Levels in SMT: Combining Approaches to Model Morphology, Syntax and Lexical Choice. In *Proceedings of the 15th Conference of the European Chapter of the Association for Computational Linguistics (EACL)*, pages 625-630, Valencia, Spain, April

- Aleš Tamchyna, **Alexander Fraser**, Ondřej Bojar, Marcin Junczys-Dowmunt (2016). Target-side Context for Discriminative Models in Statistical Machine Translation. In *Proceedings of the 54th Annual Conference of the Association for Computational Linguistics (ACL)*, pages 1704-1714, Berlin, Germany, August

- Anita Ramm, **Alexander Fraser** (2016). Modeling Verbal Inflection for English to German SMT. In *Proceedings of the First Conference on Machine Translation (WMT)*, pages 21-31, Berlin, Germany, August

- Marion Weller-Di Marco, **Alexander Fraser**, Sabine Schulte im Walde (2016). Modeling Complement Types in Phrase-based SMT. In *Proceedings of the First Conference on Machine Translation (WMT)*, pages 43-53, Berlin, Germany, August

- Fabienne Braune, **Alexander Fraser**, Hal Daumé III, Aleš Tamchyna (2016). A Framework for Discriminative Rule Selection in Hierarchical Moses. In *Proceedings of the First Conference on Machine Translation (WMT)*, pages 92-101, Berlin, Germany, August

- Fabienne Braune, Nina Seemann, **Alexander Fraser** (2015). Rule Selection with Soft Syntactic Features for String-to-Tree Statistical Machine Translation. In *Proceedings of the Conference on Empirical Methods in Natural Language Processing (EMNLP)*, pages 1095-1101, Lisbon, Portugal, September

- Thomas Müller, Ryan Cotterell, **Alexander Fraser**, Hinrich Schütze (2015). Joint Lemmatization and Morphological Tagging with Lemming **(honorable mention for best short paper)**. In *Proceedings of the Conference on Empirical Methods in Natural Language Processing (EMNLP)*, pages 2268-2274, Lisbon, Portugal, September

- Martin Wunderlich, **Alexander Fraser**, P.S. Langeslag (2015). God Wat Þæt Ic Eom God - An Exploratory Investigation Into Word Sense Disambiguation in Old English. In *Proceedings of the German Society for Computational Linguistics & Language Technology (GSCL)*, pages 39-48, Essen, Germany, September

- Ryan Cotterell, Thomas Müller, **Alexander Fraser**, Hinrich Schütze (2015). Labeled Morphological Segmentation with Semi-Markov Models. In *Proceedings of the SIGNLL Conference on Computational Natural Language Learning (CoNLL)*, pages 164-174, Beijing, China, July

- Marion Weller, **Alexander Fraser**, Sabine Schulte im Walde (2015). Target-side Generation of Prepositions for SMT. In *Proceedings the 18th Annual Conference of the European Association for Machine Translation (EAMT)*, pages 177 to 184, Antalya, Turkey, May

- Marion Weller, Sabine Schulte im Walde, **Alexander Fraser** (2014). Using Noun Class Information to Model Selectional Preferences for Translating Prepositions in SMT. In *Proceedings of the Eleventh Biennial Conference of the Association for Machine Translation in the Americas (AMTA)*, pages 275-287, Vancouver, BC, Canada, October

- Nadir Durrani, Philipp Koehn, Helmut Schmid, **Alexander Fraser** (2014). Investigating the Usefulness of Generalized Word Representations in SMT. In *Proceedings of the 25th Annual Conference on Computational Linguistics (COLING)*, pages 421-432, Dublin, Ireland, August

- Marion Weller, **Alexander Fraser**, Ulrich Heid (2014). Combining Bilingual Terminology Mining and Morphological Modeling for Domain Adaptation in SMT. In *Proceedings of the Seventeenth Annual Conference of the European Association for Machine Translation (EAMT)*, pages 11-18, Dubrovnik, Croatia, June

- Fabienne Cap, **Alexander Fraser**, Marion Weller, Aoife Cahill (2014). How to Produce Unseen Teddy Bears: Improved Morphological Processing of Compounds in SMT. In *Proceedings of the 14th Conference of the European Chapter of the Association for Computational Linguistics (EACL)*, pages 579-587, Goteborg, Sweden, April

- **Alexander Fraser**, Marion Weller, Aoife Cahill, Fabienne Cap (2013). NECTAR: Modeling Inflection and Word-Formation in SMT. In *Proceedings of the International Conference of the German Society for Computational Linguistics and Language Technology (GSCL), Nectar Track*, 6 pages, Darmstadt, Germany, September

- Nadir Durrani, **Alexander Fraser**, Helmut Schmid, Hieu Hoang, Philipp Koehn (2013). Can Markov Models Over Minimal Translation Units Help Phrase-Based SMT? In *Proceedings of the 51st Annual Meeting of the Association for Computational Linguistics (ACL)*, pages 399-405, Sofia, Bulgaria, August

- Marion Weller, **Alexander Fraser**, Sabine Schulte im Walde (2013). Using Subcategorization Knowledge To Improve Case Prediction For Translation To German. In *Proceedings of the 51st Annual Meeting of the Association for Computational Linguistics (ACL)*, pages 593-603, Sofia, Bulgaria, August

- Nadir Durrani, **Alexander Fraser**, Helmut Schmid (2013). Model With Minimal Translation Units, But Decode With Phrases. In *Proceedings of the Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies (NAACL-HLT)*, pages 1-11, Atlanta, Georgia, USA, June

- Hassan Sajjad, **Alexander Fraser**, Helmut Schmid (2012). A Statistical Model for Unsupervised and Semi-supervised Transliteration Mining. In *Proceedings of the 50th Annual Meeting of the Association for Computational Linguistics (ACL)*, pages 469-477, Jeju Island, Korea, July

- Fabienne Braune, Anita Gojun, **Alexander Fraser** (2012). Long-distance Reordering During Search for Hierarchical Phrase-based SMT. In *Proceedings of the 16th Annual Conference of the European Association for Machine Translation (EAMT)*, pages 177-184, Trento, Italy, May

- **Alexander Fraser**, Marion Weller, Aoife Cahill, Fabienne Cap (2012). Modeling Inflection and Word-Formation in SMT. In *Proceedings of the 13th Conference of the European Chapter of the Association for Computational Linguistics (EACL)*, pages 664-674, Avignon, France, April

- Anita Gojun, **Alexander Fraser** (2012). Determining the Placement of German Verbs in English-to-German SMT. In *Proceedings of the 13th Conference of the European Chapter of the Association for Computational Linguistics (EACL)*, pages 726-735, Avignon, France, April

- Hassan Sajjad, Nadir Durrani, Helmut Schmid, **Alexander Fraser** (2011). Comparing Two Techniques for Learning Transliteration Models Using a Parallel Corpus. In *Proceedings of The 5th International Joint Conference on Natural Language Processing (IJCNLP)*, pages 129-137, Chiang Mai, Thailand, November

- Nadir Durrani, Helmut Schmid, **Alexander Fraser** (2011). A Joint Sequence Translation Model with Integrated Reordering. In *Proceedings of the 49th Annual Meeting of the Association for Computational Linguistics (ACL)*, pages 1045-1054, Portland, Oregon, USA, June

- Hassan Sajjad, **Alexander Fraser**, Helmut Schmid (2011). An Algorithm for Unsupervised Transliteration Mining with an Application to Word Alignment. In *Proceedings of the 49th Annual Meeting of the Association for Computational Linguistics (ACL)*, pages 430-439, Portland, Oregon, USA, June

- Fabienne Braune, **Alexander Fraser** (2010). Improved Unsupervised Sentence Alignment for Parallel and Asymmetrical Bilingual Corpora. In *Proceedings of the 23rd International Conference on Computational Linguistics (COLING) - Posters*, pages 81-89, Beijing, China, August

- Nadir Durrani, Hassan Sajjad, **Alexander Fraser**, Helmut Schmid (2010). Hindi-to-Urdu Machine Translation Through Transliteration. In *Proceedings of the 48th Annual Meeting of the Association for Computational Linguistics (ACL)*, pages 465-474, Uppsala, Sweden, July

- Florian Schwarck, **Alexander Fraser**, Hinrich Schütze (2010). Bitext-Based Resolution of German Subject-Object Ambiguities. In *Proceedings of the 11th Annual Conference of the North American Chapter of the Association for Computational Linguistics (NAACL-HLT) - Short Paper, Oral Presentation*, pages 737-740, Los Angeles, California, USA, June

- **Alexander Fraser**, Renjing Wang, Hinrich Schütze (2009). Rich Bitext Projection Features for Parse Reranking. In *Proceedings of the 12th Conference of the European Chapter of the Association for Computational Linguistics (EACL)*, pages 282-290, Athens, Greece, March

- Hamidreza Kobdani, **Alexander Fraser**, Hinrich Schütze (2009). Word Alignment by Thresholded Two-Dimensional Normalization. In *Proceedings of MT Summit XII - Posters*, pages 260-267, Ottawa, Canada, August

- **Alexander Fraser**, Daniel Marcu (2007). Getting the Structure Right for Word Alignment: LEAF. In *Proceedings of the 2007 Joint Conference on Empirical Methods in Natural Language Processing and Computational Natural Language Learning (EMNLP-CoNLL)*, pages 51-60, Prague, Czech Republic, June

- **Alexander Fraser**, Daniel Marcu (2006). Semi-Supervised Training for Statistical Word Alignment. In *Proceedings of the 21st International Conference on Computational Linguistics and 44th Annual Meeting of the Association for Computational Linguistics (COLING/ACL)*, pages 769-776, Sydney, Australia, July

- Dragos Stefan Munteanu, **Alexander Fraser**, Daniel Marcu (2004). Improved Machine Translation Performance via Parallel Sentence Extraction from Comparable Corpora. In *Proceedings of the Human Language Technology and North American Association for Computational Linguistics Conference (HLT/NAACL)*, pages 265-272, Boston, Massachusetts, USA, May

- Franz Josef Och, Daniel Gildea, Sanjeev Khudanpur, Anoop Sarkar, Kenji Yamada, **Alex Fraser**, Shankar Kumar, Libin Shen, David Smith, Katherine Eng, Viren Jain, Zhen Jin, Dragomir Radev (2004). A Smorgasbord of Features for Statistical Machine Translation. In *Proceedings of the Human Language Technology and North American Association for Computational Linguistics Conference (HLT/NAACL)*, pages 161-168, Boston, Massachusetts, USA, May

- Jinxi Xu, **Alexander Fraser**, Ralph Weischedel (2002). Empirical Studies in Strategies for Arabic Retrieval. In *Proceedings of the 25th Annual International Conference on Research and Development in Information Retrieval (SIGIR)*, pages 269-274, Tampere, Finland, August

Refereed Conference Presentations

- Yi-Chen Liu, Yuri Pettinicchi, **Alexander Fraser** (2020). Improving SHARE translation verification. In BigSurv20, November

Refereed Conference Demonstrations

- Anita Ramm, Sharid Loáiciga, Annemarie Friedrich, **Alexander Fraser** (2017). Annotating Tense, Mood and Voice for English, French and German. In *Proceedings of the 55th Annual Meeting of the Association for Computational Linguistics (ACL)*, demo session, pages 1-6, Vancouver, Canada, July

- **Alex Fraser**, Laurie Gerber, Kevin Knight, Daniel Marcu, Franz Josef Och, William Wong (2004). Language Weaver Arabic-to-English Demo. In *Human Language Technology and North American Association for Computational Linguistics Conference - Demonstrations*. Boston, Massachusetts, USA, May

Invited Conference Publication

- **Alexander Fraser**, Renjing Wang, Hinrich Schütze (2009). A progress report on bitext parsing. In *2nd International Conference on the Future of Information Sciences (InFuture): Digital Resources and Knowledge Sharing*, pages 43-52, Zagreb, Croatia, November. Invited Paper

<h1 style="text-align:center">Invited Conference Presentation</h1>

- **Alexander Fraser** (2013). Machine Learning in the Statistical Natural Language Processing Group of the University of Stuttgart. In *Joint Third EUCogIII Members Conference and PASCAL2 Review Meeting*, Palma de Mallorca, Spain, April. Invited Flashtalk and Poster

<h1 style="text-align:center">Shared Task Conference Publications</h1>

- Lukas Edman, Lisa Bylinina, Faeze Ghorbanpour, **Alexander Fraser** (2024). Are BabyLMs Second Language Learners? In *Proceedings of the BabyLM Challenge at the 28th Conference on Computational Natural Language Learning*, Miami, FL, USA, November.

- Marion Weller-Di Marco, **Alexander Fraser** (2022). Findings of the WMT 2022 Shared Tasks in Unsupervised MT and Very Low Resource Supervised MT. In *Proceedings of the Seventh Conference on Machine Translation (WMT)*, pages 801-805, Abu Dhabi, UAE, December.

- Marion Weller-Di Marco, **Alexander Fraser** (2022). Test Suite Evaluation: Morphological Challenges and Pronoun Translation. In *Proceedings of the Seventh Conference on Machine Translation (WMT)*, pages 458-468, Abu Dhabi, UAE, December.

- Jindřich Libovický, **Alexander Fraser** (2021). Findings of the WMT 2021 Shared Tasks in Unsupervised MT and Very Low Resource Supervised MT. In *Proceedings of the Sixth Conference on Machine Translation (WMT)*, pages 726-732, Punta Cana, Dominican Republic, November

- Wen Lai, Jindřich Libovický, **Alexander Fraser** (2021). The LMU Munich System for the WMT 2021 Large-Scale Multilingual Machine Translation Shared Task. In *Proceedings of the Sixth Conference on Machine Translation (WMT)*, pages 412-417, Punta Cana, Dominican Republic, November.

- Jindřich Libovický, **Alexander Fraser** (2021). The LMU Munich Systems for the WMT21 Unsupervised and Very Low-Resource Translation Task. In *Proceedings of the Sixth Conference on Machine Translation (WMT)*, pages 989-994, Punta Cana, Dominican Republic, November.

- **Alexander Fraser** (2020). Findings of the WMT 2020 Shared Tasks in Unsupervised MT and Very Low Resource Supervised MT. In *ACL 2020 Fifth Conference on Machine Translation (WMT)*, pages 765-771, November

- Alexandra Chronopoulou, Dario Stojanovski, Viktor Hangya, **Alexander Fraser** (2020). The LMU Munich System for the WMT 2020 Unsupervised Machine Translation Shared Task. In *ACL 2020 Fifth Conference on Machine Translation (WMT)*, pages 2703-2711, November

- Jindřich Libovický, Viktor Hangya, Helmut Schmid, **Alexander Fraser** (2020). The LMU Munich System for the WMT20 Very Low Resource Supervised MT Task. In *ACL 2020 Fifth Conference on Machine Translation (WMT)*, pages 1104-1111, November

- Dario Stojanovski, Viktor Hangya, Matthias Huck, **Alexander Fraser** (2019). The LMU Munich Unsupervised Machine Translation System for WMT19. In *Proceedings of the Fourth Conference on Machine Translation (WMT)*, pages 592-598, Florence, Italy, August

- Dario Stojanovski, **Alexander Fraser** (2019). Combining Local and Document-Level Context: The LMU Munich Neural Machine Translation System at WMT19. In *Proceedings of the Fourth Conference on Machine Translation (WMT)*, pages 599-605, Florence, Italy, August

- Viktor Hangya, **Alexander Fraser** (2018). An Unsupervised System for Parallel Corpus Filtering. In *Proceedings of the Third Conference on Machine Translation (WMT)*, pages 882-887, Brussels, Belgium, October

- Matthias Huck, Dario Stojanovski, Viktor Hangya, **Alexander Fraser** (2018). LMU Munich's Neural Machine Translation Systems at WMT 2018. In *Proceedings of the Third Conference on Machine Translation (WMT)*, pages 648-654, Brussels, Belgium, October

- Dario Stojanovski, Viktor Hangya, Matthias Huck, **Alexander Fraser** (2018). The LMU Munich Unsupervised Machine Translation Systems. In *Proceedings of the Third Conference on Machine Translation (WMT)*, pages 513-521, Brussels, Belgium, October

- Matthias Huck, Fabienne Braune, **Alexander Fraser** (2017). LMU Munich's Neural Machine Translation Systems for News Articles and Health Information Texts. In *Proceedings of the Second ACL Conference on Machine Translation (WMT)*, pages 315-322, Copenhagen, Denmark, September

- Matthias Huck, **Alexander Fraser**, Barry Haddow (2016). The Edinburgh/LMU Hierarchical Machine Translation system for WMT 2016. In *Proceedings of the First Conference on Machine Translation (WMT)*, pages 311-318, Berlin, Germany, August

- Jan-Thorsten Peter, Tamer Alkhouli, Hermann Ney, Matthias Huck, Fabienne Braune, **Alexander Fraser**, Aleš Tamchyna, Ondřej Bojar, Barry Haddow, Rico Sennrich, Frédéric Blain, Lucia Specia, Jan Niehues, Alex Waibel, Alexandre Allauzen, Lauriane Aufrant, Franck Burlot, Elena Knyazeva, Thomas Lavergne, François Yvon, Stella Frank, Mārcis Pinnis (2016). The QT21/HimL Combined Machine Translation System. In *Proceedings of the First Conference on Machine Translation (WMT)*, pages 341-352, Berlin, Germany, August

- Aleš Tamchyna, Roman Sudarikov, Ondřej Bojar, **Alexander Fraser** (2016). CUNI-LMU submissions in WMT2016: Chimera constrained and beaten. In *Proceedings of the First Conference on Machine Translation (WMT)*, pages 382-387, Berlin, Germany, August

Conference Project Description

- B. Haddow, A. Birch, O. Bojar, F. Braune, C. Davenport, **A. Fraser**, M. Huck, M. Kaspar, K. Kovarikova, J. Plch, A. Ramm, J. Ried, J. Sheary, A. Tamchyna, D. Variš, M. Weller, P. Williams (2017). HimL: Health in my Language. In *EAMT 2017 User Studies and Project/Product Descriptions*, page 33, Prague, Czech Republic, May

Refereed Workshop Publications

- Shu Okabe, **Alexander Fraser** (2025). Bilingual Sentence Mining for Low-Resource Languages: a Case Study on Upper and Lower Sorbian. *Eighth Workshop on the Use of Computational Methods in the Study of Endangered Languages (ComputEL), to appear.*

- Andreas Dimmelmeier, Hendrik Doll, Malte Schierholz, Emily Kormanyos, Maurice Fehr, Bolei Ma, Jacob Beck, **Alexander Fraser**, Frauke Kreuter (2024). Informing Climate Risk Analysis Using Textual Information - A Research Agenda. In *Proceedings of the 1st Workshop on Natural Language Processing Meets Climate Change (ClimateNLP 2024)*, pages 12-26, Bangkok, Thailand, August.

- Yaqi Zhang, Viktor Hangya, **Alexander Fraser** (2024). A Study of the Class Imbalance Problem in Abusive Language Detection. *Proceedings of the Workshop on Online Abuse and Harms (WOAH)*, pages 38-51, Mexico City, Mexico, June.

- Viktor Hangya, Silvia Severini, Radoslav Ralev, **Alexander Fraser**, Hinrich Schütze (2023). Multilingual Word Embeddings for Low-Resource Languages using Anchors and a Chain of Related Languages. In Proceedings of the 3rd Workshop on Multi-lingual Representation Learning (MRL), pages 95-105, Singapore, December.

- Alexandra Chronopoulou, Dario Stojanovski, **Alexander Fraser** (2023). Language-Family Adapters for Low-Resource Multilingual Neural Machine Translation. *Workshop on Technologies for Machine Translation of Low-Resource Languages (LoResMT)*, pages 59-72, Dubrovnik, Croatia, May.

- Hossain Shaikh Saadi, Viktor Hangya, Tobias Eder, **Alexander Fraser** (2022). Comparative Analysis of Cross-lingual Contextualized Word Embeddings. In *2nd Multilingual Representation Learning Workshop (MRL)*, pages 64-75, Abu Dhabi, UAE, December.

- Silvia Severini, Viktor Hangya, Masoud Jalili Sabet, **Alexander Fraser**, Hinrich Schütze (2022). Don't Forget Cheap Training Signals Before Building Unsupervised Bilingual Word Embeddings. In *BUCC, 15th Workshop on Building and Using Comparable Corpora*, pages 15-22, Marseille, France, June.

- Jindřich Libovický, **Alexander Fraser (2022)**. Neural String Edit Distance. In *ACL 2022 6th Workshop on Structured Prediction for NLP (SPNLP)*, pages 52-66, Dublin, Ireland, May.

- Lisa Woller, Viktor Hangya, **Alexander Fraser** (2021). Do not neglect related languages: The case of low-resource Occitan cross-lingual word embeddings. In *Proceedings of the 1st Workshop on Multilingual Representation Learning*, pages 41-50, Punta Cana, Dominican Republic.

- Irina Bigoulaeva, Viktor Hangya, **Alexander Fraser** (2021). Cross-Lingual Transfer Learning for Hate Speech Detection. In *Proceedings of the First Workshop on Language Technology for Equality, Diversity and Inclusion (LT-EDI)*, pages 15-25, April.

- Dario Stojanovski, **Alexander Fraser** (2021). Addressing Zero-Resource Domains Using Document-Level Context in Neural Machine Translation. In *Proceedings of the Second Workshop on Domain Adaptation for NLP (AdaptNLP)*, pages 80-93, April.

- Silvia Severini, Viktor Hangya, **Alexander Fraser**, Hinrich Schütze (2020). LMU Bilingual Dictionary Induction System with Word Surface Similarity Scores for BUCC 2020. In *Proceedings of the 13th Workshop on Building and Using Comparable Corpora (BUCC)*, pages 49-55, May

- Matthias Huck, Diana Dutka, **Alexander Fraser** (2019). Cross-lingual Annotation Projection Is Effective for Neural Part-of-Speech Tagging. In *Proceedings of the Sixth Workshop on NLP for Similar Languages, Varieties and Dialects (VarDial)*, pages 223-233, Minneapolis, Minnesota, USA, June

- Viktor Hangya, Fabienne Braune, Yuliya Kalasouskaya, **Alexander Fraser** (2018). Unsupervised Parallel Sentence Extraction from Comparable Corpora. In *Proceedings of the International Workshop on Spoken Language Translation (IWSLT)*, pages 7-13, Bruges, Belgium, October

- Marion Weller, Fabienne Cap, Stefan Müller, Sabine Schulte im Walde, **Alexander Fraser** (2014). Distinguishing Degrees of Compositionality in Compound Splitting for Statistical Machine Translation. In *Proceedings of the First Workshop on Computational Approaches to Compound Analysis (ComaComa) at COLING*, pages 81-90, Dublin, Ireland, August

- Fabienne Fritzinger, **Alexander Fraser** (2010). How to Avoid Burning Ducks: Combining Linguistic Analysis and Corpus Statistics for German Compound Processing. In *Proceedings of the ACL 2010 Fifth Workshop on Statistical Machine Translation and MetricsMATR*, pages 224-234, Uppsala, Sweden, July

Refereed Workshop Presentations

- Marion Weller, **Alexander Fraser**, Sabine Schulte im Walde (2015). Predicting Prepositions for SMT. In *Proceedings of Ninth Workshop on Syntax, Semantics and Structure in Statistical Translation (SSST)*, pages 55-56, Denver, Colorado, USA, June

- **Alexander Fraser**, Marion Weller, Aoife Cahill, Fabienne Cap (2011). Morphological Generation of German for SMT. *Abstract accepted for presentation at Machine Translation and Morphologically-rich Languages - Research Workshop of the Israel Science Foundation.* University of Haifa, Israel, January

- Daniel Marcu, **Alex Fraser**, William Wong, Kevin Knight (2004). Language Weaver Arabic->English MT (demo). In *Semitic 2004: Proceedings of the Workshop on Computational Approaches to Arabic Script-based Languages*, page 89, Geneva, Switzerland, August

Shared Task Workshop Publications

- Viktor Hangya, **Alexander Fraser** (2023). LMU at HaSpeeDe3: Multi-Dataset Training for Cross-Domain Hate Speech Detection. In *Proceedings of EVALITA*, Parma, Italy, September.

- Fabienne Cap, Marion Weller, Anita Ramm, **Alexander Fraser** (2015). CimS - The CIS and IMS Joint Submission to WMT 2015 addressing morphological and syntactic differences in English to German SMT. In *Proceedings of the Tenth Workshop on Statistical Machine Translation*, pages 84-91, Lisbon, Portugal, September

- Fabienne Cap, Marion Weller, Anita Ramm, **Alexander Fraser** (2014). CimS – The CIS and IMS joint submission to WMT 2014: translating from English into German. In *Proceedings of the ACL Ninth Workshop on Statistical Machine Translation*, pages 71-78, Baltimore, Maryland, USA, June

- Nadir Durrani, Helmut Schmid, **Alexander Fraser**, Hassan Sajjad, Richárd Farkas (2013). Munich-Edinburgh-Stuttgart Submissions of OSM Systems at WMT13. In *Proceedings of the ACL Eighth Workshop on Statistical Machine Translation*, pages 122-127, Sofia, Bulgaria, August

- Hassan Sajjad, Svetlana Smekalova, Nadir Durrani, **Alexander Fraser**, Helmut Schmid (2013). QCRI-MES Submission at WMT13: Using Transliteration Mining to Improve Statistical Machine Translation. In *Proceedings of the ACL Eighth Workshop on Statistical Machine Translation*, pages 219-224, Sofia, Bulgaria, August

- Marion Weller, Max Kisselew, Svetlana Smekalova, **Alexander Fraser**, Helmut Schmid, Nadir Durrani, Hassan Sajjad, Richárd Farkas (2013). Munich-Edinburgh-Stuttgart Submissions at WMT13: Morphological and Syntactic Processing for SMT. In *Proceedings of the ACL Eighth Workshop on Statistical Machine Translation*, pages 232-239, Sofia, Bulgaria, August

- **Alexander Fraser** (2009). Experiments in Morphosyntactic Processing for Translating to and from German. In *Proceedings of the EACL Fourth Workshop on Statistical Machine Translation*, pages 115-119, Athens, Greece, March

- **Alexander Fraser**, Daniel Marcu (2005). ISI's Participation in the Romanian-English Alignment Task. In *Proceedings of the ACL Workshop on Building and Using Parallel Texts*, pages 91-94, Ann Arbor, Michigan, USA, June

- **Alexander Fraser**, Jinxi Xu, Ralph Weischedel (2002). TREC 2002 Cross-lingual Retrieval at BBN. In *Voorhees, E.M. & Harman, D.K. (Eds.) The Eleventh Text Retrieval Conference (TREC)*, pages 102-106, NIST Special Publication 500-251

- Jinxi Xu, **Alexander Fraser**, Ralph Weischedel (2001). TREC 2001 Cross-lingual Retrieval at BBN. In *Voorhees, E.M. & Harman, D.K. (Eds.) The Tenth Text Retrieval Conference (TREC)*, pages 68 to 77, NIST Special Publication 500-250

Technical Reports

- Wen Lai, **Alexander Fraser**, Ivan Titov (2025). Joint Localization and Activation Editing for Low-Resource Fine-Tuning. arXiv

- Yingli Shen, Wen Lai, Shuo Wang, Xueren Zhang, Kangyang Luo **Alexander Fraser**, Maosong Sun (2025). DCAD-2000: A Multilingual Dataset across 2000+ Languages with Data Cleaning as Anomaly Detection. arXiv

- Felix Friedrich, Katharina Hämmerl, Patrick Schramowski, Jindřich Libovický, Kristian Kersting, **Alexander Fraser** (2024). Multilingual Text-to-Image Generation Magnifies Gender Stereotypes and Prompt Engineering May Not Help You. arXiv

- Wen Lai, Viktor Hangya, **Alexander Fraser** (2023). Extending multilingual machine translation through imitation learning. arXiv

- Katharina Hämmerl, Björn Deiseroth, Patrick Schramowski, Jindřich Libovický, **Alexander Fraser**, Kristian Kersting (2022). Do Multilingual Language Models Capture Differing Moral Norms? arXiv

- Marion Weller-Di Marco, Matthias Huck, **Alexander Fraser** (2022). Modeling Target-Side Morphology in Neural Machine Translation: A Comparison of Strategies. arXiv

- Irina Bigoulaeva, Viktor Hangya, Iryna Gurevych, **Alexander Fraser** (2022). Addressing the Challenges of Cross-Lingual Hate Speech Detection. arXiv

- Dario Stojanovski, **Alexander Fraser** (2020). Addressing Zero-Resource Domains Using Document-Level Context in Neural Machine Translation. arXiv

- Jindřich Libovický, Rudolf Rosa, **Alexander Fraser** (2019). How Language-Neutral is Multilingual BERT? arXiv

- Anita Ramm, Ekaterina Lapshinova-Koltunski, **Alexander Fraser** (2019). Pragmatic Information in Translation: a Corpus-based Study of Tense and Mood in English and German. Technical Report, CIS/LMU Munich, September

- Marine Carpuat, Hal Daume III, **Alexander Fraser**, Chris Quirk, Fabienne Braune, Ann Clifton, Ann Irvine, Jagadeesh Jagarlamudi, John Morgan, Majid Razmara, Aleš Tamchyna, Katharine Henry, Rachel Rudinger (2012). Domain Adaptation in Machine Translation: Final Report. Johns Hopkins University 2012 Summer Workshop on Language Engineering, Center for Language and Speech Processing, Baltimore, Maryland, USA, December

- **Alexander Fraser**, Daniel Marcu (2006). Measuring Word Alignment Quality for Statistical Machine Translation. Technical Report ISI-TR-616, ISI/University of Southern California, May

- Franz Josef Och, Daniel Gildea, Sanjeev Khudanpur, Anoop Sarkar, Kenji Yamada, **Alex Fraser**, Shankar Kumar, Libin Shen, David Smith, Katherine Eng, Viren Jain, Zhen Jin, Dragomir Radev (2003). Syntax for Statistical Machine Translation. Johns Hopkins University 2003 Summer Workshop on Language Engineering, Center for Language and Speech Processing, Baltimore, Maryland, USA, August

Granted Patent

- **Alexander Fraser**, Daniel Marcu. Semi-Supervised Training for Statistical Word Alignment. Patent US8433556, United States Patent and Trademark Office, filed 11/2006, granted 04/2013