

Alexander M. Fraser

Curriculum Vitae



Personal Data

Current Position	Professor of Information and Language Processing, LMU Munich
Work Address	Ludwig-Maximilians-Universität München Centrum für Informations- und Sprachverarbeitung Oettingenstraße 67 80538 Munich, Germany
	Phone +49 89 2180-9710 Email fraser@cis.uni-muenchen.de Web www.cis.uni-muenchen.de/~fraser
Citizenship	USA
Family Status	Married, no children
Languages	English (native) German, French, Spanish (fluent) Arabic (proficient)

Research Interests

Natural Language Processing	Statistical and hybrid machine translation, statistical parsing of syntax, cross-language information retrieval, morphological analysis and generation, discourse, text mining
Machine Learning	Structured prediction, semi-supervised and unsupervised learning, hybrid discriminative and generative models, domain adaptation, distributed training algorithms

Education

12/2007	PhD in Computer Science University of Southern California, Los Angeles (CA), USA <ul style="list-style-type: none">Thesis Title: Improved Word Alignments for Statistical Machine TranslationAdvisor: Daniel MarcuCommittee: Eduard Hovy, Gareth James, Kevin Knight, Daniel Marcu, Paul Rosenbloom
---------	---

- 12/2002 **Masters of Science in Computer Science**
University of Southern California, Los Angeles (CA), USA
▪ GPA 3.86/4.00
- 06/1995 **BA in Computer Science and Economics (Concentration in Linguistics)**
University of Massachusetts, Boston (MA), USA
▪ GPA 3.72/4.00

Professional Experience

- Since
09/2017 **W2 Professor of Information and Language Processing (~Full, tenured)**
Center for Information and Language Processing (CIS), (tenure track 2017-2020)
Ludwig-Maximilians-Universität München, Germany
- 03/2014 –
08/2014 **Visiting Professor for the W3 Professor of Linguistic Computer Science**
Institute of Computational Linguistics (ICL),
Ruprecht-Karls-Universität Heidelberg
- 08/2013 -
08/2017 **Akademischer Rat (~ Assistant professor, permanent)**
Center for Information and Language Processing (CIS),
Ludwig-Maximilians-Universität München, Germany
- Since
04/2013 **Leader of Statistical Machine Translation Group**
Center for Information and Language Processing (CIS),
Ludwig-Maximilians-Universität München, Germany
- 10/2012 –
09/2017 **Special Faculty Member**
Department of Computer Science, University of Maryland,
College Park (MD), USA
- 10/2009 –
04/2013 **Senior Researcher, Leader of Statistical Machine Translation Group**
Institute for Natural Language Processing (IfNLP/IMS),
University of Stuttgart, Germany
- 11/2007 –
09/2009 **Researcher**, Institute for Natural Language Processing (IfNLP/IMS),
DFG Collaborative Research Centre 732: Incremental Specification in Context
University of Stuttgart, Germany
- 09/2000 –
09/2007 **Graduate Research Assistant** (except dates listed below)
University of Southern California, Information Science Institute (ISI),
Marina del Rey (CA), USA
- 06/2003 –
05/2004 **Research Scientist**, Language Weaver, Marina del Rey (CA), USA
- 06 – 08/2002 **Research Scientist**, Department of Speech and Language Processing,
05 – 08/2001 BBN Technologies, Cambridge (MA), USA
02 – 08/2000
- 09/1995 – 01/1998 **Technical Director**, SatelLife, Boston (MA), USA (last position held)
03/1994 – 12/1994
- 05/1987 – 01/1990 **Programmer / Analyst**, The Forum Corporation, Boston (MA), USA (last position held)

Publications

Up-to-date citation counts and bibliometric information are available from my Google Scholar page.

Refereed Journal Publications

- 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 and **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

- Katharina Hämmerl, Jindřich Libovický, **Alexander Fraser** (2022). Combining static and contextualised multilingual embeddings. In *Findings of ACL 2022*, to appear.
- Jindřich Libovický, **Alexander Fraser** (2022). Why don't people use character-level machine translation? In *Findings of ACL 2022*, to appear.
- 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 to 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 to 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 to 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 to 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 to 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 to 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 to 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 to 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 to 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 to 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 to 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 to 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 to 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 to 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 to 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 to 1234, Florence, Italy, July
- Matthias Huck, Viktor Hangya, and **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 to 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 to 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 to 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 to 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 to 193, New Orleans, Louisiana, USA, June
- Costanza Conforti, Matthias Huck, and **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 to 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 to 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 to 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 to 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 to 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 to 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 to 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 to 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 to 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 to 2274, Lisbon, Portugal, September

- Martin Wunderlich, **Alexander Fraser**, P.S. Langeslag (2015). God Wat Pæt 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 to 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 to 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 to 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 to 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 to 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 to 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 to 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 to 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 to 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 to 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 to 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 to 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 to 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 to 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 to 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 to 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 to 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 to 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 to 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 to 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 to 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 to 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 to 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 to 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 to 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 to 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 and **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 to 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 to 52, Zagreb, Croatia, November. Invited Paper

Invited Conference Presentation

- **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

Shared Task Conference Publications

- 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 to 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 to 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 to 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 (WMT20)*, pages 765 to 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 (WMT20)*, pages 2703 to 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 (WMT20)*, pages 1104 to 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*, pages 592 to 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*, pages 599 to 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*, pages 882 to 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*, pages 648 to 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*, pages 513 to 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*, pages 315 to 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*, pages 311 to 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 Burtot, Elena Knyazeva, Thomas Lavergne, François Yvon, Stella Frank, and Mărcis Pinnis (2016). The QT21/HimL Combined Machine Translation System. In *Proceedings of the First Conference on Machine Translation*, pages 341 to 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*, pages 382 to 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, and 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

- Lisa Woller, Viktor Hangya, and **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 to 50, Punta Cana, Dominican Republic.
- Irina Bigoulaeva, Viktor Hangya, **Alexander Fraser** (2021). Cross-Lingual Transfer Learning for Hate Speech Detection. Proceedings of the First Workshop on Language Technology for Equality, Diversity and Inclusion (LT-EDI), pages 15 to 25, April.
- Dario Stojanovski, **Alexander Fraser** (2021). Addressing Zero-Resource Domains Using Document-Level Context in Neural Machine Translation. Proceedings of the Second Workshop on Domain Adaptation for NLP (AdaptNLP), pages 80 to 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 to 55, May
- Matthias Huck, Diana Dutka, and **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 to 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 to 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 to 90, Dublin, Ireland, August
- Fabienne Fritzing, **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 to 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 to 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

- 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 to 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 to 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 to 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 to 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 to 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 to 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 to 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 to 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

- 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

Honors / Awards

- 1000 Euro award for open source, distributed implementation of LEAF alignment model (from dissertation) at Euromatrix Second Machine Translation Marathon. 2009
- Full sponsorship for participation in Euromatrix Second Machine Translation Marathon (one week). 2009
- Full sponsorship to spend 8 weeks in Baltimore at the Johns Hopkins University CLSP Summer Workshop 2003: Syntax for Statistical Machine Translation, led by Franz Och and Daniel Gildea. 2003
- SIGIR 2002 Student Travel Award. 2002
- University of Southern California - Charles Lee Powell Graduate Fellowship, full support and additional stipend for 3 years of PhD studies. 2000-2003
- University of Massachusetts, Boston - Graduated magna cum laude. 1995
- University of Massachusetts, Boston - Juan Carlos Merlo Prize, awarded annually to a single graduating mathematics or computer science student for outstanding achievement. 1995
- University of Massachusetts, Boston - Honors in Computer Science. 1995
- University of Massachusetts, Boston - Honors in Economics. 1995

Grants

- Principal Investigator (with Helmut Schmid), DFG Project (Sachbeihilfe), Machine Translation of German without Parallel Corpora, 2020, 581700 Euros
- Principal Investigator, Cambridge-LMU Strategic Partnership, 2019, funded two joint workshops at Cambridge (three days) and at LMU (two days) for my ERC project DASMT and Prof. Anna Korhonen's ERC project LEXICAL (Cambridge)

- Principal Investigator, LMUexcellent Investment Fund, 1 year, 2018-2019, 50000 Euros
- Principal Investigator, ERC Starting Grant (awarded by PE6 (Computer Science and Informatics) in the 2014 call, started in December 2015), 5 years, 1.2 million Euros
- Principal Investigator, Horizon2020 Innovation Action “Health in my Language (HimL)” with U. Edinburgh, Charles U. Prague, 3 other organizations, 2015-2018, 500000 Euros
- Principal Investigator (with Knight, Koehn, Schmid, Uszkoreit) on Dagstuhl Seminar on Statistical Techniques for Translating to Morphologically Rich Languages, 2014, facilitates meeting of 60 researchers in Germany for one week
- Principal Investigator (with Schütze) on DFG Sachbeihilfe, “Models of Morphosyntax for Statistical Machine Translation - Phase II”, 2 years, 2013-2015, 300000 Euros
- Principal Investigator on Johns Hopkins Summer Workshop on Domain Adaptation for Machine Translation (with Daume, Carpuat, Quirk), 2012, facilitated research by 13 people for 6 weeks on site
- Deputy Investigator, European Commission FP7 STREP, “TTC: Terminology Extraction, Translation Tools and Comparable Corpora”. Coordinator Beatrice Daille, Nantes; Stuttgart PI Ulrich Heid, 2010-2012, 1 full-time position allocated to me for 1 1/2 years
- Principal Investigator (with Schütze) on DFG Sachbeihilfe, “Models of Morphosyntax for Statistical Machine Translation”, 3 years, 2009-2012, 517000 Euros
- Principal Investigator on Small Business Innovation Research grant to Language Weaver, “Linguistic Error Analysis of Statistical Machine Translation”. 50000 USD, 2004
- Principal Investigator (sole implementing PI) on grant for design, implementation and quality assurance of Language Weaver Arabic/English hybrid statistical machine translation system, 2 years, 2003-2004, amount confidential, approx. 5 full-time positions

Teaching

I am the coordinator of the Masters degree in Computational Linguistics at LMU Munich. I typically teach a standard full teaching load (9 SWS), but after becoming a professor I received a reduction (during the rest of my ERC Starting Grant). I also organize the institute reading group (together with Schütze), which is currently on deep learning approaches for natural language processing and machine translation.

- “Informationsextraktion”, LMU München, Lecture, **Every Winter Semester from 2014/2015 to 2021/2022**
- “Seminar zur Informationsextraktion (I)”, LMU München, Seminar/Exercise, **Every Winter Semester from 2014/2015 to 2021/2022**
- “Seminar zur Informationsextraktion (II)”, LMU München, Seminar/Exercise, **Every Winter Semester from 2014/2015 to 2021/2022**
- “Erweiterungsmodul: Machine Translation”, Lecture and Exercise, Summer Semester 2021
- “Erweiterungsmodul: Machine Translation”, Lecture and Exercise, Summer Semester 2020
- “Einführung in die Computerlinguistik”, with Robert Zangenfeind, LMU München, Lecture and Exercise, Winter Semester 2019/2020

- “Lexicon, Syntax, Semantics: Structured Prediction and Domain Adaptation”, LMU München, Seminar, Winter Semester 2017/2018
- “Erweiterungsmodul: Machine Translation”, LMU München, Lecture and Exercise, Summer Semester 2017
- “Computational Morphology and Electronic Dictionaries”, LMU München, Lecture and Exercise, Summer Semester 2017
- “Lexicon, Syntax, Semantics: Statistical Machine Translation and Neural Machine Translation”, LMU München, Seminar, Winter Semester 2016/2017
- “Seminar: Advanced Neural Machine Translation”, LMU München, Seminar, Summer Semester 2016
- “Computational Morphology and Electronic Dictionaries”, LMU München, Lecture and Exercise, Summer Semester 2016
- “Seminar: Neural Machine Translation”, LMU München, Seminar, Winter Semester 2015/2016
- “Lexicon, Syntax, Semantics: Word Sense Disambiguation and Statistical Machine Translation”, LMU München, Seminar, Winter Semester 2015/2016
- “Intensive Introduction to Neural Machine Translation”, LMU München, Seminar, Summer Semester 2015
- “Lexicon, Syntax, Semantics: Word Sense Disambiguation and Statistical Machine Translation”, LMU München, Seminar, Winter Semester 2014/2015
- “Information Extraction”, DAAD International Summer School in Advanced Language Engineering, University of Colombo, Sri Lanka, September 2014
- “Programmieren II: Fortgeschrittene (Java)”, Universität Heidelberg, Lecture and Exercise, Summer Semester 2014
- “Statistical Machine Translation”, Universität Heidelberg, Lecture, Summer Semester 2014
- “Linguistic Knowledge for Statistical Machine Translation”, Universität Heidelberg, Proseminar, Summer Semester 2014
- “Informationsextraktion”, LMU München, Lecture, Winter Semester 2013/2014
- “Seminar zur Informationsextraktion”, LMU München, Seminar, Winter Semester 2013/2014
- “Statistical Machine Translation”, DAAD Summer School in Advanced Language Engineering, Kathmandu University, Nepal, September 2012
- “Statistical Machine Translation”, with Andreas Maletti, Seminar (Hauptseminar), University of Stuttgart, Winter Semester 2011/2012
- “Machine Learning for Natural Language Processing” Seminar (Hauptseminar), with Stefan Riezler and Hinrich Schütze, Joint Seminar: Heidelberg University and University of Stuttgart, Winter Semester 2010/2011
- “Statistical Machine Translation”, with Helmut Schmid, Seminar (Lektürekurs), University of Stuttgart, Summer Semester 2010
- “Parsing and Machine Translation”, with Helmut Schmid, Seminar (Hauptseminar), University of Stuttgart, Summer Semester 2009

- “Data-driven and Hybrid Approaches to Machine Translation”, with Martin Forst, intensive two week course at German Society for Linguistics Fall School in Konstanz, 2009
- One week intensive course on Statistical Machine Translation at 9th Summer School of the European Masters in Language and Speech Technology Programme, Stuttgart, 2008

Advising

Post-docs

- Dr. Viktor Hangya, 01/2017-, research on multilingual embeddings for domain adaptation
- Dr. Marion Weller-Di Marco, 09/2019-, research on morphological generalization in neural machine translation
- Dr. Jindřich Libovický, 09/2019-12/2021, research on character-based machine translation and multilingual representations
- Dr. Matthias Huck, 04/2016 to 03/2020, research on domain adaptation and translation to morphologically rich languages
- Dr. Fabienne Braune, 01/2016 to 12/2017, research on multilingual embeddings for mining comparable corpora
- Dr. Aleš Tamchyna, 01/2017 to 04/2017, research on translation to morphologically rich languages (jointly supervised with Prof. Jan Hajic, Charles University Prague)
- Dr. Tsuyoshi Okita, 01/2016 to 12/2016, research on neural models for automatic post-editing
- Dr. Liane Guillou, 04/2016 to 09/2016, research on discourse for translation

Ph.D. Advising

- Irina Bigoulaeva (with Iryna Gurevych (TU Darmstadt) and Viktor Hangya). 10/2021-, TU Darmstadt funding; Transfer Learning and Hate Speech Detection
- Katharina Hämmerl (with Jindrich Libovicky, now Charles U Prague). 07/2021-, DFG Unsup; Multilingual Representations
- Prisca Piccirilli (with Sabine Schulte im Walde, U Stuttgart). 07/2021-, Joint Stuttgart and LMU DFG funding; Metaphor Translation
- Wen (Lavine) Lai. 04/2021-, LMU/CSC; Low Resource NMT, Domain Adaptation
- Sophie Henning 01/2021-, Bosch, with Annemarie Friedrich, research on calibration
- Alexandra Chronopoulou, 09/2019-, research on transfer and multilingual learning
- Dr. rer. nat. Dario Stojanovski, 06/2017-06/2021, Modeling Contextual Information in Neural Machine Translation. First permanent position after graduation: Microsoft.
- Dr. phil. Anita Ramm (née Gojun) (with Sabine Schulte im Walde), 05/2013 to 03/2018, Modeling the position and inflection of verbs in English to German machine translation. First position after graduation: BeOne Stuttgart GmbH.
- Dr. phil. Marion Weller-Di Marco (with Sabine Schulte im Walde), 12/2011 to 09/2017, Using Morpho-Syntactic and Semantic Information to Improve Statistical Machine Translation. First position after graduation: Post-doc at University of Amsterdam with Khalil Sima'an.

- Dr. rer. nat. Fabienne Braune (with Andreas Maletti), 10/2009-11/2015, Decoding Strategies for Syntax-based Statistical Machine Translation. Winner of the **European Association for Machine Translation Best Thesis Award** for the best machine translation dissertation in Europe in 2015. First position after graduation: Post-doc at LMU Munich with Alexander Fraser.
- Dr. phil. Fabienne Cap (née Fritzing), 10/2009-10/2014, Morphological Processing of Compounds for Statistical Machine Translation. First position after graduation: Post-doc (VINNMER Marie Curie Incoming) at Uppsala University with Joakim Nivre.
- Dr. rer. nat. Nadir Durrani (with Helmut Schmid), 04/2008-11/2012, A Joint Translation Model with Integrated Reordering. Winner of the **German Society for Computational Linguistics award** for the best natural language processing / computational linguistics dissertation for the years 2011 to 2013. First position after graduation: Post-doc at University of Edinburgh with Philipp Koehn.
- Dr. rer. nat. Hassan Sajjad (with Helmut Schmid), 04/2008-10/2012, Statistical Models for Unsupervised, Semi-supervised and Supervised Transliteration Mining. First position after graduation: Qatar Computing Research Institute.

Visiting Ph.D. Students

- Denis Peskov, University of Maryland, College Park, 01/2020 to 10/2020
- Simon Rieß, LMU Data Science, 10/2019 to 12/2019
- Viktor Hangya, University of Szeged, 10/2016 to 12/2016
- Aleš Tamchyna, Charles University Prague, 10/2015 to 05/2016
- Marija Brkic, University of Zagreb, Croatia. 09/2010 to 02/2011

External Ph.D. Committees

- Ph.D. Committee Member: Gongbo Tan. Uppsala University, Sweden. 12/2020 (Joakim Nivre)
- Ph.D. Committee Member: Andreas Guta. RWTH Aachen, Germany. 09/2020 (Hermann Ney)
- Ph.D. Committee Member: Minh-Quang Pham. LIMSI, Paris, France. 06/2020 (Francois Yvon)
- Ph.D. Reviewer: Duygu Ataman. University of Trento, Italy. 09/2019 (Marcello Federico)
- Ph.D. Reviewer: Ekaterina Vylomova. University of Melbourne, Australia. 11/2018 (Tim Baldwin)
- Ph.D. Committee Member: Matthias Huck. RWTH Aachen, Germany. 08/2018 (Hermann Ney)
- Ph.D. Committee Member: Peyman Passban. Dublin City University, Ireland. 09/2017 (Andy Way)
- Ph.D. Committee Member: Arianna Bisazza. University of Trento, Italy. 04/2013 (Marcello Federico)
- Ph.D. Committee Member: Jagadeesh Jagarlamudi. University of Maryland, College Park, MD, USA. 04/2013 (Hal Daume)
- Examiner, “Licentiate of Philosophy” (similar to American Ph.D. qualification exam): Sara Stymne, Department of Computer and Information Science, Linköping University, Sweden. 12/2009

Masters/Bachelors Advising

I have advised 27 Bachelors level and 23 Masters level students through Winter Semester 2020/2021, mostly at LMU Munich and the University of Stuttgart, but also at University of Ljubljana and Université de Fribourg. I have acted as Prüfer for many more. Marion Weller's Bachelors Thesis (Studienarbeit) in Computational Linguistics, *Separate Morphologiebehandlung als Methode zur Verbesserung statistischer maschineller Übersetzung (Separate Morphological Processing as a Method for Improving Statistical Machine Translation)*, completed 05/2009, was the winner of **Informatik-Verbund-Stuttgart Award** (given yearly for the two best Bachelors Theses in computer science-related fields in Stuttgart). A good number of the student works I have supervised have resulted in joint conference publications which passed peer review, see my publications list.

Invited Presentations

- Invited Talk, TUM and TUM Heilbronn, Germany, 2022
- Invited Talk, German Society for Computational Linguistics and Language Technology, 2021
- Invited Talk, University of Cambridge, UK, 2021
- Invited Talk, University of Bamberg, Germany, 2020
- Invited Talk, Uppsala University, Sweden, 2020
- Invited Talk, Karlsruhe Institute of Technology, Germany, 2020
- Invited Talk, TU Wien, Austria, 2020
- Invited Talk, University of Regensburg, Germany, 2020
- Invited Talk, University of Vienna, Austria, 2019
- Panel, "Open Problems in Machine Translation", ACL Conference on Machine Translation, Florence, Italy, 2019
- Invited Talk, TU Wien, Austria, 2019
- Invited Talk, University of Potsdam (Digital Engineering), Germany, 2019
- Keynote, "Translation to Morphologically Rich Languages", Baltic Human Language Technologies, University of Tartu, Estonia, 2018
- Invited Talk, Heidelberg University, Germany, 2018
- Invited Talk, Dublin City University, Ireland, 2017
- Invited Talk, University of Potsdam (Linguistics), Germany, 2017
- Keynote, "Translation to Morphologically Rich Languages", European Association for Machine Translation (EAMT 2016), Riga, Latvia, 2016
- Invited Talk, University of Hamburg, Germany, 2015
- Invited Talk, Universität Zürich, Switzerland, 2014
- Invited Talk, Machine Translation Marathon, Prague, Czech Republic, 2013
- Keynote, "Improving Translation to Morphologically Rich Languages", La 20ème conférence du Traitement Automatique du Langage Naturel (TALN), Sables d'Olonnes, France, 2013
- Invited Talk, Karlsruher Institut für Technologie, Germany, 2013
- Invited Talk, Institute for Natural Language Processing, University of Stuttgart, Germany, 2012

- Invited Talk, Ubiquitous Knowledge Processing Lab, TU Darmstadt, Germany, 2012
- Panel, “Arabic Natural Language Processing Technologies Round-Table”, Qatar Computing Research Institute, Doha, Qatar, 2011
- Invited Talk, Institute for Language, Cognition and Computation Seminar, University of Edinburgh, 2010
- Keynote, “Morphosyntactic Correspondences and a Progress Report on Bitext Parsing”, InFuture 2009, Zagreb, Croatia, 2009
- Invited Talk, German Society for Computational Linguistics & Language Technology Workshop “Costs and Uses of Machine Translation”, 2009
- Invited Talk, Research Track of Euromatrix Second Machine Translation Marathon, 2008
- Invited Talk, Institute for Computational Linguistics, Heidelberg University, 2008
- Invited Talk, Linguistics Department, University of Potsdam, 2007
- Invited Talk, Language Technology Laboratory, DFKI (German Research Center for Artificial Intelligence), 2006
- Invited Talk, Institute for Natural Language Processing (IMS), University of Stuttgart, 2006
- Panel, “Issues in Arabic to English Machine Translation”, MT Summit IX Workshop on Machine Translation for Semitic Languages, 2003
- Invited Talk, US National Institute of Standards Machine Translation Evaluation Workshop, 2003

Professional Activities

Senior PC

- NAACL 2021: Area Chair in Machine Translation and Multilinguality
- EACL 2021: Area Chair in Machine Translation
- ACL 2020: Area Chair in Natural Language Processing Applications
- ACL 2019: Area Chair in Machine Translation
- EAMT 2017: Program Chair (research track)
- EMNLP 2016: (Senior) Area Co-Chair: Machine Translation
- NAACL 2016: (Senior) Area Co-Chair: Multilinguality and Morphology
- EACL 2014: (Senior) Area Co-Chair: Machine Translation
- ACL 2014: Tutorials Co-Chair

Journal Reviewer

- Computational Linguistics (CL)
- Transactions of the Association for Computational Linguistics (TACL)
- Journal of Artificial Intelligence Research (JAIR)
- Artificial Intelligence (AI)

Conference/Workshop Reviewer

- Regular reviewer at ACL, NAACL, EACL, EMNLP
- Regular reviewer at Conference on Machine Translation
- Regular reviewer at ACL/NAACL/EACL Workshops on Machine Translation
- NIPS 2010-2011
- COLING 2008, COLING 2010
- MT Summit 2009
- IJCNLP 2008
- Theoretical and Methodological Issues in Machine Translation (TMI) 2004
- Selected Workshop reviews: ACL Workshop on Building and Using Parallel Texts; ACL Workshop on Syntax, Semantics and Structure in Statistical Translation; EACL/ACL Workshops on Computational Approaches to Semitic Languages; Workshop on Hybrid Approaches to Machine Translation; IWSLT International Workshop on Spoken Language Translation; COLING Workshop on Human Judgements in Computational Linguistics; Analyzing and interpreting neural networks for NLP; MT Summit IX Workshop on Machine Translation for Semitic Languages

Other Reviewing

- European Research Council - Proposal Review
- German Research Foundation (DFG) - Proposal Reviewer
- U.S. National Science Foundation - Proposal Reviewer
- Neural Information Processing Series, "Learning Machine Translation", MIT Press
- Grace Hopper Celebration of Women in Computing, Anita Borg Institute

Other Activities

- Designed and managed creation of the first commercially available statistical machine translation system (Language Weaver Arabic/English), available almost two years before Google Translate
- Engineered top ranked systems in cross-language information retrieval, information retrieval, word alignment and machine translation
 - Best English to German machine translation, WMT 2017 (news), constrained track
 - Best German to English machine translation, ACL-WMT 2009 constrained track
 - Word alignments used in 1st or 2nd best ranked Chinese to English and Arabic to English machine translation systems, NIST-MT 2006, 2008, 2009
 - Best Romanian/English word alignment, ACL-WPT 2005 constrained track
 - Best Arabic to English machine translation, NIST-MT 2003 constrained and unconstrained tracks
 - Best English to Arabic cross-language retrieval, TREC 2001, TREC 2002
 - Best Arabic information retrieval, TREC 2001

- Software provided to the research community:
 - Implementation of a Discriminative Lexicon in Moses, with Kristina Toutanova and Ondrej Bojar and team at Open Source track of the Eighth Machine Translation Marathon 2013
 - Integration of Vowpal Wabbit, a high-speed learner of classification models, as a lexicon model into the open source Moses statistical machine translation engine, with team at the Johns Hopkins Summer Workshop on Domain Adaptation for Machine Translation 2012
 - Cube pruning algorithm implemented in open source Moses statistical machine translation engine, with team at Open Source track of Euromatrix Second Machine Translation Marathon 2009
 - Word alignment graphical user interface implemented in Java Swing, used by 7 research labs, 2007
- Data provided to the research community:
 - Standard data sets for domain-adaptation experimentation in machine translation created at the Johns Hopkins Summer Workshop on Domain Adaptation for Machine Translation
 - Probabilistic Dictionary (with Jinxi Xu) for use in TREC-2002 Cross-Lingual Information Retrieval Task (using English queries to request Arabic documents)
 - BBN release of the Arabic/English UN corpus (with Jinxi Xu, Ralph Weischedel), distributed by the Linguistic Data Consortium
 - Sentence aligned the Ummah corpus, distributed by the Linguistic Data Consortium
- Organizing of Stuttgart Reading Group on Parsing and Machine Translation
- Stuttgart Coordinator for Pascal2 EU Network of Excellence (Pascal2: Pattern Analysis, Statistical Modelling and Computational Learning)