

Katrin E. Erk
Linguistics Department
The University of Texas at Austin

Professional appointments

University of Texas at Austin

Department of Linguistics

Professor

from Sep 2018

Associate professor

Sep 2012 - Aug 2018

Assistant professor

Aug 2006 - Aug 2012

Saarland University

2002-Aug 2006

Research scientist, Institute for Computational Linguistics

Education

Saarland University, Germany

2002

Dr.-Ing (PhD in Engineering)

Universität Koblenz-Landau, Germany

1998

Diplom-Informatiker (*Diploma in Computer Science*),

minor: Computational Linguistics

mit Auszeichnung (with distinction)

University of Georgia

1994-1995

Institute for Artificial Intelligence

Publications

Books

Erk, K. (2002). *Parallelism constraints in underspecified semantics*. Saarbrücken Dissertations in Computational Linguistics and Language Technology, vol. 14.

Erk, K., and L. Priese (2000). *Theoretische Informatik. Eine umfassende Einführung*. First edition 2000, second revised edition 2002, third revised edition 2008, Springer-Verlag, Berlin.

Peer-reviewed Journal Articles

Erk, K. (2016). What do you know about an alligator when you know the company it keeps? *Semantics and Pragmatics* 9, 1-63.

Beltagy, I., S. Roller, P. Cheng, K. Erk, and R. Mooney (2016). Representing meaning with a combination of logical form and vectors. *Computational Linguistics* 42:4, 763808. Special issue on formal distributional semantics.

McCarthy, D., M. Apidianaki, and K. Erk (2016). Word sense clustering and clusterability. *Computational Linguistics* 42(2), 245275.

Erk, K., D. McCarthy, and N. Gaylord (2013). Measuring word meaning in context. *Computational Linguistics* 39(3), 511-554.

Moon, T., and K. Erk (2013). An inference-based model of word meaning in context as a paraphrase distribution. *ACM Transactions on Intelligent Systems and Technology* 4(3), special issue on paraphrasing.

Erk, K. (2012). Vector space models of word meaning and phrase meaning: a survey. *Language and Linguistics Compass* 6(10), 635-653.

Erk, K., S. Padó, and U. Padó (2010). A flexible, corpus-driven model of regular and inverse selectional preferences. *Computational Linguistics* 36(4), 723-763.

Palmer, A., T. Moon, J. Baldridge, K. Erk, E. Campbell, and T. Can (2010). Computational strategies for reducing annotation effort in language documentation: A case study in creating interlinear texts for Uspanteko. *Linguistic Issues in Language Technologies* 3(4).

Čulo, O., K. Erk, S. Padó and S. Schulte im Walde (2008). Comparing and combining semantic verb classifications. *Language Resources and Evaluation* 42(3):265-291.

Erk, K., and J. Niehren (2007). Dominance constraints in stratified context unification. *Information Processing Letters* 101(4), 141-147.

Erk, K., A. Koller, and J. Niehren (2003). Processing underspecified semantic representations in the Constraint Language for Lambda Structures. *Journal of Research on Language and Computation* 1/2003, 127-169.

Peer-reviewed Book and Encyclopedia Chapters

Erk, K. (2018). Computational semantics. In M. Aronoff (ed.), *Oxford Research Encyclopedia of Linguistics* New York: Oxford University Press. (36 pages).

Garrette, D., K. Erk, and R. Mooney (2013). A formal approach to linking logical form and vector-space lexical semantics. In *Computing Meaning*, Vol. 4, H. Bunt, J. Bos and S. Pulman (Eds.), Springer, Berlin, pp. 27-48.

Burchardt, A., K. Erk, A. Frank, A. Kowalski, S. Padó, and M. Pinkal (2009). Using FrameNet for the semantic analysis of German: annotation, representation, and automation. In Boas, H. (ed.): *Multilingual FrameNets in Computational Lexicography: Methods and Applications*. Mouton de Guyter.

Papers in Peer-Reviewed Conference Proceedings (Including Accepted)

Wang, S., E. Holgate, G. Durrett, and K. Erk (2018). *Picking Apart Story Salads*. Proceedings of EMNLP. (11 pages)

Cheng, P., and K. Erk (2018). *Implicit argument prediction with event knowledge*. Proceedings of NAACL. (11 pages)

Rosenfeld, A., and K. Erk (2018). *Deep Neural Models of Semantic Shift*. Proceedings of NAACL. (11 pages)

Wang, S., G. Durrett, and K. Erk (2018). *Modeling semantic plausibility by injecting world knowledge*. Proceedings of NAACL. (6 pages)

Wang, S., S. Roller, and K. Erk (2017). *Distributional modeling on a diet: one-shot learning from text only*. Proceedings of IJCNLP. (10 pages)

Roller, S. and K. Erk (2016). *Relations such as Hypernymy: Identifying and Exploiting Hearst Patterns in Distributional Vectors for Lexical Entailment*. Proceedings of EMNLP.

(11 pages)

Roller, S. and K. Erk (2016). *PIC a different word: a simple model for lexical substitution in context*. Proceedings of NAACL. (4 pages)

Beltagy I., and K. Erk (2015). *On the Proper Treatment of Quantifiers in Probabilistic Logic Semantics*. Proceedings of IWCS. (11 pages)

Roller, S., K. Erk, and G. Boleda (2014). *Inclusive yet Selective: Supervised Distributional Hypernymy Detection*. Proceedings of CoLing. (12 pages)

Beltagy, I., K. Erk and R. Mooney (2014). *Probabilistic Soft Logic for Semantic Textual Similarity*. Proceedings of ACL. (10 pages)

Kremer, G., K. Erk, S. Padó, and S. Thater (2014). *What substitutes tell us - analysis of an "all-words" lexical substitution corpus*. Proceedings of the Conference of the European Chapter of the Association for Computational Linguistics. (10 pages)

Beltagy, I., Chau Kim C., G. Boleda, D. Garrette, K. Erk, and R. Mooney (2013). *Montague meets Markov: Deep semantics with probabilistic logical form*. Proceedings of STARSEM. (11 pages)

Erk, K. (2013) *Towards a semantics for distributional representations*. Proceedings of the 10th International Conference on Computational Semantics (IWCS), Potsdam, Germany. (11 pages.)

Garrette, D., K. Erk, and R. Mooney (2011). *Integrating logical representations with probabilistic information using Markov Logic*. Proceedings of the 9th International Conference on Computational Semantics (IWCS), Oxford, UK. (10 pages.)

Moon, T., K. Erk, and J. Baldridge (2010). *Crouching Dirichlet, Hidden Markov Model: Unsupervised POS tagging with context local tag generation*. Proceedings of the Conference on Empirical Methods in Natural Language Processing (EMNLP), MIT, MA. (11 pages.)

Erk, K., and S. Padó (2010). *Exemplar-based models for word meaning in context*. Proceedings of the 48th Annual Meeting of the Association for Computational Linguistics (ACL), Uppsala. (5 pages.)

Ponvert, E., K. Erk, and J. Baldridge. *Simple unsupervised identification of low-level constituents*. Proceedings of the Fourth IEEE International Conference on Semantic Computing. Pittsburg, PA, 2010. (8 pages.)

Erk, K., and D. McCarthy (2009). *Graded word sense assignment*. Proceedings of the Conference on Empirical Methods in Natural Language Processing (EMNLP), Singapore. (10 pages.)

Moon, T., K. Erk, and J. Baldridge (2009). *Unsupervised morphological segmentation and clustering with document boundaries*. Proceedings of the Conference on Empirical Methods in Natural Language Processing (EMNLP), Singapore. (10 pages.)

Erk, K., D. McCarthy, and N. Gaylord (2009). *Investigations on word senses and word usages*. Proceedings of the Joint conference of the 47th Annual Meeting of the Association for Computational Linguistics and the 4th International Joint Conference on Natural Language Processing of the Asian Federation of Natural Language Processing (ACL-IJCNLP), Singapore. (9 pages)

Erk, K. (2009). *Representing words as regions in vector space*. Proceedings of the Thirteenth Conference on Computational Natural Language Learning (CoNLL), Boulder, CO. (9 pages)

Erk, K. (2009). *Supporting inferences in semantic space: representing words as regions*. Proceedings of the International Conference on Computational Semantics (IWCS), Tilburg, The Netherlands. (12 pages)

- Erk, K., and S. Padó (2008). *A structured vector space model for word meaning in context*. Proceedings of the Conference on Empirical Methods in Natural Language Processing (EMNLP), Waikiki. (10 pages)
- Erk, K. (2007). *A simple, similarity-based model for selectional preferences*. Proceedings of the 45th Annual Meeting of the Association for Computational Linguistics (ACL), Prague, Czech Republic. (8 pages)
- Padó, S., U. Padó and K. Erk (2007). *Flexible, corpus-based modelling of human plausibility judgements*. Proceedings of the Conference on Empirical Methods in Natural Language Processing (EMNLP), Prague, Czech Republic. (8 pages)
- Erk, K. (2006). *Unknown word sense detection as outlier detection*. Proceedings of the Human Language Technology conference - North American chapter of the Association for Computational Linguistics annual meeting (HLT-NAACL), New York City. (8 pages)
- Erk, K., and S. Padó (2005). *Analysing models for semantic role assignment using confusability*. Proceedings of the Human Language Technology Conference and Conference on Empirical Methods in Natural Language Processing (HLT/EMNLP-05), Vancouver, Canada. (8 pages)
- Frank, A., and K. Erk (2004). *Towards an LFG syntax-semantics interface for frame semantics annotation*. Fifth International Conference on Intelligent Text Processing and Computational Linguistics (CICLing), Seoul. (12 pages)
- Erk, K., A. Kowalski, S. Padó and M. Pinkal (2003). *Towards a resource for lexical semantics: A large German corpus with extensive semantic annotation*. Proceedings of the 41st Annual Meeting of the Association for Computational Linguistics (ACL), Sapporo. (8 pages)
- Erk, K., and J. Niehren (2003). *Well-nested parallelism constraints for ellipsis resolution*. Proceedings of the 10th Conference of the European Chapter of the Association for Computational Linguistics (EACL), Budapest. (8 pages)
- Erk, K. and A. Koller (2001). *VP ellipsis by tree surgery*. Proceedings of the 13th Amsterdam Colloquium, Amsterdam. (8 pages)
- Egg, M., and K. Erk (2001). *A compositional account of VP ellipsis*. Proceedings of the Conference on HPSG, Trondheim. (6 pages)
- Bodirsky, M., K. Erk, A. Koller, and J. Niehren (2001). *Underspecified beta reduction*. Proceedings of the 39th Annual Meeting of the Association for Computational Linguistics (ACL), Toulouse. (8 pages)
- Bodirsky, M., K. Erk, A. Koller, and J. Niehren (2001). *Beta reduction constraints*. Proceedings of the 12th International Conference on Rewriting Techniques and Applications (RTA), Utrecht. (8 pages)
- Erk, K., and J. Niehren (2000). *Parallelism constraints*. Proceedings of the 11th International Conference on Rewriting Techniques and Applications (RTA), Norwich. (8 pages)

Papers in Peer-Reviewed Workshop Proceedings

- Ferracane, E., I. Marshall, B. Wallace, and K. Erk (2016). *Leveraging coreference to identify arms in medical abstracts: An experimental study*. Seventh International Workshop on Health Text Mining and Information Analysis (LOUHI) at EMNLP 2016, Austin, TX. (11 pages)
- Boleda, G., and K. Erk (2015). *Distributional Semantic Features as Semantic Primitives or Not*. Knowledge Representation and Reasoning: Integrating Symbolic and Neural Ap-

- proaches. Papers from the AAAI Spring Symposium. Palo Alto, California: AAAI Press. (4 pages)
- Beltagy, I., K. Erk and R. Mooney (2014). *Semantic Parsing using Distributional Semantics and Probabilistic Logic*. Extended abstract. Proceedings of ACL 2014 Workshop on Semantic Parsing (SP-2014). (5 pages)
- Speriosu, M., T. Brown, T. Moon, J. Baldridge, and K. Erk (2010). *Connecting language and geography with region-topic models*. Proceedings of the 1st Workshop on Computational Models of Spatial Language Interpretation. Portland, Oregon. (8 pages.)
- Herdağdelen, A., M. Baroni, and K. Erk (2009). *Measuring semantic relatedness with vector space models and random walks*. Proceedings of the ACL-IJCNLP Workshop on Graph-based Methods for Natural Language Processing (TextGraphs), Singapore. (4 pages)
- Erk, K., and S. Padó (2009). *Paraphrase assessment in structured vector space: Exploring parameters and datasets*. Proceedings of the EACL Workshop on Geometrical Methods for Natural Language Semantics (GEMS). (9 pages)
- Baldridge, J., and K. Erk (2008). *Teaching computational linguistics to a large, diverse student body: courses, tools, and interdepartmental interaction*. Proceedings of the Third Workshop on Issues in Teaching Computational Linguistics (TeachCL). Columbus, OH. (9 pages)
- Moon, T., and K. Erk (2008). *Minimally supervised lemmatization scheme induction through bilingual parallel corpora*. Proceedings of the International Conference on Global Interoperability for Language Resources. Hong Kong. (8 pages)
- Erk, K., and S. Padó (2007). *Towards a computational model of gradience in word sense*. Proceedings of the International Workshop on Computational Semantics (IWCS), Tilburg, The Netherlands. (4 pages)
- Palmer, A., and K. Erk (2007). *IGT-XML: an XML format for interlinearized glossed texts*. Proceedings of the Linguistics Annotation Workshop (LAW-07). Prague, Czech Republic, 2007. (8 pages)
- Erk, K., and S. Padó (2006). *Shalmaneser – a toolchain for shallow semantic parsing*. Proceedings of the Fifth International Conference on Language Resources and Evaluation (LREC 2006), Genoa. (6 pages)
- Burchardt, A., K. Erk, A. Frank, A. Kowalski, S. Padó, and M. Pinkal (2006). *The SALSA corpus: a German corpus resource for lexical semantics*. Proceedings of the Fifth International Conference on Language Resources and Evaluation (LREC 2006), Genoa. (6 pages)
- Burchardt, A., K. Erk, A. Frank, A. Kowalski, and S. Padó (2006). *SALTO – a versatile multi-level annotation tool*. Proceedings of the Fifth International Conference on Language Resources and Evaluation (LREC 2006), Genoa. (6 pages)
- Erk, K. (2005). *Frame assignment as word sense disambiguation*. Proceedings of the International Workshop on Computational Semantics (IWCS), Tilburg, The Netherlands. (3 pages)
- Schulte im Walde, S., and K. Erk (2005). *A comparison of German semantic verb classifications*. Proceedings of the International Workshop on Computational Semantics (IWCS), Tilburg, The Netherlands. (12 pages)
- Padó, S., and K. Erk (2005). *To cause or not to cause: Cross-lingual semantic matching for paraphrase modelling*. Proceedings of the Cross-Language Knowledge Induction Workshop, Cluj-Napoca, Romania. (8 pages)
- Burchardt, A., K. Erk, and A. Frank (2005). *A WordNet detour to FrameNet*. Proceedings

of the GermaNet II workshop, Bonn. (15 pages)

Erk, K., and S. Padó (2004). *A powerful and versatile XML format for representing role-semantic annotation*. Proceedings of the Fourth International Conference on Language Resources and Evaluation (LREC), Lisbon. (4 pages)

Heid, U., H. Voormann, J. Milde, U. Gut, K. Erk, and S. Padó (2004). *Querying both time-aligned and hierarchical corpora with NXT search*. Proceedings of the Fourth International Conference on Language Resources and Evaluation (LREC), Lisbon. (6 pages)

Ellsworth, M., K. Erk, P. Kingsbury, and S. Padó (2004). *PropBank, SALSA, and FrameNet: How design determines product*. Proceedings of the Workshop on Building Lexical Resources from Semantically Annotated Corpora, Lisbon.

Erk, K., A. Kowalski, and S. Padó (2003). *The SALSA Annotation Tool*. Proceedings of the Sixth Lorraine-Saarland Workshop, Nancy. (4 pages)

Erk, K., and G. M. Kruijff (2002). *A constraint-programming approach to parsing with resource-sensitive Categorical Grammar*. Proceedings of the Seventh International Workshop on Natural Language Understanding and Logic Programming (NLULP), Copenhagen.

Non-Peer-Reviewed Papers

Erk, K. (2014). *Who evoked that frame? Some thoughts on context effects and event types*. Invited abstract, Proceedings of Frame Semantics in NLP: A Workshop in Honor of Chuck Fillmore (1929-2014).

Beltagy, I., S. Roller, G. Boleda, K. Erk, and R. Mooney (2014). *UTexas: Natural Language Semantics using Distributional Semantics and Probabilistic Logic*. Proceedings of SemEval. (6 pages)

Erk, K. (2010). *What is word meaning, really? (And how can distributional models help us describe it?)*. Proceedings of the workshop on Geometrical Models of Natural Language Semantics (GEMS). Uppsala, Sweden, 2010. (10 pages.)

Baker, C., M. Ellsworth, and K. Erk (2007). *SemEval-2007 Task 19: frame semantic structure extraction*. Proceedings of the Fourth International Workshop on Semantic Evaluations (SemEval-2007), Prague, Czech Republic. (5 pages)

Baldewein, U., K. Erk, S. Padó, and D. Prescher (2004). *Semantic role labeling with similarity-based generalization using EM-based clustering*. Proceedings of the Third International Workshop on the Evaluation of Systems for the Semantic Analysis of Text (Senseval-3), Barcelona. (5 pages)

Baldewein, U., K. Erk, S. Padó, and D. Prescher (2004). *Semantic role labeling for chunk sequences*. Proceedings of the Conference on Computational Natural Language Learning (CoNLL), Shared Task Session, Boston. (4 pages)

Grants

Defense Advanced Research Projects Agency, Active Interpretation of Disparate Alternatives (AIDA) program PI. Title: *Narrative Based Hypothesis Generation Using Event Coherence and Probabilistic Inference (TA3)*. \$815,054. 2018-2022

- National Science Foundation** PI (NSF 1523637). Title: *RI: Small: Deep Natural Language Understanding with Probabilistic Logic and Distributional Similarity*. \$408,287 2015-2018
- Defense Advanced Research Projects Agency, Deep Exploration and Filtering of Text (DEFT) Program** co-PI. Title: *Statistical Relational Learning and Script Induction for Textual Inference*. PI: Dr. R. Mooney. \$1,302,318 2012 - 2017
- Alexander von Humboldt Foundation, TransCoop program** PI. Title: *Linguistically Adequate Empirical Representation of Word Meaning*. Also PI: Dr. Sebastian Padó. EUR 28,575 2011-2014
- Morris Memorial Grant, New York Community Trust** co-PI. Title: *Spatial and Temporal Analysis of Multilingual Texts*. PI: Dr. J. Baldrige, co-PI: Dr. D. Beaver. \$120,000 2010-2012
- National Science Foundation CAREER program** PI (NSF IIS 0845925). Title: *CAREER: Word meaning: beyond dictionary senses*. \$433,449 2009-2014
- Morris Memorial Grant, New York Community Trust** co-PI. Title: *Research on Multilingual Text Interpretation*. PI: Asst. Prof. J. Baldrige, co-PI: Assoc. Prof. D. Beaver. \$120,000 2008-2010
- Documenting endangered languages, BCS, National Science Foundation** co-PI (NSF BCS 0651988). Title: *Reducing annotation effort in the documentation of languages using machine learning and active learning*. PI: Asst. Prof. J. Baldrige. \$79,106 2007-2008
- Liberal Arts Instructive Technology grant, College of Liberal Arts, UT Austin** PI. Title: *Automatic semantic analysis software for classroom experimentation*. \$42,740 2007-2008

Scholarships and awards

- CSLI Fellow, Stanford**. Visiting scholar at the Center for the Study of Language and Information, Stanford, May 8-26 2017
- College Research Fellowship, UT Austin**. Award for one semester research leave Fall 2013
- Big XII Faculty Fellowship, Dean of Graduate Studies, UT Austin**. Award for research travel to the University of Colorado. \$1000 2009
- Dean's Fellowship, Dean of Graduate Studies, UT Austin**. Award for one semester research leave Fall 2008
- Summer Research Assignment, Dean of Graduate Studies, UT Austin** (two months summer salary) for conducting research on word sense. The University of Texas at Austin Summer 2007
- Doctoral dissertation scholarship of DFG (German Research Foundation)**. Graduiertenkolleg Kognitionswissenschaft (post-graduate program in cognitive science). 1998-2001

Scholarship of Studienstiftung des deutschen Volkes (German national academic foundation) 1995-1998

Scholarship of DAAD (German academic exchange service) 1994-1995

Invited talks

“Implicit argument prediction with event knowledge.” Google Assistant & Dialog Faculty Workshop. Mountainview, CA, May 2018.

“Learning word meaning from a single appearance in context.” Abschlusskolloquium (final colloquium) SFB 732 “Incremental Specification in Context”, Universität Stuttgart. Stuttgart, Germany, March 2018.

“What do you know about an alligator when you know the company it keeps?” STARSEM Joint Conference on Lexical and Computational Semantics. Vancouver, Canada. July 2017.

“In-depth natural language understanding with probabilistic semantics and distributional models.” Computer Science Department, Technische Universität Darmstadt. Darmstadt, Germany. July 2017.

“One-shot learning of word meaning with distributional models”, Conference on logic and machine learning, Gothenburg, June 2017.

“Distributional model on a diet: Experiments in distributional one-shot learning”, Google, May 2017.

“One-shot learning of word meaning with distributional models”, Stanford Linguistics department colloquium, Stanford University, May 2017.

“The probabilistic samovar: an attempt at explaining how people can learn from distributional data”, workshop “Distributional Semantics and Linguistic Theory” (DSALT), Bolzano, Italy, August 2016.

“More precise fuzziness, more fuzzy precision”, COMPOSES end-of-project workshop, Bolzano, Italy, August 2016.

“In-depth text understanding with formal and distributional semantics: A to-do list”, workshop “Representation Learning for Natural Language Processing” (RepL4NLP), Berlin, Germany, August 2016.

“Representing meaning with a combination of logical and distributional models”, Allen Institute for Artificial Intelligence, Seattle, WA, June 2016.

“Deep natural language understanding with probabilistic logic and distributional similarity”, Wright State University, Dayton, Ohio, March 2016.

“Are distributional dimensions semantic features?” Meaning in Context Symposium. Munich, Germany, September 2015.

“Semantics as a heterogeneous mess (and how to reason over it)”. Centre for Linguistic Theory and Studies in Probability. Gothenburg, Sweden, August 2015.

“What do you know about an alligator when you know the company it keeps?”. Workshop “Advances in Distributional Semantics”, London, UK, April 2015.

“Who evoked that frame? Some thoughts on context effects and event types”, at “Frame Semantics in Natural Language Processing: A Workshop in Honor of Chuck Fillmore”, Baltimore, MD, June 2014.

“You shall know a logical form by the company it keeps”, Rutgers University, March 2014.

“Inference with messy data: a story of distributional semantics and model theory”, Dagstuhl Seminar “Computational Models of Language Meaning in Context”, Dagstuhl, Germany, November 2013.

“You shall know a logical form by the company it keeps”, Trento University, Italy, November 2013.

“Combining logic-based and distributional representations for inferences over text”, workshop “Towards a Formal Distributional Semantics”. Potsdam, Germany, March 2013.

“How geometric should our semantic models be?”, Johns Hopkins University, September 2012.

“What is word meaning, really? (And how can distributional models help us describe it?)”, workshop on Geometrical Models of Natural Language Semantics (GEMS). Uppsala, Sweden, July 2010.

“Semantic space models for word meaning in context”, University of Kentucky, Lexington KY, November 2009.

“Investigating graded meaning representations”, University of Colorado at Boulder, Institute of Cognitive Science colloquium, Boulder, CO, October 2009.

“Semantic space models for word meaning in context”, University of Colorado at Boulder, Computer Science department colloquium, Boulder, CO, October 2009.

“Word meaning and vector space representations”, University of North Texas, Denton, TX, October 2008.

“Word meaning and vector space representations”, Georgetown University, Washington, DC, September 2008.

“Vector space representations and word meaning”, Saarbrücken University, Germany, August 2008.

“Limited compositionality phenomena in SALSA”, Unified Linguistics Annotation Workshop, University of Colorado at Boulder, March 2008.

“The SALSA experience: semantic role annotation”, Workshop on Treebanks at the Annual Conference of the North American Chapter of the Association for Computational Linguistics (NAACL-HLT 2007), April 2007.

“Experiments in wide-coverage word sense and semantic role analysis”, Columbia University, New York, March 2007.

“SALSA: Annotating a German corpus with Frame-semantic information”, Nancy University, France, September 2005.

“SALSA: The Saarbrücken Lexical Semantics Annotation & Analysis Project”, Joint SIGSEM/ISO meeting at the Sixth International Workshop on Computational Semantics, January 2005.

Invited talks at the University of Texas at Austin

“What do you know about an alligator when you know the company it keeps?”. Linguistics Department colloquium, February 2015.

“Representing gradience in word meaning”, Texas Linguistics Society conference, June 2012.

“Polysemy and word sense annotation”, Cognition and Perception seminar series, April 2008.

“Diving into the data: a lexical semantics annotation project”, Texas Linguistics Society

conference, November 2007.

“Automatic meaning analysis of free text: small steps towards a big goal”, Forum for Artificial Intelligence, December 2006.

“Detecting outliers: useful for word sense assignment – and for aiding manual annotation?”, Texas Linguistics Society conference, November 2006.

Teaching

Graduate courses at the University of Texas at Austin

Computational Linguistics I/Introduction to Computational Linguistics

Course covering core algorithms and data structures in computational linguistics.

Computational Linguistics II

Course covering the main statistical Natural Language Processing techniques.

Working with Corpora

Course covering programming basics, text search techniques and corpus annotation methodology for students with a liberal arts background.

Analyzing Linguistic Data

Course providing an introduction to statistical analysis and to analyses of linguistic data with the R statistics package

Computational Semantics

Course covering logic-based approaches to representing and computing with sentence meaning, and distributional approaches to modeling similarity at the word and phrase level

Research in Computational Linguistics

Course for graduate students doing research in computational linguistics, covers fundamental methods in the field (including Bayesian inference, graphical models, and deep learning), prominent recent research in the field, research done by students at UT, and techniques for research and writing.

Graduate seminars

Graduate seminars on computational lexical semantics and on word meaning research across the disciplines

Undergraduate courses at the University of Texas at Austin

Introduction to Computational Linguistics

This course covers core problems and core algorithms in computational linguistics.

Analyzing Linguistic Data

This course provides an introduction to statistical analysis and to analyses of linguistic data, focusing on real-world data sets and using the R statistics package. The class involves student research projects.

Computational semantics

This course discusses both context-based approaches to modeling word meaning, and logic-based approaches to representing and computing with sentence meaning. The class involves student research projects.

Language and Computers

This course introduces core computational linguistics methods from the point of view of

language technology applications. It discusses important algorithms in a pen-and-paper fashion, without programming.

Courses at Saarland University, Germany

Seminars

Graduate seminars on computational lexical semantics.

Computational linguistics algorithms in Prolog

Undergraduate course covering core computational linguistics algorithms and their implementation in Prolog

Courses at Universität Koblenz-Landau, Germany

Theoretical Computer Science

Undergraduate course on automata theory and formal languages

Advising

Ph.D. committee chair:

Pengxiang Cheng

Alex Rosenfeld

Stephen Roller graduated 2017

Islam Beltagy (with R. Mooney) graduated 2016

Nicholas Gaylord (with C. Bannard) graduated 2013

Taesun Moon graduated 2011

Alexis Palmer (with J. Baldridge) graduated 2009

Graduate advisees not yet advanced to candidacy:

Elisa Ferracane, NSF Graduate Research Fellowship Program (GRFP) Fellowship 2017

Eric Holgate

Su Wang

Ph.D. committee member:

Justin Cope

Guy Emerson (Cambridge University, UK)

Aice Lai (University of Illinois)

Chris Brown graduated 2018

Nazneen Rajani (computer science) graduated 2018

Karl Pichotta (computer science) graduated 2017

Patrick Schultz (English) graduated 2017

Jacqueline Larsen Serigos (Spanish and Portuguese) graduated 2017

Axel Bohmann (English) graduated 2016

Zachary Childers graduated 2016

Jason Mielens graduated 2016

Benjamin Wing graduated 2015

Ian Gerg (Music) graduated 2015

Leah Velleman graduated 2014

Nicholas Bacuez (French) graduated 2013
Anja Möhring (German) graduated 2013
Michael Speriosu graduated 2013
Joseph Reisinger (computer science) graduated 2012
David Chen (computer science) graduated 2011
Michael Glass (computer science) graduated 2011
Elias Ponvert graduated 2011
Susan Windisch Brown (University of Colorado at Boulder) graduated 2010
Josh Iorio graduated 2009
Makiko Nakayama graduated 2008
Joey Frazee
Thomas Laborde (German)

Master's thesis reader:

Grant DeLozier (second reader) 2016
Vidhoon Visvanathan (computer science; second reader) 2015
Simon Hafner (second reader) 2014
James Evans (second reader) 2014
Stephen Hilderbrand (first reader) 2011
Weiwei Ding (second reader) 2011
Joshua Brewster (second reader) 2010
Robert Elwell (second reader) 2008
Inge de Bleecker (first reader) 2007

Undergraduate honors thesis advisor:

Urmilla Banerjee (Turing Scholars, computer science)
Meghana Chappidi (Dean's Scholar honors program, computer science)

Emily Tagtow (Linguistics honors thesis) 2015; Rapoport-King Scholarship for 2014/2015
Winston Wu (Turing Scholars) 2015
Jessika Roesner (Linguistics honors thesis) 2010; Dean's Distinguished Graduate in the College of Liberal Arts; NSF Graduate Research Fellowship Program (GRFP) Fellowship 2011, declined
Andrew Young (Linguistics honors thesis) 2009
William Brands

Second reader, undergraduate honors theses:

Kaitlin Adams (Linguistics honors thesis) 2016
Tyler Yates (Turing Scholars, computer science) 2016
Anthony Munoz (Linguistics honors thesis) 2011
Trevor Fountain (computer science honors thesis) 2008

Other graduate advising:

Dan Garrette 2009-2010; National Defense Science & Engineering Fellowship 2010

Other undergraduate advising:

Kelsey Ball 2015-2016
Jeane Wenzel 2012-2014
Travis Nguyen 2012-2014

Kristen Marshello 2012-2013

Aidan Coyne 2011-2013

Kelly Smith 2011

Service

Program co-chair , Annual Conference of the Association for Computational Linguistics (ACL)	2016
Program co-chair , European Summer School for Logic, Language and Information (ESSLLI), area Language and Computation	2016
Action editor , Transactions of the Association for Computational Linguistics (TACL) journal	2015-2019
Secretary , SIGSEM	2010-present
Executive board , North American Chapter of the Association for Computational Linguistics (NAACL)	2013-2014
Program chair , International Conference on Computational Semantics (IWCS), Potsdam, Germany	2013
Editorial board , Computational linguistics	2010-2012
Information officer , SIGSEM, the Special Interest Group on Semantics of the Association for Computational Linguistics	2007-2010
Member of the executive board , SIGLEX, the Special Interest Group on the Lexicon of the Association for Computational Linguistics	2004-2013
Co-organizer , workshop “SemEval-2 Evaluation Exercises on Semantic Evaluation” in Uppsala, Sweden	2010
Co-organizer , workshop ”Multilingual semantic annotation: theory and applications” in Saarbrücken, Germany	2006
Co-organizer , ”Interdisciplinary Workshop on the Identification and Representation of Verb Features and Verb Classes” in Saarbrücken, Germany	2005
Co-organizer , workshop “Building Lexical Resources from Semantically Annotated Corpora” at the 4th International Conference on Language Resources and Evaluation	2004
Journal reviewing: Computational Linguistics, Transactions of the Association of Computational Linguistics, Journal of Artificial Intelligence Research, Natural Language Engineering, Cognitive Science, Language Resources and Evaluation, Linguistic Issues in Language Technology, Transactions on Speech and Language Processing, Linguistics, Semantics and Pragmatics, Computational Linguistics special issue on Semantic Role Labeling, Journal of Logic and Computation special issue	

Grant reviewing: National Science Foundation; Research Foundation Flanders; European Research Council; the British Academy; Fonds de recherche du Québec; US-Israel Binational Science Foundation

Conference reviewing: ACL, EACL, NAACL, EMNLP, IWCS, ICSC, TLS, HJCL, CogSci, CoNLL

Outreach

Presentation at First Bytes summer camp, a one-week residential camp program for high school girls, June 2017.

“When will my computer understand me?”, Presentation at the Blue Knot Austin See the Supercomputer event, December 2014.