

## Academic background

### University of Edinburgh

Institute for Language, Cognition and Computation  
School of Informatics  
Edinburgh, Scotland, UK

September 2014–Present

Postdoctoral researcher with [Mark Steedman](#).

*Research topic:* computational modeling of grammar and meaning

### Carnegie Mellon University

Language Technologies Institute  
School of Computer Science  
Pittsburgh, PA

August 2008–August 2014

Ph.D. work in statistical natural language processing, advised by [Noah Smith](#).

*Research topics:* machine learning for NLP, semantics, named entity recognition, corpus annotation, social media text, nonnative writing, Arabic NLP, text visualization

*Dissertation:* [Lexical Semantic Analysis in Natural Language Text](#). (Defended June 19, 2014; submitted September 30, 2014.)

### Information Sciences Institute,

University of Southern California  
Los Angeles, CA

June 2012–August 2012

Summer research intern with [Kevin Knight](#).

*Research topic:* semantics for machine translation

### University of California, Berkeley

Berkeley, CA

August 2004–May 2008

B.A. with High Honors; Double Major in Computer Science (with Honors), Linguistics

*Honors:* Phi Beta Kappa; Computer Science Faculty Commendation

*Research topics:* construction grammar formalisms, Hebrew morphology, semantics, metonymy recognition, statistical news analysis

*Thesis:* “Hebrew Verb Morphology and Embodied Construction Grammar”

## Industry experience

### Microsoft

Redmond, WA

May 2007–August 2007

*Role:* Software Development Intern

In the Office Sharepoint Search group, worked on designing and implementing a new feature for the next version of Office

### Convergys

Cincinnati, OH

June 2006–August 2006

*Role:* Software Intern

Involved in various aspects of the planning, development, and testing phases of customizations of a Java web application

*Role:* Software Intern  
Web design work, including use of Macromedia Flash

## Teaching & mentorship

### **Foundations of Natural Language Processing (INFR09028)**

Spring 2016

University of Edinburgh

I expect to co-teach (along with Sharon Goldwater) this course for third-year undergraduates.

### **Edinburgh MSc thesis projects supervised in 2015**

- Marco Damonte, “Machine Translation with Coarse Lexical Semantics” (co-supervised with Alexandra Birch)
- Nora Hollenstein, “Inconsistency Detection in Semantic Annotation” (co-supervised with Bonnie Webber)
- Ye Yang, “Recognizing Annotator Behavior in Crowdsourcing” (co-supervised with Bonnie Webber)
- Felisia Loukou, “Light Verb Constructions in Distributional Entailment Graphs” (co-supervised with Mark Steedman and Omri Abend)

### **Committee member**

- Liane Guillou (Edinburgh Informatics Ph.D. student 3rd year review, 2015-01-22)
- Federico Fancellu (Edinburgh Informatics Ph.D. student 1st year review, 2015-09-30)

### **Natural Language Processing (11-411)**

Spring 2011

Carnegie Mellon University

*Role:* Teaching assistant  
*Lectures:* Lexical Semantics; Semantic Role Labeling and Relational Semantics

### **Advanced NLP Seminar (11-713)**

Spring 2010

Carnegie Mellon University

*Role:* Teaching assistant

### **Neural Basis of Thought and Language (CS C182)**

Spring 2008

University of California, Berkeley

*Role:* Course reader  
*Lecture:* A Symbolic Formalism for Embodied Grammar

### **Structure and Interpretation of Computer Programs (CS 61A)**

Spring 2006

University of California, Berkeley

*Role:* Lab assistant

## Tutorials

### **The Logic of AMR: Practical, Unified, Graph-Based Sentence Semantics for NLP**

2015-05-31

NAACL-HLT conference, Denver, CO

*Co-presenters:* Jeffrey Flanigan, Tim O’Gorman  
*Length:* 3 hours

*Co-presenters:* Collin Baker, Miriam R. L. Petruck, Michael Ellsworth  
*Length:* 3 hours

## Presentations

- **Dealing With Off-the-Charts Polysemy in English Prepositions: Representation, Annotation, and Disambiguation.** Saarland University FEAST series (2015-10-20).
- [Struggling with English prepositional verbs](#). Talk at ICLC 2015, Newcastle, England.
- **Hacking a Way Through the Twitter Language Jungle: Syntactic Annotation, Tagging, and Parsing of English Tweets.** Keynote at [Workshop on Natural Language Processing for Informal Text](#), Rotterdam (2015-06-23).
- **Semantic Abstraction at Scale: Opportunities and Challenges for NLP with the FrameNet and AMR Representations.**
  - Invited talk at the [Swedish FrameNet++ Workshop](#), University of Gothenburg Språkbanken Group (2014-11-19).
  - Google Mountain View (2015-02-06).
  - Bielefeld University (2015-04-28).
- **Bridging the Gap: Integrated Development of Linguistic Resources and Analyzers for NLP.**
  - Johns Hopkins University CLSP (2014-04-02).
  - University of Toronto Computational Linguistics Meeting (2014-08-08).
  - University of Edinburgh ILCC Seminar (2014-10-24).
  - Bielefeld University (2015-04-27).
  - University of Copenhagen Centre for Language Technology (2015-04-30).
- Broad-coverage analysis of English MWEs: from annotation to sequence tagging. Poster at PARSEME 4th General Meeting (March 2015), Valleta, Malta.
- [Lexical semantics with WordNet + Python](#). Python Linguistics Group, University of Pittsburgh (2013-10-21).
- [Discriminative Lexical Semantic Segmentation with Gaps: Running the MWE Gamut](#).
  - Natural Language Processing course, CMU (2014-04-24).
  - Berkeley NLP Group (2013-09-04).
  - LTI Student Research Symposium (2013-08-21). **Best Presentation Award**
- [Exploiting and expanding corpus resources for frame-semantic parsing](#). Presented at International FrameNet Workshop 2013.
- [English Understanding: From Annotation to AMRs](#). ISI NLP Group (2012-08-28).
- [Casting a wider 'Net: NLP for the Social Web](#). CMU Qatar Computer Science, Doha (2011-10-05).
- Introduction to Computational Linguistics and Language Technologies.
  - Introduction to Linguistics course, University of Pittsburgh (2010-04-15).
  - Introduction to Arabic Linguistics course, University of Pittsburgh (2012-04-11).
- [English morphology in construction grammar](#). Poster at CSDL 2010, San Diego.
- Probabilistic frame-semantic parsing. ICSI, Berkeley (2010-02-05).

## Service

Co-organizer: SemEval 2016 shared task on [Detecting Minimal Semantic Units and their Meanings](#) (with Dirk Hovy, Anders Johannsen, and Marine Carpuat)

Co-proposer: Theme session on Computational Aspects of Frames and Constructions for the International Conference on Construction Grammar 2016 (with Miriam R. L. Petruck)

Mentor: NAACL Student Research Workshop 2015

Reviewer for the journals *Computational Linguistics*, *Transactions of the ACL*, and *Language Resources and Evaluation*; and conferences/workshops including ACL, NAACL, EACL, EMNLP, COLING, CoNLL, ICML, and AAAI.

Area chair: Language Resources and Evaluation, ACL 2016 (with Nianwen Xue)

Invited panelist at the Linguistic Annotation Workshop, June 6, 2015 in Denver (topic: syntactic annotation of non-canonical language).

Invited facilitator at the FrameNet Workshop, September 9–13, 2013 in Berkeley.

*At Edinburgh:*

- Organizer: ILCC Research Tutorial and Hackathon (2015-05-08)
- Organizer: Arguments and Adjuncts Reading Group (Fall 2014)

*At CMU:*

- Organizer: Linguistics Reading Group
- Co-organizer: CL+NLP Lunch Seminar, [LTI Colloquium](#) Committee, Student Activities Committee
- Volunteer: North American Computational Linguistics Olympiad and International Linguistics Olympiad

[Abstract Meaning Representation \(AMR\)](#) design team, Summer 2012–present

## Publications

### Refereed Journal Articles

- Nathan Schneider, Emily Danchik, Chris Dyer, and Noah A. Smith (2014). [Discriminative lexical semantic segmentation with gaps: running the MWE gamut](#). *Transactions of the Association for Computational Linguistics*. Presentation at ACL 2014.
- Dipanjan Das, Desai Chen, André F. T. Martins, Nathan Schneider, and Noah A. Smith (2014). [Frame-semantic parsing](#). *Computational Linguistics* 40(1).

### Reviews & position papers

- Nathan Schneider (2015). [What I've learned about annotating informal text \(and why you shouldn't take my word for it\)](#). *Proceedings of the Linguistic Annotation Workshop*.
- Nathan Schneider and Reut Tsarfaty (2013). [Design Patterns in Fluid Construction Grammar](#), Luc Steels (editor). *Computational Linguistics* 39(2).

### Refereed Conference Papers

- Nathan Schneider and Noah A. Smith (2015). [A corpus and model integrating multiword expressions and supersenses](#). *Proceedings of NAACL-HLT*.

- Meghana Kshirsagar, Sam Thomson, Nathan Schneider, Jaime Carbonell, Noah A. Smith, and Chris Dyer (2015). [Frame-semantic role labeling with heterogeneous annotations](#). *Proceedings of ACL-IJCNLP*.
- Lizhen Qu, Gabriela Ferraro, Liyuan Zhou, Weiwei Hou, Nathan Schneider, and Timothy Baldwin (2015). Big data small data, in domain out-of domain, known word unknown word: the impact of word representations on sequence labelling tasks. *Proceedings of CoNLL*.
- Lingpeng Kong, Nathan Schneider, Swabha Swayamdipta, Archana Bhatia, Chris Dyer, and Noah A. Smith (2014). [Dependency parsing for tweets](#). *Proceedings of EMNLP*.
- Archana Bhatia, Chu-Cheng Lin, Nathan Schneider, Yulia Tsvetkov, Fatima Talib Al-Raisi, Laleh Roostapour, Jordan Bender, Abhimanu Kumar, Lori Levin, Mandy Simons, and Chris Dyer (2014). [Automatic classification of communicative functions of definiteness](#). *Proceedings of COLING*.
- Nathan Schneider, Spencer Onuffer, Nora Kazour, Emily Danchik, Michael T. Mordowanec, Henrietta Conrad, and Noah A. Smith (2014). [Comprehensive annotation of multiword expressions in a social web corpus](#). *Proceedings of LREC*.
- Yulia Tsvetkov, Nathan Schneider, Dirk Hovy, Archana Bhatia, Manaal Faruqui, and Chris Dyer (2014). [Augmenting English adjective senses with supersenses](#). *Proceedings of LREC*.
- Nathan Schneider, Behrang Mohit, Chris Dyer, Kemal Oflazer, and Noah A. Smith (2013). [Supersense tagging for Arabic: the MT-in-the-middle attack](#). *Proceedings of NAACL-HLT*.
- Olutobi Owoputi, Brendan O'Connor, Chris Dyer, Kevin Gimpel, Nathan Schneider, and Noah A. Smith (2013). [Improved part-of-speech tagging for online conversational text with word clusters](#). *Proceedings of NAACL-HLT*.
- Nathan Schneider, Behrang Mohit, Kemal Oflazer, and Noah A. Smith (2012). [Coarse lexical semantic annotation with supersenses: an Arabic case study](#). *Proceedings of ACL*.
- Behrang Mohit, Nathan Schneider, Rishav Bhowmick, Kemal Oflazer, and Noah A. Smith (2012). [Recall-oriented learning of named entities in Arabic Wikipedia](#). *Proceedings of EACL*.
- Kevin Gimpel, Nathan Schneider, Brendan O'Connor, Dipanjan Das, Daniel Mills, Jacob Eisenstein, Michael Heilman, Dani Yogatama, Jeffrey Flanigan, and Noah A. Smith (2011). [Part-of-speech tagging for Twitter: annotation, features, and experiments](#). *Proceedings of ACL-HLT*.
- Dipanjan Das, Nathan Schneider, Desai Chen, and Noah A. Smith (2010). [Probabilistic frame-semantic parsing](#). *Proceedings of NAACL-HLT*.
- Nathan Schneider (2010). [Computational cognitive morphosemantics: modeling morphological compositionality in Hebrew verbs with Embodied Construction Grammar](#). *Proceedings of BLS*.

## Refereed Workshop Papers

- Nathan Schneider, Vivek Srikumar, Jena D. Hwang, and Martha Palmer (2015). [A hierarchy with, of, and for preposition supersenses](#). *Proceedings of Linguistic Annotation Workshop*.
- Sam Thomson, Brendan O'Connor, Jeffrey Flanigan, David Bamman, Jesse Dodge, Swabha Swayamdipta, Nathan Schneider, Chris Dyer, and Noah A. Smith (2014). [CMU: Arc-Factored, Discriminative Semantic Dependency Parsing](#). *Proceedings of SemEval* (Task 8: Broad-Coverage Semantic Dependency Parsing).
- Nathan Schneider, Brendan O'Connor, Naomi Saphra, David Bamman, Manaal Faruqui, Noah A. Smith, Chris Dyer, and Jason Baldridge (2013). [A framework for \(under\)specifying dependency syntax without overloading annotators](#). *Proceedings of Linguistic Annotation Workshop*.
- Laura Banarescu, Claire Bonial, Shu Cai, Madalina Georgescu, Kira Griffitt, Ulf Hermjakob, Kevin Knight, Philipp Koehn, Martha Palmer, and Nathan Schneider (2013). [Abstract Meaning Representation for sembanking](#). *Proceedings of Linguistic Annotation Workshop*.

- Yulia Tsvetkov, Naama Twitto, Nathan Schneider, Noam Ordan, Manaal Faruqui, Victor Chahuneau, Shuly Wintner, and Chris Dyer (2013). [Identifying the L1 of non-native writers: the CMU-Haifa system](#). *Proceedings of the 8th Workshop on Innovative Use of NLP for Building Educational Applications*.
- Desai Chen, Nathan Schneider, Dipanjan Das, and Noah A. Smith (2010). [SEMAFOR: Frame argument resolution with log-linear models](#). *Proceedings of SemEval* (Task 10: Linking Events and their Participants in Discourse).

## Reports

- Nathan Schneider, Rebecca Hwa, Philip Gianfortoni, Dipanjan Das, Michael Heilman, Alan W. Black, Frederick L. Crabbe, and Noah A. Smith (July 2010). [Visualizing topical quotations over time to understand news discourse](#). Technical Report CMU-LTI-10-013.
- Dipanjan Das, Nathan Schneider, Desai Chen, and Noah A. Smith (April 2010). [SEMAFOR 1.0: A probabilistic frame-semantic parser](#). Technical Report CMU-LTI-10-001.
- Reza Bosagh Zadeh and Nathan Schneider (December 2008). [Unsupervised approaches to sequence tagging, morphology induction, and lexical resource acquisition](#). LS2 course literature review.

## Technical and personal skills

- Programming languages used on a regular basis: Python (text processing and statistical modeling) and HTML/CSS/Javascript/PHP/MySQL (web development and user interfaces for annotation); also fluent in Java
- Open source software projects: in addition to sharing code from published research, contributions to [NLTK](#) (including the [SemCor](#) and [FrameNet](#) modules) and [Zotero](#)
- Languages: English (native); studied conversational Hebrew, Arabic, French, and German
- Music: classical violin