

NATHAN SCHNEIDER

www.nathan.cl
nschneid@cs.cmu.edu

Ph.D. student in computational linguistics/natural language processing

Education

Carnegie Mellon University

August 2008–Present

Language Technologies Institute, School of Computer Science
Pittsburgh, PA

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

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

University of California, Berkeley

August 2004–May 2008

Berkeley, CA

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

May 2007–August 2007

Redmond, WA

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

June 2006–August 2006

Cincinnati, OH

Role: Software Intern

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

CAD Centric Systems

June 2003–July 2003

Cincinnati, OH

Role: Software Intern

Web design work, including use of Macromedia Flash

Teaching experience

Natural Language Processing (11-411)

Spring 2011

UC Berkeley

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

Invited presentations

- [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 (Ling 1000), University of Pittsburgh (2010-04-15).
- Probabilistic frame-semantic parsing. ICSI, Berkeley (2010-02-05).

Service

Reviewer for the journal *Computational Linguistics* and conferences/workshops including ACL, EMNLP, and AACL.

Volunteer for the North American Computational Linguistics Olympiad and International Linguistics Olympiad.

Within LTI: organizer of Linguistics Reading Group; co-organizer of CL+NLP Lunch Seminar.

Publications

Refereed

- Behrang Mohit, Nathan Schneider, Rishav Bhowmick, Kemal Oflazer, and Noah A. Smith (2012, to appear). [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*.
- Desai Chen, Nathan Schneider, Dipanjan Das, and Noah A. Smith (2010). [SEMAFOR: Frame argument resolution with log-linear models](#). *Proceedings of SemEval*.
- Dipanjan Das, Nathan Schneider, Desai Chen, and Noah A. Smith (2010). [Probabilistic frame-semantic parsing](#). *Proceedings of NAACL-HLT*.
- Nathan Schneider (2010). [English morphology in construction grammar](#). Poster at CSDL.
- Nathan Schneider (2010). [Computational cognitive morphosemantics: modeling morphological compositionality in Hebrew verbs with Embodied Construction Grammar](#). *Proceedings of BLS*.

Reports

- Behrang Mohit, Nathan Schneider, Rishav Bhowmick, Kemal Oflazer, and Noah A. Smith (August 2011). [Recall-oriented learning for named entity recognition in Wikipedia](#). Technical Report CMU-LTI-11-012.
- 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); Java (statistical modeling); HTML/CSS/Javascript/PHP/MySQL (web development)
- Open source software projects: small contributions to [NLTK](#) and [Zotero](#)
- Languages: English (native); studied Hebrew, Arabic, French, and German
- Hobbies: playing violin, following American politics and Constitutional law, reading