Jason M. Eisner
Curriculum Vitae

<table>
<tr>
<td>CONTACT</td>
<td>Department of Computer Science, NEB 224
Johns Hopkins University
3400 N. Charles St.
Baltimore, MD 21218-2691
U.S.A.</td>
<td>Phone: (410) 516-8438
Fax: (410) 516-6134

jason@cs.jhu.edu
http://cs.jhu.edu/~jason</td>
</tr>
</table>

EDUCATION AND HONORS

Ph.D. in Computer Science   2001
University of Pennsylvania
    Graduate Teaching Award
    Thesis: *Smoothing a Probabilistic Lexicon via Syntactic Transformations*
    Advisor: Mitch Marcus

B.A./M.A. in Mathematics   1993
University of Cambridge
    First-class honours
    Note: Second undergraduate degree.

A.B. in Psychology, Cognitive Science track   1990
Harvard University
    Summa cum laude
    Junior-year election to Phi Beta Kappa
    GPA: 3.9/4.0; GRE: 800/800/800
    Thesis: *Dynamical-Systems Behavior in Recurrent and Non-Recurrent Connectionist Nets*

PROFESSIONAL EXPERIENCE

Johns Hopkins University   July 2000–
Assistant Professor of Computer Science
Joint appointment in Cognitive Science (Jan. 2003–)

University of Rochester   Jan. 2000–June 2001
Assistant Professor of Computer Science
Secondary appointment in Linguistics

iReactor Inc., Philadelphia, PA   1994–
Consultant

AT&T Bell Labs, Murray Hill, NJ   summers 1989–1992
Artificial Intelligence Research Department

Microsoft Corporation, Seattle, WA   summer 1988
Programmer

IBM Research Center, Yorktown Heights, NY   1987–1988
Consultant

PROFESSIONAL ACTIVITIES	President, ACL SIG for Computational Morphology and Phonology (2001–); executive committee (1998–)
	Editorial board member, *Research in Language and Computation* (2006–2011).
	Planning committee, Computational Linguistics Olympiad (2006–).
	Advisory board member, cogito.org (2005–).
	Graduate program chair, CS Department, Johns Hopkins University (2002–).
	Program area chair, EMNLP (2006, machine learning).
	Program area co-chair, HLT-NAACL (2006, syntax/grammar/morphology).
	Program area chair, EACL (2006, phonology/morphology/finite-state/tagging/ segmentation).
	Publications chair, ACL (2005). Co-chair for HLT/NAACL Tutorial and Workshop Programs (2003).
	Special issue editor, *Cognitive Science* (2002).
	Organizer, SIGPHON Workshop on Finite-State Phonology at COLING (2000).
	Journal reviewer, *Computational Linguistics* (2005, 2004, 2002, 2001, 2000), *ACM Transactions on Speech and Language Processing* (2005), *Cognition* (2002), *Language and Speech* (1999), *J. of Algorithms* (1997).
	Program committee for ACL (2005, 2004, 2003, 2002), EACL (2006, 2003, 1999), EMNLP (2006, 2003, 2002), FSMNLP (2005, 2001), ICML (2004), MITWPL (1999), NAACL (2006), SIGPHON (2006, 2004, 2002, 2000, 1998), TNLP (2005, 2002), Formal Approaches to Slavic Linguistics 8 (1999); reviewer for IJCAI (2007).
	NSF panel reviewer (2004, 2003).
	Lecturer, NAACL Summer School in Human Language Technology (2002–2006).
	Member, EU/NSF joint working group: "ePhilology: Emerging Language Technologies and the Rediscovery of the Past" (2002).

FELLOWSHIPS AND AWARDS	Robert B. Pond, Sr. Excellence in Teaching Award Johns Hopkins University, Whiting School of Engineering	2005
	Nominated for best paper award (EMNLP, ACL)	2002, 2005
	NSF Graduate Research Fellowship (computer science)	1993–1996
	Herchel Smith Harvard Scholarship (mathematics)	1991–1993
	Fulbright Scholarship (creative writing), South Africa	1990–1991
	Harvard National Scholarship	1986–1990

GRANTS	JHU WSE-APL Partnership Fund: Learning with Less (PI, $68K)	2006–2007
	NSF PIRE: Investigation of Meaning Representations in Language Understanding for Machine Translation Systems (co-PI, $2.5M)	2005–2010
	NSF CAREER: Finite-State Machine Learning on Strings and Sequences (PI, $500K)	2004–2009
	NSF ITR: Weighted Dynamic Programming for Statistical Natural Language Processing (PI, $425K)	2003–2007
	ONR: Improving Statistical Translation Models Via Text Analyzers Trained From Parallel Corpora (co-PI, $4.3M)	2001–2006
	NSF ITR/IM+PE+SY: Summer Workshops on Human Language Technology (co-PI, $2.35M)	2001–2006

PUBLICATIONS AND PRESENTATIONS

Invited talks

Bootstrapping without the boot. MITRE Corporation, August 2006; IPAM Document Space Workshop, January 2006.

The Dyna language. MIT, November 2006; IBM Yorktown Heights, May 2006; Microsoft Research, August 2005; University of Washington, August 2005.

Parameterized finite-state machines and their training. U. of Saarland, Germany, March 2004; AT&T Research, October 2002.

Inferring transformations. Mathematics of Language Conference (MoL8), Bloomington, June 2003.

Learning natural-language grammars using a Bayesian prior. Rochester Institute of Technology, May 2000; Johns Hopkins University, February 2000; UCLA Linguistics Department, June 1999; Stanford, 1999; U. of Texas at Austin, 1999; U. of Toronto, 1999; U. of Rochester, 1999; U. of Chicago, 1999.

Doing OT in a straitjacket. Johns Hopkins Cognitive Science Dept., 2002; U. of Rochester Linguistics Dept., 2000; UCLA Linguistics Dept., 1999; Stanford Linguistics Dept., 1999; MIT Linguistics Dept., 1997.

Journal Articles

John Eng and Jason M. Eisner. Radiology report entry with automatic phrase completion driven by language modeling. *Radiographics*, 24(5):1493–1501, 2004.

Jason Eisner. Discovering syntactic deep structure via Bayesian statistics. *Cognitive Science*, 26(3):255–268, 2002.

Jason Eisner. Introduction to the special section on linguistically apt statistical methods. *Cognitive Science*, 26(3):235–237, 2002.

Book Chapters

Jason Eisner. Bilexical grammars and their cubic-time parsing algorithms. In Harry Bunt and Anton Nijholt, editors, *Advances in Probabilistic and Other Parsing Technologies*, pages 29–62. Kluwer Academic Publishers, 2000.

Book Reviews

Jason Eisner. Review of *Optimality Theory* by René Kager. *Computational Linguistics*, 26(2):286–290, 2000.

Ph.D. Thesis

Jason Eisner. *Smoothing a Probabilistic Lexicon via Syntactic Transformations*. PhD thesis, University of Pennsylvania, 2001. 318 pages.

Refereed Conference and Workshop Proceedings

Joshua Mason, Kathryn Watkins, Jason Eisner, and Adam Stubblefield. A natural-language approach to automated cryptanalysis of two-time pads. In *Proceedings of the ACM Conference on Computer and Communications Security (ACM CCS)*, 2006. 10 pages.

Jason Eisner and John Blatz. Program transformations for optimization of parsing algorithms and other weighted logic programs. In Paola Monachesi, Gerald Penn, Giorgio Satta, and Shuly Wintner, editors, *Pre-proceedings of the 11th Conference on Formal Grammar (FG-2006)*, pages 39–59, 2006. Revised version to appear in the post-proceedings from CSLI Publications.

Markus Dreyer and Jason Eisner. Better informed training of latent syntactic features. In *Proceedings of the Conference on Empirical Methods in Natural Language Processing (EMNLP)*, pages 317–326, 2006.

David A. Smith and Jason Eisner. Minimum-risk annealing for training log-linear models. In *Proceedings of the International Conference on Computational Linguistics and the Association for Computational Linguistics (COLING-ACL), Companion Volume*, pages 787–794, 2006.

Noah A. Smith and Jason Eisner. Annealing structural bias in multilingual weighted grammar induction. In *Proceedings of the International Conference on Computational Linguistics and the Association for Computational Linguistics (COLING-ACL)*, pages 569–576, 2006.

Jason Eisner and Roy W. Tromble. Local search with very large-scale neighborhoods for optimal permutations in machine translation. In *Proceedings of the HLT-NAACL Workshop on Computationally Hard Problems and Joint Inference in Speech and Language Processing*, pages 57–75, 2006.

David A. Smith and Jason Eisner. Quasi-synchronous grammars: Alignment by soft projection of syntactic dependencies. In *Proceedings of the HLT-NAACL Workshop on Statistical Machine Translation*, pages 23–30, 2006.

Roy W. Tromble and Jason Eisner. A fast finite-state relaxation method for enforcing global constraints on sequence decoding. In *Proceedings of the Human Language Technology*

Conference of the North American Association for Computational Linguistics (HLT-NAACL), pages 423–430, 2006.

Jason Eisner and Noah A. Smith. Parsing with soft and hard constraints on dependency length. In *Proceedings of the International Workshop on Parsing Technologies (IWPT)*, pages 30–41. Association for Computational Linguistics, 2005.

Jason Eisner and Damianos Karakos. Bootstrapping without the boot. In *Proceedings of Human Language Technology Conference and Conference on Empirical Methods in Natural Language Processing (HLT-EMNLP)*, pages 395–402. Association for Computational Linguistics, 2005.

Jason Eisner, Eric Goldlust, and Noah A. Smith. Compiling comp ling: Weighted dynamic programming and the Dyna language. In *Proceedings of Human Language Technology Conference and Conference on Empirical Methods in Natural Language Processing (HLT-EMNLP)*, pages 281–290. Association for Computational Linguistics, 2005.

Noah A. Smith and Jason Eisner. Guiding unsupervised grammar induction using contrastive estimation. In *International Joint Conference on Artificial Intelligence (IJCAI) Workshop on Grammatical Inference Applications*, pages 73–82, 2005.

Noah A. Smith and Jason Eisner. Contrastive estimation: Training log-linear models on unlabeled data. In *Proceedings of the 43rd Annual Meeting of the Association for Computational Linguistics (ACL)*, pages 354–362, 2005.

André Kempe, Jean-Marc Champarnaud, Jason Eisner, Franck Guingne, and Florent Nicart. A class of rational n-WFSM auto-intersections. In *Proceedings of the Tenth International Conference on Implementation and Application of Automata (CIAA-2005)*, number 3845 in Lecture Notes in Computer Science, pages 189–200. Springer-Verlag, 2005.

André Kempe, Jean-Marc Champarnaud, and Jason Eisner. A note on join and auto-intersection of n-ary rational relations. In Loek Cleophas and Bruce Watson, editors, *Proceedings of the Eindhoven FASTAR Days (Computer Science Technical Report 04-40)*, pages 64–78. Department of Mathematics and Computer Science, Technische Universiteit Eindhoven, Netherlands, 2004.

Jason Eisner, Eric Goldlust, and Noah A. Smith. Dyna: A declarative language for implementing dynamic programs. In *Proceedings of the 42nd Annual Meeting of the Association for Computational Linguistics (ACL), Companion Volume*, pages 218–221, 2004.

Noah A. Smith and Jason Eisner. Annealing techniques for unsupervised statistical language learning. In *Proceedings of the 42nd Annual Meeting of the Association for Computational Linguistics (ACL)*, pages 486–493, 2004.

Jason Eisner. Learning non-isomorphic tree mappings for machine translation. In *Proceedings of the 41st Annual Meeting of the Association for Computational Linguistics (ACL), Companion Volume*, pages 205–208, 2003.

Jason Eisner. Simpler and more general minimization for weighted finite-state automata. In *Proceedings of the Joint Meeting of the Human Language Technology Conference and the North American Chapter of the Association for Computational Linguistics (HLT-NAACL)*, pages 64–71, 2003.

Jason Eisner. Parameter estimation for probabilistic finite-state transducers. In *Proceedings of the 40th Annual Meeting of the Association for Computational Linguistics (ACL)*, pages 1–8, 2002.

Jason Eisner. Comprehension and compilation in Optimality Theory. In *Proceedings of the 40th Annual Meeting of the Association for Computational Linguistics (ACL)*, pages 56–63, 2002.

Jason Eisner. An interactive spreadsheet for teaching the forward-backward algorithm. In Dragomir Radev and Chris Brew, editors, *Proceedings of the ACL Workshop on Effective Tools and Methodologies for Teaching NLP and CL*, pages 10–18, 2002.

Jason Eisner. Transformational priors over grammars. In *Proceedings of the Conference on Empirical Methods in Natural Language Processing (EMNLP)*, pages 63–70, 2002.

Jason Eisner. Expectation semirings: Flexible EM for finite-state transducers. In Gertjan van Noord, editor, *Proceedings of the ESSLLI Workshop on Finite-State Methods in Natural Language Processing (FSMNLP)*, 2001. Extended abstract (5 pages).

Jason Eisner. Easy and hard constraint ranking in Optimality Theory: Algorithms and complexity. In Jason Eisner, Lauri Karttunen, and Alain Thériault, editors, *Finite-State Phonology: Proceedings of the 5th Workshop of the ACL Special Interest Group in Computational Phonology (SIGPHON)*, pages 22–33, 2000.

Jason Eisner. Directional constraint evaluation in Optimality Theory. In *Proceedings of the 18th International Conference on Computational Linguistics (COLING 2000)*, pages 257–263, 2000.

Jason Eisner and Giorgio Satta. A faster parsing algorithm for lexicalized tree-adjoining grammars. In *Proceedings of the 5th Workshop on Tree-Adjoining Grammars and Related Formalisms (TAG+5)*, pages 14–19, 2000.

Jason Eisner and Giorgio Satta. Efficient parsing for bilexical context-free grammars and head-automaton grammars. In *Proceedings of the 37th Annual Meeting of the Association for Computational Linguistics (ACL)*, pages 457–464, 1999.

Jason Eisner. FOOTFORM decomposed: Using primitive constraints in OT. In Benjamin Bruening, editor, *Proceedings of SCIL VIII*, number 31 in MIT Working Papers in Linguistics, pages 115–143, 1998.

Jason Eisner. Bilexical grammars and a cubic-time probabilistic parser. In *Proceedings of the 5th International Workshop on Parsing Technologies (IWPT)*, pages 54–65, 1997.

Jason Eisner. Efficient generation in primitive Optimality Theory. In *Proceedings of the 35th Annual Meeting of the Association for Computational Linguistics (ACL)*, pages 313–320, 1997.

Jason Eisner. Three new probabilistic models for dependency parsing: An exploration. In *Proceedings of the 16th International Conference on Computational Linguistics (COLING-96)*, pages 340–345, 1996.

Jason Eisner. Efficient normal-form parsing for combinatory categorial grammar. In *Proceedings of the 34th Annual Meeting of the Association for Computational Linguistics (ACL)*, pages 79–86, 1996.

Breck Baldwin, Jeff Reynar, Mike Collins, Jason Eisner, Adwait Ratnaparkhi, Joseph Rosenzweig, Anoop Sarkar, and Srinivas. Description of the University of Pennsylvania entry in the MUC-6 competition. In *Proceedings of the Sixth Message Understanding Conference*, pages 177–191, 1995.

Jason Eisner. ∀-less in Wonderland? Revisiting *any*. In Janet Fuller, Ho Han, and David Parkinson, editors, *Proceedings of ESCOL 11 (October 1994)*, pages 92–103. DMLL Publications, 1995.

Mark A. Jones and Jason M. Eisner. A probabilistic parser applied to software testing documents. In *Proceedings of National Conference on Artificial Intelligence (AAAI-92)*, pages 322–328, 1992.

Mark A. Jones and Jason M. Eisner. A probabilistic parser and its application. In Carl Weir, editor, *Statistically-Based Natural Language Processing Techniques: Papers from the 1992 Workshop*, pages 20–27. Menlo Park: AAAI Press, 1992. Technical Report W-92-01.

Refereed Presentations

Jason Eisner, Michael Kornbluh, Gordon Woodhull, Raymond Buse, Samuel Huang, Constantinos Michael, and George Shafer. Visual navigation through large directed graphs and hypergraphs. In *Electronic Proceedings of the IEEE Symposium on Information Visualization (InfoVis'06), Poster/Demo Session*, 2006.

Jason Eisner. What constraints should OT allow? Talk handout available online (22 pages), Linguistic Society of America (LSA), Chicago, 1997.

Invited Papers

Damianos Karakos, Sanjeev Khudanpur, Jason Eisner, and Carey E. Priebe. Unsupervised classification via decision trees: An information-theoretic perspective. In *Proceedings of the 2005 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, volume 5, pages 1081–1084, 2005. Invited talk.

Technical Reports

Jia Cui and Jason Eisner. Finite-state Dirichlet allocation: Learned priors on finite-state models. Technical Report 53, Center for Language and Speech Processing, Johns Hopkins University, 2006.

Jan Hajič, Martin Čmejrek, Bonnie Dorr, Yuan Ding, Jason Eisner, Daniel Gildea, Terry Koo, Kristen Parton, Gerald Penn, Dragomir Radev, and Owen Rambow. Natural language generation in the context of machine translation. Technical report, Center for Language and Speech Processing, Johns Hopkins University, 2004. Final report from 2002 CLSP summer workshop (87 pages).

Jason Eisner. State-of-the-art algorithms for minimum spanning trees: A tutorial discussion. Manuscript available online (78 pages), University of Pennsylvania, 1997.

Jason Eisner. An empirical comparison of probability models for dependency grammar. Technical Report IRCS-96-11, Institute for Research in Cognitive Science, Univ. of Pennsylvania, 1996. Available online (18 pages).

Jason Eisner. Indirect STV election: A voting system for South Africa. White paper, University of Cape Town, 1991. Available online (16 pages).

Jason Eisner. Dynamical-systems behavior in recurrent and non-recurrent connectionist nets. Undergraduate honors thesis, Harvard University, 1990. Available online (57 pages).

Outreach (general audience)

Jason Eisner. The science of language: Computational linguistics. *Imagine Magazine*, 7(4):14–15, 2000.

Jason Eisner. Cognitive science and the search for intelligence. Invited paper presented to the Socratic Society, University of Cape Town, South Africa, 1991. Available online (24 pages).

PATENTS

Jeffrey C. Reynar, Fred Herz, Jason Eisner, and Lyle Ungar. A Lempel-Ziv data compression technique utilizing a dictionary pre-filled with frequent letter combinations, words and/or phrases. U.S. Patent #5,951,623 issued 9/14/1999, filed 1996.

Frederick Herz, Lyle Ungar, and Jason M. Eisner. System for the automatic determination of customized prices and promotions. Patent pending, filed 1996.

Frederick S. M. Herz, Jason M. Eisner, and Lyle H. Ungar. System for generation of object profiles for a system for customized electronic identification of desirable objects. U.S. Patent #5,835,087 issued 11/10/1998, filed 1995.

Frederick S. M. Herz, Jason M. Eisner, Lyle H. Ungar, and Mitchell P. Marcus. System for generation of user profiles for a system for customized electronic identification of desirable objects. U.S. Patent #5,754,939 issued 5/19/1998, filed 1995.

Frederick S. M. Herz, Jason Eisner, and Marcos Salganicoff. Pseudonymous server for system for customized electronic identification of desirable objects. U.S. Patent #5,754,938 issued 5/19/1998, filed 1995.

Department of Computer Science, Johns Hopkins University
Robert B. Pond, Sr. Excellence in Teaching Award, Whiting School of Engineering

– *Natural Language Processing*   2001–2004, 2006

A mixed graduate-undergraduate class that teaches a synthesis of statistical models, formal grammars, and linguistic theory, with associated algorithms. It is reputed to be one of the most challenging classes in the Computer Science department, requiring both rigor and intellectual flexibility.

Faculty at several other universities have asked to use the extensive online course materials.

Enrollment: about 30.

– *Declarative Methods*   2005, 2006

A new course for juniors, seniors, and graduate students. It surveys computational problems that tend to pop up frequently in different guises (e.g., constraint satisfaction); the specification languages used to describe instances of these problems; general toolkits for solving these instances; and the algorithms run by these toolkits.

Enrollment: about 35.

– *Seminar in Natural Language Processing*   every semester

A weekly reading and discussion group, exploring important current research in natural language processing and potentially relevant material from related fields. Topics are chosen by the group; each lasts 3–4 weeks.

Attendance: 5–15.

– *Totally Random*   2004, 2005

A 4-class discussion unit about random numbers and the uses of randomness in computer science. Part of the department's new freshman experience course.

Enrollment: 8-10.

– *Data Structures*   2003, 2004

A sophomore-level class, the third and last in the programming sequence for majors. Covers basic data structures and algorithms, basic analysis of algorithms, and object-oriented programming style. Online "warmups" and highly interactive classes stimulated the students to come up with designs and variations. The challenging weekly assignments often used real-world data.

Faculty at a dozen other universities have asked to use the course materials.

Enrollment: about 50.

– *Finite-State Methods in Natural Language Processing*   2000–2001

A graduate class on semiring-weighted finite-state transducers. Covers theory and practice, including the theory of formal power series, the extended regular expression calculus, and a range of applications to natural language. Rigorous assignments.

Attendance: about 20.

- *Statistical Language Learning* 2002

  A graduate class about past and present research that has attempted to induce the structure of language from raw data such as text. Lectures are intermixed with reading and discussion of the primary literature.

  Attendance: about 10.

Lecturer, NAACL Summer School in Human Language Technology   summers 2002–2006

- Lectures on NLP and applied probability.
- Laboratory exercise in competitive grammar writing.

Department of Computer Science, University of Rochester

- *Statistical Learning of Natural Language* 2000
- *Graduate Problem Seminar* 2000

  Boot camp for new Ph.D. students. Students learn research skills by teaming up to tackle a series of open-ended engineering problems that touch on research in the department. (I made them build systems for face orientation detection, distributed calendar management, and information retrieval.) Several written and oral presentations are required and receive extensive feedback. The class also includes career advice, familiarization with departmental resources, presentations by other faculty, and a final research project.

  Enrollment: 10.

Department of Computer Science, University of Pennsylvania
Graduate Teaching Award

- TA in *Introduction to Programming* 1994–1995

Ph.D. research students (as primary advisor only)

Jason Smith	2006–
Omar Zaidan	2005–
John Blatz	2004–
Eric Goldlust *(Wolman Fellow; Muuss Research Award;*	2004–
Hon. Mention for CRA Outstanding Undergraduate Award)	
Markus Dreyer *(Wolman Fellow)*	2003–
David A. Smith *(NSF Fellow)*	2002–
Roy Tromble *(NSDEG Fellow)*	2002–
Noah A. Smith *(Hertz Foundation Fellow; now Asst. Prof., CMU)*	2001–2006

B.S./M.S.E. research students

Asheesh Laroia	2006–
Samuel Huang *(Pistritto Fellow)*	2006–
Constantinos Michael	2005–
George Shafer	2005–2006
John Graettinger *(Outstanding Research Award)*	2005–2006
Michael Kornbluh *(Pistritto Fellow, Outstanding Senior Award)*	2004–2005
Chalaporn Hathaidharm	2002–2004

Ph.D. thesis committees

Gaja Jarosz Snover	2006
Jonathan Allen	2006
Gideon Mann	2006
Charles Schafer	2006
Myroslava Dzikovska (Univ. of Rochester)	2004
Silviu Cucerzan	2003
Radu "Hans" Florian	2002
Jun Wu	2002
Richard Wicentowski	2002
Grace Ngai	2000

October 17, 2006