

RICHARD WILLIAM SPROAT

February 29, 2008

Richard Sproat
Department of Linguistics
Department of Electrical and Computer Engineering
(by courtesy) Department of Computer Science
(by courtesy) Department of Psychology
Beckman Institute
University of Illinois at Urbana-Champaign

Foreign Languages Building 4103,
707 South Mathews Avenue, MC-168
Urbana, IL, 61801
Phone: +1-217-244-4120 (Beckman)
Fax: +1-217-333-3466 (Linguistics)
rws@uiuc.edu

EDUCATION

1981-1985. Massachusetts Institute of Technology, Cambridge, Massachusetts, Department of Linguistics and Philosophy. Ph.D. September 1985. Thesis Title: *On Deriving the Lexicon*. Thesis Supervisor: Kenneth Locke Hale. Minor: Artificial Intelligence.

1977-1981. University of California at San Diego, La Jolla, California. B.A. June 1981. *Summa cum Laude*, with Department Honors in Linguistics with Highest Distinction. Minors: Music, Classical Greek.

EMPLOYMENT

August 2003 —

Professor, Department of Linguistics, Department of Electrical and Computer Engineering, University of Illinois at Urbana-Champaign.

Courtesy (adjunct) appointment in Psychology (January 2007–).

Courtesy (adjunct) appointment in Computer Science (January 2005–).

Fulltime faculty, Beckman Institute.

December 2005 —

Visiting Scientist, Google Labs.

March 1999 — August 2003.

Technology Consultant (AT&T Labs' Distinguished Member of the Technical Staff job title), AT&T Labs – Research,

- Until February 2002: Human–Computer Interaction Research Department
- From February 2002: Information Systems and Analysis Research Department

June 1997 — March 1999.

Distinguished Member of the Technical Staff, Language Modeling Research Department, Bell Laboratories, Lucent Technologies.

January 1996 — June 1997.

Member of the Technical Staff, Language Modeling Research Department, Bell Laboratories, Lucent Technologies.

March 1987 — January 1996.

Member of the Technical Staff, Linguistics Research Department, AT&T Bell Laboratories.

June 1986 — March 1987.

Member of the Technical Staff, Linguistics and Artificial Intelligence Research Department, AT&T Bell Laboratories.

September 1985-May 1986.

Post-doctoral Member of the Technical Staff, Linguistics and Artificial Intelligence Research Department, AT&T Bell Laboratories. Supervisor: Mark Y. Liberman.

June 1984-September 1984.

Member of the Technical Staff, Linguistics and Artificial Intelligence Department, AT&T Bell Laboratories. Supervisor: Kenneth W. Church.

TEACHING EXPERIENCE

Students for whom I am the primary advisor.

- Cecilia Ovesdotter Alm (Linguistics) “Emotions in Text and Speech.” Exp: 2007.
- Hahn Koo (Linguistics) “Learning Non-adjacent Phonotactic Constraints from Brief Experience” . Exp: 2007.
- Ayelet Weiss (Linguistics) Error correction for second-language learners of Hebrew. Exp: 2007.
- Eugene Chung (Linguistics) Semantics of spatial expressions. Exp: 2008.
- Su-youn Yoon (Linguistics); Kyoung-young Kim (Linguistics); Alla Rozovskaya (Linguistics); Boon-pang Lim (ECE); Suma Pramod (ECE).

Courses at UIUC

Spring 2008. LING270, Language, Technology and Society

Spring 2008. LING406, Introduction to Computational Linguistics.

Spring 2007. LING270, Language, Technology and Society

Fall 2006. ECE398RS, Speech Technology.

Fall 2006. LING402, Tools and Techniques for Speech and Language Processing.

Spring 2006. LING406, Introduction to Computational Linguistics.

Spring 2006. LING408, Writing Systems.

Fall 2005. LING506, Topics in Computational Linguistics.

Fall 2005. LING402, Tools and Techniques for Speech and Language Processing.

Spring 2005. LING406, Introduction to Computational Linguistics

Fall, 2004. LING 490, Tools and Techniques for Speech and Language Processing.

Fall, 2004. ECE 598, Speech Synthesis.

Spring, 2004. LING 406, Topics in Computational Linguistics.

Spring, 2004. LING 403C, Writing Systems.

Fall, 2003. LING 306, Introduction to Computational Linguistics.

Named in “Research Memorandum #271: Teachers Ranked as Excellent by their Students, Spring 2004”, *Daily Illini*, September 8, 2004.

Courses and Lectures Prior to Coming to UIUC

August 2002. “Corpus-Based Methods in Chinese Morphology”, COLING 2002, Taipei, Taiwan.

July 2001. “Corpus-based Methods in Chinese Morphology and Phonology”, (with Chilin Shih), 2001 LSA Institute, Santa Barbara, CA.

August 1997. “Finite-state Methods in Morphology, Text Analysis and the Analysis of Writing Systems,” Tutorial given at ROCLING X, Academia Sinica, Taiwan, ROC.

August 1996. “Algorithms for speech recognition and language processing,” Tutorial given at COLING '96, Copenhagen, Denmark (with Mehryar Mohri and Michael Riley).

July 1994. Australian Linguistic Institute, 2-week course in computational morphology.

July 1994. “Text Analysis Tools in Spoken-Language Processing,” Tutorial given at the Association for Computational Linguistics, 32nd Annual Meeting, New Mexico State University (with Michael Riley).

June 1989. “Tutorial on Morphology and Computational Morphology,” Tutorial given at the Association for Computational Linguistics, 27th Annual Meeting, University of British Columbia.

February 1988-May 1988. Visiting Assistant Professor, Program in Linguistics, Rutgers University. Title of course: Linguistic Analysis.

January 1985. Organized and taught a course in LISP to fellow graduate students in linguistics at the Massachusetts Institute of Technology.

September 1984-December 1984. Teaching Assistant for Prof. Moira Yip, Massachusetts Institute of Technology. Title of Course: The Study of Language.

September 1983-December 1983. Teaching Assistant for Prof. Wayne O’Neil, Massachusetts Institute of Technology. Title of Course: The Study of Language.

Other Teaching-Related Experience Prior to Coming to UIUC

1987, 1993, 1994, 1995, 1998, 2000, 2002. Supervision of summer projects for students in Bell Laboratories Summer Research Program, Graduate Research Program for Women, and University Relations programs, and the AT&T Labs Summer Research Programs.

1992 External member of thesis committee for Jill Burstein, CUNY, (thesis title: *The Stress and Syntax of Compound Nominals*).

2002 External member of thesis committee for Scobie Smith, Harvard, (thesis title: *Hurrian Interference in Nuzi Akkadian: Comparative Graphemics*).

2002 External member of thesis committee for Martin Jansche, Ohio State University, (thesis title: *Inference of String Mappings for Speech Technology*)

2003 External Examiner, Swarthmore Honors Program.

OTHER PROFESSIONAL ACTIVITIES

Area Chair—phonology/morphology/finite-state technology/tagging/word segmentation—ACL 2007.

Co-organizer (with Steve Farmer) of *Scripts, Non-scripts and (Pseudo)decipherment*, a workshop held in conjunction with the 2007 LSA Summer Institute, Stanford, CA, July 11, 2007.

<http://catarina.ai.uiuc.edu/2007Workshop/>.

Technical co-chair (with Yuji Matsumoto) of the 21st International Conference on Computer Processing of Oriental Languages, December 17–19, 2006, <http://www.iccpol-06.org/>.

Co-organizer (with Dan Roth) of the Third Midwest Computational Linguistics Colloquium, May 20–21, University of Illinois at Urbana-Champaign, <http://l2r.cs.uiuc.edu/cogcomp/mclc/>.

External member on Parsing Arabic Dialects project, Workshop '05 (July 11 – August 19, 2004), Center for Language and Speech Processing, Johns Hopkins University.

Team Leader for the Chinese Dialectal Speech Recognition project, Workshop '04 (July 6 – August 18, 2004), Center for Language and Speech Processing, Johns Hopkins University.

Guest editor, *Journal of Computational Linguistics and Chinese Language Processing* special issue on “Word Formation and Chinese Language Processing.”

Co-organizer, AT&T Speech Days Symposium, October 14 – 16, 2002, Florham Park, NJ, <http://www.research.att.com/conf/spchday2002/>

Chair for SIGHAN (ACL Special Interest Group on Chinese Language Processing), 2003.

Vice-Chair and Treasurer for SIGHAN (ACL Special Interest Group on Chinese Language Processing), 2002.

Archivist for the Association for Computational Linguistics, 2000 – present.

Member of the W3C Voice Browser Working Group, 1999 – present.

Team Leader for the Text Normalization project, Workshop '99 (July 12 – August 20, 1999), Center for Language and Speech Processing, Johns Hopkins University.

Tutorials Chair, ACL 1999.

Co-organizer (with Mari Ostendorf and Andrew Hunt) of the *NSF Workshop for Discussing Research Priorities and Evaluation Strategies in Speech Synthesis*, Arlington, VA, August 6–7, 1998.

International Co-Chair, ROCLING (Republic of China Computational Linguistics Society), 1997.

Chair, LSA Committee on Computing, 1997.

Editorial Board, *COLIPS*, 1996– .

Editorial Board, *Computational Linguistics and Chinese Language Processing*, 1996– .

Visiting scholar, Max Planck Institute for Psycholinguistics (Nijmegen, the Netherlands), August 1996.

Chair of Program Committee, Second SIGPHON Workshop, June 1996.

Editor, *The Finite String* (newsletter of the Association for Computational Linguistics), 1996–1997.

Executive Committee, Special Interest Group in Phonology (SIGPHON), 1995–.

Associate Editor, *Language*, 1994–1996.

Co-organizer, Second ESCA/IEEE Workshop on Speech Synthesis, September 1994.

Visiting scholar, Max Planck Institute for Psycholinguistics (Nijmegen, the Netherlands), April 1994.

Editorial Board, *Natural Language and Linguistic Theory*, 1993–1995.

Editorial Board, *Computational Linguistics*, 1991–1993.

Program Committee, *Association for Computational Linguistics*, 1993.

Program Committee, *American Association for Artificial Intelligence*, 1992.

Prepublication reviewer (subset):

Journals: *Artificial Intelligence*, *Cognition*, *Computational Linguistics*, *Computer Speech and Language*, *Journal of East Asian Linguistics*, *Language*, *Linguistic Inquiry*, *Natural Language and Linguistic Theory*, *Yearbook of Morphology*, inter alia.

Book manuscripts: University of Chicago Press, MIT Press, Kluwer, inter alia.

Conferences: AAI, GLOW, IJCAI, LabPhon, North East Linguistics Society, West Coast Conference on Formal Linguistics, inter alia.

RESEARCH INTERESTS

Accent Prediction, Articulatory Phonetics, Chinese Linguistics, Computational Morphology, Computational modeling of language change, Corpus-Based Methods, Morphology, Pronunciation Modeling, Speech Data Mining, Syntax, Text Analysis/Normalization, Text-to-Scene Conversion, Text-to-Speech Synthesis, Writing Systems.

For details on some of these, see <http://www.research.att.com/~rws/newindex/research.html>.

PATENTS

1. Awarded Patents

- (a) "E-Mail Signature Block Analysis" (with H. Chen and J. Hu) April 16, 2002, U.S. Pat. Num. 6,373,985.
- (b) "E-Mail Signature Block Segmentation" (with H. Chen and J. Hu), March 19, 2002, U.S. Pat. Num. 6,360,010.
- (c) "A Method and Apparatus for Measuring the Degree of Polysemy in Polysemous Words" (with J. van Santen), July 3, 2001, U.S. Pat. Num. 6,256,629.
- (d) "Compilation of Weighted Finite-State Transducers from Decision Trees," September 8, 1998, U.S. Pat. Num. 5,806,032.
- (e) "Grapheme-to-phoneme conversion of digit strings using weighted finite state transducers to apply grammar to powers of a number basis" (with F. Pereira and M. Riley), July 14, 1998, U.S. Pat. Num. 5,781,884.
- (f) "Method for conversion of phonetic Chinese to character Chinese," December 14, 1993, U.S. Pat. Num. 5,270,927.
- (g) "Text to Scene Conversion" (with B. Coyne). March 21, 2006, U.S. Pat. Num. 7,016,828

2. Pending Patents

- (a) "A System and Method Using Meta-Data for Spoken Dialog Systems" (with Michiel Bacchiani, Sameer Maskey and Brian Roark), U. S. Prov. Application No: 60/515,896.
- (b) "Inferring the Environment in a Text-to-Scene Conversion System", provisional filing 5/2001.
- (c) "A Method for Orthography-Phoneme Alignment", filed 8/2000.

FUNDING

1. Richard Sproat (PI), Mark Hasegawa-Johnson (co-PI), Chilin Shih (co-PI), Kay Bock (co-PI), Brian Ross (co-PI), Fred Davidson (Senior Researcher), "DHB: An Interdisciplinary Study of the Dynamics of Second Language Fluency", NSF-IIS-0623805, February 2007 — January 2010, **\$680,781**.
2. Richard Sproat (PI), "Segmental Awareness and Indic Writing Systems", NIH, 1R03HD048644-01, August 2006–August 2008, **\$108,259**.
3. Richard Sproat (PI), Mark Hasegawa-Johnson (co-PI), Chilin Shih (co-PI), Kay Bock (co-PI), Brian Ross (co-PI), Dan Roth (co-PI) "Automated Methods for Second-Language Fluency Assessment." University of Illinois Critical Research Initiative. August 2005 – July 2007, **\$147,019**; with possibility of extension to a third year.
4. Richard Sproat (PI), "Segmental Awareness and Indic Writing Systems", University of Illinois Research Board, January 2006 – May 2006. **\$6120**.
5. Richard Sproat (PI), "Machine transliteration for underresourced languages", Google, Inc, February 2006 –, **\$150,000**.

6. P. Bryan Heidorn (PI), Richard Sproat (co-PI), "Spatial language and spatial reasoning for understanding geo-reference information in natural history collections." University of Illinois Research Board, August 2005 – May 2006. **\$6120.**
7. Richard Sproat (PI), Elabbas Benmamoun (co-PI), "Language Adaptation for Colloquial Arabic", Supplemental NSF grant from National Security Agency, Grant Code A6206, September 2004 — August 2005, **\$125,237.**
8. Richard Sproat (PI), Dan Roth (co-PI), Elabbas Benmamoun (co-PI), Chengxiang Zhai (co-PI), "Named Entity Recognition and Transliteration for 50 Languages", Department of Interior/National Security Agency, NBCHC040176, 30 September 2004 – 29 September 2006, **\$377,930.**
9. Jan van Santen (PI), Alan Black (co-PI), John-Paul Hosom, Richard Sproat (co-PI). "Prosody Generation for Child Oriented Speech Synthesis." NSF#0205731 \$2.75 Million, September 2002–August 2007.

PUBLICATIONS

Books

1. Richard Sproat. *Language, Technology and Society*. Oxford, Oxford University Press, Forthcoming 2009.
2. Brian Roark, Richard Sproat. *Computational Approaches to Syntax and Morphology*. Oxford, Oxford University Press, 2007.
3. Martin Neef, Anneke Neijt and Richard Sproat (editors). *The Relation of Writing to Spoken Language*. Tübingen: Niemeyer (= Linguistische Arbeiten 460), 2002.
4. Richard Sproat, *A Computational Theory of Writing Systems*, (ACL Studies in Natural Language Processing Series), Cambridge, Cambridge University Press, 2000.
5. Richard Sproat (editor), *Multilingual Text-to-Speech Synthesis*, Boston, MA, Kluwer Academic Publishers, 1997.
6. Jan van Santen, Richard Sproat, Joseph Olive and Julia Hirschberg (editors), *Progress in Speech Synthesis*, New York, Springer, 1997.
7. Richard Sproat, *Morphology and Computation*, Cambridge, MA, MIT Press, 1992.

Refereed Journals

1. Richard Sproat. "Brahmi-Derived Scripts, Script Layout, and Phonological Awareness." *Written Language and Literacy*, 9(1), 45–66, 2006.
2. Steve Farmer, Richard Sproat and Michael Witzel. "The Collapse of the Indus-Script Thesis: The Myth of a Literate Harappan Civilization." *Electronic Journal of Vedic Studies*, 11(2), 2004
3. Michiel Bacchiani, Brian Roark, Michael Riley and Richard Sproat "MAP Adaptation of Stochastic Grammars". *Computer Speech and Language*, 20(1):41-68, 2006.
4. Narasimhan, Bhuvana, Sproat, Richard, and Kiraz, George. "Schwa-deletion in Hindi Text-to-Speech Synthesis," *International Journal of Speech Technology*, 7(4), 319–333, 2004.
5. Richard Sproat, Alan Black, Stanley Chen, Shankar Kumar, Mari Ostendorf, and Christopher Richards. "Normalization of non-standard words." *Computer Speech and Language*, 15(3), 287–333, 2001.
6. Richard Sproat. "Multilingual Text Analysis for Text-to-Speech Synthesis", *Natural Language Engineering*, 2(4), 369–380, 1997.
7. Chilin Shih and Richard Sproat, "Issues in Text-to-Speech Conversion for Mandarin," *Computational Linguistics and Chinese Language Processing*, 1996.
8. Richard Sproat, Chilin Shih, William Gale and Nancy Chang, "A Stochastic Finite-State Word-Segmentation Algorithm for Chinese," *Computational Linguistics*, 22(3), 1996.

9. Harald Baayen and Richard Sproat, "Estimating Lexical Priors for Low-Frequency Morphologically Ambiguous Forms," *Computational Linguistics*, 22(2), 1996.
10. Richard Sproat and Chilin Shih. "A Corpus-Based Analysis of Mandarin Nominal Root Compounds." *Journal of East Asian Linguistics*, 5, 49–71, 1996.
11. Richard Sproat and Joseph Olive, "Text to Speech Synthesis," *AT&T Technical Journal*, 74(2), 35–44, 1995.
12. Richard Sproat, "English Noun-Phrase Accent Prediction for Text-to-Speech." *Computer Speech and Language*, 8, 79–94, 1994.
13. Richard Sproat and Chilin Shih, "Why Mandarin Morphology is not Stratum-Ordered." *Yearbook of Morphology*, 185–217, 1993.
14. Richard Sproat and Osamu Fujimura, "Allophonic variation in English /l/ and its implications for phonetic implementation," *Journal of Phonetics*, 21, 291–311, 1993.
15. Gail McKoon, Gregory Ward, Roger Ratcliff and Richard Sproat, "Morphosyntactic and pragmatic factors affecting the accessibility of discourse entities," *Journal of Memory and Language*, 32, 1–20, 1993.
16. David Roe, Pedro Moreno, Richard Sproat, Fernando Pereira, Michael Riley and Alejandro Macarrón. "A Spoken Language Translator for Restricted-Domain Context-Free Languages," *Speech Communication*, 11, 1–9, 1992.
17. Richard Sproat, "Unhappier is not a Bracketing Paradox," *Linguistic Inquiry*, 23, 347–352, 1992.
18. Gregory Ward, Richard Sproat and Gail McKoon, "A Pragmatic Analysis of So-Called Anaphoric Islands," *Language*, 67, 439–474, 1991.
19. Richard Sproat and Chilin Shih, "A Statistical Method for Finding Word Boundaries in Chinese Text," *Computer Processing of Chinese and Oriental Languages*, 4, 336–351, 1990.
20. Richard Sproat, "Welsh Syntax and VSO Structure," *Natural Language and Linguistic Theory*, 3, 173–216, 1985.
21. Richard Sproat, "Southern California Reflexives: An Example of Translation Borrowing?" *Journal of California and Great Basin Anthropology: Papers in Linguistics*, 3, 77–94, 1981.

Refereed Collections

1. Sproat, Richard. "The Consistency of the Orthographically Relevant Level in Dutch." in Martin Neef, Anneke Neijt and Richard Sproat (editors). *The Relation of Writing to Spoken Language*. Tübingen: Niemeyer (= Linguistische Arbeiten 460), 2002.
2. Jianying Hu, Richard Sproat and Hao Chen. "Electronic Mail", in *McGraw-Hill Yearbook of Science and Technology, 2001*, 144–146, McGraw-Hill, New York, 2000.
3. R. Harald Baayen, Robert Schreuder and Richard Sproat. "Morphology in the Mental Lexicon: A Computational Model for Visual Word Recognition", in F. Van Eynde and D. Gibbon (eds.) *Lexicon Development for Speech and Language Processing*, 267–293, Dordrecht, Kluwer Academic Publishers, 2000.
4. Richard Sproat, Christer Samuelsson, Jennifer Chu-Carroll and Bob Carpenter. "Computational Linguistics", in Mark Aronoff and Janie Rees-Miller (eds.) *The Blackwell Handbook of Linguistics*, Oxford, Blackwell, 2000.
5. Richard Sproat, "Lexical Analysis", in Robert Dale, Hermann Moisl, and Harold Somers (eds.) *Handbook of Natural Language Processing*, 37–58, New York, Dekker, 2000.
6. Richard Sproat and Joseph Olive, "Text-to-Speech Synthesis," in L. Rabiner and R. Cox (eds.), *IEEE/CRC Press Handbook of Signal Processing*, Cleveland, CRC Press, 1997.
7. Richard Sproat, "Text Interpretation for TtS Synthesis", in Ron Cole (ed.) *Survey of the State of the Art in Human Language Technology*, <http://www.cse.ogi.edu/CSLU/HLTsurvey/HLTsurvey.html>, 1996.

8. Richard Sproat and Joseph Olive, "A Modular Architecture for Multilingual Text-to-Speech," in J. van Santen, R. Sproat, J. Olive and J. Hirschberg (eds.), *Progress in Speech Synthesis*, New York, Springer, 1997.
9. Richard Sproat and Joseph Olive, "An Approach to Text-to-Speech Synthesis," in W. Kleijn and K. Paliwal (eds.), *Speech Coding and Synthesis*, 611–633, Amsterdam, Elsevier, 1995.
10. Julia Hirschberg and Richard Sproat, "Pitch Accent Prediction from Text Analysis," in J. Cole, G. Green and J. Morgan (eds.), *Linguistics and Computation*, 281–296, CSLI Publications, Chicago, University of Chicago Press, 1995.
11. Richard Sproat, "Looking into Words," in S. Hargus and E. Kaisse (eds.), *Studies in Lexical Phonology*, 173–195, San Diego, Academic Press, 1993.
12. Mark Liberman and Richard Sproat, "The Stress and Structure of Modified Noun Phrases in English," in I. Sag (ed.), *Lexical Matters*, 131–181, CSLI Publications, Chicago, University of Chicago Press, 1992.
13. Richard Sproat and Chilin Shih, "The Cross-Linguistic Distribution of Adjective Ordering Restrictions," in C. Georgopoulos and R. Ishihara (eds.), *Interdisciplinary Approaches to Language: Essays in Honor of S.-Y. Kuroda*, 565–593, Dordrecht, Kluwer Academic Publishers, 1990.
14. Richard Sproat, "Bracketing Paradoxes, Cliticization and Other Topics: The Mapping Between Syntactic and Phonological Structure," in M. Everaert, A. Evers, R. Huybregts and M. Trommelen (eds.), *Morphology and Modularity: In Honour of Henk Schultink*, 339–360, Dordrecht, Foris, 1988.
15. Richard Sproat, "On Anaphoric Islandhood," in M. Hammond and M. Noonan (eds.), *Theoretical Morphology*, 291–301, San Diego, Academic Press, 1988.

Conference Proceedings

1. Xuanhui Wang, ChengXiang Zhai, Xiao Hu, and Richard Sproat, "Mining Correlated Bursty Topic Patterns from Coordinated Text Streams", *Proceedings of the 2007 ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD'07)*, pages 784-793
2. Su-Youn Yoon, Kyoung-young Kim and Richard Sproat. "Multilingual Transliteration Using Feature based Phonetic Method". *ACL*, June, 2007, Prague.
3. Alla Rozovskaya, Richard Sproat, "Multilingual Word Sense Discrimination: A Comparative Cross-Linguistic Study", *Balto-Slavonic Natural Language Processing*. June, 2007, Prague.
4. Boon Pang Lim, Richard Sproat: Using Latent Semantics for NE Translation. *ICCPOL 2006*: 457-465
5. Tao Tao, Su-Youn Yoon, Andrew Fister, Richard Sproat and ChengXiang Zhai. "Unsupervised Named Entity Transliteration Using Temporal and Phonetic Correlation." *EMNLP*, July 22-23, 2006, Sydney, Australia.
6. Richard Sproat, Tao Tao and ChengXiang Zhai. "Named Entity Transliteration with Comparable Corpora". *ACL 2006*, July 17-21, 2006, Sydney, Australia.
7. Alla Rozovskaya, Richard Sproat and Elabbas Benmamoun. "Language Modeling of Arabic Dialects". Colloquium on Arabic Language Processing, June 5-7, Rabat, Morocco.
8. Cecilia Ovesdotter Alm, Dan Roth and Richard Sproat "Emotions from text: machine learning for text-based emotion prediction." *HLT/EMNLP 2005*. October 6-8, 2005, Vancouver.
9. Alm, Cecilia and Sproat, Richard. "Emotional sequencing and development in fairy tales." *First International Conference on Affective Computing and Intelligent Interaction*, Beijing, China, Oct. 22–24, 2005.
10. Alm, Cecilia and Sproat, Richard. "Perceptions of emotions in expressive storytelling." *InterSpeech 2005*, Lisbon, Portugal, Sep. 4–8, 2005.
11. Yanli Zheng, Richard Sproat, Liang Gu, Izhak Shafran, Haolang Zhou, Yi Su, Dan Jurafsky, Rebecca Starr, Su-Youn Yoon. "Accent Detection and Speech Recognition for Shanghai-Accented Mandarin." *InterSpeech 2005*, Lisbon, Portugal, Sep. 4–8, 2005.

12. Sameer Maskey, Michiel Bacchiani, Brian Roark, and Richard Sproat. "Improved Name Recognition with Meta-Data Dependent Name Networks". ICASSP 2004, Montréal, May 2004.
13. Murat Saraclar, Richard Sproat. "Lattice-Based Search for Spoken Utterance Retrieval." HLT-NAACL 04, Boston, May 2004.
14. Richard Sproat, "A Formal Computational Analysis of Indic Scripts," *International Symposium on Indic Scripts: Past and Future*, Tokyo. December 2003.
15. Michael Riley, Brian Roark and Richard Sproat. "Good-Turing Estimation from Word Lattices for Unsupervised Language Model Adaptation." IEEE Automatic Speech Recognition and Understanding Workshop. Fisherman's Reef, U.S. Virgin Islands. December 2003.
16. Jason Zhang, Alan Black and Richard Sproat, "Identifying Speakers in Children's Stories for Speech Synthesis," *Proceedings of EUROSPEECH 2003*, Geneva, September 2003.
17. Richard Sproat and Thomas Emerson, "The First International Chinese Word Segmentation Bakeoff", *The Second SIGHAN Workshop on Chinese Language Processing*, Sapporo, Japan, July 2003.
18. Srinivas Bangalore, Owen Rambow, Tahir Butt, Alexis Nasr and Richard Sproat, "Creating a Finite-State Parser with Application Semantics", *COLING*, Taipei, Taiwan, August 2002.
19. Richard Sproat, "Inferring the Environment in a Text-to-Scene Conversion System", *First International Conference on Knowledge Capture (K-CAP '01)*, Victoria, BC, Canada, 2001.
20. Richard Sproat, "Pmtools: A Pronunciation Modeling Toolkit", *Proceedings of the Fourth ISCA Tutorial and Research Workshop on Speech Synthesis*, Blair Atholl, Scotland, 2001.
21. Bob Coyne and Richard Sproat, "WordsEye: An Automatic Text-to-Scene Conversion system", *SIGGRAPH 2001*, Los Angeles, CA, 2001.
22. Richard Sproat, "Corpus-Based Methods and Hand-Built Methods", *International Conference on Spoken Language Processing*, 2000.
23. Bhuvana Narasimhan, Richard Sproat and George Kiraz. "Schwa deletion in Hindi Text-to-Speech Synthesis." Workshop on Computational Linguistics in South Asian Languages, 21st SALA, October 2001, Konstanz.
24. Jan van Santen and Richard Sproat. "High-accuracy automatic segmentation." *Proceedings of EUROSPEECH 99*, Budapest, Hungary, 1999.
25. Richard Sproat, Jianying Hu, Hao Chen, "EMU: An E-mail Preprocessor for Text-to-Speech," IEEE Signal Processing Society 1998 Workshop on Multimedia Signal Processing, Los Angeles, CA.
26. Hao Chen, Jianying Hu and Richard Sproat. "E-mail Signature Block Analysis", ICPR'98, Brisbane, Australia.
27. Richard Sproat, Andrew Hunt, Mari Ostendorf, Paul Taylor, Alan Black, Kevin Lenzo, Mike Edgington, "SABLE: A Standard for TTS Markup", *International Conference on Spoken Language Processing*, 1998. Also presented at the ESCA/COCOSDA Speech Synthesis Workshop. Jenolan Caves, Australia; and at the W3C Meeting on Voice Browsers, Cambridge, MA, October 13, 1998.
28. Richard Sproat, Jan van Santen, "Automatic Ambiguity Detection", *International Conference on Spoken Language Processing*, 1998.
29. Richard Sproat, Paul Taylor, Michael Tanenblatt and Amy Isard, "A Markup Language for Speech Synthesis," *Proceedings of EUROSPEECH*, 1997.
30. Bernd Möbius, Richard Sproat, Jan van Santen and Joseph Olive, "A German Text-to-Speech System," *Proceedings of EUROSPEECH*, 1997.
31. Elena Pavlova, Chilin Shih and Richard Sproat, "A Text-to-Speech System for Russian," *Proceedings of EUROSPEECH*, 1997. (Also published as http://www.kcn.ru/tat_en/science/fccl/pavlova.htm)

32. Chilin Shih and Richard Sproat, "Mandarin Text-to-Speech Synthesis," *Proceedings of the Eighth North American Conference on Chinese Linguistics*, 1997.
33. Richard Sproat. "Multilingual Text Analysis for Text-to-Speech Synthesis", *International Conference on Spoken Language Processing*, 1996.
34. Richard Sproat. "Multilingual Text Analysis for Text-to-Speech Synthesis", *ECAI '96 Workshop on Extended Finite State Models of Language*, Budapest, Hungary, 75–81, 1996.
35. Richard Sproat and Michael Riley. "Compilation of Weighted Finite-State Transducers from Decision Trees", *34th Annual Meeting of the Association for Computational Linguistics: Proceedings of the Conference*, 1996.
36. Mehryar Mohri and Richard Sproat. "An Efficient Compiler for Weighted Rewrite Rules", *34th Annual Meeting of the Association for Computational Linguistics: Proceedings of the Conference*, 1996.
37. Richard Sproat. "A Finite-State Architecture for Tokenization and Grapheme-to-Phoneme Conversion for Multilingual Text Analysis," *Proceedings of the EACL SIGDAT Workshop*, 65–72, Dublin, Ireland. Association for Computational Linguistics, 1995.
38. Richard Sproat and Chilin Shih, "A Corpus-Based Analysis of Mandarin Nominal Root Compounds," *Proceedings of the Sixth North American Conference on Chinese Linguistics*, 1995.
39. Benjamin Ao, Chilin Shih and Richard Sproat, "A Corpus-Based Mandarin Text-to-Speech Synthesizer," *International Conference on Spoken Language Processing*, 1771–1774, 1994.
40. Richard Sproat and Joseph Olive, "A Modular Architecture for Multilingual Text-to-Speech," *Proceedings of The Second ESCA/IEEE Workshop on Speech Synthesis*, 187–190, 1994.
41. Richard Sproat, Chilin Shih, William Gale and Nancy Chang, "A Stochastic Finite-State Word-Segmentation Algorithm for Chinese," *32nd Annual Meeting of the Association for Computational Linguistics: Proceedings of the Conference*, 66–73, 1994.
42. Fernando Pereira, Michael Riley and Richard Sproat, "Weighted Rational Transductions and their Application to Human Language Processing," *ARPA Workshop on Human Language Technology*, 249–254, 1994.
43. Richard Sproat, Julia Hirschberg and David Yarowsky, "A corpus-based synthesizer," *International Conference on Spoken Language Processing*, 563–566, 1992.
44. Richard Sproat and Chilin Shih, "On the sources of some constraints on Mandarin morphology," *Proceedings of the Third International Symposium on Chinese Languages and Linguistics*, 20–37, 1992.
45. David Roe, Fernando Pereira, Richard Sproat, Michael Riley, Pedro Moreno and Alejandro Macarrón, "Efficient Grammar Processing for a Spoken Language Translation System," *ICASSP-92(1)*, 213–216, 1992.
46. David Roe, Fernando Pereira, Richard Sproat, Michael Riley, Pedro Moreno and Alejandro Macarrón, "Toward a Spoken Language Translator for Restricted-Domain Context-Free Languages," *EUROSPEECH 91(3)*, 1063–1066, 1991.
47. Richard Sproat, "An Application of Statistical Optimization with Dynamic Programming to Phonemic-Input-to-Character Conversion for Chinese," *Proceedings of R.O.C. Computational Linguistics Conference III*, 379–390, 1990.
48. Richard Sproat, "Stress Assignment in Complex Nominals for English Text-to-Speech," *Proceedings of ESCA Workshop on Speech Synthesis*, 129–132.
49. Richard Sproat and Chilin Shih, "Prenominal Adjectival Ordering in English and Mandarin," *Proceedings of the 18th Annual Meeting of the North East Linguistics Society*, 465–489, 1988.
50. Richard Sproat and Barbara Brunson, "Constituent-Based Morphological Parsing: A New Approach to the Problem of Word-Recognition," *25th Annual Meeting of the Association for Computational Linguistics: Proceedings of the Conference*, 65–72, 1987.

51. Richard Sproat and Mark Liberman, "Toward Treating English Nominals Correctly," *25th Annual Meeting of the Association for Computational Linguistics: Proceedings of the Conference*, 140–146, 1987.
52. Richard Sproat and Gregory Ward, "Pragmatic Considerations in Anaphoric Island Phenomena," *Papers from the 23rd Annual Meeting of the Chicago Linguistic Society*, 321–335, 1987.
53. Richard Sproat, "Malayalam Compounding: A Non-Stratum-Ordered Account," *Proceedings of the Fifth West Coast Conference on Formal Linguistics*, 268–288, 1986.
54. Richard Sproat, "The Projection Principle and the Syntax of Synthetic Compounds," *Proceedings of the 16th Annual meeting of the North East Linguistics Society*, 462–475, 1986.
55. Richard Sproat, "INFL and the Configurability of VSO Languages," *Proceedings of the 14th Annual meeting of the North East Linguistics Society*, 418–431, 1984.

Invited Review Articles and Commentaries

1. Richard Sproat, "Morphology as Component or Module," invited review article in A. Spencer and A. Zwicky (eds.), *Handbook of Morphology*, Oxford, Basil Blackwell Ltd., 2000.
2. Richard Sproat, "Text Interpretation for Text-to-Speech Synthesis," in R. Cole, J. Mariani, H. Uszkoreit, A. Zaenen and V. Zue (eds.), *Survey of the State of the Art in Human Language Technology*, Dordrecht, Kluwer, 1996.
3. Richard Sproat, "Computational Interpretations of Neurolinguistic Observations," in M. Gazzaniga (ed.), *The Cognitive Neurosciences*, Cambridge, MIT Press, 1994, 931–942.
4. Richard Sproat, "Commentary on Bird and Klein." *Computational Linguistics*, 493, 1994.
5. Richard Sproat, "The Lexicon in Generative Grammar," invited review article in W. Bright (ed.), *Oxford International Encyclopedia of Linguistics*, New York, Oxford University Press, 1991.
6. Richard Sproat, "Competence, Performance and Agrammatism: A Reply to Grodzinsky," *Brain and Language*, 27, 160–167, 1986.

Invited Book Reviews

1. Chilin Shih, Richard Sproat, "Review of Horne (ed.) *Prosody: Theory and Experiment. Studies Presented to Gösta Bruce*," *Computational Linguistics*, 27(3), 2001, to appear.
2. Richard Sproat, "Review of Packard (ed.) *New Approaches to Chinese Word Formation*." *Yearbook of Morphology*, 1999.
3. Richard Sproat, "Review of Daniels and Bright, *The World's Writing Systems*," *Written Language and Literacy*, 1(1), 129–137, 1998.
4. Richard Sproat, "Review of Beard, *Lexeme-Morpheme Base Morphology*," *Journal of Linguistics*, 1997.
5. Richard Sproat, "Review of Miller *Complex Verb Formation*," *Language*, 70, 361–365, 1994.
6. Richard Sproat, "Review of Ritchie et al. *Computational Morphology*," *Language*, 69, 152–155, 1993.
7. Richard Sproat, "Review of Carstairs: *Current Morphology*," *Phonology Yearbook*, 9, 353–357, 1993.
8. Richard Sproat, "Morphological Non-Separation Revisited: a Review of R. Lieber's *Deconstructing Morphology*." *Yearbook of Morphology*, 235–258, 1992.
9. Richard Sproat, "Review of *The Syntax of the Modern Celtic Languages*." *Lingua*, 347–370, 1992.
10. Richard Sproat, "Review of *PC-KIMMO*," *Computational Linguistics*, 17, 229–231, 1991.

Other Book Reviews

1. Richard Sproat, "Review of Packard: *The Morphology of Chinese: A Linguistic and Cognitive Approach*", LINGUIST List: Vol-12-11.

2. Richard Sproat, "Review of Bright and Daniels: *The World's Writing Systems*", LINGUIST List: Vol-7-400.

Other Publications

1. Steve Farmer, Richard Sproat and Michael Witzel. "The Collapse of the Indus Script Thesis: The Myth of a Literate Harappan Civilization" in R.S. Bisht and M.P. Joshi, eds. *The Roots of the South Asian Civilisations: The Harappans*. 2005.
2. Bernd Möbius and Richard Sproat, "Linguistische Analyse in einem multilingualen Sprachsynthesystem." *Spektrum der Wissenschaften* 1996, (German Edition of Scientific American)
3. Richard Sproat, *On Deriving the Lexicon*. MIT Working Papers in Linguistics, Cambridge, MA. 1985. (Published version of Ph.D. dissertation)
4. Richard Sproat, "On Deriving the Lexicon," *The Linguistic Review* (Dissertation Abstracts), 5, 185–196, 1985.
5. Richard Sproat, "A Note on Rebracketing in Morphology," in D. Archangeli, A. Barss and R. Sproat (eds.), *MIT Working Papers in Linguistics*, 6, 199–205, 1985.
6. Richard Sproat, "Review of Schank's 'Identification of Conceptualizations Underlying Natural Language'," in B. Levin (ed.), *Lexical Semantics in Review*, Lexicon Project Working Papers, number 1, Cambridge, MA, Center for Cognitive Science, Massachusetts Institute of Technology, 1985.
7. Richard Sproat, "A Scientific Curriculum in Linguistics for Use In Secondary Schools." In *Doing Science: Constructing Scientific Theories as an Introduction to Scientific Method*, Technical Report, Cambridge, MA, Educational Technology Center, Harvard Graduate School of Education, 1985.
8. Richard Sproat, "On Bracketing Paradoxes," in M. Speas and R. Sproat (eds.), *MIT Working Papers in Linguistics*, 7, 110–130, 1984.
9. Richard Sproat, "VSO Languages and Welsh Configurationality," in I. Haik and D. Massam (eds.), *MIT Working Papers in Linguistics*, 5, 243–276, 1983.
10. Richard Sproat, letter to the editor, December issue of *Scientific American*: a critique of Derek Bickerton's paper entitled "Creole Languages," *Scientific American*, December, 1983.
11. John Newman, with the assistance of Richard Sproat, *Old High German Reader*, Advocate Publishing Group, Reynoldsburg, OH, 1981.

MISCELLANEOUS INVITED TALKS AND CONFERENCE PRESENTATIONS

1. Invited panelist, "Symposium on Finite-State Methods in Natural Language Analysis," American Academy of Sciences, Boston, MA, February 18, 2002.
2. Invited panelist at "Symposium on Distributed Morphology," 1994 Annual Meeting of the Linguistic Society of America, Boston, MA.
3. "Recent work on TTS at Bell Labs," Telecommunications Laboratories, Taiwan, August 1993.
4. "Computational Interpretations of Neurolinguistic Observations." McDonnell Summer Institute on Cognitive Neuroscience, July, 1993.
5. "Recent work on TTS at Bell Labs," Kiel University, Germany, October 1992.
6. "VEST: A Restricted-Domain Spoken Language Translation System." First International Workshop on Speech Translation, Schloß Dagstuhl, Germany, October 1992.
7. "Pitch Accent Prediction from Text Analysis," (with Julia Hirschberg) presented at the University of Illinois Workshop on Computational Linguistics, June 13-15, 1991.
8. "A pragmatic analysis of so-called anaphoric islands," (with Gregory Ward and Gail McKoon) University of Pennsylvania, November, 1990.

9. "Phonetic and pragmatic evidence that interpretive components can 'see' morphological structure," NSF Workshop on Lexical Phonology, University of Washington, Seattle, WA, June 8-10, 1990.
10. "Light and dark /l/ are not categorical in English: Articulatory and physiological evidence for phonetic implementation," (with Osamu Fujimura) University of Texas at Austin, October, 1989; Cornell University, October, 1989.
11. "The Articulation of /l/ in English: an Investigation of the Conditioning of Allophony," (with Osamu Fujimura) Tianjin Conference on Phonetics and Phonology, Tianjin, People's Republic of China, May 1989.
12. "Computational morphology," The Massachusetts Institute of Technology, April 1989.
13. "Connectionist Networks and Natural Language Morphology," (with Dania Egedi) University of Maryland Conference on Grammar and Language Processing, December, 1988.
14. "Connectionist Networks and Natural Language Morphology," (with Dania Egedi) SUNY Stonybrook, October 1987; University of Toronto, November 1987; Yale University, April 1988.
15. "The Analysis of English Compounds for Text-to-Speech," The Massachusetts Institute of Technology, May 1988.
16. "Constituent-Based Morphological Parsing," University of North Carolina, October 1986; Northwestern University, October 1986.
17. "On Bracketing Paradoxes: The Mapping Between S-Structure and PF," Workshop on Morphology and Modularity, University of Utrecht, June, 1986.
18. "More on Celtic Consonant Mutation," Kentucky Foreign Language Conference, University of Kentucky, Lexington, April, 1986.
19. "Constituent-Based Morphological Parsing," Bell Communications Research, Morristown, NJ, February 1985.
20. "Parsing Morphology," University of California at Irvine, October 1984; University of California at Los Angeles, October 1984.
21. "INFL and the Configurationality of VSO Languages," University of California at San Diego, October 1984; University of Southern California, October 1984.
22. "Some Issues in Current Morphological Theory," AT&T Bell Laboratories, Murray Hill, NJ, February 1984.

OTHER UNPUBLISHED CONFERENCE PRESENTATIONS

1. "Mandarin second-tone sandhi is not a phonological rule," (with Chilin Shih) International Conference on Chinese Linguistics, Singapore, June 1992.
2. "Mandarin morphology is not stratum-ordered," (with Chilin Shih) LSA Winter Meeting, 1992. Also presented at the University of Delaware, April, 1992, and Düsseldorf University, Germany, October, 1992.
3. "A pragmatic analysis of outbound anaphora and *vice-versa*," (with Gregory Ward) presented at the NSF-Sponsored Conference on 'Grammatical Foundations of Prosody and Discourse', July, 1991.
4. "Articulatory evidence for the non-categoricalness of English /l/ allophones," (with Osamu Fujimura) LSA Winter Meeting, 1989.
5. "An Investigation of Tag-Intonation in English," (with Shirley Steele) Acoustical Society of America, Miami, FL, November 1987.
6. "On the Resyllabification of /l/ in English," (with Toni Borowsky) Acoustical Society of America, Indianapolis, IN, May 1987.
7. "Stressing English Compounds Correctly," (with Mark Liberman) Acoustical Society of America, Anaheim, CA, December 1986.

8. "On the Phonological and Morphological Characterization of Celtic Consonant Mutation," The International Linguistics Association Meeting, New York, NY, March 1985.
9. "VSO Languages and Welsh Configurationality," Harvard Celtic Colloquium and the Massachusetts Institute of Technology, April 1983.
10. "Redundancy Rules and Welsh Mutation," Workshop in Morphology, Massachusetts Institute of Technology, January 1983.

LANGUAGES

Speaking and reading knowledge of German, French, Spanish and Mandarin. Reading knowledge of Dutch, Italian, Russian, Welsh, Old English, Classical Greek, Old Icelandic, Latin.

COMPUTER LANGUAGES AND SYSTEMS

Expertise in C, C++, Lisp, S/R, Python, Linux/Unix.

CURRENT SOCIETY MEMBERSHIPS

Linguistic Society of America, Association for Computational Linguistics.

UNIVERSITY SERVICE

Executive Committee, College of Liberal Arts and Sciences, University of Illinois at Urbana Champaign, 2007–2009.

Executive Committee, Graduate College, University of Illinois at Urbana Champaign, 2007–2009.

Program Advisory Committee, Beckman Institute, University of Illinois at Urbana Champaign, 2007–2010.

Department of Electrical and Computer Engineering, Graduate Committee, 2007–2008

Independent Program of Study Committee, College of LAS, University of Illinois at Urbana Champaign, 2005–.

Chair, UIUC Language and Speech Certificate Program, 2005–

Department of Electrical and Computer Engineering, Curriculum Committee, 2006–2007

Department of Linguistics, Institutional Review Board Committee, 2006–

Department of Linguistics, Promotions and Tenure Committee, 2006–2007

Department of Linguistics, Curriculum Committee, 2006–2007

UI College of LAS, Independent Program of Study Committee, 2005–2007

Department of Linguistics, Departmental Webpage Committee, 2004–

Department of Linguistics, Advisory Committee, 2004–2006

AWARDS AND HONORS

University Scholar, University of Illinois at Urbana Champaign, 2007–2008.

Associate, Center for Advanced Studies, University of Illinois at Urbana Champaign, Fall 2007.

UIUC College of Liberal Arts and Sciences Alumni Discretionary Award in recognition of exceptional service, July 2005.

Distinguished Member of the Technical Staff, Bell Laboratories, June 1997.

National Science Foundation Graduate Fellowship Award (awarded in 1982).

Phi Beta Kappa (awarded June 1981).

University of California President's Undergraduate Fellowship Award (awarded in 1979).

EARLY BACKGROUND

Born in National City, California on November 18, 1960. Raised in San Diego, California, subsequently in Romsey, England, and finally again in San Diego.

REFERENCES

Available upon request.