

RICHARD WILLIAM SPROAT

December 13, 2025

Richard Sproat
Sakana.ai
Toranomom Hills Business Tower 15F
1-17-1 Toranomom, Minato-ku
Tokyo 105-6415 Japan

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

September 2024 —

Research Scientist, Sakana.ai.

November 2015 — August 2024

Senior Staff Research Scientist, Google.

October 2012 — November 2015

Staff Research Scientist, Google.

January 2009 — October 2012

Professor, Center for Spoken Language Understanding (CSLU) and Division of Biomedical Computer Science – Department of Science and Engineering, Oregon Health & Science University (OHSU).

August 2003 — January 2009

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 have been the primary advisor.

- Cecilia Ovesdotter Alm (Linguistics) ‘Affect in Text and Speech.’ UIUC Linguistics, 2007.
- Hahn Koo (Linguistics) “Learning Non-adjacent Phonotactic Constraints from Brief Experience” . UIUC Linguistics, 2007.
- Suma Pramod (ECE). “Estimation Problems in Speech and Natural Language.” UIUC ECE. 2009.
- Su-youn Yoon. “Automated Assessment of Speech Fluency for L2 English Learners.” UIUC Linguistics, 2009.
- Boon-pang Lim. “Computational Differences in Whispered and Non-Whispered Speech.” UIUC ECE. 2010.
- Kyoung-young Kim. “Opinion Topic, Holder and Polarity in Texts: Exploration and Automatic Identification from Cross-Lingual Data.” UIUC Linguistics, 2011.

Courses at OHSU

Fall 2009. Computational Linguistics.

Spring 2010. Seminar on Speech and Language Processing for Augmentative and Alternative Communication (with Brian Roark).

Fall 2010. Practical Linguistics.

Spring 2011. Problem solving with large clusters (with Zak Shafran).

Fall 2011. Text Normalization (with Steven Bedrick).

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

Visiting scholar, The Views Project, <https://viewsproject.wordpress.com/>, Cambridge University, November–December, 2025.

Senior Area Chair, ACL Rolling Review, 2025.

President, pro tempore, *ACL Special Interest Group on Writing Systems and Written Language* <https://sigwrit.org/>, 2023–2024.

Action Editor, *Transactions on Computational Linguistics*, 2019 —

Editor in Chief, *ACM Transactions on Asian Language Information Processing*, June 2013 — June 2015

International Advisory Committee, International Speech Communication Association, 2012 —

General chair, InterSpeech 2012.

Local sponsorship chair, ACL-HLT 2011

Workshops Co-chair, NAACL 2010.

Tutorial (with Kevin Knight) on “Writing Systems, Transliteration and Decipherment”, NAACL 2009, Boulder, May 31, 2009.

Team Leader for “Multilingual Spoken Term Detection: Finding and Testing New Pronunciations” Workshop ’08 (July 7 – August 14, 2008), Center for Language and Speech Processing, Johns Hopkins University.

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/mcllc/>.

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 – 2003.

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) “Text to Scene Conversion” (with B. Coyne). December 27, 2011, U.S. Pat. Num. 8,086,028
- (b) “System and Method of Using Meta-Data in Speech Processing” (with M. Bacchiani, S. Maskey, B. Roark). August 9, 2011, U.S. Pat. Num. 7,996,224
- (c) “Text to Scene Conversion” (with B. Coyne). June 14, 2011, U.S. Pat. Num. 7,962,329
- (d) “System and Method of Lattice-Based Search for Spoken Utterance Retrieval” (with M. Saraçlar) March 22, 2011, U.S. Pat. Num. 7,912,699
- (e) “Text to Scene Conversion” (with B. Coyne). February 16, 2010, U.S. Pat. Num. 7,664,313
- (f) “Text to Scene Conversion” (with B. Coyne). May 13, 2008, U.S. Pat. Num. 7,373,292
- (g) “Text to Scene Conversion” (with B. Coyne). March 21, 2006, U.S. Pat. Num. 7,016,828
- (h) “E-Mail Signature Block Analysis” (with H. Chen and J. Hu) April 16, 2002, U.S. Pat. Num. 6,373,985.
- (i) “E-Mail Signature Block Segmentation” (with H. Chen and J. Hu), March 19, 2002, U.S. Pat. Num. 6,360,010.
- (j) “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.
- (k) “Compilation of Weighted Finite-State Transducers from Decision Trees”, September 8, 1998, U.S. Pat. Num. 5,806,032.
- (l) “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.

- (m) “Method for conversion of phonetic Chinese to character Chinese,” December 14, 1993, U.S. Pat. Num. 5,270,927.
- (n) “Semiotic class normalization” (with Kyle Gorman and Ke Wu), December 26, 2017, U.S. Pat. Num. 9,852,123.
- (o) “Conversion of Input Text Strings” (with Sascha Brawer and Martin Jansche), filed August 26, 2010. US 2014/0163952 A1..

FUNDING

1. DARPA BOLT (SRI prime, with Brian Roark), “Activity A: Genre-Independent Translation and Information Retrieval System”, October 2011–.
2. Sproat, Richard (PI), NIH-K25DC011308, “Natural Language Processing for Augmentative and Alternative Communication in Adults,” August, 2011–July, 2014, \$487K.
3. Sproat, Richard (PI et al.), NIH-R01-DC012033, “Computational characterization of language use in autism spectrum disorder,” August, 2011–July, 2016, \$3.7M.
4. Sproat, Richard (PI). NSF-BCS-1137631, “Corpora of Non-Linguistic Symbol Systems, and Statistical Analysis”, 04/01/2011 – 03/31/2012, \$114,140
5. Sproat, Richard (PI, et al.). Intelligence Community Postdoc, 2010*1043506*00, “Stochastic Text Normalization and Automatic Genre Adaptation for Structured Text Processing” 07/12/2010 — 07/11/2013, \$360,000.
6. Roark, Brian (PI, et al.). NSF-IIS-0964102, “RI: Medium: Semi-Supervised Discriminative Training of Language Models”, 07/01/2010 — 06/30/2013, \$500,000.
7. Julia Hirschberg (PI), Owen Rambow (co-PI), Richard Sproat (co-PI). “RI Medium: Collaborative Research:From Text to Pictures,” NSF-RI, July 2009 — June 2012, \$259,862.
8. 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**.
9. Richard Sproat (PI), “Segmental Awareness and Indic Writing Systems”, NIH, 1R03HD048644-01, August 2006–August 2008, **\$108,259**.
10. 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.
11. Richard Sproat (PI), “Segmental Awareness and Indic Writing Systems”, University of Illinois Research Board, January 2006 – May 2006. **\$6120**.
12. Richard Sproat (PI), “Machine transliteration for underresourced languages”, Google, Inc, February 2006 –, **\$150,000**.
13. 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**.
14. 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**.

15. 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**.
16. 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. Brian Roark, Richard Sproat and Suyoun Yoon. *Tools of the Scribe: How writing systems, technology, and human factors interact to affect the act of writing*. Cham, SpringerNature, Forthcoming, 2025.
2. Richard Sproat. *Symbols: An Evolutionary History from the Stone Age to the Future*. Cham, SpringerNature, 2023.
3. Kyle Gorman and Richard Sproat. *Finite-State Text Processing*. Williston, VT, Morgan Claypool, 2021.
4. Fabian Brackhane, Richard Sproat and Jürgen Trouvain (editors). *Wolfgang von Kempelen’s Mechanismus der menschlichen Sprache*. New edition with English translation, and commentary. Technische Universität Dresden Press, 2017.
5. Richard Sproat. *Language, Technology and Society*. Oxford, Oxford University Press, 2010.
6. Brian Roark, Richard Sproat. *Computational Approaches to Syntax and Morphology*. Oxford, Oxford University Press, 2007.
7. Martin Neef, Anneke Neijt and Richard Sproat (editors). *The Relation of Writing to Spoken Language*. Tübingen: Niemeyer (= Linguistische Arbeiten 460), 2002.
8. Richard Sproat, *A Computational Theory of Writing Systems*, (ACL Studies in Natural Language Processing Series), Cambridge, Cambridge University Press, 2000.
9. Richard Sproat (editor), *Multilingual Text-to-Speech Synthesis*, Boston, MA, Kluwer Academic Publishers, 1997.
10. Jan van Santen, Richard Sproat, Joseph Olive and Julia Hirschberg (editors), *Progress in Speech Synthesis*, New York, Springer, 1997.
11. Richard Sproat, *Morphology and Computation*, Cambridge, MA, MIT Press, 1992.

Videos

1. Richard Sproat. *Introduction to Writing Systems: How Writing Encodes Language*, SpringerNature, 2024. <https://link.springer.com/video/10.1007/978-3-031-72193-9>.

Refereed Journals

1. Llion Jones, Richard Sproat, Haruko Ishikawa and Alexander Gutkin. “Helpful Neighbors: Leveraging Neighbors in Geographic Feature Pronunciation.” *Transactions of the Association for Computational Linguistics*, To appear.
2. Richard Sproat. “Last Words: Boring Problems are Sometimes the Most Interesting.” *Computational Linguistics*, 48(2): 483–490.

3. Richard Sproat and Alexander Gutkin. "The Taxonomy of Writing Systems: How to Measure how Logographic a System is." *Computational Linguistics*, 47(3), 477–528, 2021.
4. Richard Sproat. "Review of Zev Handel. 2019. Sinography: The Borrowing and Adaptation of the Chinese Script." *Asiatische Studien/Études Asiatiques*, 74(4): 1103–1115, 2021.
5. Juliette Blevins, Richard Sproat. "Statistical Evidence for the Proto-Indo-European-Euskarian Hypothesis: A word-list approach integrating phonotactics". *Diachronica*, 2021.
6. Hao Zhang, Richard Sproat, Axel Ng, Felix Stahlberg, Xiaochang Peng, Kyle Gorman, Brian Roark. "Neural models of text normalization for speech applications." *Computational Linguistics*, 45(2), 2019: 293–337.
7. Richard Sproat. "Review of Daniels *An Exploration of Writing*", *Written Language and Literacy*, 21(2), 2018: 269–278.
8. Richard Sproat. "A Computational Model of the Discovery of Writing", *Written Language and Literacy*, 20(2), 2017: 194–226.
9. Richard Sproat, "Language typology in speech and language technology", *Linguistic Typology*, 20(3), 2016.
10. Kyle Gorman and Richard Sproat. "Minimally supervised models for number normalization". *Transactions of the Association for Computational Linguistics*. 2016.
11. Richard Sproat. "On misunderstandings and misrepresentations: A reply to Rao et al." *Language*, 91(4), 2015.
12. Peter Ebdon and Richard Sproat. "The Kestrel TTS Text Normalization System." *Journal of Natural Language Engineering*, December 2014.
13. Richard Sproat, Mahsa Yarmohammadi, Izhak Shafran and Brian Roark. "Applications of Lexicographic Semirings to Problems in Speech and Language Processing," *Computational Linguistics*, 40(4), 2014.
14. Richard Sproat, "A statistical comparison of written language and non-linguistic symbol systems," *Language*, 90(2), 2014.
15. Richard Sproat, "A note on Unger's 'What linguistic units do Chinese characters represent?'" *Written Language & Literacy*, 16(1), 107–111, 2013.
16. Jan van Santen, Richard Sproat and Alison Presmanes Hill, "Quantifying Repetitive Speech in Autism Spectrum Disorders and Language Impairment," *Autism Research*, 2013.
17. Padmapriya Kandhadai and Richard Sproat, "Impact of spatial ordering of graphemes in alphasyllabic scripts on phonemic awareness in Indic languages," *Writing Systems Research*, 2(2), 105–116, 2010.
18. Richard Sproat. "Reply to Rao et al. and Lee et al." *Computational Linguistics*, 36:4, 2010.
19. Richard Sproat. "Last Words: Ancient symbols, computational linguistics, and the reviewing practices of the general science journals" *Computational Linguistics*, 36:3, 2010.
20. Aamir Wali, Richard Sproat, Prakash Padakannaya and Bhuvaneshwari, B. "Model for phonemic awareness in readers of Indian script." *Written Language and Literacy*, 12:2, 161–169, 2009.
21. Su-Youn Yoon, Lisa Pierce, Amanda Huensch, Eric Juul, Samantha Perkins, Richard Sproat, and Mark Hasegawa-Johnson, Construction of a rated speech corpus of L2 learners' speech, *CALICO journal*, 2009
22. Richard Sproat. "Brahmi-Derived Scripts, Script Layout, and Phonological Awareness." *Written Language and Literacy*, 9(1), 45–66, 2006.

23. 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
24. Michiel Bacchiani, Brian Roark, Michael Riley and Richard Sproat "MAP Adaptation of Stochastic Grammars". *Computer Speech and Language*, 20(1):41-68, 2006.
25. 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.
26. 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.
27. Richard Sproat. "Multilingual Text Analysis for Text-to-Speech Synthesis", *Natural Language Engineering*, 2(4), 369–380, 1996.
28. Chilin Shih and Richard Sproat, "Issues in Text-to-Speech Conversion for Mandarin," *Computational Linguistics and Chinese Language Processing*, 1996.
29. Richard Sproat, Chilin Shih, William Gale and Nancy Chang, "A Stochastic Finite-State Word-Segmentation Algorithm for Chinese," *Computational Linguistics*, 22(3), 1996.
30. Harald Baayen and Richard Sproat, "Estimating Lexical Priors for Low-Frequency Morphologically Ambiguous Forms," *Computational Linguistics*, 22(2), 1996.
31. Richard Sproat and Chilin Shih. "A Corpus-Based Analysis of Mandarin Nominal Root Compounds." *Journal of East Asian Linguistics*, 5, 49–71, 1996.
32. Richard Sproat and Joseph Olive, "Text to Speech Synthesis," *AT&T Technical Journal*, 74(2), 35–44, 1995.
33. Richard Sproat, "English Noun-Phrase Accent Prediction for Text-to-Speech." *Computer Speech and Language*, 8, 79–94, 1994.
34. Richard Sproat and Chilin Shih, "Why Mandarin Morphology is not Stratum-Ordered." *Yearbook of Morphology*, 185–217, 1993.
35. Richard Sproat and Osamu Fujimura, "Allophonic variation in English /l/ and its implications for phonetic implementation," *Journal of Phonetics*, 21, 291–311, 1993.
36. 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.
37. 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.
38. Richard Sproat, "Unhappier is not a Bracketing Paradox," *Linguistic Inquiry*, 23, 347–352, 1992.
39. Gregory Ward, Richard Sproat and Gail McKoon, "A Pragmatic Analysis of So-Called Anaphoric Islands," *Language*, 67, 439–474, 1991.
40. 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.
41. Richard Sproat, "Welsh Syntax and VSO Structure," *Natural Language and Linguistic Theory*, 3, 173–216, 1985.

42. 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. Christo Kirov and Richard Sproat. "Computational Morphology". *Probabilistic Linguistics*, CSLI, to appear.
2. Richard Sproat. "Writing Systems". *Oxford Handbook of the History of Phonology*, to appear.
3. Gregory Ward and Richard Sproat. "Anaphoric Islands: A History of a Misleading Geographical Metaphor". *The Wiley-Blackwell Companion to Morphology*. Oxford: Wiley-Blackwell, to appear.
4. Sproat, Richard. "English among the writing systems of the world", in Vivian Cook and Des Ryan, *Handbook of the English Writing System*, Routledge.
5. Sproat, Richard and Prakash Padakannaya. "Script Indices". in N. Srinivasan, A.K. Gupta and J. Pandey, *Advances in Cognitive Science*, Volume 1. New Delhi, Sage: 2008.
6. 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.
7. 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.
8. 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.
9. 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.
10. Richard Sproat, "Lexical Analysis", in Robert Dale, Hermann Moisl, and Harold Somers (eds.) *Handbook of Natural Language Processing*, 37–58, New York, Dekker, 2000.
11. 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.
12. 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.
13. 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.
14. 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.
15. 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.
16. Richard Sproat, "Looking into Words," in S. Hargus and E. Kaisse (eds.), *Studies in Lexical Phonology*, 173–195, San Diego, Academic Press, 1993.
17. 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.

18. 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.
19. 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.
20. Richard Sproat, “On Anaphoric Islandhood,” in M. Hammond and M. Noonan (eds.), *Theoretical Morphology*, 291–301, San Diego, Academic Press, 1988.

Conference Proceedings, Presentations

1. Shigeki Karita, Richard Sproat and Haruko Ishikawa. “Lenient Evaluation of Japanese Speech Recognition: Modeling Naturally Occurring Spelling Inconsistency.” First ACL Workshop on Computation and Written Language, Toronto, July 14, 2023.
2. Abhirut Gupta, Ananya B. Sai, Richard Sproat, Yuri Vasilevski, James S. Ren, Ambarish Jash, Sukhdeep S. Sodhi and Aravindan Raghuvier. “Bi-Phone: Modeling Inter Language Phonetic Influences in Text.” Association for Computational Linguistics, Toronto, July 9-14, 2023.
3. Richard Sproat. “A Computational Model of the Invention of Writing.” Anchoring the Invention of Writing. Leiden, December 15–17, 2022.
4. Alexander Gutkin, Cibu Johny, Raiomond Doctor, Brian Roark and Richard Sproat. “Beyond Arabic: Software for Perso-Arabic Script Manipulation,” *WANLP*, 2022.
5. Christo Kirov, Richard Sproat and Alexander Gutkin. “Mockingbird at the SIGTYP 2022 Shared Task: Two Types of Models for the Prediction of Cognate Reflexes,” *ACL SIGTYP*, July, 2022.
Raiomond Doctor, Alexander Gutkin, Cibu Johny, Brian Roark and Richard Sproat “Graphemic Normalization of the Perso-Arabic Script,” *Grapholinguistics in the 21st Century*, June 2022.
6. Raiomond Doctor, Alexander Gutkin, Cibu Johny, Brian Roark and Richard Sproat “Graphemic Normalization of the Perso-Arabic Script,” *Grapholinguistics in the 21st Century*, June 2022.
7. Kyle Gorman and Richard Sproat. “On the Persistent Conflation of Writing and Language,” *Grapholinguistics in the 21st Century*, June 2022.
8. Richard Sproat. “Computational Methods in the Analysis of Graphical Symbol Systems,” Keynote Presentation, *Grapholinguistics in the 21st Century*, June 2022.
9. Kyle Gorman, Christo Kirov, Brian Roark and Richard Sproat. “Structured abbreviation expansion in context”, *Findings of EMNLP*, 2021.
10. Jae Ro, Hao Zhang and Richard Sproat. “Semi-supervised URL Segmentation with Recurrent Neural Networks Pre-trained on Knowledge Graph Entities”. *COLING*, 2020.
11. Alexander Gutkin and Richard Sproat, “NEMO: Frequentist Inference Approach to Constrained Linguistic Typology Feature Prediction in SIGTYP 2020 Shared Task”. *SIGTYP Workshop*, 2020.
12. Alena Butryna, Shan-Hui Cathy Chu, Isin Demirsahin, Alexander Gutkin, Linne Ha, Fei He, Martin Jansche, Cibu Johny, Anna Katanova, Oddur Kjartansson, Chenfang Li, Tatiana Merkulova, Yin May Oo, Knot Pipatsrisawat, Clara Rivera, Supheakmongkol Sarin, Pasindu de Silva, Keshan Sodimana, Richard Sproat, Theeraphol Wattanavekin and Jaka Aris Eko Wibawa. “Google Crowdsourced Speech Corpora and Related

- Open-Source Resources for Low-Resource Languages and Dialects: An Overview". 2019 *UNESCO International Conference Language Technologies for All (LT4All): Enabling Linguistic Diversity and Multilingualism Worldwide*, 4-6 December, Paris, France, 2019. <https://arxiv.org/abs/2010.06778>"><https://arxiv.org/abs/2010.06778>
13. Ajda Gokcen, Hao Zhang and Richard Sproat. "Dual Encoder Classifier Models as Constraints in Neural Text Normalization." *Interspeech*, 2019.
 14. Sandy Ritchie, Richard Sproat, Kyle Gorman, Daan van Esch, Christian Schallhart, Nikos Bampounis, Benoît Brard, Jonas Fromseier Mortensen, Millie Holt, and Eoin Mahon. "Unified Verbalization for Speech Recognition & Synthesis Across Languages." *Interspeech*, 2019.
 15. Keshan Sodimana, Pasindu De Silva, Richard Sproat, Theeraphol Wattanavekin, Alexander Gutkin and Knot Pipatsrisawat. "Text Normalization for Bangla, Khmer, Nepali, Javanese, Sinhala and Sundanese Text-to-Speech Systems." *SLTU*. 2018.
 16. Hao Zhang, Ke Wu and Richard Sproat "Fast and Accurate Reordering with ITG Transition RNN." *COLING*. 2018.
 17. Ke Wu, Kyle Gorman and Richard Sproat. "Minimally supervised written-to-spoken text normalization." *ASRU*. 2017.
 18. Fabian Brackhane, Richard Sproat and Jürgen Trouvain. "Editing Kempelen's *Mechanismus der menschlichen Sprache*: Experiences and Findings," *History of Speech Communication Research*, 2017.
 19. Richard Sproat and Navdeep Jaitly. "An RNN Model of Text Normalization," *Proceedings of Interspeech*, 2017.
 20. Alexander Gutkin and Richard Sproat. "Areal and Phylogenetic Features for Multilingual Speech Synthesis," *Proceedings of Interspeech*, 2017.
 21. Daan van Esch and Richard Sproat. "An Expanded Taxonomy of Semiotic Classes for Text Normalization," *Proceedings of Interspeech*, 2017.
 22. Alexander Gutkin, Linne Ha, Martin Jansche, Oddur Kjartansson, Knot Pipatsrisawat, Richard Sproat, "Building statistical parametric multi-speaker synthesis for Bangladeshi Bangla", *IEEE Workshop on Spoken Language Technology (SLT)*, 2016.
 23. Alexander Gutkin, Linne Ha, Martin Jansche, Knot Pipatsrisawat and Richard Sproat , "TTS for Low Resource Languages: A Bangla Synthesizer." *LREC 2016*.
 24. Masoud Rouhizadeh, Emily Prud'hommeaux, Jan van Santen, Richard Sproat. "Measuring idiosyncratic interests in children with autism spectrum disorder". *Association for Computational Linguistics*, 2015, Beijing.
 25. Masoud Rouhizadeh, Jan van Santen, Richard Sproat. "Similarity Measures for Quantifying Restrictive and Repetitive Behavior in Conversations of Autistic Children". *CLPsych*, 2015, Denver.
 26. Emily Prud'hommeaux, Eric Morley, Masoud Rouhizadeh, Laura Silverman, Jan van Santen, Brian Roark, Richard Sproat, Sarah Kauper, Rachel DeLaHunta. "Computational analysis of trajectories of linguistic development in autism". *IEEE Workshop on Spoken Language Technology*, 2014, South Lake Tahoe.
 27. Richard Sproat and Keith Hall. "Applications of Maximum Entropy Rankers to Problems in Spoken Language Processing." *Interspeech 2014*.
 28. Brian Roark and Richard Sproat. "Hippocratic abbreviation expansion." *ACL 2014*.

29. Masoud Rouhizadeh, Emily Prud'hommeaux, Jan van Santen, Richard Sproat "Detecting linguistic restricted interests in autism using distributional semantic models," *CLPsych*, 2014, Baltimore, MD
30. Richard Sproat, Bruno Cartoni, HyunJeong Choi, Linne Ha, David Huynh, Ravindran Rajakumar and Evelyn Wenzel-Grondie. "A Database for Measuring Linguistic Information Content." *LREC 2014*.
31. Keith Hall and Richard Sproat. "Russian Stress Prediction using Maximum Entropy Ranking", *EMNLP 2013*.
32. Bob Coyne, Alex Klapheke, Masoud Rouhizadeh, Richard Sproat and Daniel Bauer. 2012. "Annotation Tools and Knowledge Representation for a Text-To-Scene System." *COLING*, Mumbai, India, December, 2012.
33. Steven Bedrick, Russell Beckley, Brian Roark and Richard Sproat. "Robust kaomoji detection in Twitter". *Language and Social Media*, Montreal, Canada, June, 2012.
34. Margaret Mitchell and Richard Sproat. "Discourse-Based Modeling for AAC". *Third Workshop on Speech and Language Processing for Assistive Technologies* Montreal, Canada, June, 2012.
35. Brian Roark, Michael Riley, Cyril Allauzen, Terry Tai and Richard Sproat. "The OpenGrm open-source finite-state grammar software libraries". *ACL 2012*, Jeju Island, Korea, July.
36. Richard Sproat, Katherine Wu, Jennifer Solman, Ruth Linehan. "Corpora of Non-Linguistic Symbol Systems". *Linguistic Society of America*, Portland, Oregon, January, 2012.
37. Terry Tai, Wojciech Skut and Richard Sproat. "Thrax: An Open Source Grammar Compiler Built on OpenFst". *ASRU 2011*, Waikoloa Resort, Hawaii, December.
38. Izhak Shafran, Richard Sproat, Mahsa Yarmohammadi and Brian Roark. "Efficient Determinization of Tagged Word Lattices using Categorical and Lexicographic Semirings". *ASRU 2011*, Waikoloa Resort, Hawaii, December.
39. Masoud Rouhizadeh, Bob Coyne, and Richard Sproat, "Collecting Semantic Information for Locations in the Scenario-Based Lexical Knowledge Resource of a Text-to-Scene Conversion System", *15th International Conference on Knowledge-Based and Intelligent Information & Engineering Systems (KES 2011)*, Kaiserslautern, Germany, September, 2011.
40. Brian Roark, Andrew Fowler, Richard Sproat, Christopher Gibbons and Melanie Fried-Oken, "Towards technology-assisted co-construction with communication partners". *SLPAT 2011*, Edinburgh. 2011.
41. Masoud Rouhizadeh, Daniel Bauer, Bob Coyne, Owen Rambow and Richard Sproat. 2011. "Collecting Spatial Information for Locations in a Text-to-Scene Conversion System." *CoSLI 2011 Workshop*, Boston, MA. 2011.
42. Brian Roark, Richard Sproat and Izhak Shafran, "Lexicographic Semirings for Exact Automata Encoding of Sequence Models". *ACL-HLT 2011*, Portland, OR. 2011. **Best short paper award.**
43. Alexander Kotov, ChengXiang Zhai and Richard Sproat, "Mining Named Entities with Temporally Correlated Bursts from Multilingual Web News Streams", *WSDM 2011*. 2011.
44. Masoud Rouhizadeh, Margit Bowler, Richard Sproat and Bob Coyne. "Collecting Semantic Data from Amazon's Mechanical Turk for a Lexical Knowledge Resource in a Text to Picture Generating System." *International Conference on Computational Semantics (IWCS 2011)*, Oxford, January, 2011.
45. Aaron M. Cohen, Kyle Ambert, Jianji Yang, Robert Felder, Richard Sproat, Brian Roark, Kristy Hollingshead and Kari Baker. "OHSU/Portland VAMC Team Participation in the 2010 i2b2/VA Challenge Tasks," *Fourth i2b2/VA NLP Workshop*, Washington, December 2010.

46. Masoud Rouhizadeh, Margit Bowler, Richard Sproat and Bob Coyne. "Data Collection and Normalization for Building the Scenario-Based Lexical Knowledge Resource of a Text-to-Scene Conversion System," *SMAP 2010: 5th International Workshop on Semantic Media Adaptation and Personalization*. Limassol, Cyprus, 2010.
47. Richard Sproat, "Lightly Supervised Learning of Text Normalization: Russian Number Names," *IEEE Workshop on Spoken Language Technology*, Berkeley, CA, 2010.
48. Su-Youn Yoon, Mark Hasegawa-Johnson, and Richard Sproat, "Landmark-based Automated Pronunciation Error Detection," *Proceedings of Interspeech 2010*, Makuhari, Japan, 2010.
49. Suma Bhat, Mark Hasegawa-Johnson, Richard Sproat. "Automatic Fluency Assessment by Signal-Level Measurement of Spontaneous Speech". *2010 INTERSPEECH Satellite Workshop on Second Language Studies: Acquisition, Learning, Education and Technology*, Makuhari, Japan, 2010.
50. Bob Coyne, Richard Sproat, Julia Hirschberg. "Spatial Relations in Text-to-Scene Conversion." *Computational Models of Spatial Language Interpretation*, Workshop at *Spatial Cognition 2010*, Mt. Hood, Oregon.
51. Bob Coyne, Owen Rambow, Julia Hirschberg, and Richard Sproat, "Frame Semantics in Text-to-Scene Generation". *14th International Conference on Knowledge-Based and Intelligent Information & Engineering Systems*, Cardiff, Wales, 2010.
52. Ting Qian, Kristy Hollingshead, Su-youn Yoon, Kyoung-young Kim, and Richard Sproat. "A Python Toolkit for Universal Transliteration". *LREC*, Malta, 2010.
53. Suma Bhat, Richard Sproat, Mark Hasegawa-Johnson and Fred Davidson, "Automatic fluency assessment using thin-slices of spontaneous speech," *LTRC*, 2010.
54. Su-youn Yoon, Mark Hasegawa-Johnson and Richard Sproat. "Automated Pronunciation Scoring using Confidence Scoring and Landmark-based SVM", *InterSpeech*, 2009
55. Martin Jansche, Richard Sproat, "Named Entity Transcription with Pair n-Gram Models", *Named Entities Workshop: Shared Task on Transliteration (NEWS 2009)*, ACL-IJCNLP 2009, pp. 32–35.
56. Suma Bhat and Richard Sproat. "Knowing the Unseen: Estimating Vocabulary Size over Unseen Samples", *ACL*, 2009.
57. 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
58. Su-Youn Yoon, Kyoung-young Kim and Richard Sproat. "Multilingual Transliteration Using Feature based Phonetic Method". *ACL*, June, 2007, Prague.
59. Alla Rozovskaya, Richard Sproat, "Multilingual Word Sense Discrimination: A Comparative Cross-Linguistic Study", *Balto-Slavonic Natural Language Processing*. June, 2007, Prague.
60. Boon Pang Lim, Richard Sproat: Using Latent Semantics for NE Translation. *ICCPOL 2006*: 457-465
61. 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.
62. Richard Sproat, Tao Tao and ChengXiang Zhai. "Named Entity Transliteration with Comparable Corpora". *ACL 2006*, July 17-21, 2006, Sydney, Australia.
63. Alla Rozovskaya, Richard Sproat and Elabbas Benmamoun. "Language Modeling of Arabic Dialects". *Colloquium on Arabic Language Processing*, June 5-7, Rabat, Morocco.

64. 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.
65. 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.
66. Alm, Cecilia and Sproat, Richard. “Perceptions of emotions in expressive storytelling.” *InterSpeech 2005*, Lisbon, Portugal, Sep. 4–8, 2005.
67. 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.
68. Sameer Maskey, Michiel Bacchiani, Brian Roark, and Richard Sproat. “Improved Name Recognition with Meta-Data Dependent Name Networks”. ICASSP 2004, Montréal, May 2004.
69. Murat Saraclar, Richard Sproat. “Lattice-Based Search for Spoken Utterance Retrieval.” *HLT-NAACL 04*, Boston, May 2004.
70. Richard Sproat, “A Formal Computational Analysis of Indic Scripts,” *International Symposium on Indic Scripts: Past and Future*, Tokyo. December 2003.
71. 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.
72. Jason Zhang, Alan Black and Richard Sproat, “Identifying Speakers in Children’s Stories for Speech Synthesis,” *Proceedings of EUROSPEECH 2003*, Geneva, September 2003.
73. Richard Sproat and Thomas Emerson, “The First International Chinese Word Segmentation Bakeoff”, *The Second SIGHAN Workshop on Chinese Language Processing*, Sapporo, Japan, July 2003.
74. Srinivas Bangalore, Owen Rambow, Tahir Butt, Alexis Nasr and Richard Sproat, “Creating a Finite-State Parser with Application Semantics”, *COLING*, Taipei, Taiwan, August 2002.
75. 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.
76. Richard Sproat, “*Pmtools*: A Pronunciation Modeling Toolkit”, *Proceedings of the Fourth ISCA Tutorial and Research Workshop on Speech Synthesis*, Blair Atholl, Scotland, 2001.
77. Bob Coyne and Richard Sproat, “WordsEye: An Automatic Text-to-Scene Conversion system”, *SIGGRAPH 2001*, Los Angeles, CA, 2001.
78. Richard Sproat, “Corpus-Based Methods and Hand-Built Methods”, *International Conference on Spoken Language Processing*, 2000.
79. 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.
80. Jan van Santen and Richard Sproat. “High-accuracy automatic segmentation.” *Proceedings of EUROSPEECH 99*, Budapest, Hungary, 1999.
81. 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.

82. Hao Chen, Jianying Hu and Richard Sproat. "E-mail Signature Block Analysis", ICPR'98, Brisbane, Australia.
83. 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.
84. Richard Sproat, Jan van Santen, "Automatic Ambiguity Detection", *International Conference on Spoken Language Processing*, 1998.
85. Richard Sproat, Paul Taylor, Michael Tanenblatt and Amy Isard, "A Markup Language for Speech Synthesis," *Proceedings of EUROSPEECH*, 1997.
86. Bernd Möbius, Richard Sproat, Jan van Santen and Joseph Olive, "A German Text-to-Speech System," *Proceedings of EUROSPEECH*, 1997.
87. 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)
88. Chilin Shih and Richard Sproat, "Mandarin Text-to-Speech Synthesis," *Proceedings of the Eighth North American Conference on Chinese Linguistics*, 1997.
89. Richard Sproat. "Multilingual Text Analysis for Text-to-Speech Synthesis", *International Conference on Spoken Language Processing*, 1996.
90. 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.
91. 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.
92. 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.
93. 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.
94. 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.
95. Benjamin Ao, Chilin Shih and Richard Sproat, "A Corpus-Based Mandarin Text-to-Speech Synthesizer," *International Conference on Spoken Language Processing*, 1771–1774, 1994.
96. 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.
97. 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.
98. 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.
99. Richard Sproat, Julia Hirschberg and David Yarowsky, "A corpus-based synthesizer," *International Conference on Spoken Language Processing*, 563–566, 1992.

100. 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.
101. 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.
102. 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.
103. 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.
104. Richard Sproat, "Stress Assignment in Complex Nominals for English Text-to-Speech," *Proceedings of ESCA Workshop on Speech Synthesis*, 129–132.
105. 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.
106. 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.
107. 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.
108. 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.
109. Richard Sproat, "Malayalam Compounding: A Non-Stratum-Ordered Account," *Proceedings of the Fifth West Coast Conference on Formal Linguistics*, 268–288, 1986.
110. 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.
111. Richard Sproat, "INFL and the Configurationality of VSO Languages," *Proceedings of the 14th Annual meeting of the North East Linguistics Society*, 418–431, 1984.

Invited Review Articles and Commentaries

1. Amalia Gnanadesikan and Richard Sproat. "Writing Systems". *Oxford Bibliographies*, 2018.
2. Richard Sproat, "Phonemic diversity and the out-of-Africa theory". *Linguistic Typology*, 15, 199–206, 2011.
3. 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.
4. 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.
5. Richard Sproat, "Computational Interpretations of Neurolinguistic Observations," in M. Gazzaniga (ed.), *The Cognitive Neurosciences*, Cambridge, MIT Press, 1994, 931–942.
6. Richard Sproat, "Commentary on Bird and Klein." *Computational Linguistics*, 493, 1994.

7. 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.
8. Richard Sproat, "Competence, Performance and Agrammatism: A Reply to Grodzinsky," *Brain and Language*, 27, 160–167, 1986.

Invited Book Reviews

1. Richard Sproat, "Review of Shouhui Zhao & Richard B. Baldauf, Jr., *Planning Chinese Characters: Reaction, Evolution or Revolution*. (Language Policy, 9.)," *Written Language and Literacy*, 11(2), 2008.
2. Chilin Shih, Richard Sproat, "Review of Horne (ed.) *Prosody: Theory and Experiment. Studies Presented to Gösta Bruce*, *Computational Linguistics*, 27(3), 2001.
3. Richard Sproat, "Review of Packard (ed.) *New Approaches to Chinese Word Formation*." *Yearbook of Morphology*, 1999.
4. Richard Sproat, "Review of Daniels and Bright, *The World's Writing Systems*," *Written Language and Literacy*, 1(1), 129–137, 1998.
5. Richard Sproat, "Review of Beard, *Lexeme-Morpheme Base Morphology*," *Journal of Linguistics*, 1997.
6. Richard Sproat, "Review of Miller *Complex Verb Formation*," *Language*, 70, 361–365, 1994.
7. Richard Sproat, "Review of Ritchie et al. *Computational Morphology*," *Language*, 69, 152–155, 1993.
8. Richard Sproat, "Review of Carstairs: *Current Morphology*," *Phonology Yearbook*, 9, 353–357, 1993.
9. Richard Sproat, "Morphological Non-Separation Revisited: a Review of R. Lieber's *Deconstructing Morphology*." *Yearbook of Morphology*, 235–258, 1992.
10. Richard Sproat, "Review of *The Syntax of the Modern Celtic Languages*." *Lingua*, 347–370, 1992.
11. 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. Richard Sproat, Tianyu Zhao and Llion Jones. "TransEvalnia: Reasoning-based Evaluation and Ranking of Translations", <https://arxiv.org/abs/2507.12724>. 2025.
2. Kyle Gorman and Richard Sproat. How to get superior text processing in Python with Pynini. O'Reilly. 2016.
<https://www.oreilly.com/ideas/how-to-get-superior-text-processing-in-python-with-pynini>.
3. M. Rouhizadeh, J. van Santen, R. Sproat. "Computational Semantic Analysis of Restrictive and Repetitive Behavior in Language Samples of Children with Autism" *IMFAR*, 2015, Salt Lake City.
4. M. Rouhizadeh, J. van Santen, R. Sproat, K. Gorman, P. Heeman, A. P. Hill, S. Bedrick, E. T. Prud'hommeaux and G. Kiss. "Children's Differing Patterns of Discourse Marker Use in ASD and Typical Development" *IMFAR*, 2014, Atlanta.

5. Richard Sproat, Lois Black, Emily Tucker-Prud'hommeaux, Jan van Santen, Brian Roark. "Automated Analysis of Natural Language Samples: Comparison of Children with Autism Spectrum Disorders, Developmental Language Disorders, and Typical Development." *IMFAR*, 2011, San Diego.
6. 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.
7. Bernd Möbius and Richard Sproat, "Linguistische Analyse in einem multilingualen Sprachsynthesystem." *Spektrum der Wissenschaften* 1996, (German Edition of Scientific American)
8. Richard Sproat, *On Deriving the Lexicon*. MIT Working Papers in Linguistics, Cambridge, MA. 1985. (Published version of Ph.D. dissertation)
9. Richard Sproat, "On Deriving the Lexicon," *The Linguistic Review* (Dissertation Abstracts), 5, 185–196, 1985.
10. 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.
11. 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.
12. 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.
13. Richard Sproat, "On Bracketing Paradoxes," in M. Speas and R. Sproat (eds.), *MIT Working Papers in Linguistics*, 7, 110–130, 1984.
14. Richard Sproat, "VSO Languages and Welsh Configurationality," in I. Haik and D. Massam (eds.), *MIT Working Papers in Linguistics*, 5, 243–276, 1983.
15. 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.
16. 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. Keynote speaker, EMNLP 2009, Singapore, August 6-7, 2009.
2. Invited lecturer, Winter School in Speech and Audio Processing, January, 2008, Chennai.
3. Invited speaker, Second Workshop on Computational Approaches to Arabic Script-Based Languages, Stanford University, July 21-22, 2007.
4. Invited Workshop on Text-to-Speech Synthesis, National University of Computer and Emerging Sciences, Lahore, Pakistan, June, 2005.
5. Keynote speaker, Second Midwest Computational Linguistics Colloquium, Ohio State University, May 14-15, 2005.
6. Invited speaker, PorTAL', Faro, Portugal, June 23-26, 2002.

7. Invited panelist, "Symposium on Finite-State Methods in Natural Language Analysis," American Academy of Sciences, Boston, MA, February 18, 2002.
8. Invited panelist at "Symposium on Distributed Morphology," 1994 Annual Meeting of the Linguistic Society of America, Boston, MA.
9. "Recent work on TTS at Bell Labs," Telecommunications Laboratories, Taiwan, August 1993.
10. "Computational Interpretations of Neurolinguistic Observations." McDonnell Summer Institute on Cognitive Neuroscience, July, 1993.
11. "Recent work on TTS at Bell Labs," Kiel University, Germany, October 1992.
12. "VEST: A Restricted-Domain Spoken Language Translation System." First International Workshop on Speech Translation, Schloß Dagstuhl, Germany, October 1992.
13. "Pitch Accent Prediction from Text Analysis," (with Julia Hirschberg) presented at the University of Illinois Workshop on Computational Linguistics, June 13-15, 1991.
14. "A pragmatic analysis of so-called anaphoric islands," (with Gregory Ward and Gail McKoon) University of Pennsylvania, November, 1990.
15. "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.
16. "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.
17. "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.
18. "Computational morphology," The Massachusetts Institute of Technology, April 1989.
19. "Connectionist Networks and Natural Language Morphology," (with Dania Egedi) University of Maryland Conference on Grammar and Language Processing, December, 1988.
20. "Connectionist Networks and Natural Language Morphology," (with Dania Egedi) SUNY Stonybrook, October 1987; University of Toronto, November 1987; Yale University, April 1988.
21. "The Analysis of English Compounds for Text-to-Speech," The Massachusetts Institute of Technology, May 1988.
22. "Constituent-Based Morphological Parsing," University of North Carolina, October 1986; Northwestern University, October 1986.
23. "On Bracketing Paradoxes: The Mapping Between S-Structure and PF," Workshop on Morphology and Modularity, University of Utrecht, June, 1986.
24. "More on Celtic Consonant Mutation," Kentucky Foreign Language Conference, University of Kentucky, Lexington, April, 1986.
25. "Constituent-Based Morphological Parsing," Bell Communications Research, Morristown, NJ, February 1985.
26. "Parsing Morphology," University of California at Irvine, October 1984; University of California at Los Angeles, October 1984.

27. "INFL and the Configurationality of VSO Languages," University of California at San Diego, October 1984; University of Southern California, October 1984.
28. "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

At various points in my life I have had 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

Many years' experience programming in C, C++, Python, Lisp.

C++ code I have written is running in production at Google, and in open-sourced software (<http://www.openfst.org/twiki/bin/view/GRM/Thrax>).

Many years' experience with Linux/Unix.

Knowledge of R.

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

Fellow of the Association of Computational Linguistics. 2012.

NIH Career Development Award (K25), 2011-14.

Best Short Paper Award (with Brian Roark and Izhak Shafran), ACL 2011.

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. Raised in San Diego, California, subsequently in Romsey, England, and finally again in San Diego.

REFERENCES

Available upon request.