

# Michael R. Brent

[brent@wustl.edu](mailto:brent@wustl.edu)  
314-286-0210

Campus Box 8810  
Washington University  
4444 Forest Park Blvd.  
Saint Louis, MO 63108

## Education

1991 Ph.D. in Computer Science, Massachusetts Institute of Technology  
1988 S.M. in Computer Science, Massachusetts Institute of Technology  
1985 S.B. in Mathematics, Massachusetts Institute of Technology

## Academic Appointments

2006 – The Henry Edwin Sever Professor of Engineering, Washington U.  
Joint in Biomedical Engineering and Genetics  
2004 – Professor in Computers Science, Washington University,  
Joint in Biomedical Engineering and Genetics  
1999 – 2004 Associate Professor in Computer Science, Washington University,  
Joint in Biomedical Engineering and Genetics  
1997 – 1999 Associate Professor in Cognitive Science, Johns Hopkins University  
1991 – 1997 Assistant Professor in Cognitive Science, Johns Hopkins University

## Academic Leadership

2008 – Associate Editor, EURASIP J. on Bioinformatics and Systems Biology  
2005 – Editorial Board, *BMC Bioinformatics*  
2004 – 2007 Editorial Board, *Genome Research*  
2003 – Chair, Bioinformatics Working Group, Mammalian Gene Collection  
2003 – Director, Genome Analysis Training Program (\$1.7M training grant)  
2000 – Director, Laboratory for Computational Genomics  
1997 – 1999 President, ACL Special Interest Group on Natural Language Learning  
1996 – 1999 Editorial board member and guest editor of *Cognition*

## Teaching and Curriculum Development

CS241: Algorithms and Data Structures (undergraduate core course) Taught fall 2000, spring 2001, and fall 2001. Total enrollment: ~260 students. Improved the match between the lecture material and the skills students are expected to develop.

CS511A: Artificial Intelligence I Taught in fall 2002 and two weeks in fall 2001.

CS587A: Algorithms for Computational Biology Taught spring 2002-2006, fall 2006-2007. Redesigned to focus on current technology in genome sequencing and analysis. Merged with Bio5495 in fall of 2006.

Bio4342: Research Explorations in Genomics (undergraduate) Taught one week covering gene prediction and genome annotation in spring 2003, 2004, and 2005.

Bio5495/BME 537 Computational Molecular Biology Taught one week covering gene-structure prediction in spring 2001, and fall 2001-2005.

Other Curriculum development at Washington University Helped design the Ph.D. program in Computational Biology in the Division of Biology and Biomedical Sciences. Continue to serve on the steering committee, mentor rotation students, teach Bio5495/CSE587A, and serve on oral qualifying and thesis committees. Completely redesigned the Genome Analysis Training Program (GATP, supported by an NIH training grant).

Curriculum development at Johns Hopkins University Played a leading role in redesign of the Ph.D. program in Cognitive Science; developed Formal Methods in Cognitive Science, a three-semester sequence, and taught one semester on formal approaches to artificial intelligence (logic and probability); designed and taught two undergraduate courses on artificial intelligence.

## **Research Supervision and Training**

### Post-doctoral fellows trained

Current: Dr. Jeltje van Baren, Dr. Sooraj Kuttykrishnan

Former: Dr. Randall Brown (supported by a mentored NIH F33 award, now Research Associate Professor at Washington University), Dr. Mikhail Velikanov (now Research Scientist at Uppsala University), Dr. Ian Korf (supported by \$1 million NIH faculty transition award, now Assistant Professor at the University of California, Davis), Dr. Jicheng “Daniel” Duan (now Senior Bioinformatics Engineer at Invitrogen Corporation), Dr. Xiaopeng Tao (now Assistant Professor at Fudan University), and Dr. Delphine Dahan (now Assistant Professor at University of Pennsylvania).

### Graduate researchers trained

Current Ph.D. students: Aaron Tenney (expected graduation 8/08), Brian Haynes, David Lu, Ezekiel Meier, Cole Johnson.

Ph.D.’s supervised: Timothy Cartwright, 1996 (JHU). Paul Flicek, 2004; Chaochun Wei, 2006; Evan Keibler, 2006.

M.S. research supervised: Matthew Snover (5/02), Chaochun Wei (5/02), Evan Keibler (8/03), Min Wang (12/03), David Shteynberg (6/04), Aaron Tenney (12/04), Robert Zimmermann (12/06), Suman Chilivaru (8/07), Akshat Shrivastava (5/08).

### Undergraduate researchers trained

Former (WU): Sam Gross (now a Ph.D. student in CS, Stanford University), Gaja Jarosz (now Asst. Prof. at Yale), Evan Keibler (at Microsoft Research), Phil Kirschner (now an IT consultant), Matthew Snover (now a Ph.D. student at

University of Maryland), Richard Souvenir (now faculty at NC State), Charlie Vaske (now a Ph.D. student at UCSC).  
Former (JHU): Sarah Carricaburu, Jamie Herslow (now Ph.D. student at Berkeley), Anna Kanevsky (a Ph.D. student at JHU), Kelly Lowe (now a Ph.D. student at SUNY Albany), Mathew Shomphe (now a Ph.D. student at University of Southern California), and Marni Soupcoff (graduated Stanford Law School).

## FUNDING

### Note

Dollar amounts are direct and indirect costs for the entire project period. If there are multiple investigators involved only the money for my group is shown in front; the total for all investigators is provided at the end in parentheses.

### **Funded Grants: Ongoing**

- 10/07 – 9/11 \$750,000 Integrated Human Genome Annotation: generation of a reference gene set, NIH, with T. Hubbard of The Sanger Institute. (~\$10 million)
- 5/07 – 4/11 \$2,217,000 Comprehensive characterization of the Drosophila transcriptome, NIH, with S. Celniker of Lawrence Berkeley Labs.
- 9/05 – 8/08 \$307,000 Developing Software for Protein Based Gene Finding, NIH, with M. Giddings of University of North Carolina, Chapel Hill
- 4/06 – 4/09 \$900,000 Gene-structure Prediction by Vertebrate Genome Comparison, funded by NIH (Human Genome)
- 1/02 – 12/10 \$921,000 Exploiting genome comparison for gene-structure prediction in plants, funded by Monsanto

### **Funded Grants: Completed**

- 9/05 – 8/07 \$229,000 Ab Initio Gene Finding in Maize, NSF, with W. Barbazuk of Danforth Plant Sciences Center (\$545,000)
- 9/03 – 7/06 \$157,000 The Encyclopedia of Genes and Gene Variants, funded by NIH (Human Genome), with R. Guigo of IMIM and others (\$1,484,000)
- 5/04 – 4/06 \$1,400,000 Bioinformatics for the Mammalian Gene Collection, NIH contract, with R. Gibbs of Baylor College of Medicine (\$1,500,000)
- 3/03 – 2/06 \$300,000 Identification of Antifungal Targets Using Proteomics, funded by NIH (Allergy and Infectious Diseases), with J Lodge of SLU (\$974,000)

- 4/02 – 3/06 \$915,000 Gene-structure Prediction by Vertebrate Genome Comparison, funded by NIH (Human Genome)
- 4/02 – 5/05 \$249,000 Comparative Gene-structure Prediction in Invertebrates, funded by NSF (Biology Directorate). \$30,000 supplement added in 2003.
- 3/01 – 2/03 \$150,000 Complete Gene-structure Prediction, funded by NSF (Biology Directorate)
- 7/01 – 3/02 \$20,000 Genome Analysis of *C. elegans*, funded by NIH (Human Genome) through the Washington University Genome Sequencing Center
- 7/01 – 3/02 \$70,000 Genome Analysis of *C. neoformans*, funded by NIH (Allergy and Infectious Diseases) through The Institute for Genomic Research
- 8/96 – 7/01 \$600,000 Mechanisms of Speech Segmentation and Word Discovery, funded by NIH (Deafness and Communication Disorders)
- 7/95 – 6/96 \$27,000 Segmentation and Word Identification in the Presence of Multiple Pronunciations, funded by ONR as through the JHU Center for Speech and Language processing
- 8/99 – 7/04 Problem-centered Research Training: Integrating Formal and Empirical Methods in the Cognitive Science of Language, funded by the NSF (IGERT program). Co-PI with P. Smolensky, B. Rapp, P. Jusczyk, and G. Legendre (\$4,110,000; remained at JHU)
- 7/97 – 8/00 Optimization in Language and Language Learning, funded by NSF (Learning and Intelligent Systems). Co-PI with P. Smolensky, R. Frank, P. Jusczyk, and G. Legendre (\$1,611,000; remained at JHU)

### **Pending Grants and contracts**

#### **Grants for training in my lab on which I am mentor & sponsor**

- 2/01 – 2/06 \$1,000,000 Comparative Genomics and Gene Prediction, funded by NIH (Human Genome). K22 to Ian Korf, Ph.D. Two years spent in my lab
- 12/02 – 12/04 \$165,000 Maximum Entropy Models for Gene Prediction, funded by NIH (Human Genome). F33 to Randall Brown, Ph.D., who spent it in my lab

### **PUBLICATIONS**

#### Note on change of research areas

In 2000 I began a new research program in computational biology after many years of working in computational linguistics. I have focused exclusively on computational biology since August, 2002. Publications for each field are presented separately.

## **Papers in peer-reviewed journals: Computational biology**

### Papers on which I am senior author

**Brent**, M.R. 2008. Steady progress and recent breakthroughs in the accuracy of automated genome annotation. *Nat Rev Genet* **9**: 62-73.

Siepel, A., M. Diekhans, B. Brejova, L. Langton, M. Stevens, C.L. Comstock, C. Davis, B. Ewing, S. Oommen, C. Lau,...M. **Brent**. 2007. Targeted discovery of novel human exons by comparative genomics. *Genome Res* **17**: 1763-1773.

**Brent**, M.R. 2007. How does eukaryotic gene prediction work? *Nat Biotechnol* **25**: 883-885.

Tenney, A.E., J.Q. Wu, L. Langton, P. Klueh, R. Quatrano, and M.R. **Brent**. 2007. A tale of two templates: automatically resolving double traces has many applications, including efficient PCR-based elucidation of alternative splices. *Genome Res* **17**: 212-218.

Keibler, E., M. Arumugam, and M. **Brent**. 2007. The Treeterbi and Parallel Treeterbi algorithms: Efficient, optimal decoding for ordinary, generalized, and Pair HMMs. *Bioinformatics* **23**: 545-554.

Arumugam, M., C. Wei, R.H. Brown, and M.R. **Brent**. 2006. Pairagon+N-SCAN\_EST: A Model-based Gene Annotation Pipeline. *Genome Biol*: **7 Suppl 1**: p. S5 1-10.

Flicek, P. and M.R. **Brent**. 2006. Using several pair-wise informant sequences for de novo prediction of alternatively spliced transcripts. *Genome Biol*: **7 Suppl 1**: p. S8 1-9.

Gross, S.S. and M.R. **Brent**. 2006. Using Multiple Alignments To Improve Gene Prediction. *Journal of Computational Biology* **13**: 379-393.

van Baren, M.J. and M.R. **Brent**. 2006. Iterative gene prediction and pseudogene removal improves genome annotation. *Genome Res* **16**: 678-685.

Wei, C. and M.R. **Brent**. 2006. Using ESTs to improve the accuracy of de novo gene prediction. *BMC Bioinformatics* **7**: 327.

Brown, R.H., S.S. Gross, and M.R. **Brent**. 2005. Begin at the Beginning: Predicting Genes with 5' UTRs. *Genome Res* **15**: 742-747.

Wei, C., P. Lamesch, M. Arumugam, J. Rosenberg, P. Hu, M. Vidal, and M.R. **Brent**. 2005. Closing in on the *C. elegans* ORFeome by Cloning TWINSKAN predictions. **Genome Res** **15**: 577-582.

Eyras, E., A. Reymond, R. Castelo, J.M. Bye, F. Camara, P. Flicek, E.J. Huckle, G. Parra, D.D. Shteynberg, C. Wyss, J. Rogers, S.E. Antonarakis, E. Birney, R. Guigo, and M.R. **Brent**. 2005. Gene finding in the chicken genome. **BMC Bioinformatics** **6**: 131.

Tenney, A.E., R.H. Brown, C. Vaske, J.K. Lodge, T.L. Doering, and M.R. Brent. 2004. Gene prediction and verification in a compact genome with numerous small introns. **Genome Res** **14**: 2330-2335.

Wu, J.Q., D. Shteynberg, M. Arumugam, R.A. Gibbs, and M.R. **Brent**. 2004. Identification of rat genes by TWINSKAN gene prediction, RT-PCR, and direct sequencing. **Genome Res** **14**: 665-671.

Flicek, P., E. Keibler, P. Hu, I. Korf, and M.R. **Brent**. 2003. Leveraging the mouse genome for gene prediction in human: From whole-genome shotgun reads to a global synteny map. **Genome Res** **13**: 46-54.

Guigó, R., E.T. Dermitzakis, P. Agarwal, C. Ponting, G. Parra, A. Reymond, J.F. Abril, E. Keibler, R. Lyle, C. Ucla, S.E. Antonarakis, and M.R. **Brent**. 2003. Comparison of mouse and human genomes followed by experimental verification yields an estimated 1,019 additional genes. **Proc Natl Acad Sci USA** **100**: 1140-1145.

Keibler, E. and M.R. **Brent**. 2003. Eval: A software package for analysis of genome annotations. **BMC Bioinformatics** **4**: 50. (Designated a "Highly accessed" paper by BioMed Central.)

Korf, I., P. Flicek, D. Duan, and M.R. **Brent**. 2001. Integrating genomic homology into gene structure prediction. **Bioinformatics** **17**: S140-148.

#### Papers on which I am a contributing author

Cahan, P., L.E. Godfrey, P.S. Eis, T.A. Richmond, R.R. Selzer, M. **Brent**, H.L. McLeod, T.J. Ley, and T.A. Graubert. 2008. wuHMM: a robust algorithm to detect DNA copy number variation using long oligonucleotide microarray data. **Nucleic Acids Res** **36**: e41.

Clark, A.G. M.B. Eisen D.R. Smith C.M. Bergman B. Oliver T.A. Markow T.C. Kaufman M. Kellis W. Gelbart V.N. Iyer..... **M.R. Brent** et al. 2007. Evolution of genes and genomes on the *Drosophila* phylogeny. **Nature** **450**: 203-218.

Birney, E. J.A. Stamatoyannopoulos A. Dutta R. Guigo T.R. Gingeras E.H. Margulies Z. Weng M. Snyder E.T. Dermitzakis... **M.R. Brent** et al. 2007. Identification

- and analysis of functional elements in 1% of the human genome by the ENCODE pilot project. *Nature* **447**: 799-816.
- Vermeirssen, V., B. Deplancke, M.I. Barrasa, J.S. Reece-Hoyes, H.E. Arda, C.A. Grove, N.J. Martinez, R. Sequerra, L. Doucette-Stamm, **M.R. Brent**, and A.J. Walhout. 2007. Matrix and Steiner-triple-system smart pooling assays for high-performance transcription regulatory network mapping. *Nat Methods* **4**: 659-664.
- Bajic, V.B., M.R. **Brent**, R.H. Brown, A. Frankish, J. Harrow, U. Ohler, V.V. Solovyev, and S.L. Tan. 2006. Performance assessment of promoter predictions on ENCODE regions in the EGASP experiment. *Genome Biol* **7 Suppl 1**: S3 1-13.
- Giannakis, M., T.S. Stappenbeck, J.C. Mills, D.G. Leip, M. Lovett, S.W. Clifton, J.E. Ippolito, J.I. Glasscock, M. Arumugam, M.R. **Brent**, and J.I. Gordon. 2006. Molecular properties of adult mouse gastric and intestinal epithelial progenitors in their niches. *J Biol Chem*.
- Loftus, B.J., et al. (including M.R. **Brent** as author). 2005. The Genome of the Basidiomycetous Yeast and Human Pathogen *Cryptococcus Neoformans*. *Science* **307**: 1221-1324.
- The International Chicken Genome Sequencing Consortium (including M.R. **Brent** as author). 2004. Sequence and comparative analysis of the chicken genome provide unique perspectives on vertebrate evolution. *Nature* **432**: 695-716.
- The Rat Genome Sequencing Consortium (including M.R. **Brent** as author). 2004. Evolution of the Mammalian Genome: Sequence of the Genome of the Brown Norway Rat. *Nature*, **428**: 493-521.
- The MGC Project Team (including M.R. **Brent** as author). 2004. The Status, Quality, and Expansion of the NIH Full-Length cDNA Project: The Mammalian Gene Collection (MGC). *Genome Res* **14**: 2121-2127.
- Stein L.D., Z. Bao, D. Blasiar, T. Blumenthal, M.R. **Brent**, et al. 2003. The Genome Sequence of *Caenorhabditis briggsae*: A Platform for Comparative Genomics. *PLoS Biology* **1**:166-192
- Hillier, L.W. et. al. 2003. The DNA Sequence of Human Chromosome 7. *Nature* **424**: 157-164
- Waterston, R.H. K. Lindblad-Toh E. Birney J. Rogers J.F. Abril P. Agarwal R. Agarwala R. Ainscough M. Alexandersson P. An S.E. Antonarakis J. Attwood R. Baertsch J. Bailey K. Barlow S. Beck E. Berry B. Birren T. Bloom P. Bork M. Botcherby N. Bray M.R. **Brent**, et al., 2002. Initial sequencing and comparative analysis of the mouse genome. *Nature* **420**: 520-562. I led the

comparative gene prediction group and wrote one journal page of the text.  
My contribution is acknowledged in "Authors' Contributions"

### **Papers in peer-reviewed journals: Computational linguistics**

- Ganger, J., and M.R. Brent (2004). Re-Examining the Vocabulary Spurt. *Developmental Psychology*, **40**:621-632.
- Brent, M.R. and J.M. Siskind. 2001. The role of exposure to isolated words in early vocabulary development. *Cognition* **81**: B33-44
- Brent, M.R. 1999. Speech segmentation and word discovery: a computational perspective. *Trends Cogn Sci* **3**: 294-301.
- Dahan, D. and M.R. Brent. 1999. On the discovery of novel wordlike units from utterances: an artificial-language study with implications for native-language acquisition. *J Exp Psychol Gen* **128**: 165-185.
- Cartwright, T.A. and M.R. Brent. 1997. Syntactic categorization in early language acquisition: formalizing the role of distributional analysis. *Cognition* **63**: 121-170.
- Brent, M.R. 1996. Advances in the computational study of language acquisition. *Cognition* **61**: 1-38.
- Brent, M.R. and T.A. Cartwright. 1996. Distributional regularity and phonotactic constraints are useful for segmentation. *Cognition* **61**: 93-125.
- Brent, M.R. 1994. Surface cues and robust inference as a basis for the early acquisition of subcategorization frames. *Lingua* **92**: 433-470.
- Brent, M.R. 1993. From grammar to lexicon: Unsupervised learning of lexical syntax. *Computational Linguistics* **19**: 243-262.

### **Papers in review at peer-reviewed journals**

#### **Review articles and book chapters**

- Brent**, M.R. 2005. Genome annotation past, present and future: How to define an ORF at each locus. *Genome Res* **15**: 1777-1786.
- Brent**, M.R. and R. Guigo. 2004. Recent advances in gene structure prediction. *Curr Opin Struct Biol* **14**: 264-272.
- Wang, M., J. Buhler, and M.R. **Brent**. 2004. The effects of evolutionary distance on TWINSCAN, an algorithm for pairwise comparative gene prediction. In *68<sup>th</sup>*

*Symposium: The Genome of Homo Sapiens* (pp. 125-130), Cold Spring Harbor, NY: Cold Spring Harbor Press.

Hu, P. and M.R. **Brent**. 2003. Using TWINSKAN to Predict Gene Structures in Genomic DNA Sequences. In *Current Protocols in Bioinformatics* (ed. A.D. Baxevanis). New York: Wiley.

**Brent**, M.R. 2002. Predicting full-length transcripts. *Trends Biotechnol* **20**: 273-275.

**Brent**, M.R. 1996. Advances in the computational study of language acquisition. Reprinted in M.R. **Brent** (Ed.), *Computational Approaches to Language Acquisition*. Cambridge, MA:MIT Press. Reprinted from *Cognition*.

**Brent**, M.R., and T.A. Cartwright. 1996. Distributional regularity and phonotactics are useful for segmentation. Reprinted in M.R. **Brent** (Ed.), *Computational Approaches to Language Acquisition*. Cambridge, MA:MIT Press. Reprinted from *Cognition*.

**Brent**, M.R. 1994. Surface cues and robust inference as a basis for the early acquisition of subcategorization frames. Reprinted in L. Gleitman and B. Landau (Eds.), *The Acquisition of the Lexicon* (pp. 433--471). Cambridge, MA:MIT Press. Reprinted from *Lingua*.

**Brent**, M.R. 1993. From grammar to lexicon: Unsupervised learning of lexical syntax. Reprinted in S. Armstrong (Ed.) *Using Large Corpora* (pp. 203-22). Cambridge, MA: MIT Press. Reprinted from *Computational Linguistics*.

### **Papers in peer-reviewed archival conference proceedings**

#### Note on publication in peer-reviewed archival conference proceedings

Papers in this section are 3000-4000 words, reviewed by at least two referees, and printed in archival annual proceedings. Associated talks are not listed separately.

Gross, S.S. and M.R. **Brent**. 2005. Using Multiple Alignments To Improve Gene Prediction. In *9th Annual International Conference, RECOMB 2005* (eds. S. Miyano S. Kasif P. Pevzner J. Mesirov S. Istrail, and M. Waterman), pp. 374-388. Springer, Boston.

**Brent**, M.R. and X.P. Tao. 2001. Chinese Text Segmentation with MBDP-1: Making the most of training corpora. In *Proceedings of the Association for Computational Linguistics*, **39**: 82-89.

Snover, M.G. and M.R. **Brent**. 2001. A Bayesian Model for Morpheme and Paradigm Discovery. *Proceedings of the Association for Computational Linguistics*, **39**: 482-490.

- Hollich, G., P.W. Jusczyk, and M.R. **Brent**. 2000. How infants use the words they know to learn new words. In *Proceedings of the 24<sup>th</sup> Annual Boston University Conference on Language Development*, 353-364.
- Snedeker, J., L.R. Gleitman and M.R. **Brent**. 1998. The successes and failures of word-to-world mapping. In *Proceedings of the 23<sup>rd</sup> Annual Boston University Conference on Language Development*, 654-665.
- Brent**, M.R., and T. A. Cartwright. 1996. Lexical categorization: Fitting template grammars by incremental MDL optimization. In *Grammatical inference: ICGI96*, 84-94.
- Brent**, M.R., S. K. Murthy, and A. Lundberg. 1995. Discovering morphemic suffixes: A case study in minimum description length induction. In *Proceedings of the Fifth International Workshop on Artificial Intelligence and Statistics*.
- Brent**, M.R., A. Gafos, and T.A. Cartwright. 1994. Phonotactics and the lexicon: Beyond bootstrapping. In *Proceedings of the Twenty-sixth Stanford Child Language Research Forum*, 11-21.
- Cartwright, T. A., and M.R. **Brent**. 1994. Segmenting speech without a lexicon: Evidence for a bootstrapping model of lexical acquisition. In *Proceedings of the Sixteenth Annual Conference of the Cognitive Science Society*, 148-152.
- Cartwright, T. A., and M.R. **Brent**. 1994. Segmenting speech without a lexicon: The roles of phonotactics and speech source. In *Proceedings of the First Meeting of the ACL Special Interest Group in Computational Phonology*, 83 -91.
- Brent**, M.R. 1993. Minimal generative models: A middle ground between neurons and triggers. In *Proceedings of the Fifteenth Annual Conference of the Cognitive Science Society*, 28 -36.
- Brent**, M.R. 1991. Automatic acquisition of subcategorization frames from untagged text. In *Proceedings of the 29th Annual Meeting of the ACL*, 209-214.
- Brent**, M.R. 1991. Semantic classification of verbs from their syntactic contexts: An implemented classifier for stativity. In *Proceedings of the 5th European ACL Conference*, 222-226.
- Brent**, M.R. 1990. A simplified theory of tense representations and their composition. In *Proceedings of the 28th Annual Meeting of the Association for Computational Linguistics*, 119-126.
- Brent**, M R. 1990. Semantic classification of verbs from their syntactic contexts: Automated lexicography with implications for child language acquisition. In

*Proceedings of the 12th Annual Conference of the Cognitive Science Society, 426-437.*

**Brent, M.R.** 1989. Earning dividends on lexical knowledge: How the rich can get richer. In *Proceedings of the First Annual Workshop on Lexical Acquisition, International Joint Conference on Artificial Intelligence.*

**Brent, M.R.** 1989. Temporal/Causal connectives: Syntax and lexicon. In *Proceedings of the 11th Annual Conference of the Cognitive Science Society, Cognitive Science Society, 90-98.*

**Brent, M.R.** 1989. Why can't I smear paint at the wall. In *Proceedings of the Sixth Israeli Conference on Artificial Intelligence.*

**Brent, M.R.** 1989. Temporal/Causal connectives: Syntax and lexicon. In *Proceedings of the 11th Annual Conference of the Cognitive Science Society, Cognitive Science Society, 90-98.*

## **SCIENTIFIC PRESENTATIONS: COMPUTATIONAL BIOLOGY**

### **Colloquium series and keynote addresses**

Penn Bioinformatics Forum, University of Pennsylvania, 2005  
New York University, 2006  
University of North Carolina, 2007  
Washington University Genetics Colloquium, 2008

### **Ad hoc seminars, workshops and conferences, and other external invited talks**

Monsanto research review (speaker and external reviewer), 2004  
Annual Computational Genomics conference, MIT, 2003  
Third Cryptococcus neoformans Genome Meeting, Vancouver, 2003  
68<sup>th</sup> Symposium: The Genome of *Homo Sapiens*, Cold Spring Harbor, 2003  
Genoscope, Paris, France, 2003  
La Caixa human genome symposium, Barcelona, Spain, 2003  
Monsanto-Washington University research symposium, 2003  
Cold Spring Harbor course on Computational Genomics, 2003  
Baylor College of Medicine, Houston, 2003  
Advances in Genome Biology and Biotechnology, 2002  
Genomic Signal Processing and Statistics (GENSIPS), 2002  
The Institute for Systems Biology, Seattle, 2002  
Second annual ORFEOME meeting, Harvard, 2002  
Monsanto-Washington University research symposium, 2002  
Workshop on analysis of the mouse genome, MIT, 2001  
NSF Workshop on Language Modeling of Biological Data, 2001  
Second Cryptococcus neoformans Genome Meeting, 2001

## **Selected conference talks based on refereed short abstracts**

Cold Spring Harbor Genome Sequencing and Biology Meeting (2002); Cold Spring Harbor Genome Sequencing and Biology Meeting (2001), Wellcome Trust Conference on Genome Informatics Meeting (2001, given by P. Flicek); Annual Meeting of the Biomedical Engineering Society (2001, given by P. Flicek)

## **SCIENTIFIC PRESENTATIONS: COMPUTATIONAL LINGUISTICS**

### **2001**

Max Planck Institute Workshop on Speech Recognition; AAAS, Symposium on Tools Infants Use to Learn Language; Society for Research in Child Development, Bayesian modeling symposium; Linguistics Society of America, Symposium on Statistical Methods

### **2000**

Language Development in Ontogeny, Max Plank Institute for Anthropology, Germany; Bayesian Modeling Symposium, Annual Meeting of the Cognitive Science Society; Finding the Words workshop, Stanford University; Institute for Mathematical Statistics, Computational Linguistics Symposium

### **1999**

Massachusetts Institute of Technology; Stanford; University of Chicago; Northwestern University; McGill University; Brown University; University of Michigan; Washington University; University of Southern California; Georgia Institute of Technology; University of New Mexico; University of Illinois, Urbana-Champaign; ICML Workshop on Automata Induction; Annual meeting of the Cognitive Science Society

### **1997**

Columbia University; New York University; University of Pennsylvania; University of Maryland; Rutgers University

### **1996**

Cambridge University, England; University of Edinburgh, Scotland; Katholieke Universiteit Brabant, the Netherlands; Max Plank Institute for Psycholinguistics, the Netherlands; University of Massachusetts, Amherst; City University of New York Graduate Center.

### **1995**

University of Pennsylvania; University of Delaware; University of Maryland at College Park; University of Toronto; NEC Research Institute; Princeton, NJ; Consequences of Phonological Diversity meeting, Trieste, Italy

## **International Speaking Invitations**

Note: These international invitations paid all expenses. Also listed in "Invited Talks".

Max Planck Institute for Psycholinguistics workshop, Nijmegen, the Netherlands  
Max Plank Institute for Anthropology, Leipzig, Germany  
NSF planning workshop for joint EAGLES/NSF funding initiative  
Visiting Scientist, Max Planck Institute for Psycholinguistics, Nijmegen, the Netherlands  
Cambridge University, Cambridge, England  
Consequences of Phonological Diversity workshop, Trieste, Italy

### **Selected conference talks based on refereed short abstracts**

Stanford Child Language Research Forum (1999); Computational Psycholinguistics (1997); Architectures and Mechanisms of Language Processing (1996); CUNY Conference on Human Sentence Processing (1996); Workshop on Cognitive Models of Language Acquisition (1994); Boston University Conference on Language Development (1994)

### **OTHER SERVICE AND RECOGNITION**

#### **Attention to Scientific Results by the Popular Press and Scientific Press**

After giving a talk and news conference on mothers' speech to infants at the annual meeting of the American Association for the Advancement of Science, research results were cited in the front page of the St. Louis Post Dispatch (2/20/01) Time magazine (2/28/01) The San Francisco Examiner (2/20/01) The San Francisco Chronicle (2/20/01) The London Daily Telegraph (2/20/01) The Ottawa Citizen (2/21/01) Vancouver Sun (2/21/01), Voice of America, Healthscout.com, and ScienceDaily.com

A 2003 Viewpoint in *Science* (287: 1977-1978) covered my gene prediction work and quoted me.

#### **Contributions to Public Research Infrastructure**

Provide a web server for our TWINSCAN gene-structure prediction software (<http://genes.cse.wustl.edu/>) to which about 100 genomic sequences are submitted each week.

Provide open source software for gene prediction and genome annotation.

Recorded and transcribed the largest existing database of mothers' spontaneous speech to their infants. This database is available through the Child Language Data Exchange System (<http://childes.psy.cmu.edu/data/audlinked/>).

#### **Contributions to Conference organization**

Steering Committee, International Conference on Grammar Induction

## Program Committees

Workshop on Unsupervised Learning in Natural Language Processing 1999  
International Conference on Grammar Induction 1996 and 1998  
Computational Psycholinguistics meeting 1997  
International Conference on Spoken Language Processing 1996

## Ad hoc reviewing (selected journals)

Biology Science, Nature, Nature Genetics, Nature Biotechnology, PLOS Biology, Genome Research, Nucleic Acids Research, PLoS Computational Biology, Bioinformatics, BMC Bioinformatics, Genomics, Trends in Genetics, Briefings in Bioinformatics, Gene

Cognitive Psychology Science, Child Development, Cognitive Psychology, Cognitive Science, Journal of Experimental Psychology: General, Journal of Memory and Language, Perception and Psychophysics

Artificial Intelligence Computational Linguistics, Machine Learning,

## Service for Funding Agencies

Panelist, NHGRI workshop on selecting genomes to sequence  
Panelist, NHGRI workshop on the proposed Encode program  
Panelist, NHGRI workshop to develop a new 5-year plan for the Institute  
Chair, Bioinformatics Working Group of the Mammalian Gene Collection project  
Panelist for the National Science Foundation  
Referee for the Indiana 21st Century Research and Technology Fund  
Referee for the Human Frontiers Foundation  
Referee for Fundacion BBVA (a Spanish Foundation)  
European Advisory Group for Language Engineering Standards (EAGLES)

## Service at Washington University

2007 -- Chair, Computer Science & Engineering Leadership Committee  
2006 -- Committee on admission of faculty to the Division of Biological and Biomedical Sciences (DBBS) umbrella graduate program  
2000 – Steering committee for the Computational Biology Ph.D. program in the Division of Biology and Biomedical Sciences  
2000 – Faculty Search Committee for Biomedical Engineering  
2000 – Interim co-chair of the Computational Biology Ph.D. program in Biomedical Engineering  
2001 – Faculty Search Committee for Computer Science  
2000 – 2002 Outreach and Continuing Education Committee, Computer Science  
2001 – 2002 Chair, Ph.D. Program Committee for Computer Science  
2000 – 2001 Member, Computer Science Ph.D. Program Committee  
2000 – 2001 Chair of the Computer Science Committee on Graduate Recruitment