

## Daniel P. W. Ellis - Curriculum Vitæ

### *office address:*

Schapiro CEPSR room 718  
Department of Electrical Engineering  
Columbia University  
New York, NY 10027

*phone:* (212) 854-8928  
*fax:* (815) 366-8005  
*email:* dpwe@ee.columbia.edu  
*web:* <http://www.ee.columbia.edu/dpwe/>

### *Research interests:*

- Signal processing and machine learning for analysis of general audio, speech, and music.
- Audio source separation in underconstrained conditions.
- Computational models of human sound processing and organization.
- Automatic speech recognition in adverse environments.
- Visualization and browsing tools for audio and speech databases.

### *Positions held:*

**Associate Professor of Electrical Engineering**, Columbia University, NY Jul 2005-present  
**Assistant Professor of Electrical Engineering** Sep 2000-Jun 2005

Leading LabROSA, the laboratory for the Recognition and Organization of Speech and Audio (<http://labrosa.ee.columbia.edu/>), investigating all aspects of intelligent sound processing from auditory scene analysis to robust speech recognition to music recommendation. Teaching includes introductory Digital Signal Processing and Speech and Audio Processing.

**Senior Research Scientist**, ICSI, Berkeley CA May 1996-Aug 2000

Researching automatic speech recognition robustness with Prof. Nelson Morgan, including applications of perceptual models. Advised graduate students and managed grants.

**Research Assistant**, M.I.T. Media Lab Perceptual Computing, Cambridge MA 1989-1996

Researching audio signal processing, sound synthesis and analysis under Prof. Barry L. Vercoe, while pursuing degrees at M.I.T.

**Intern**, Interval Research Corporation, Palo Alto CA June-August, 1994

Developed sound analysis/synthesis techniques based on auditory models.

**Technical staff**, AWARE Inc., Cambridge MA 1991-1993

Research and development of psychoacoustic-based compression schemes for high-quality audio including interactive tools. Contributed to MPEG audio 'committee code'.

### *Degrees:*

**Ph.D.**, Dept. of Elec. Eng. and Comp. Science, M.I.T. June 1996

Dissertation title: "A prediction-driven model of computational auditory scene analysis," advisor Prof. Barry L. Vercoe, readers Prof. Louis Braidai, Dr. Ben Gold.

**Master of Science**, Dept. of Elec. Eng. and Comp. Science, M.I.T. February 1992

Thesis title: "A perceptual representation of audio," advisors Prof. Barry L. Vercoe and Dr. Thomas F. Quatieri.

**Bachelor of Arts** (honors), Dept. of Engineering, Cambridge University, U.K. June 1987

Electrical & Information Sciences Tripos. First class with distinction (ranked top). Final year project: A microcomputer-based linear predictive coding system for musical applications.

*Projects at LabROSA:*

**Data-Driven Music Understanding** (0713334) Sep 2007–Aug 2010  
(PI Ellis, \$450,000 from NSF IIS to LabROSA)

PI on collaboration with Columbia’s Computer Music Center to analyze the high-level content of music audio with a view to discovering the latent structure that ‘defines’ music. Also includes education outreach component to local middle and high schools.

**Trustworthy Media: Media Forensics for Content Integrity Verification** Sep 2007–Aug 2010  
(PI Chang (Columbia); \$350,000 from NSF; LabROSA portion \$173,000)

Collaboration with Digital Video/Multimedia group on analyzing audio-video recordings to detect possible forgeries. Approaches focus on synchrony and correlation between audio and video elements, as well as ‘device signatures’ in the audio stream.

**Separating Speech from Speech Noise** Jan 2006–Dec 2008  
(PI Ellis; \$747,000 from NSF IIS; LabROSA portion \$180,000)

PI on cross-disciplinary collaboration between engineers and psychologists seeking to apply signal separation in real-world conditions with demonstrable improvements for human listeners. Involves rotating students between engineering and psychology labs, including a collaborator in Paris.

**Audio LifeLogs** Jan 2006–Dec 2006  
(PI Ellis; \$50,000 from Microsoft Research)

Gift plus specialized equipment and software to develop techniques for gathering and indexing everyday personal recordings.

**NIGHTINGALE** Sep 2005 - Aug 2010  
(PI Israel (SRI); \$10,000,000 from DARPA GALE; LabROSA portion \$563,976)

Large consortium addressing the Global Autonomous Language Exploitation requirements, spanning speech recognition, machine translation, information extraction, access and presentation. LabROSA will collaborate with Hirschberg in CS to identify prosodic factors to support translation and extraction e.g. phrase segments, stressed words.

**Music Engineering Art Collaboration** Oct 2004 - Aug 2007  
(PI Ellis; \$332,319 from Columbia Academic Quality Fund; LabROSA portion \$173,970)

Interdisciplinary collaboration with Columbia Computer Music Center to develop music signal analysis and organization tools, teach a mixed engineering/art projects class, and establish a more permanent entity, using multimedia art installations as a motivating goal.

**Music Similarity Modeling** Sep 2003 - Aug 2005  
(PI Ellis; \$63,494 from Google, Inc.)

Sponsorship of one student looking at automatic analysis of music signals to predict listener preference, with the goal of developing intelligent music browsing and recommendation for little-known music.

**The Listening Machine: Sound Organization for Multimedia** Feb 2003 - Jan 2008  
(PI Ellis; \$558,207 from NSF Career program and Columbia SEAS.)

Project to develop intelligent analysis of real-world sound mixtures in terms of their component sources using techniques drawn from speech recognition, machine learning, and human auditory modeling. Current results include a novel analysis, the “spectral deformation model”, with applications in speech recognition and source separation.

**Pushing the Envelope: Rethinking Speech Signal Representation** Apr 2002 - Aug 2005  
(PI Morgan; Ellis co-PI; \$6,000,000 from DARPA EARS; LabROSA portion \$729,430)

Large consortium to develop “novel approaches” to dislodge speech recognition from its diminishing incremental returns. Our contribution has been a novel model for temporal energy modulations in the

speech signal (based on the dual of linear prediction), and current work includes methods to transform informal, hasty speech into more careful pronunciation.

**Audio Signal Analysis**

Oct 2001 - Aug 2003

(PI Ellis; \$104,626 from NEC Research Lab)

Sponsorship of one student working on models for music similarity and developing and evaluating a browsing tool to navigate large music collections based solely on signal properties rather than manually-provided metadata.

**Mapping Meetings: Language Technology for Natural Interactions**

Sep 2001 - May 2005

(PI Morgan; Ellis co-PI; \$1,402,851 from NSF ITR program; LabROSA portion \$125,674)

Project to analyze a large corpus of natural meetings resulting from a project I set up while at ICSI. Our work included a variety of techniques to analyze these recordings, including locating interesting events (stressed utterances, laughter), recovering speaker turns from multiple distant mic channels, and inferring ‘talkativity’ models to characterize each participant.

*Teaching:*

Spring 2008:

**Speech and Audio Processing and Recognition** (ELEN E6820, 9 students, in progress)

Latest version of my novel course straddling the LabROSA research interests. Weekly meetings include some traditional lecturing, paper presentation by student participants, and free-discussion ‘brainstorming’ on an example practical problem relevant to that week’s topic. 50% of grade from semester project; part of grade comes from peer assessment of project presentations. All the materials for this course are freely available at <http://www.ee.columbia.edu/~dpwe/e6820/>; I have had numerous messages of thanks from many sources who have downloaded them.

Fall 2007:

**Digital Signal Processing** (ELEN E4810, 67 students, overall student rating 4.4/5)

Fundamental senior/masters level class emphasizing a combined mathematical and intuitive understanding of the Fourier domain, and providing tools for filter design. Extensive use of live in-class demos; 30% of grade on individual final projects; all notes, demos, and other materials are available online: <http://www.ee.columbia.edu/~dpwe/e4810/>.

**Music Engineering Art Project** (8 students)

This project class is our ongoing activity within the EMAC project described above. With an even balance of engineering and music students, all of whom are passionate about both technology and music, we are developing interactive audio/music analysis/synthesis systems, including MEAPsoft. <http://labrosa.ee.columbia.edu/meapsoft/>

Spring 2007:

**Speech and Audio Processing and Recognition** (ELEN E6820, 15 students, overall student rating 4.3/5)

Relatively large enrollment with several extremely good projects; several went on to be published.

**Music Engineering Art Project** (10 students)

Spent a lot of time discussing the user interface to MEAPsoft, what core tasks it is or could be used for, and ideas for how to improve it; one student was simultaneously taking a user interface class in CS.

**Music Signal Processing** (ELEN E4998/E4896, 12 students)

By popular demand, we ran a version of this class originally created by Prof. Eleftheriadis who was on sabbatical. Students presented papers on digital audio music effects algorithms, analyzed from an engineering perspective, and demonstrated and discussed practical implementation. Final project involved applying specially-developed effects algorithm to a real musical multitrack mixing task.

Fall 2006:

**Digital Signal Processing** (ELEN E4810, 50 students, overall rating 3.8/5)

Improved use of online tools such as discussion boards and online videos to help connect and empower students.

**Music Engineering Art Project** (7 students)

This project class is our ongoing activity within the EMAC project described above. With an even balance of technologically-oriented engineering and music students, we are developing interactive audio/music analysis/synthesis systems. We are planning an interactive art installation on the Columbia campus using our MEAPsoft software (see below).

Spring 2006:

**Speech and Audio Processing and Recognition** (ELEN E6820, 10 students, overall student rating 4.8/5)

With a smaller than average enrollment, we experimented with a much more interactive classroom format, devoting about half of each week to discussion and brainstorming.

**Music Engineering Art Project** (10 students)

After one year of meeting for this project, we finally got down to building something - the experimental audio/music reorganization package, MEAPsoft version 1.0, was released over the summer. <http://labrosa.ee.columbia.edu/meapsoft/>

Fall 2005:

**Digital Signal Processing** (ELEN E4810, 68 students, overall student rating 3.4/5)

Continuing revisions to content and ordering to improve intelligibility.

**Music Engineering Art Project** (9 students)

Follow-on project class incorporated more structure, with each student (and both faculty!) responsible for making a presentation one week during the semester.

Spring 2005:

**Speech and Audio Processing and Recognition** (ELEN E6820, 20 students; rating 4.4/5)

Added more material on signal separation and large audio database analysis based on our recent research work. Half the grade comes from a final project, which amounts to a mini independent semester project with each student.

**Music Engineering Art Project** (8 students)

Our first activity within the EMAC project. Initial investigation into practicality and issues of running a class made up of both engineering and music graduate students.

**Machine Learning Signal Processing Seminar** (5-15 students)

I manage this weekly informal reading group for the Ph.D. students in the signal processing area. This semester we looked at neurophysiological underpinnings of perception and some other novel machine learning techniques.

Fall 2004:

**Digital Signal Processing** (ELEN E4810, 50 students, overall student rating 3.5/5)

Fewer students than average; some sessions extensively revised to improve flow of ideas.

**Machine Learning Signal Processing Seminar** (5-15 students)

In-depth coverage of normalized cuts and spectral clustering theory and applications.

Spring 2004:

**Speech and Audio Processing and Recognition** (ELEN E6820, 9 students, rating 4.5/5)

Since it was offered the same semester as a new adjunct class on speech recognition, I added new units on signal separation and musical signal analysis to replace the earlier detail on speech recognition.

**Machine Learning Signal Processing Seminar** (ELEN E9701, 8 registered students)

Most of the semester was spent working through the Scholkopf/Smola book on kernel methods. We ran

the reading group as a formal course number with registration as an experiment.

Fall 2003:

**Digital Signal Processing** (ELEN E4810, 72 students, rating 3.5/5)

Largest ever enrollment; now a required course for the Biomedical Engineering Imaging MS.

Spring 2003:

**Speech and Audio Processing and Recognition** (ELEN E6820, 23 students, rating 4.0/5)

One senior undergrad went on to publish his final project with me at an IEEE conference.

**Musical Content Analysis by Machine Learning** (9 hour short course, 20 students)

A one-week seminar taught by invitation at the Music Technology Group of Barcelona's Pompeu Fabra University (UPF), combining pattern recognition background with applications in music analysis. Also a one-day session at the Johns Hopkins University Center for Spoken Language Processing Summer School, for which I created a self-paced practical. Materials available at: <http://www.ee.columbia.edu/~dpwe/muscontent/>

Fall 2002:

**Digital Signal Processing** (ELEN E4810, 58 students, rating 4.0/5)

Moved entire course over to Powerpoint after complaints about my handwriting.

Spring 2002:

**Speech and Audio Processing and Recognition** (ELEN E6820, 12 students, rating 4.3/5)

Rearranged material and added a new section on spatial audio.

Fall 2001:

**Digital Signal Processing** (ELEN E4810, 42 students, rating 3.8/5)

Course open to distance students (watching videos of lectures) for the first time.

Spring 2001:

**Speech and Audio Processing and Recognition** (ELEN E6820, 11 students, rating 3.9/5)

First offering of this new course which I based on my research interests.

Fall 2000:

**Digital Signal Processing** (ELEN E4810, 41 students, rating 3.6/5)

My first time teaching a full course; it took me a while to get the pacing right.

*Other teaching:*

Supervision and advising of UC Berkeley graduate students at ICSI, 1997-2001.

Tutor, MIT OME Tutorial Services Room, working with individual students on probability and signal processing, 1993-95.

Occasional guest lecturer for courses Digital Audio Signal Processing, and Introduction to Perception (Auditory Organization), MIT, 1991-1995.

*Students supervised:*

**Completed Ph.D. students (Columbia):**

Marios Athineos, Ph.D., graduated Sep 2007, postdoc at ICSI, Berkeley.

Adam Berenzweig, Ph.D., graduated Sep 2006, research scientist at Google Research.

Manuel Reyes, Ph.D., graduated Apr 2005, currently working at Microsoft Research.

**Completed Ph.D. students (other schools):**

Brian Whitman, Ph.D. from MIT Media Lab Mar 2005, who I effectively co-supervised with my former Ph.D. supervisor from MIT, Barry Vercoe, currently running his own music search startup in Boston.

Patricia Scanlon, Ph.D. from University College Dublin Sep 2005, spend 18 months at LabROSA from 2002-2003 with me as her local advisor, currently working for Lucent in Ireland.

**Current Ph.D. students (Columbia):**

Graham Poliner, supervised since Jun 2004, expected to graduate May 2008.

Ron Weiss, supervised since Sep 2004, expected to graduate Aug 2008.

Keansub Lee, supervised since Sep 2003, expected to graduate Nov 2008.

Michael Mandel (SEAS Presidential Fellow), supervised since Sep 2004, expected to graduate 2009.

Xanadu Halkias, supervised since Jan 2005, expected to graduate 2009.

Christine Smit (NSF Graduate Fellow), supervised since Sep 2006, expected to graduate 2010.

Courtenay Cotton, supervised since June 2007, expected to graduate 2011.

**Visitors (stays of 1-12 months):** Mads Christensen (Aalborg, Denmark), Mikkel Schmidt (DTU, Denmark), Jesper Jensen (Aalborg, Denmark), Jouni Paulus (TUT Finland), Kofi Boakye (UC Berkeley), Elvira Perez (U. Liverpool), Thomas Blumensath (Queen Mary, U. London), Sunil Sivadas (OGI).

**MS students:** Jim Ogle, Nathan Lesser, John Arroyo, Sam Keene, Jack Chan, Alex Sheh, Joseph Hazboun, Banky Omodunbi, Lyndon Kennedy, Rob Turetsky, Uday Arya, Chi Wong, Wayzen Lin, Johanna Devaney (Music).

**Undergraduates:** Jerry Liu, Suman Ravuri, Clement Tong, Rob Spinella, David Wilmot, Ezra Schneck.

**High school students:** Jeff Bauer, Ben Chang, Angel Umpierre.

I have also served on, or am currently serving on, the following Ph.D. committees:

Michel Galley (McKeown, CS), Jackson Lipscombe (Hirschberg, CS), Kai Li (Wang), Raju Hormis (Wang), Boonsit Yimwadsana (Coffman), Winston Hsu (Chang), Lexing Xie (Chang), Dongqing Zhang (Chang), Ana Benitez (Chang), Di Zhong (Chang), Danny Hong (Eleftheriadis), Lai-Tee Cheok (Eleftheriadis), Ari Klein (Tsividis), Aya Aner (Kender, CS), Alex Klein (Pederson, App. Phys.), Lisa Margulis (Lerdahl, Music), Peter LeVoci (Longman, ME), Scott Otterson (Ostendorf, U. Washington), Mitch Parry (Essa, Georgia Tech), Jitendra Ajmera (Bourlard, IDIAP), Sofia Cavaco (Lewicki, CMU), Tuomas Virtanen (TUT Finland), Olivier Gillet (ENST France), Nayeemulla Khan (Yegnanarayanan, IIT Madras), Andrij Temko (Nadeau, UPC Barcelona).

*Distinctions/Awards:*

NSF Career Award, 2003-2008.

Speech Communication Best Paper Award for 2001-2003 (one paper chosen from around 200 published in Speech Communication over two years).

Co-developer of best-ranked system for Audio Artist Identification MIREX Evaluations at the International Conference on Music Information Retrieval, 2004 and 2005 (among 8 teams including University of Montreal, Austrian Research Institute for AI, HP Labs, and Sun Microsystems).

Developer of best-ranked system for Audio Cover Song Identification at MIREX-2006 (among 8 submissions including Stanford University and Fraunhofer IDMT).

Co-developer of best-performing system in the ETSI Aurora-2 Noisy Speech Recognition evaluation at the Eurospeech Special Event, 2001 (among 17 teams including IBM, Microsoft, Bell Labs, and UCLA).

*Journal papers:*

(note: **citation counts** are from scholar.google.com, updated 2008-03-17, and shown when  $\geq 20$ .)

Overall Google-based h-index (Hirsch number) including conference papers: **26.**)

1. R. Weiss & D. Ellis. "Speech separation using speaker-adapted Eigenvoice speech models," accepted by Computer Speech & Language, Mar 2008 (18 pages).
2. M. Slaney, D. P. W. Ellis, M. Sandler, M. Goto, & M. M. Goodwin, "Introduction to the Special Issue on Music Information Retrieval," IEEE Tr. Audio, Speech & Lang. Proc., 16 (2), 253-254,

Feb 2008.

3. M. Athineos & D. Ellis. "Autoregressive modeling of temporal envelopes," *IEEE Tr. Sig. Proc.*, 15(11), 5237-5245, Nov 2007.
4. G. Poliner, D. Ellis, A. Ehmann, E. Gomez, S. Streich, & B. Ong (2007) "Melody Transcription from Music Audio: Approaches and Evaluation," *IEEE Tr. Audio, Speech & Lang. Proc.*, 15 (4), 1247-1256, May 2007.
5. D. Ellis. "Beat Tracking by Dynamic Programming," *J. New Music Research, Special Issue on Algorithms for Beat Tracking and Tempo Extraction*, 36(1), 51–60, Mar 2007.
6. P. Scanlon, D. Ellis, & R. Reilly (2007) "Using Broad Phonetic Group Experts for Improved Speech Recognition," *IEEE Tr. Audio, Speech & Lang. Proc.* 15 (3), 803-812, Mar 2007.
7. G. Poliner & D. Ellis (2007) "A Discriminative Model for Polyphonic Piano Transcription," *EURASIP Signal Processing Journal*, 2007 (2007), Article ID 48317, 9 pages.
8. D. Ellis & K.-S. Lee (2006) "Segmenting and Clustering Personal Audio Archives," *IEEE Multimedia* 13 (4), 30-38, Oct-Dec 2006.
9. D. Ellis & G. Poliner (2006) "Classification-based Melody Transcription," *Machine Learning* 65 (2-3), 439-456, Dec 2006.
10. M. Mandel, G. Poliner, & D. Ellis (2006) "Support Vector Machine Active Learning for Music Retrieval," *Multimedia Systems Journal* 12 (1), 3-13, Aug 2006.
11. D. Ellis (2006) "Extracting Information from Music Audio", *Communications of the ACM* 49(8), 32-37, Aug 2006.
12. X. Halkias & D. Ellis (2006) "Call detection and extraction using Bayesian inference," *Applied Acoustics*, 67(11-12), 1164-1174, Nov-Dec 2006.
13. D. Ellis, B. Raj, J. Brown, M. Slaney, & P. Smaragdis (2006) "Editorial - Special Section on Statistical and Perceptual Audio Processing," *IEEE Tr. Audio, Speech & Language Proc.*, 14(1), 2-4, Jan. 2006.
14. N. Morgan, Q. Zhu, A. Stolcke, K. Sönmez, S. Sivasdas, T. Shinozaki, M. Ostendorf, P. Jain, H. Hermansky, D. Ellis, G. Doddington, B. Chen, Ö. Çetin, H. Bourlard, & M. Athineos (2005) "Pushing the envelope – aside," *IEEE Signal Processing* 22(5), 81-88, Sep. 2005.
15. J. Barker, M. Cooke, & D. Ellis (2005) "Decoding speech in the presence of other sources," *Speech Communication* 45(1), 5-25, Jan 2005. **(Cited by 33)**
16. A. Berenzweig, B. Logan, D. Ellis, & B. Whitman (2004) "A Large-Scale Evaluation of Acoustic and Subjective Music-Similarity Measures," *Computer Music Journal (MIT Press)* 28(2), 63-76, June 2004. **(Cited by 87)**
17. A. Robinson, G. Cook, D. Ellis, E. Fosler-Lussier, S. Renals, & G. Williams (2002) "Connectionist speech recognition of Broadcast News," *Speech Communication* 37(1-2), 27-45, 2002. **(Cited by 39)**
18. M. Cooke & D. Ellis (2001) "The auditory organization of speech in listeners and machines," *Speech Communication* 35(3-4), 141-177, 2001 (winner of *Speech Communication's Best Paper Award* 2001-2003). **(Cited by 74)**
19. D. Ellis (1999) "Using knowledge to organize sound: The prediction-driven approach to computational auditory scene analysis, and its application to speech/nonspeech mixtures," *Speech Communication* 27(3-4), 281-298, 1999.

*Theses:*

1. D. Ellis “Prediction-driven Computational Auditory Scene Analysis”, Ph.D. dissertation, Dept. of Elec. Eng. and Comp. Sci. MIT, April 1996. **(Cited by 214)**
2. D. Ellis “A Perceptual Representation of Audio”, M.S. Thesis, Dept. of Elec. Eng. and Comp. Sci. MIT, Feb. 1992.

*Book chapters:*

1. D. Ellis “Model-Based Acoustic Scene Analysis”, chap. 4 in: D. Wang & G. Brown (eds.) *Computational Auditory Scene Analysis: Principles, Algorithms, and Applications* (Wiley/IEEE Press), pp. 115-146, 2006.
2. D. Ellis “Modeling the auditory organization of speech,” chap. 24 in: S. Greenberg & W. Ainsworth (eds.) *Listening to speech: An auditory perspective* (Lawrence Erlbaum), pp. 393-407, 2006.
3. D. Ellis “Evaluating Speech Separation Systems”, in: P. Divenyi (ed.) *Speech Separation by Humans and Machines* (Kluwer), pp. 295-304, 2004.
4. D. Ellis & D. Rosenthal “Mid-Level representations for Computational Auditory Scene Analysis,” in: D. Rosenthal & H. Okuno (eds.) *Computational Auditory Scene Analysis* (Lawrence Erlbaum, Mahwah), pp. 257-272, 1998. **(Cited by 22)**

*Software/Online resources:*

(all available via <http://www.ee.columbia.edu/~dpwe/>).

- **Online Course Materials** including complete slidepacks for the Speech and Audio Processing and Recognition course. These have been used at multiple other universities including MIT, U. Montreal, Washington University (St. Louis), U. Mississippi, U. Porto (Portugal).
- A collection of **Audio Processing Examples in Matlab** including the Phase Vocoder, PLP-RASTA speech feature calculation, Dynamic Time Warping, and Sinusoidal Analysis/Synthesis. Many of these grew out of examples used in my classes. Posting these materials online has led to their widespread use, as measured by the more than 100 email messages I have received regarding them.
- **Music Similarity Ground Truth** data, distilled from projects and online surveys we have conducted in this area. This data was used as the basis for the first international evaluation of music similarity algorithms at the International Symposium on Music Information Retrieval in October 2004. We also created “uspop2002”, a standard database of basic features for 8700 pop music tracks, for researchers getting started in Music IR; we have sent out more than 20 copies of this 3 DVD set.
- The **AUDITORY list** and homepage - archives and related materials for this research discussion list I have operated since 1992, currently with over 1500 subscribers in 33 countries, and about 70 postings a month.

*Peer-reviewed proceedings papers:*

(Notes: **citation counts** are from scholar.google.com, updated 2008-03-26, and shown when  $\geq 20$ .)

1. S. Ravuri & D. Ellis (2008) “Stylization of pitch with syllable-based linear segments,” Proc. ICASSP-08, pp. 3985–3988, Las Vegas, April 2008.
2. D. Ellis, C. Cotton, & M. Mandel (2008) “Cross-correlation of beat-synchronous representations for music similarity,” Proc. ICASSP-08, pp. 57–60, Las Vegas, April 2008.

3. J. H. Jensen, M. G. Christensen, D. Ellis, & S. H. Jensen (2008) "A tempo-insensitive distance measure for cover song identification based on chroma features," Proc. ICASSP-08, pp. 2209–2212, Las Vegas, April 2008.
4. K. Lee & D. Ellis (2008) "Detecting music in ambient audio by long-window autocorrelation," Proc. ICASSP-08, pp. 9–12, Las Vegas, April 2008.
5. M. Mandel & D. Ellis (2007) "EM localization and separation using interaural level and phase cues," Proc. IEEE Workshop on Apps. of Sig. Proc. to Acous. and Audio WASPAA-07, pp. 275–278, Mohonk NY, October 2007.
6. R. Weiss & D. Ellis (2007) "Monaural speech separation using source-adapted models," Proc. IEEE Workshop on Apps. of Sig. Proc. to Acous. and Audio WASPAA-07, pp. 114–117, Mohonk NY, October 2007.
7. C. Smit & D. Ellis (2007) "Solo voice detection via optimal cancelation," Proc. IEEE Workshop on Apps. of Sig. Proc. to Acous. and Audio WASPAA-07, pp. 207–210, Mohonk NY, October 2007.
8. G. Poliner & D. Ellis (2007) "Improving generalization for polyphonic piano transcription," Proc. IEEE Workshop on Apps. of Sig. Proc. to Acous. and Audio WASPAA-07, pp. 86–89, Mohonk NY, October 2007.
9. S.-F. Chang, D. Ellis, W. Jiang, K. Lee, A. Yanagawa, A. Loui, J. Luo (2007) "Large-scale multimodal semantic concept detection for consumer video," Multimedia Information Retrieval workshop, ACM Multimedia, Augsburg, Germany, Sep 2007. (10pp, to appear)
10. D. Ellis (2007) "Classifying Music Audio with Timbral and Chroma Features," Proc. Int. Conf. on Music Info. Retr. ISMIR-07, pp. 339–340, Vienna, Austria, Sep 2007.
11. M. Mandel & D. Ellis (2007) "A Web-Based Game for Collecting Music Metadata," Proc. Int. Conf. on Music Info. Retr. ISMIR-07, pp. 365–366, Vienna, Austria, Sep 2007.
12. J. H. Jensen, D. Ellis, M. G. Christensen, S. H. Jensen (2007) "Evaluation Distance Measures Between Gaussian Mixture Models of MFCCs," Proc. Int. Conf. on Music Info. Retr. ISMIR-07, pp. 107-108, Vienna, Austria, Sep 2007.
13. J. Ogle & D. Ellis (2007) "Fingerprinting to Identify Repeated Sound Events in Long-Duration Personal Audio Recordings," Proc. ICASSP-07, pp. I-233–236, Honolulu HI, Apr 2007.
14. D. Ellis & G. Poliner (2007) "Identifying Cover Songs With Chroma Features and Dynamic Programming Beat Tracking," Proc. ICASSP-07, pp. IV-1429–1432, Honolulu HI, Apr 2007.
15. M. Mandel, D. Ellis, & T. Jebara (2006) "An EM algorithm for localizing multiple sound sources in reverberant environments," Adv. NIPS, pp. 953–960, Vancouver BC, Dec 2006.
16. K. Lee & D. Ellis (2006) "Voice activity detection in personal audio recordings using autocorrelogram compensation," Proc. Interspeech, pp. 1970-1973, Pittsburgh PA, Oct 2006.
17. M. Mandel & D. Ellis (2006) "A probability model for interaural phase difference," Proc. Workshop on Statistical & Perceptual Audition SAPA-06, pp. 1-6, Pittsburgh PA, Oct 2006.
18. R. Weiss & D. Ellis (2006) "Estimating single-channel source separation masks: Relevance Vector Machine classifiers vs. pitch-based masking," Proc. Workshop on Statistical & Perceptual Audition SAPA-06, pp. 31-36, Pittsburgh PA, Oct 2006.
19. D. Ellis & R. Weiss (2006) "Model-Based Monaural Source Separation Using a Vector-Quantized Phase-Vocoder Representation," Proc. ICASSP-06, pp. V-957-960, Toulouse, May 2006.
20. X. Halkias & D. Ellis (2006) "Estimating the Number of Marine Mammals using Recordings of Clicks from One Microphone," Proc. ICASSP-06, pp. V-769-772, Toulouse, May 2006.

21. K. Dobson, B. Whitman, & D. Ellis (2005) "Learning Auditory Models of Machine Voices", IEEE Workshop on Apps. of Sig. Proc. to Acous. & Audio WASPAA-05, pp.339-342, Mohonk NY, October 2005.
22. G. Poliner & D. Ellis (2005) "A Classification Approach to Melody Transcription", Int. Conf. on Music Info. Retrieval ISMIR-05, pp. 161-166, London, September 2005.
23. M. Mandel & D. Ellis (2005) "Song-Level Features and Support Vector Machines for Music Classification", Int. Conf. on Music Info. Retrieval ISMIR-05, pp. 594-599, London, September 2005. **(Cited by 37)**
24. N. Lesser & D. Ellis (2005) "Clap Detection and Discrimination for Rhythm Therapy", Proc. ICASSP-05, pp. III-37-40, Philadelphia, March 2005.
25. C.-P. Chen, J. Bilmes, & D. Ellis (2005) "Speech Feature Smoothing for Robust ASR", Proc. ICASSP-05, pp. I-525-528, Philadelphia, March 2005.
26. M. Reyes-Gomez, N. Jovic, & D. Ellis (2005) "Deformable Spectrograms", AI & Statistics 2005, pp. 285-292, Barbados, Jan 2005.
27. D. Ellis & K.S. Lee (2004) "Minimal-Impact Audio-Based Personal Archives", Proc. 1st ACM workshop on Continuous Archiving and Recording of Personal Experiences CARPE-04, pp.39-47, New York, Oct 2004. **(Cited by 29)**
28. D. Ellis & J. Arroyo (2004) "Eigenrhythms: Drum pattern basis sets for classification and generation", International Symposium on Music Information Retrieval ISMIR-04, pp. 101-106, Barcelona, Oct 2004.
29. B. Whitman & D. Ellis (2004) "Automatic Record Reviews", International Symposium on Music Information Retrieval ISMIR-04, pp. 86-93, Barcelona, Oct 2004.
30. M. Athineos, H. Hermansky, & D. Ellis (2004) "LP-TRAP: Linear predictive temporal patterns", International Conference on Spoken Language Processing ICSLP-04, pp. 949-952, Jeju, Korea, Oct 2004.
31. M. Athineos, H. Hermansky, & D. Ellis (2004) "PLP<sup>2</sup>: Autoregressive modeling of auditory-like 2-D spectro-temporal patterns", ISCA Tutorial and Research Workshop on Statistical and Perceptual Audio Processing SAPA-04, pp. 25-30, Jeju, Korea, Oct 2004.
32. M. Reyes-Gomez, N. Jovic, & D. Ellis (2004) "Towards single-channel unsupervised source separation of speech mixtures: The layered harmonics/formants separation-tracking model", ISCA Tutorial and Research Workshop on Statistical and Perceptual Audio Processing SAPA-04, pp. 37-42, Jeju, Korea, Oct 2004.
33. D. Ellis & K.S. Lee (2004) "Features for Segmenting and Classifying Long-Duration Recordings of Personal Audio", ISCA Tutorial and Research Workshop on Statistical and Perceptual Audio Processing SAPA-04, pp. 1-4, Jeju, Korea, Oct 2004.
34. D. Ellis & J. Liu (2004) "Speaker turn segmentation based on between-channel differences", NIST ICASSP Meeting Recognition Workshop, pp. 112-117, Montreal, May 2004.
35. L. Kennedy & D. Ellis (2004) "Laughter Detection in Meetings", NIST ICASSP Meeting Recognition Workshop, pp. 118-121, Montreal, May 2004.
36. M.J. Reyes-Gomez, N. Jovic, & D. Ellis (2004) "Detailed graphical models for source separation and missing data interpolation in audio", Learning Workshop, Snowbird, 2004.
37. M.J. Reyes-Gomez, D. Ellis, & N. Jovic (2004) "Multiband Audio Modeling for Single Channel Acoustic Source Separation", Proc. ICASSP-04, pp. V-641-644, Montreal, May 2004. **(Cited by 21)**

38. L. Kennedy & D. Ellis (2003) "Pitch-based emphasis detection for characterization of meeting recordings", Automatic Speech Recognition and Understanding Workshop IEEE ASRU 2003, pp. 243-284, St. Thomas, December 2003.
39. M. Athineos & D. Ellis (2003) "Frequency-domain linear prediction for temporal features", Automatic Speech Recognition and Understanding Workshop IEEE ASRU 2003, pp. 261-266, St. Thomas, December 2003.
40. A. Sheh & D. Ellis (2003) "Chord Segmentation and Recognition using EM-Trained Hidden Markov Models", 4th International Symposium on Music Information Retrieval ISMIR-03, pp. 185-191, Baltimore, October 2003. **(Cited by 57)**
41. R. Turetsky & D. Ellis (2003) "Ground-Truth Transcriptions of Real Music from Force-Aligned MIDI Syntheses", 4th International Symposium on Music Information Retrieval ISMIR-03, pp. 135-141, Baltimore, October 2003. **(Cited by 22)**
42. A. Berenzweig, B. Logan, D. Ellis, & B. Whitman (2003) "A large-scale evaluation of acoustic and subjective music similarity measures", 4th International Symposium on Music Information Retrieval ISMIR-03, pp. 103-109, Baltimore, October 2003. **(Cited by 62)**
43. M.J. Reyes-Gomez, B. Raj, & D. Ellis (2003) "Multi-channel Source Separation by Beamforming Trained with Factorial HMMs", Proc. IEEE Workshop on Apps. of Sig. Proc. to Acous. and Audio, pp. 13-16, Mohonk NY, October 2003.
44. B. Logan, D. Ellis, & A. Berenzweig (2003) "Toward evaluation techniques for music similarity", Keynote address, Workshop on the Evaluation of Music Information Retrieval (MIR) Systems at SIGIR 2003, Toronto, August 2003.
45. P. Scanlon, D. Ellis, & R. Reilly (2003) "Using Mutual Information to design class-specific phone recognizers", Proc. Eurospeech-03, pp. 857-860, Geneva, September 2003.
46. M. Reyes-Gomez & D. Ellis (2003) "Selection, Parameter Estimation, and Discriminative Training of Hidden Markov Models for General Audio Modeling", Proc. ICME-03, pp. I-73-76, Baltimore, July 2003.
47. A. Berenzweig, D. Ellis, & S. Lawrence (2003) "Anchor Space for Classification and Similarity Measurement of Music", Proc. ICME-03, pp. I-29-32, Baltimore, July 2003. **(Cited by 33)**
48. S. Renals & D. Ellis (2003) "Audio Information Access from Meeting Rooms", Proc. ICASSP-03, pp. IV-744-747, Hong Kong, April 2003 (invited). **(Cited by 27)**
49. M. Athineos & D. Ellis (2003) "Sound Texture Modelling with Linear Prediction in both Time and Frequency Domains", Proc. ICASSP-03, pp. V-648-651, Hong Kong, April 2003.
50. M. Reyes, B. Raj, & D. Ellis (2003) "Multi-channel Source Separation by Factorial HMMs", Proc. ICASSP-03, pp. I-664-667, Hong Kong, April 2003. **(Cited by 21)**
51. A. Janin, D. Baron, J. Edwards, D. Ellis, D. Gelbart, N. Morgan, B. Peskin, T. Pfau, E. Shriberg, A. Stolcke, & C. Wooters (2003) "The ICSI Meeting Corpus", Proc. ICASSP-03, pp. I-364-367, Hong Kong, April 2003. **(Cited by 147)**
52. D. Ellis, B. Whitman, A. Berenzweig, & S. Lawrence (2002) "The Quest for Ground Truth in Musical Artist Similarity", Proc. ISMIR-02, pp. 170-177, Paris, October 2002. **(Cited by 49)**
53. A. Berenzweig, D. Ellis, & S. Lawrence (2002) "Using Voice Segments to Improve Artist Classification of Music", Proc. AES-22 Intl. Conf. on Virt., Synth., and Ent. Audio. Espoo, Finland, June 2002. **(Cited by 43)**
54. M. Reyes-Gomez & D. Ellis (2002) "Error visualization for tandem acoustic modeling on the Aurora task," Proc. IEEE Int. Conf. on Acous., Speech & Sig. Proc., Student Session, Orlando, May 2002.

55. T. Pfau, D. Ellis & A. Stolcke (2001) "Multispeaker Speech Activity Detection for the ICSI Meeting Recorder," Proc. ASRU-01, 4 pp., Italy, December 2001. **(Cited by 39)**
56. A. Berenzweig & D. Ellis (2001) "Locating Singing Voice Segments within Music Signals," Proc. IEEE Workshop on Apps. of Sig. Proc. to Acous. and Audio, Mohonk, pp. 119-122, October 2001. **(Cited by 64)**
57. D. Ellis (2001) "Detecting Alarm Sounds," Proc. CRAC workshop, pp. 59-62, Aalborg, September 2001.
58. J. Barker, M. Cooke, & D. Ellis (2001) "Integrating bottom-up and top-down constraints to achieve robust ASR: The multisource decoder," Proc. CRAC workshop, pp. 63-66, Aalborg, September 2001.
59. D. Ellis & M. Reyes (2001) "Investigations into Tandem Acoustic Modeling for the Aurora Task," Proc. Eurospeech-01 (Special Event on Noise Robust Recog.), pp. 189-192, Denmark, September 2001. **(Cited by 22)**
60. D. Ellis, R. Singh, & S. Sivasdas (2001) "Tandem acoustic modeling in large-vocabulary recognition," Proc. IEEE Int. Conf. on Acous., Speech & Sig. Proc., pp. I-517-520, Salt Lake City, May 2001. **(Cited by 34)**
61. N. Morgan, D. Baron, J. Edwards, D. Ellis, D. Gelbart, A. Janin, T. Pfau, E. Shriberg, & A. Stolcke (2001) "The Meeting Project at ICSI" Proc. Human Lang. Tech. Conf., pp. 1-7, San Diego, March 2001. **(Cited by 87)**
62. D. Ellis & J. Bilmes (2000) "Using mutual information to design feature combinations," Proc. Int. Conf. Spoken Lang. Proc., Beijing, October 2000. **(Cited by 29)**
63. J. Barker, M. Cooke, & D. Ellis (2000) "Decoding speech in the presence of other sound sources," Proc. Int. Conf. Spoken Lang. Proc., Beijing, October 2000. **(Cited by 24)**
64. J. Ferreiros & D. Ellis (2000) "Using acoustic condition clustering to improve acoustic change detection on Broadcast News," Proc. Int. Conf. Spoken Lang. Proc., Beijing, October 2000.
65. H. Hermansky, D. Ellis, & S. Sharma (2000) "Tandem connectionist feature extraction for conventional HMM systems," Proc. IEEE Int. Conf. on Acous., Speech & Sig. Proc., pp. III-1365-1368, Istanbul, June 2000. **(Cited by 123)**
66. S. Sharma, D. Ellis, S. Kajarekar, P. Jain, & H. Hermansky (2000) "Feature extraction using non-linear transformation for robust speech recognition on the Aurora database," Proc. IEEE Int. Conf. on Acous., Speech & Sig. Proc., pp. II-1117-1120, Istanbul, June 2000. **(Cited by 57)**
67. D. Genoud, D. Ellis, & N. Morgan (1999) "Combined speech and speaker recognition with speaker-adapted connectionist models," Proc. IEEE Automatic Speech Recognition & Understanding Workshop, Keystone CO, December 1999.
68. D. Abberley, S. Renals, T. Robinson, & D. Ellis (1999) "The THISL SDR system at TREC-8," Proc. Text Retrieval Conference TREC-8, Washington, November 1999.
69. G. Williams & D. Ellis (1999) "Speech/music discrimination based on posterior probability features," Proc. Eurospeech-99, Budapest, September 1999. **(Cited by 52)**
70. A. Janin, D. Ellis, & N. Morgan (1999) "Multistream: Ready for prime-time?" Proc. Eurospeech-99, Budapest, September 1999. **(Cited by 50)**
71. D. Ellis & N. Morgan (1999) "Size matters: An empirical study of neural network training for large vocabulary continuous speech recognition," Proc. IEEE Int. Conf. on Acous., Speech & Sig. Proc., Phoenix, April 1999.

72. N. Morgan, D. Ellis, E. Fosler-Lussier, A. Janin, & B. Kingsbury (1999) "Reducing errors by increasing the error rate: MLP acoustic modeling for Broadcast News transcription," DARPA Broadcast News Transcription and Understanding Workshop, Herndon VA, February 1999.
73. G. Cook, J. Christie, D. Ellis, E. Fosler-Lussier, Y. Gotoh, B. Kingsbury, N. Morgan, S. Renals, A. Robinson, & G. Williams (1999) "An overview of the SPRACH system for the transcription of Broadcast News," DARPA Broadcast News Transcription and Understanding Workshop, Herndon VA, February 1999. **(Cited by 28)**
74. D. Ellis (1997) "Computational Auditory Scene Analysis exploiting speech recognizer knowledge," Proc. IEEE Workshop on Apps. of Sig. Proc. to Acous. and Audio, Mohonk, Oct 1997.
75. D. Ellis (1997) "The Weft: a representation for periodic sounds," Proc. IEEE Int. Conf. on Acous., Speech & Sig. Proc., Munich, April 1997. **(Cited by 20)**
76. S. Greenberg, J. Hollenback, & D. Ellis (1996) "Insights into spoken language gleaned from phonetic transcriptions of the Switchboard corpus," Proc. Int. Conf. on Spoken Lang. Proc., Philadelphia, October 1996. **(Cited by 86)**
77. D. Ellis (1996) "Prediction-driven computational auditory scene analysis for dense sound mixtures," Proc. ESCA Workshop on the Auditory Basis of Speech Perception, Keele, July 1996.
78. D. Ellis (1995) "Underconstrained stochastic representations for top-down computational auditory scene analysis," Proc. IEEE Workshop on Apps. of Sig. Proc. to Acous. and Audio, Mohonk, October 1995.
79. D. Ellis (1994) "A computer implementation of psychoacoustic grouping rules," Proc. 12th Int. Conf. on Pattern Recog., Jerusalem, October 1994. **(Cited by 41)**
80. D. Ellis (1993) "Hierarchic models for sound analysis and reconstruction," Proc. IEEE Workshop on Apps. of Sig. Proc. to Acous. and Audio, Mohonk, October 1993.
81. D. Ellis (1992) "Timescale modifications and wavelet representations," Proc. Int. Computer Music Conf., San José , pp. 6-9, June 1992.
82. D. Ellis & B. Vercoe (1991) "A perceptual representation of audio for co-channel source separation," Proc. IEEE Workshop on Apps. of Sig. Proc. to Acous. and Audio, Mohonk, 1991.
83. D. Ellis & B. Vercoe (1991) "A wavelet-based sinusoid model of sound for auditory signal separation," Proc. Int. Computer Music Conf., Montréal, September 1991
84. B. Vercoe & D. Ellis (1990) "Real-time CSound: Software Synthesis with Sensing and Control," Proc. Int. Computer Music Conf., Glasgow, pp. 209-211, September 1990.

*Patents:*

1. "System for identifying cover versions of musical recordings", sole inventor, provisional filed Sep 2006.
2. "Relevance feedback and active learning for music playlist generation", co-inventor with Graham Poliner and Michael Mandel, covers music similarity work based on audio alone, filed 2006.
3. "Computer implemented methods and systems for modeling and recognition of speech", co-inventor with Marios Athineos and Hynek Hermansky, covers the FDLP and PLP<sup>2</sup> techniques for extracting and exploiting temporal envelopes in speech recognition, filed 2004.
4. "Nonlinear feature preprocessing for speech recognition", co-inventor with Nelson Morgan and Hynek Hermansky, covers the Tandem technique for using a neural-net classifier as a front-end to a conventional speech recognizer, filed 2000.

*Invited talks:*

(slides viewable at <http://www.ee.columbia.edu/~dpwe/talks/>)

“Searching for Similar Phrases in Music Audio,” invited talk, Digital Music Research Network workshop, Queen Mary, University of London, 2007-12-18.

“Using source models in speech separation,” invited presentation at the Next-Generation Statistical Models for Speech and Audio Signal Processing workshop, Radcliffe Institute for Advanced Study, 2007-11-09.

“Extracting and Using Music Audio Information,” invited ECE department seminar, UC San Diego, 2007-11-02.

“Analysis of everyday sounds,” invited talk, Kodak Research Labs, Rochester NY, 2007-07-24.

“Using sound source models to organize mixtures,” invited talk, ASIP-NET.DK hearing aid industry seminar, Denmark Technical University, 2007-05-24.

“Beat-synchronous chroma representations for music analysis,” invited talk, Intelligent Sound Workshop '08, Karlslunde, Denmark, 2007-05-23.

“Sound Organization by Source Models in Humans and Machines,” invited talk, NIPS Workshop on Advances in Models of Acoustic Processing, Whistler BC, 2006-12-09.

“Extracting Information from Music Audio,” invited talk at joint Danish Technical University-Aalborg University Intelligent Sound Workshop, Sæby, Denmark, 2006-05-22.

“Auditory Scene Analysis in Humans and Machines,” tutorial at the AES Convention, Paris, 2006-05-20.

“VQ Source Models: Perceptual & Phase Issues,” invited talk at special session on source separation organized by Prof. Shoji Makino, IEEE ICASSP-06, Toulouse, 2006-05-16.

“Using Learned Source Models to Organize Sound Mixtures,” invited talk at New Ideas In Hearing workshop organized by Dr. Alain de Cheveigné, Ecole Normale Supérieure, Paris, 2006-05-12.

“Model-Based Separation in Humans and Machines,” invited talk at the special session on approaches to audio separation organized by Emmanuel Vincent, ICA-2006, Charleston SC, 2006-03-08.

“Music Information Extraction,” invited talk as guest of Prof. Ozgur Izmirli at Connecticut College, 2006-02-13.

“Speech Separation in Humans and Machines,” **opening keynote**, IEEE Automatic Speech Recognition and Understanding ASRU-05, San Juan Puerto Rico, 2005-11-28.

“Computational Auditory Scene Analysis” and “Model-Based Scene Analysis”, invited talks at the Hearing Aid Developers” Forum organized by Prof. Birger Kollmeier, Univ. Oldenburg, Germany, 2005-06-30.

“Audio signal recognition for speech, music, and environmental sounds,” invited talk, Special session on classification, 146th meeting of the Acoustical Society of America, Austin, 2003-11-13.

*Professional affiliations:*

Member of the Audio Engineering Society since 2004.

Member of the International Speech Communications Association since 1999.

Member of the Institute of Electrical and Electronic Engineers since 1994.

Member of the Acoustical Society of America since 1993.

*Other information:*

Since 1993, administrator of the AUDITORY list, an email discussion list for researchers in auditory organization (currently over 1500 participants in 33 countries). <http://www.auditory.org/>

General co-chair of the 9th International Conference on Music Information Retrieval (ISMIR-08), Philadel-

phia, Sep 2008 (100+ papers and 200+ participants expected). <http://www.ismir2008.org/>

Co-guest editor of Special Issue of IEEE Transactions on Speech and Audio Processing on Music Information Retrieval, to appear 2008.

Co-chair of the ISCA Tutorial and Research Workshop on Statistical and Perceptual Audition SAPA-2006, held as a satellite to ICSLP-2006 in Pittsburgh PA, September 2006 (12 papers presented, 37 participants). <http://www.sapa2006.org/>

Co-guest editor of Special Issue of IEEE Transactions on Speech and Audio Processing on Statistical and Perceptual Audio Processing (12 papers, Jan 2006).

Co-organizer of Workshop on Music Information Processing Systems, held as a satellite to the Neural Information Processing Systems conference, Whistler, BC, December 2004 (8 papers presented, 25 participants). <http://www.iro.umontreal.ca/~eckdoug/mips/>

Co-organizer of AFOSR/NSF-sponsored 2nd Montreal Workshop on Speech Separation, November 2004. 40 invited participants. <http://labrosa.ee.columbia.edu/Montreal2004/>

Co-chair of the ISCA Tutorial and Research Workshop on Statistical and Perceptual Audio Processing SAPA-2004, held as a satellite to ICSLP-2004 in Jeju-da, Korea, November 2004 (22 papers, 40 participants). <http://www.sapa2004.org/>

Co-guest editor of Special Issue of Speech Communication on Recognition and Organization of Real-World Sounds (vol 43 issue 4, published September 2004).

Co-organizer of NSF-sponsored workshop on Speech Separation, held in Montreal in November 2003. Twenty invited participants attended. <http://www.ebire.org/speechseparation/>

Invited Participant of NSF workshop on Next-Generation Automatic Speech Recognition, Atlanta, Oct 2003. <http://users.ece.gatech.edu/~chl/ngasr03/>

Co-chair of workshop on Consistent and Reliable Acoustic Cues for sound analysis (CRAC2001), held as a satellite event to Eurospeech in Denmark in September 2001 (24 papers presented, over 40 participants). <http://www.ee.columbia.edu/crac/>

Co-organizer of the first Workshop on Computational Auditory Scene Analysis in Montreal, 1995, and the following meeting in Nagoya, 1997.

Associate Editor for IEEE Multimedia Magazine (since 2005) and IEEE Signal Processing Magazine (Lecture Notes column, since 2006).

Co-chair of Audio Engineering Society Technical Committee on Semantic Audio Analysis (since 2004).

Treasurer of the 1997 IEEE Workshop on Apps. of Sig. Proc. to Acous. and Audio (Mohonk).

Reviewer for IEEE ICASSP, IEEE Tr. Speech & Audio Proc., Speech Communication, Computer Speech & Language, IEEE Mohonk WASPAA, AES Convention, IEEE Trans. Neural Networks, Free Speech Journal (online), J. VLSI Sig. Proc. Sys.

Regular contributor of software and book reviews published in IEEE Spectrum magazine.

Awarded a Harkness Fellowship of the Commonwealth Fund of New York to study at MIT.

Author of numerous public-domain sound analysis and processing tools for Unix including “dspB” (a soundfile viewer) and ‘pva’ (a phase vocoder/short-time Fourier analyzer); maintainer of the ‘SPRACH-core’ set of connectionist speech recognition tools.