CURRICULUM VITA

Richard M. Stern, Jr.

Department of Electrical and Computer Engineering
Carnegie Mellon University
Pittsburgh, PA 15213
Phone: (412) 268-2535

513 Emerson Street
Pittsburgh, PA 15206
Cell phone: (412) 916-7386

Email: rms@cmu.edu
Citizenship: U.S.A.

www.ece.cmu.edu/~rms

PROFESSIONAL INTERESTS

Automatic speech recognition, auditory perception, acoustics, signal processing, biomedical instrumentation

EDUCATION

Ph.D. (1977)	Electrical Engineering and Computer Science Massachusetts Institute of Technology, Cambridge, MA
M.S. (1972)	Electrical Engineering and Computer Sciences University of California, Berkeley, CA
S.B. (1970)	Electrical Engineering Massachusetts Institute of Technology, Cambridge, MA

EXPERIENCE

1995 - present	Professor of Electrical and Computer Engineering Carnegie Mellon University.
1988 - present	Associate Professor and Professor by Courtesy, Language Technologies Institute, Computer Science Department, Biomedical Engineering Department
2009 - present	Artist Lecturer, School of Music Carnegie Mellon University
1995 - 2003	Associate Director of the Information Networking Institute Carnegie Mellon University
1982 - 1995	Associate Professor of Electrical and Biomedical Engineering Carnegie Mellon University
1985	Visiting Professor in Speech and Communication Sciences, Nippon Telegraph and Telephone Electrical Communications Laboratory, Tokyo, Japan
1977 - 1982	Assistant Professor of Electrical and Biomedical Engineering Carnegie Mellon University

Amazon v. SoundClear
US Patent 11,244,675
Amazon Ex. 1006

1979 - 1981	Adjunct Assistant Professor of Otolaryngology University of Pittsburgh School of Medicine
1973 - 1976	Teaching and Research Assistant, Department of Electrical Engineering, Massachusetts Institute of Technology

PROFESSIONAL ACTIVITIES (partial listing)

Distinguished Lecturer, International Speech Communication Association, 2008-2009.

General Chair, INTERSPEECH International Conference on Spoken Language Processing, September, 2006.

Technical Program Co-Chair, IEEE Workshop on Automatic Speech Recognition and Understanding, December 2005.

Technical Program Chair, 141st meeting of the Acoustical Society of America, June 2002.

General Chair, DARPA Spoken Language Technologies Workshop, March, 1994.

Publications Chair, ARPA Spoken Language Technology and Applications Day, April, 1993.

Publications Chair, IEEE Workshop on Applications of Signal Processing to Audio and Acoustics, October, 1993.

Chair, standing DARPA Speech and Natural Language Workshop Organizing Committee, 1991 1992.

Secretary, ARPA Spoken Language Coordinating Committee, 1990 - 1995.

General Chair, DARPA Speech and Natural Language Workshop, June, 1990.

International Advisory Board, International Speech Communication Association, 2006 - present.

International Advisory Board, Center for Speech and Language Technologies, Tsinghua University, Beijing, China, 2007 - 2010.

Chair, Selection Committee for IEEE James L. Flanagan Speech & Audio Processing Award, 2006 - 2008.

IEEE Signal Processing Society Technical Committee on Audio and Electroacoustics, 1991 1995.

IEEE Signal Processing Society Technical Committee on Speech, 1993 - 1997.

Editorial board, *Journal of Computer Speech and Language*, 1994 - 2010.

Collaborative research in binaural hearing with the Department of Otolaryngology at the University of Connecticut Medical School, Farmington, CT, 1990-2019.

Member of Institute of Electrical and Electronics Engineers, Acoustical Society of America, International Speech Communication Association, Association for Research in Otolaryngology, Audio Engineering Society.

Reviewer for National Science Foundation, International Speech Communication Association, IEEE, *J. Acoust. Soc. Amer.*, *Hearing Research*, *IEEE Transactions on Signal Processing*, *IEEE Transactions on Speech and Language*, *IEEE Transactions on Systems, Man, and Cybernetics*, and *Communications of the Association of Computing Machinery.*

HONORS AND AWARDS

Fellow, Institute of Electrical and Electronics Engineers (IEEE)

Fellow, International Speech Communication Association (ISCA)

Fellow, Acoustical Society of America (ASA)

Distinguished Lecturer of the International Speech Communication Association, 2008 to 2009

Allen Newell Award for Research Excellence, Carnegie Mellon University Department of Computer Science, 1992

IEEE Student Branch Award for Teacher of the Year, Carnegie Mellon University Department of Electrical Engineering, 1979

Joel and Ruth Spira Excellence in Teaching Award (with three colleagues), Carnegie Mellon University Department of Electrical and Computer Engineering, 2018

Eta Kappa Nu Award for Teacher of the Year, Carnegie Mellon University Department of Electrical and Computer Engineering, 2021

PUBLICATIONS AND PAPERS

Papers in Archival Journals

WUTH, J., MAHU, R., COHEN, I., STERN, R. M., and BECERRA YOMA, N. (2024). "A unified beamforming and source separation model for static and dynamic human-robot interaction," *JASA Express Letters,* 1 March 2024, **4**(3)**:** 035203.

ALVARADO, E., GRAGEDA, N., LUZANTO, A., MAHU, R., WUTH, J, MENDOZA, L., STERN, R., BECERRA YOMA, N. (2023). "Automatic detection of dyspnea in real human-robot interaction scenarios," *Sensors (Basel)* 2023 Sep 1;23(17):7590.

STERN, R. M., COLBURN, H. S., BERNSTEIN, L. R., AND TRAHIOTIS, C. (2019). "The fMRI data of Thompson *et al.* (2006) do not constrain how the human midbrain represents interaural time delay," *Journal of the Association for Research in Otolaryngology* **20**:305-311.

DIETZ, M., LESTANG, J.-H., MAJDAK, P., STERN, R. M., MARQUARDT, T., EWERT, S. D., HARTMANN, W. M., and GOODMAN, D. F. M. (2017). "A Framework for Testing and Comparing Binaural Models," *Hearing Research* **360**:92-106.

DE LA CALLE SILOS, F., and STERN, R. M. (2017). "Synchrony-based feature extraction for robust automatic speech recognition," *IEEE Signal Processing Letters* **24:**1158-1162.

FREDES, J., NOVOA, J., KING, S., STERN, R. M., and BECERRA YOMA, N. (2017). "Locally-normalized filter banks applied to deep neural network-based robust speech recognition," *IEEE Signal Processing Letters* **24:**377-381.

KIM, C., and STERN, R. M. (2016). "Power-normalized cepstral coefficients (PNCC) for robust speech recognition," *IEEE Trans. on Audio, Speech, and Language Processing* **24:**1315-1329. **[Received IEEE Signal Processing Society Best Paper Award, 2019]**

CHO, B. J., KWON,H. Cho, J.-W., KIM, C., STERN, R. M. and PARK, H.-M. (2016). "A subband-based stationary-component suppression method using harmonics and power ratio for reverberant speech recognition, *IEEE Signal Processing Letters*, **23**:780-784.

ROMIGH, G. D., BRUNGART, D. S., STERN, R. M., and SIMPSON, B. D. (2015). "Efficient real spherical harmonic representation of head-related transfer functions," *IEEE Journal of Selected Topics in Signal Processing*, **9**: 921-930, August 2015.

POBLETE, V., ESPIC, F., KING, S., STERN, R. M., HUENEPAN, F., and BECERRA YOMA, N. (2015). "A perceptually-motivated low-complexity channel normalization technique applied to speaker verification." *Computer Speech and Language,* **31**1-27, 2015.

POBLETE, V., BECERRA YOMA, N., and STERN, R. M. (2014). "Optimizing the parameters characterizing sigmoidal rate-level functions based on acoustic features," *Speech Communication,* **56:**19-34, January 2014.

HERMANSKY, H., COHEN, J. R., and STERN, R. M. (2013). "Perceptual properties of current speech recognition technology," *Proc. IEEE* **101**:1968-1985, September 2013.

STERN, R. M., and MORGAN, N. (2012). "Hearing is believing: biologically-inspired methods for robust speech recognition," *IEEE Signal Processing Magazine* **29**:34-43, November, 2012.

CHIU, Y.-H. B., RAJ, B., and STERN, R. M. (2012). "Learning-based auditory encoding for robust speech recognition," *IEEE Trans. on Audio, Speech, and Language Processing* **20:**900-914, March 2012.

KIM, W., and STERN, R. M. (2011). "Mask classification for missing-feature reconstruction for robust speech recognition," *Speech Communication*, **53:**1-11, January 2011.

PARK, H.-M., and STERN, R. M. (2009). "Spatial Separation of Speech Signals using Continuosly-Variable Weighting Factors Estimated from Comparisons of Zero Crossings, *Speech Communication Journal,* **51**(1):15-25, January 2009.

SELTZER, M. L., and STERN, R. M. (2006). "Subband Likelihood-Maximizing Beamforming for Speech Recognition in Reverberant Environments," *IEEE Transactions of Speech, Language, and Audio Processing* 14(6): 2109-2121, November 2006.

RAJ, B., and STERN, R. M. (2005). "Missing-Feature Methods for Robust Automatic Speech Recognition," *IEEE Signal Processing Magazine*, September 2005.

KIM, N. S., LIM, W., and STERN, R. M. (2005). "Feature compensation based on switching linear dynamic model," *IEEE Signal Processing Letters*, **12**(6): 473-476.

SELTZER, M. L., RAJ, B., and STERN, R. M. (2004). "Likelihood-Maximizing Beamforming for Robust Hands-Free Speech Recognition," *IEEE Transactions of Speech and Audio Processing*, **12**(5): 489-498, September 2004. **[Received IEEE Signal Processing Society Best Student Paper Award, 2007]**

OBUCHI, Y., HATAOKA, N., and STERN, R. M. (2004), "Normalization of Time-Derivative Parameters for Robust Speech Recognition in Small Devices," *IEICE Trans. on Information and System*s, **87-D**(4): 1004-1011, April 2004.

RAJ, B., SELTZER, M. L., and STERN, R. M. (2004), "Reconstruction of Missing Features for Robust Speech Recognition," *Speech Communication Journal*, **43**(4): 275-296, September 2004.

SELTZER, M. L., RAJ, B,, and STERN, R. M. (2004). "A Bayesian Framework for Spectrographic Mask Estimation for Missing Feature Speech Recognition," *Speech Communication Journal,* **43**(4): 379-393, September 2004.

SINGH, R., RAJ, B., and STERN, R. M. (2001), "Automatic Generation of Sub-Word Units for Speech Recognition Systems," *IEEE Trans. on Speech and Audio Proc.* **10**(2):89-99.

HUERTA, J. M., and STERN, R. M. (2001). "Distortion-Class Modeling for Robust Speech Recognition under GSM RPE-LTP Coding," *Speech Communication Journal*, **34:**213-225 (invited paper).

MORENO, P. J., RAJ, B., and STERN, R. M. (1998). "Data-Driven Environmental Compensation for Speech Recognition: A Unified Approach," *Speech Communication Journal*, **24:** 267-85.

STERN, R. M., and SHEAR, G. D. (1996a) "Lateralization and Detection of Low- Frequency Binaural Stimuli: Effects of Distribution of Internal Delay," *J. Acoust. Soc. Amer.* **100**: 2278-2288.

STERN, R. M., and SHEAR, G. D. (1996b) "Lateralization and Detection of Low- Frequency Binaural Stimuli: Specification of the Extended Position-Variable Model," *Physics Auxiliary Publication Service*, AIP document E-JASMA-100-2278- 0.175MB via http://www.aip.org/epaps/ epaps.html.

TRAHIOTIS, C., and STERN, R. M. (1994) "Across-Frequency Interaction in Lateralization of Complex Binaural Stimuli," *J. Acoust. Soc. Amer.* **96**: 3804- 3806 (L).

STERN, R. M., ZEPPENFELD, T., and SHEAR, G. D. (1991). "Lateralization of Rectangularly Modulated Noise: An Explanation for Counterintuitive Reversals," *J. Acoust. Soc. Amer.* **90:** 1901-1907.

COAST, D. A., STERN, R. M., CANO, G. G., and BRILLER, S. A. (1990). "An Approach to Cardiac Arrhythmia Analysis Using Hidden Markov Models," *IEEE Trans. Biomed. Eng*. **37**: 826-836.

TRAHIOTIS, C., and STERN, R. M. (1989). "Lateralization of Bands of Noise: Effects of Bandwidth and Differences of Interaural Time and Phase," *J. Acoust. Soc. Amer.* **86**: 1285-1293.

RUDNICKY, A. I., and STERN, R.M. (1989). "Spoken Language Research at Carnegie Mellon," *Speech Technology Magazine* **4:** 38-43.

STERN, R. M., ZEIBERG, A. S., and TRAHIOTIS, C. (1988). "Lateralization of Complex Binaural Stimuli: A Weighted Image Model," *J. Acoust. Soc. Amer.* 84, 156-165.

STERN, R. M., and LASRY, M. J. (1987). "Dynamic Speaker Adaptation for Feature-Based Isolated Letter Recognition," I*EEE Trans. on Acoustics, Speech, and Signal Processing* **35:** 751-763.

STERN, R. M., and COLBURN, H. S. (1985). "Lateral-Position Models of Interaural Discrimination," J*. Acoust. Soc. Amer.* **77:** 753-755.

STERN, R. M., and COLBURN, H. S. (1985). "Subjective Lateral Position and Interaural Discrimination," *Physics Auxiliary Publication Service*, AIP document no. PAPS JASMA-77-753-29.

LASRY, M. J., and STERN, R. M. (1984). "A Posteriori Estimation of Correlated Jointly Gaussian Mean Vectors," *IEEE Trans. on Pattern Anal. and Mach. Intel.* **6:** 530-535.

CROWLEY, J. L., and STERN, R. M., Jr. (1984). "Fast Computation of the Difference of Low Pass (DOLP) Transform," *IEEE Transactions on Pattern Analysis and Machine Intelligence* **6:** 212-222.

STERN, R. M., Jr., SLOCUM, J. E., and PHILLIPS, M. S. (1983). "Interaural Time and Amplitude Discrimination in Noise", *J. Acoust. Soc. Amer.* **73:**1714-1722.

YOST, W. A., GRANTHAM, D. W., LUFTI, R. A., and STERN, R. M., Jr. (1982). "The Phase Angle of Addition in Temporal Masking for Diotic and Dichotic Listening Conditions," *Hearing Res.* **7:** 247-259.

MURTI, K. G., STERN, R. M., CANTEKIN, E. I. and BLUESTONE, C. D. (1982). "Classification of Spectral Patterns Obtained from Eustacian Tube Sonometry," *IEEE Trans. Biomed. Eng.* **29:** 473-477.

MURTI, K. G., STERN, R. M., Jr., CANTEKIN, E. I. and BLUESTONE, C. D. (1980). "Sonometric Evaluation of Eustachian Tube Function Using Broadband Stimuli", *Annals of Otology, Rhinology, and Laryngology*, (Suppl. 68) 89, 178-189.

RUOTOLO, B. R., STERN, R. M., Jr., and COLBURN, H. S. (1979). "Discrimination of Symmetric, Time-Intensity Traded Binaural Stimuli," *J. Acoust. Soc. Amer.*, **66:** 1733-1737.

STERN, R. M., Jr. and COLBURN, H. S. (1978). "Theory of Binaural Interaction Based on Auditory-Nerve Data. IV. A Model for Subjective Lateral Position," *J. Acoust. Soc. Amer.*, **64:** 127-140.

Critically-Reviewed Books, Book Chapters, and Theses

STERN, R. M. and MENON, A. (2020). "Binaural Technology for Machine Speech Recognition and Understanding," Chapter in *The Technology of Binaural Understanding*, J. Blauert and J. Braasch, Eds., Springer Nature.

MITRA, V., FRANCO, H., STERN, R. M., VAN HOUT, J., FERRER, L.,GRACIARENA, M., WANG, W., VERGYRI, D., ALWAN, A., and HANSEN, J.H.L. (2017). "Robust features in Deep Learning based Speech Recognition," Chapter in *New Era for Robust Speech Recognition: Exploiting Deep Learning*, S. Watanabe, M. Delcroix, F.Metze, & J. R. Hershey (eds.), pp. 165196, Springer International Publishing.

STERN, R. M. and MORGAN, N. (2013). "Features Based on Auditory Physiology and Perception," Chapter in *Noise-Robust Techniques for Automatic Speech Recognition,* T. Virtanen, R. Singh, and B. Raj, Eds., Wiley Press.

STERN, R. M., WANG, D., and BROWN, G. (2006). "Binaural Sound Localization," Chapter in *Computational Auditory Scene Analysis: Principles, Algorithms and Applications*, D. Wang and G. Brown, Eds., Wiley and IEEE Press.

STERN, R. M., TRAHIOTIS, C., and RIPEPI, A. M. (2006). "Fluctuations in Amplitude and Frequency Enable Interaural Delays to Foster the Identification of Speech-like Stimuli," Chapter in *Dynamics of Speech Production and Perception*, P. Divenyi, Ed., IOS Press.

TRAHIOTIS, C., BERNSTEIN, L. R., STERN, R. M., and BUELL, T. N. (2005). "Interaural Correlation as the Basis of a Working Model of Binaural Processing: An Introduction," Chapter in *Springer Handbook of Auditory Research: Sound Source Localization*, R. Fay and T. Popper, Eds., Springer-Verlag.

STERN, R. M. (2004). "Signal Separation Motivated by Human Auditory Perception: Applications to Automatic Speech Recognition," Chapter in *Speech Separation by Humans and Machines*, P. Divenyi, Ed., Springer-Verlag.

SINGH, R., STERN, R. M., and RAJ, B. (2002). "Signal and Feature Compensation Methods for Robust Speech Recognition," Chapter in *CRC Handbook on Noise Reduction in Speech Applications*, Gillian Davis, Ed., Boca Raton: CRC Press.

SINGH, R., RAJ, B., and STERN, R. M. (2002). "Model Compensation and Matched Condition Methods for Robust Speech Recognition," Chapter in *CRC Handbook on Noise Reduction in Speech Applications*, Gillian Davis, Ed., Boca Raton: CRC Press.

STERN, R. M., ACERO, A., LIU, F.-H., and OHSHIMA, Y. (1996). "Signal Processing for Robust Speech Recognition," Invited chapter in *Speech Recognition*, pp. 351-378, C.-H. Lee and F. Soong, Eds., Boston: Kluwer Academic Publishers.

STERN, R. M., and TRAHIOTIS, C. (1996). "Models of Binaural Perception," Invited chapter in *Binaural and Spatial Hearing in Real and Virtual Environments*, pp. 499-531, R. Gilkey and T. R. Anderson, Eds. New York: Lawrence Erlbaum Associates

STERN, R. M. (1995). "Robust Speech Recognition," Invited chapter in *Survey on the State of the Art in Speech and Natural Language Processing*, R. A. Cole et al., Ed.

STERN, R. M., and TRAHIOTIS, C. (1995). "Models of Binaural Interaction," Invited chapter in *Handbook of Perception and Cognition, Volume 6: Hearing*, pp. 347-386, B. C. J. Moore., Ed. New York: Academic Press.

STERN, R. M., Jr. (1976b). *Lateralization, Discrimination, and Detection of Binaural Pure Tones*, Ph.D. Thesis, Electrical Engineering Department, MIT, December, 1976.

Non-reviewed Book

STERN, R. M. (2020). *Selected Advanced Topics in Digital Signal Processing.* Textbook for CMU course 18-792.

Invited Conference Presentations

STERN, R. M. (2019). "Comparing Models of Binaural Lateralization," invited talk at the Macquarie University ARC Laureate Workshop, November 2019, Sydney, Australia.

STERN, R. M. (2019). "How we can apply binaural hearing principles to improve speech recognition accuracy," invited talk at the Bash Down Under, November 2019, Sydney, Terrigal, Australia.

STERN, R. M. (2017). "Predicting Binaural Lateralization, Interaural Discrimination, and Binaural Detection Using the Position-Variable Model," invited talk at the Macquarie University ARC Laureate Workshop: Creating a Sense of Auditory Space, October 2017, Sydney, Australia.

STERN, R. M, KIM, C., MOGHIMI, A.R., and MENON, A. (2016). "Binaural technology and automatic speech recognition," invited talk at the International Congress on Acoustics, September 2016, Buenos Aires, Argentina.

STERN, R. M. (2016). “Applying Models of Auditory Processing to Automatic Speech Recognition: Progress and Promise,” invited keynote talk at the 2015 Meeting of the Information Processing Sociaty of Japan: Special Interest Group in Spoken Language Processing, Toyama, Japan, February, 2016.

STERN, R. M. (2014). “Applying Models of Auditory Processing to Automatic Speech Recognition: Progress and Promise,” invited talk at the Frederick Jelinek Memorial Workshop on Meaning Representations in Language and Speech Processing, Prague, Czech Republic, July, 2014.

STERN, R. M. (2014). “Robust Automatic Speech Recognition in the 21st Century,” invited keynote talk at the 2014 AFEKA Conference for Speech Processing, Tel Aviv, Israel, July, 2014.

STERN, R. M. (2014). “Applying Models of Auditory Processing to Automatic Speech Recognition: Promise and Progress,” International Symposium on Speech Recognition, University of Zaragoza, Spain, May, 2014.

STERN, R. M. (2011). “Applying Physiologically-Motivated Models of Auditory Processing to Automatic Speech Recognition,” invited talk at the Third International Symposium on Auditory and Audiological Research, Nyborg, Denmark, August, 2011.

STERN, R. M. (2010). “The impact of the distribution of internal delays in binaural models on predictions for psychoacoustical data,” invited talk at the 161th Meeting of the Acoustical Society of America, Cancun, Mexico, November, 2010.

STERN, R. M. (2009). “New Directions in Robust Speech Recognition: What We Can Learn from Auditory Models,” invited keynote address at the Symposium on Frontiers of Research in Speech and Music, Gwalior, India, December, 2009.

STERN, R. M. (2009). “New Directions in Robust Automatic Speech Recognition,” invited keynote address at the Workshop on Image and Speech Processing, Hyderabad, India, December, 2009.

STERN, R. M. (2008). “Applying Physiologically-Motivated Models of Auditory Processing to Automatic Speech Recognition: Promises, Progress, and Problems,” Invited keynote address at the ISCA Tutorial and Research Workshop on Statistical and Perceptual Audition, Brisbane, Australia, September, 2008.

STERN, R. M, GOUVEA, E., KIM, C., KUMAR, K., and PARK, H.-M. (2008). “Binaural and Multiple-Microphone Processing for Robust Automatic Speech Recognition,” Invited keynote address at the IEEE Workshop on Hands-free Speech Communication and Microphone Arrays, Trento, Italy, May, 2008.

STERN, R. M. (2004). “Signal Processing for Sound Separation and Robust Representation,” Invited keynote address at AFOSR/NSF Symposium on Speech Separation and Comprehension in Complex Acoustic Environments, Montreal, Quebec, November 2004.

STERN, R. M. (2003). “Signal Separation Motivated by Auditory Processing: Applications to Speech Recognition,” invited review talk at the NSF Symposium on Signal Separation, Montreal, Quebec, November, 2003.

STERN, R. M. (2003). “Signal Processing for Robust Recognition,” invited talk at the NAIST International Center of Excellence Symposium, Nara, Japan, March, 2003.

STERN, R. M. (2002). “Using Computational Models of Binaural Hearing to Improve Automatic Speech Recognition Accuracy: Promise, Progress, and Problems,” AFOSR Workshop on Computational Audition, Columbus, Ohio, August, 2002.

STERN, R. M. (2000). “Robust Signal Representations for Automatic Speech Recognition,” Institute for Mathematics and Its Applications Workshop on the Mathematical Foundations of Speech Processing and Recognition, Minneapolis, Minnesota, September, 2000.

STERN, R. M. (2000). “The Language of Music,” invited keynote talk presented at the Third International Symposium on Text, Speech, and Dialog, Brno, Czech Republic, September, 2000.

STERN, R. M. (2000). “Tendencias Actuales en el Procesamiento del Lenguaje Hablado y Sistemas Conversacionales (Current Trends in Spoken Language Processing and Conversational Systems)”, invited keynote talk at the XV Simposium Internacional de Electrónica y Comunicación, Instituto Tecnológico de Estudios Superiores de Monterrey Mexico, February, 2000.

STERN, R. M. (1999). “Tendencias Actuales en el Procesamiento del Lenguaje Hablado y Sistemas Conversacionales (Current Trends in Spoken Language Processing and Conversational Systems)”, invited keynote talk at the XXIV Simposium Internacional de Sistemas Computacionales, Instituto Tecnológico de Estudios Superiores de Monterrey, Monterrey, Mexico, March, 1999.

STERN, R. M., and TRAHIOTIS, C. (1997). “Binaural Mechanisms that Emphasize Consistent Interaural Timing Information over Frequency,” invited keynote talk in *Psychophysical and Physiological Advances in Hearing,* Proceedings of the XI International Symposium on Hearing, August, 1997, Grantham, United Kingdom. A. R. Palmer, A. Rees, A. Q. Summerfield, and R. Meddis, Eds., Whurr Publishers, London, 1998.

STERN, R. M., RAJ, B., and MORENO, P. J. (1997). “Compensation for Environmental Degradation in Automatic Speech Recognition,” invited keynote talk presented at the *Proc. of the ESCA Tutorial and Research Workshop on Robust Speech Recognition for Unknown Communication Channels*, April, 1997, Pont-au-Mousson, France, pp. 33-42.

STERN, R. M. (1996). “The Current State of the Art of in Speech Recognition (Estado-da-Arte em Reconhecimento de Voz),” invited keynote talk presented at VOICETECH’96, the First Brazilian Workshop in Automatic Speech Recognition Campinas, Sao Paolo, Brazil, September, 1996.

STERN, R. M. (1996). “New Directions in Spoken Language Processing,” invited talk at the Second Joint NSF/CONACyT Workshop on Bilateral Collaboration, Jalapa, Mexico, March, 1996.

STERN, R. M. (1996). “Tendencias Actuales en el Procesamiento del Lenguaje Hablado (Current Trends in Spoken Language Processing)”, invited talk at the Universidad Veracruzana, Jalapa, Mexico, March, 1996.

STERN, R. M., and SULLIVAN, T. M. (1996). “Robust Speech Recognition Using Signal Processing Based On Binaural Perception,” invited talk presented at the First Forum Acusticum, Antwerp, Belgium, April, 1996.

STERN, R. M., MORENO, P.J., and RAJ, B. (1996). “Compensation for Speech Recognition in Degraded Acoustical Environments,” invited talk at the 132th meeting of the Acoustical Society of America, Honolulu, Hawaii, December, 1996.

STERN, R. M. (1995). “Nuevos Enfoques en Procesamiento de Lenguaje Hablado (New Directions in Spoken Language Processing),” invited talk at the Universitat Politecnica de Catalunya, Barcelona, Spain, September, 1995.

STERN, R. M. (1995). “New Directions in Spoken Language Processing,” invited talk presented at the Telefónica Investigación y Desarrollo Laboratory Symposium on Spoken Language Processing, Madrid, Spain, September 1995.

STERN, R. M. (1995). "Automatic Speech Recognition using Signal Processing based on Auditory Physiology and Perception," invited paper presented at the 129th meeting of the Acoustical Society of America, Washington, D.C., June, 1995.

MORENO, P. J., RAJ, B., and STERN, R. M. (1995). "Approaches to Environmment Compensation in Automatic Speech Recognition," invited paper presented at the 15th International Conference on Acoustics, Trondheim, Norway, Vol. III, pp. 109-112, June, 1995.

STERN, R. M., and SULLIVAN, T. M. (1994). "Robust Speech Recognition Based on Human Binaural Perception," invited paper presented at the ATR workshop on A Biological Framework for Speech Perception and Production, Kansai Science City, September, 1994. Reprinted in ATR technical report TR-H-121: *Proceedings of the ATR workshop on A Biological Framework for Speech Perception and Production*, 122 pages, (1995).

STERN, R. M. LIU, F.-H., SULLIVAN, T. M., MORENO, P. J., and ACERO, A. (1994). "Multiple Approaches to Robust Speech Recognition," invited keynote paper at the Fifth Western Pacific Regional Acoustical Conference, Seoul, Korea, August, 1994.

STERN, R. M. (1993). "Models of Binaural Interaction," invited keynote paper at the AFOSR Conference on Binaural and Spatial Hearing, Wright-Patterson Air Force Base, September, 1993.

STERN, R. M. (1993). "Psychoacoustical Basis of Machine Speech Recognition," invited talk at the Annual Meeting of the American Association for the Advancement of Science, February, 1993.

STERN, R. M. (1989). "Recent Progress in Spoken-Language Systems," invited lecture at the Second International Symposium on Artificial Intelligence, Monterrey, Mexico, October, 1989.

STERN, R. M. (1988). "Overview of Models of Binaural Perception," invited review paper at the 1988 National Research Council CHABA Symposium, Washington, D.C., October, 1988.

STERN, R. M. (1988). "Estado Actual de la Tecnología de Entradas/Salidas de Canales de Voz (Overview of Current Voice Input/Output Technologies)," invited keynote lecture at the XIII Simposium Internacional de Sistemas Computacionales, Monterrey, Mexico, March, 1988.

COLE, R. A., STERN, R. M., and LASRY, M. J. (1986). "Performing Fine Phonetic Distinctions: Templates vs. Features," invited talk, reprinted in *Invariance and Variability of Features in Spoken English Letters*, J. Perkell *et al*., eds., Lawrence Erlbaum, New York.

Critically-Reviewed Conference Presentations

LINDSEY, M., KUBALA, F., and STERN, R. M. (2025). "A Unified Metric for Simultaneous Evaluation of Error Rate and Annotation Cost," *Proc. IEEE ICASSP 2025,* April 2025, Hyderabad, India.

LINDSEY, M., VUONG, T., and STERN, R.M. (2023). "Unsupervised voice type discrimination score adaptation using X-vector clusters," *Proc. IEEE ICASSP 2023*, June 2023, Rhodes, Greece.

LINDSEY, M., ROBINSON, N.R., KUBALA, K., STERN, R.M. (2023) "Reducing the cost of spoof detection labeling using mixed-strategy active learning and pretrained models*," Proc. IEEE Automatic Speech Recognition and Understanding Workshop (ASRU)*, December 2023, Taipei, Taiwan.

VUONG, T., MADAAN, N., PUNDA, R., and STERN, R.M. (2023). "Investigating the important temporal modulations for deep-learning-based speech activity detection," Proc. *IEEE Spoken Language Technologies Workshop*, January 2023, Doha, Qatar.

ALVARADO, E., GRAGEDA, N., LUZANTO, A., MAHU, R., WUTH, J., MENDOZA, L., STERN, R. M, and YOMA, N. B. "Respiratory Distress Estimation in Human-Robot Interaction Scenario". *Proc. Interspeech 2023*, Dublin, Ireland, September 2023, Dublin Ireland.s

VUONG, T., and STERN, R. M. (2022). "Improved modulation-domain loss for neural-network-based speech enhancement," *Interspeech 2022*, September 2022, Incheon, Korea.

VUONG, T., LINDSEY, M., XIA, Y., STERN, R. (2022). "*L3DAS22: Exploring loss functions for 3D speech enhancement," Proc. L3DAS22: Machine Learning for 3D Audio Signal Processing,"* May 2022.

VUONG, T., XIA, Y., and STERN, R. M. (2021). "The application of learnable STRF kernels to the 2021 Fearless Steps Phase-03 SAD Challenge," *Interspeech 2021*, September 2021, Brno, Czech Republic.

XIA, Y., CHEN, L.—W., RUDNICKY, A., and STERN, R. M. (2021). "Temporal context in speech emotion recognition," *Interspeech 2021*, September 2021, Brno, Czech Republic.

VUONG, T., XIA, Y., and STERN, R. M. (2021). "A modulation-domain loss for neural-network-based real-time speech enhancement," *Proc. IEEE ICASSP 2021*, June 2021, Toronto, Canada.

VUONG, T., XIA, Y., and STERN, R. M. (2020). Learnable Spectro-temporal Receptive Fields for Robust Voice Type Discrimination, Interspeech *2018*, September 2020, Shanghai, China.

MENON, A., Kim, C., and STERN, R. M. (2019), "Robust recognition of reverberant and noisy speech using coherence-based processing," *IEEE International Conference on Acoustics, Speech, and Signal Processing*, May 2019, Brighton, United Kingdom.

XIA, Y. and STERN, R. M. (2018). *A Priori* SNR Estimation Based on a Recurrent Neural Network for Robust Speech Enhancement, *Interspeech 2018*, September 2018, Hyderabad, India.

MICHELSON, J., STERN, R. M., and SULLIVAN, T. M. (2018), "Automatic guitar tablature transcription from audio using inharmonicity regression and Bayesian classification," *145th Convention of the Audio Engineering Society*, October 2018, New York City.

KIM, C., MENON, A., BACCHIANI, M., and STERN, R. M. (2018), "Source separation using phase difference and reliable mask selection," *IEEE International Conference on Acoustics, Speech, and Signal Processing*, April 2018, Calgary, Alberta, Canada.

MENON, A., KIM, C., KUROKAWA, U., and STERN, R. M. (2017), "Binaural Processing for Robust Recognition of Degraded Speech," *IEEE Automatic Speech Recognition and Understanding Workshop,* December 2017, Naha, Okinawa, Japan.

MENON, A., KIM, C., and STERN, R. M. "Robust Speech Recognition Based on Binaural Auditory Processing*," Interspeech 2017*, September 2017, Stockholm, Sweden.

NOVOA, J., WUTH, J., ESCUDERO, J. P., FREDES, J., MAHU, R., STERN, R. M., and BECERRA YOMA, N., "Robustness over time-varying channels in DNN-HMM ASR-Based human-robot interaction," *Interspeech 2017,* September 2017, Stockholm, Sweden.

POBLETE, V., ESCUDERO, J. P., FREDES, J., NOVOA, J., STERN, R. M., KING, S., and BECERRA YOMA, N. (2016). The use of locally normalized cepstral coefficients (LNCC) to improve speaker recognition accuracy in highly reverberant rooms. *Proceedings of Interspeech 2016,* September 2016, San Francisco, California.

FREDES, J., NOVOA, J., POBLETE, V., KING, S., STERN, R. M., and YOMA, N. B. (2015), "Robustness to additive noise of locally-normalized cepstral coefficients in speaker verification," *Interspeech 2015*, September 2015, Dresden, Germany.

HARVILLA, M. J., and STERN, R. M. (2015). Efficient Audio Declipping Using Regularized Least Squares. *ICASSP: International Conference on Acoustics, Speech and Signal Processing,* April 2015, Brisbane Australia.

HARVILLA, M. J., and STERN, R. M. (2014). "Least squares declipping for robust speech recognition," *Interspeech 2014*, September 2014, Singapore.

MOGHIMI, A. R., RAJ, B., and STERN, R. M. (2014), "Post-masking: A hybrid approach to array processing for speech recognition," *Interspeech 2014*, September 2014, Singapore.

KIM, C., CHIN, K. K., BACCHIANI, M., and STERN, R. M. (2014), "Robust speech recognition using temporal masking and threshold algorithm," *Interspeech 2014*, September 2014, Singapore.

PARK, H.-M., MACIEJEWSKI, M., KIM, C., and STERN, R. M. (2014). "Robust speech recognition in reverberant environments using subband-based steady-state monaural and binaural suppression," *Interspeech 2014*, September 2014, Singapore.

MOGHIMI, A. R., and STERN, R. M. (2014), "An analysis of binaural spectro-temporal masking as nonlinear beamforming," *IEEE International Conference on Acoustics, Speech, and Signal Processing*, May 2014, Florence, Italy.

HARVILLA, M., and STERN, R. M. (2012). "Histogram-based subband power warping and spectral averaging for robust speech recognition under matched and multistyle training," *IEEE International Conference on Acoustics, Speech, and Signal Processing*, March 2012, Kyoto, Japan.

KIM, C. and STERN, R. M. (2012). "Power-normalized cepstral coefficients (PNCC) for robust speech recognition, *IEEE International Conference on Acoustics, Speech, and Signal Processing*, March 2012, Kyoto, Japan.

KIM, C., KHAWAND, C, and STERN, R. M. (2012). "Two-microphone source separation algorithm based on statistical modeling of angle distributions," *IEEE International Conference on Acoustics, Speech, and Signal Processing*, March 2012, Kyoto, Japan.

KIM, C., KUMAR, K., and STERN, R. M. (2011). "Binaural sound source separation motivated by auditory processing", *IEEE International Conference on Acoustics, Speech, and Signal Processing*, May 2011, Prague, Czech Republic.

KUMAR, K., KIM, C., and STERN, R. M. (2011). "Delta-spectral cepstral coefficients for robust speech recognition, *IEEE International Conference on Acoustics, Speech, and Signal Processing*, May 2011, Prague, Czech Republic.

KUMAR, K., RAJ, B., SINGH, R., and STERN, R. M. (2011)."An iterative least-squares techique for dereverberation," *IEEE International Conference on Acoustics, Speech, and Signal Processing*, May 2011, Prague, Czech Republic.

KUMAR, K., SINGH, R., RAJ, B., and STERN, R. M. (2011). "Gammatone sub-band magnitudedomain dereverberation, *IEEE International Conference on Acoustics, Speech, and Signal Processing*, May 2011, Prague, Czech Republic.

KIM, C., STERN, EOM, K., and Lee, J. (2010). "Automatic selection of thresholds for signal separation algorithms based on interaural delay," *Interspeech 2010, September 2010,* Makuhari, Japan.

KIM, C., and STERN, R. M. (2010). "Nonlinear enhancement of onset for robust speech recognition," *Interspeech 2010, September 2010,* Makuhari, Japan.

AL BAWAB, Z., RAJ, B, and STERN, R. M. (2010). "A hybrid physical and statistical dynamic articulatory framework incorporating analysis-by-synthesis for improved phone classification," *IEEE International Conference on Acoustics, Speech, and Signal Processing*, March 2010, Dallas, Texas.

CHIU, Y.-H., RAJ, B., and STERN, R. M. (2010). "Learning Based Auditory Encoding For Robust Speech Recognition," *IEEE International Conference on Acoustics, Speech, and Signal Processing*, March 2010, Dallas, Texas.

KIM, C., and STERN, R. M. (2010). "Feature Extraction For Robust Speech Recognition Based On Maximizing The Sharpness Of The Power Distribution And On Power Flooring," *IEEE International Conference on Acoustics, Speech, and Signal Processing*, March 2010, Dallas, Texas.

KUMAR, K., and STERN, R. M. (2010). "Maximum-Likelihood-Based Cepstral Inverse Filtering For Blind Speech Dereverberation," *IEEE International Conference on Acoustics, Speech, and Signal Processing*, March 2010, Dallas, Texas.

CHIU, Y.-H. B, and STERN, R. M. (2009). "Minimum variance modulation filters for robust speech recognition," *IEEE International Conference on Acoustics, Speech, and Signal Processing*, April 2009, Taipei, Taiwan.

AL BAWAB, Z., TURICCHIA, L., STERN, R. M., and RAJ, B. (2009). "Deriving vocal tract shapes from electromagnetic articulograph data via geometric adaptation and matching, *Interspeech 2009, September 2009,* Brighton, United Kingdom.

BUERA, L., MIGUEL, A., ORTEGA, E., LLEIDA, E., and STERN, R. (2009). "Unsupervised training scheme with non-stereo data for empirical feature vector compensation, *Interspeech 2009, September 2009,* Brighton, United Kingdom.

CHIU, Y.-H. B., RAJ, B., and STERN, R. M. (2009). "Learning-based auditory encoding," *Interspeech 2009, September 2009,* Brighton, United Kingdom.

GU, L., and STERN, R. M. (2009). "Speaker segmentation and clustering for sumultaneously-presented speech," *Interspeech 2009, September 2009,* Brighton, United Kingdom.

KIM, C., KUMAR, K., RAJ, B., and STERN, R. M. (2009). "Signal separation for robust speech recognition based on phase difference information obtained in the frequency domain," *Interspeech 2009, September 2009,* Brighton, United Kingdom.

KIM, C., and STERN, R. M. (2009). “Feature extraction for robust speech recognition using a power-law nonlinearity and power-bias subtraction,” *Interspeech 2009, September 2009,* Brighton, United Kingdom.

KIM, C., and STERN, R. M. (2009). "Power Function-Based Power Distribution Normalization Algorithm for Robust Speech Recognition," *IEEE Automatic Speech Recognition and Understanding Workshop*, December 2009, Merano, Italy.

KIM, C., and STERN, R. M. (2009). "Robust Speech Recognition using a Small Power Boosting Algorithm," *IEEE Automatic Speech Recognition and Understanding Workshop*, December 2009, Merano, Italy.

CHIU, Y.-H., and STERN, R. M. (2008). "Analysis of Physiologically-Motivated Signal Processing for Robust Speech Recognition," Interspeech 2008, September 2008, Brisbane, Australia.

KIM, C., and STERN, A. M. (2008). "Robust Signal-to-Noise Ratio Estimation Based on Waveform Amplitude Distribution Analysis," Interspeech 2008, September 2008, Brisbane, Australia.

AL BAWAB, Z., RAJ, B., and STERN, R. M. (2008). “Analysis-by-synthesis features for speech recognition,” IEEE International Conference on Acoustics, Speech, and Signal Processing, April 2008, Las Vegas, Nevada.

GU, L., and STERN, R. M., “Single-channel speech separation based on modulation frequency,” IEEE International Conference on Acoustics, Speech, and Signal Processing, April 2008, Las Vegas, Nevada.

KUMAR, K., and STERN, R. M. (2008). “Environment-invariant compensation for reverberation using linear post-filtering for minimum distortion,” IEEE International Conference on Acoustics, Speech, and Signal Processing, April 2008, Las Vegas, Nevada.

STERN, R. M., GOUVEA, E., and THATTAI, G. (2007). “‘Polyaural’ array processing for automatic speech recognition in degraded environments,” Interspeech 2007, August 2007, Antwerp, Belgium.

PARK, H.-M., and STERN, R. M. (2007). “Missing-feature speech recognition using dereverberation and echo supporession in reverberant environments,” IEEE International Conference on Acoustics, Speech, and Signal Processing, April 2007, Honolulu, Hawaii.

KUMAR, K., CHEN, T., and STERN, R. M. (2007). “Profile view lip reading,” IEEE International Conference on Acoustics, Speech, and Signal Processing, April 2007, Honolulu, Hawaii.

KIM, C., CHIU, Y.-H., and STERN, R. M. (2006). “Physiologically-motivated synchrony-based processing for robust automatic speech recognition,” Interspeech 2006, September 2006, Pittsburgh, Pennsylvania.

NARAYANASWAMY, B., GANGAGHARIAH, R., and STERN, R. M. (2006). “Voting for two speaker segmentation,” Interspeech 2006, September 2006, Pittsburgh, Pennsylvania.

PARK, H.-M., and STERN, R. M. (2006). “Spatial separation of speech sgnals using continuously-variable masks estimated from comparisons of zero crossings,” IEEE International Conference on Acoustics, Speech, and Signal Processing, May 2006, Toulouse, France.

KIM, W., and STERN, R. M. (2006). “Band-independent mask estimation for missing-feature reconstruction,” IEEE International Conference on Acoustics, Speech, and Signal Processing, May 2006, Toulouse, France.

KIM, W., STERN, R. M., and KO, H. (2005). “Environment-Independent Mask Estimation for Missing Feature Reconstruction,” *Proc. Eurospeech-2005* September, 2005, Lisbon, Portugal.

LI, X., and STERN, R. M. (2004). “Parallel Feature Generation Based on Maximum Normalized Acoustic Likelihood for Improved Combination Performance,” *Proc. of the International Conference of Spoken Language Processing*, October, 2004, Jeju Island, Korea.

LI, X., and STERN, R. M. (2004). “Feature Generation Based on Maximum Normalized Acoustic Likelihood for Improved Speech Recognition,” *IEEE International Conference on Acoustics, Speech, and Signal Processing*, May 2004, Montreal.

RAJ, B., SINGH, R., and STERN, R. M. (2004). “On Tracking Noise with Linear Dynamical System Models,” *IEEE International Conference on Acoustics, Speech, and Signal Processing*, May 2004, Montreal.

SELTZER, M. L., and STERN, R. M. (2003). “Parameter Sharing in Subband Likelihood-Maximizing Beamforming for Speech Recognition using Microphone Arrays,” *IEEE International Conference on Acoustics, Speech, and Signal Processing*, May 2004, Montreal.

LI, X., and STERN, R. M. (2003). “Feature Generation Based on Maximum Classification Probability for Improved Speech Recognition," *Proc. Eurospeech-2003*, September, 2003, Geneva, Switzerland.

NEDEL, J. P., and STERN, R. M. (2003). “Duration Normalization and Hypothesis Combination for Improved Spontaneous Speech Recognition,” *Proc. Eurospeech-2003*, September, 2003, Geneva, Switzerland.

OBUCHI, Y., and STERN, R. M. (2003). “Normalization of Time-Derivative Parameters using Histogram Equalization,” *Proc. Eurospeech-2003*, September, 2003, Geneva, Switzerland.

LI, X., and STERN, R. M. (2003). “Training of Stream Weights for the Decoding of Speech using Parallel Feature Streams,” *IEEE International Conference on Acoustics, Speech, and Signal Processing*, April 2003, Hong Kong.

SELTZER, M. L., and STERN, R. M. (2003). “Subband Parameter Optimization of Microphone Arrays for Speech Recognition in Reverberant Environments,” *IEEE International Conference on Acoustics, Speech, and Signal Processing*, April 2003, Hong Kong.

LI, X., SINGH, R., and STERN, R. M. (2002). "Lattice Combination for Improved Speech Recognition," *Proc. of the International Conference of Spoken Language Processing*, September, 2002, Denver, Colorado.

SELTZER, M. L., RAJ, B., and STERN, R. M. (2002). “Speech Recognizer-Based Microphone Array Processing for Robust Hands-Free Speech Recognition,” Proc. IEEE Conf. on Acoustics, Speech, and Sig. Proc., May, 2002, Orlando, Florida.

RAJ, B., SELTZER, M. L., and STERN, R. M. (2001). “Robust Speech Recognition: The Case for Restoring Missing Features,” Proc. of the Workshop on Consistent and Reliable Auditory Cues, September, 2001, Aalborg, Denmark.

SINGH, R., SELTZER, M. L., RAJ, B., and STERN, R. M. (2001). “Speech in Noisy Environments: Robust Automatic Segmentation, Feature Extraction, and Hypothesis Combination,” *Proc. IEEE Conf. on Acoustics, Speech, and Sig. Proc.*, May, 2001, Salt Lake City, Utah.

NEDEL, J. N., and STERN, R. M. (2001). "Duration Normalization for Improved Recognition of Spontaneous and Read Speech via Missing Feature Methods," *Proc. IEEE Conf. on Acoustics, Speech, and Sig. Proc.*, May, 2001, Salt Lake City, Utah.

DOH, S.-H., and STERN, R. M. (2000). "Using Class Weighting in Inter-Class MLLR," *Proc. of the International Conference of Spoken Language Processing*, October, 2000, Beijing, China.

HUERTA, J. M., and STERN, R. M. (2000). "Instantaneous Distortion-Based Weighted Acoustic Modeling for Robust Recognition of Coded Speech," *Proc. of the International Conference of Spoken Language Processing*, October, 2000, Beijing, China.

NEDEL, J. P., SINGH, R., and STERN, R. M. (2000a). "Automatic Subword Unit Refinement for Spontanteous Speech Recognition via Phoneword Splitting," *Proc. of the International Conference of Spoken Language Processing*, October, 2000, Beijing, China.

NEDEL, J. P., SINGH, R., and STERN, R. M. (2000b). "Phone Transition Acoustic Modeling: Application to Speaker Independent and Spontaneous Speech Systems," *Proc. of the International Conference of Spoken Language Processing*, October, 2000, Beijing, China.

RAJ, B., SELTZER, M. L., and STERN, R. M. (2000). "Reconstruction of Damaged Spectrographic Features for Robust Speech Recognition," *Proc. of the International Conference of Spoken Language Processing*, October, 2000, Beijing, China.

SELTZER, M. L., RAJ, B., and STERN, R. M. (2000). "Classifier-Based Mask Estimation for Missing Feature Methods of Robust Speech Recognition," *Proc. of the International Conference of Spoken Language Processing*, October, 2000, Beijing, China.

SINGH, R., RAJ, B., and STERN, R. M. (2000). "Structured Redefinition of Sound Units by Merging and Splitting for Improved Speech Recognition," *Proc. of the International Conference of Spoken Language Processing*, October, 2000, Beijing, China.

DOH, S.-J., and STERN, R M. (2000). "Inter-Class MLLR for Speaker Adaptation," *Proc. IEEE Conf. on Acoustics, Speech, and Sig. Proc.*, June, 2000, Istanbul, Turkey.

SINGH, R., RAJ, B., and STERN, R. M. (2000). "Automatic Generation of Phone Sets and Lexical Transcriptions," *Proc. IEEE Conf. on Acoustics, Speech, and Sig. Proc.*, June, 2000, Istanbul, Turkey.

DOH, S.-J., and STERN, R M. (1999). "Weighted Principal Component MLLR For Speaker Adaptation," *Proc. of the IEEE Workshop on Automatic Speech Recognition and Understanding*, December, 1999, Keystone, Colorado.

SINGH, R., RAJ, B., and STERN, R. M. (1999). "Domain Adduced State Tying for Cross-domain Acoustic Modelling," *Proc. Eurospeech-99*, September, 1999, Budapest, Hungary.

SINGH, R., RAJ, B., and STERN, R. M. (1999). "Automatic Clustering And Generation Of Contextual Questions For Tied States In Hidden Markov Models," *Proc. IEEE Conf. on Acoustics, Speech, and Sig. Proc.*, March, 1999, Phoenix, Arizona.

HUERTA, J. M., and STERN, R. M. (1999). "Distortion-class weighted acoustic modeling for Robust Speech Recognition under GSM RPE-LTP coding," *Proc. of the Workshop on Robust Methods for Speech Recognition in Adverse Conditions*, Tampere, Finland, 1999.

HUERTA, J. M., AND STERN, R. M. (1998). "Speech Recognition from GSM CODEC Parameters," *Proc. of the International Conference of Spoken Language Processing*, Sydney Australia, December, 1998.

RAJ, B., SINGH, R., and STERN, R. M. (1998). "Inference of Missing Spectrographic Features for Robust Speech Recognition," *Proc. of the International Conference of Spoken Language Processing*, Sydney Australia, December, 1998.

RAJ, B., GOUVEA, E., and STERN, R. M. (1997). "Cepstral Compensation using Statistical Linearization," *Proc. of the ESCA Tutorial and Research Workshop on Robust Speech Recognition for Unknown Communication Channels*, Pont-au-Mousson, France, April, 1997.

GOUVEA, E. B., and STERN, R. M. (1997). "Speaker Normalization through Formant-Based Warping of the Frequency Scale", *Proc. Eurospeech-97*, September, 1997, Rhodes, Greece.

HUERTA, J. M., and STERN, R. M. (1997). "Compensation for Environmental and Speaker Variability by Normalization of Pole Locations," *Proc. Eurospeech-97*, September, 1997, Rhodes, Greece.

RAJ, B, PARIKH, V., and STERN, R. M. (1997). "The Effects of Background Music on Speech Recognition Accuracy," *Proc. IEEE Conf. on Acoustics, Speech, and Sig. Proc.*, April, 1997, Munich, Germany.

RAJ, B., GOUVEA, E., MORENO, P. J., and STERN, R. M. (1996). "Cepstral Compensation by Polynomial approximation for Environment-Independent Speech Recognition," *Proc. of the International Conference on Spoken Language Processing,* Philadelphia, Pennsylvania, October, 1996.

MORENO, P. J., RAJ, B., and STERN, R. M. (1996). "A Vector Taylor Series Approach for Environment-Independent Speech Recognition," *Proc. of the IEEE International Conference on Acoustics, Speech, and Signal Processing*, Atlanta, Georgia, 1996.

MORENO, P. J., RAJ, B., and STERN, R. M. (1995). "A Unified Approach to Robust Speech Recognition," *Proc. of Eurospeech-95*, Madrid, Spain, September, 1995.

MORENO, P. J., RAJ, B., GOUVEA, E., and STERN, R. M. (1995). "Multivariate- GaussianBased Cepstral Normalization for Robust Speech Recognition," *Proc. of the IEEE International Conference on Acoustics, Speech, and Signal Processing*, Detroit, Michigan, 1995.

SIEGLER, M. A., and STERN, R. M. (1995). "On the Effects of Speech Rate in Large Vocabulary Speech Recognition Systems," *Proc. of the IEEE International Conference on Acoustics, Speech, and Signal Processing*, Detroit, Michigan, May, 1995.

STERN, R. M., LIU, F.-H., MORENO, P. J., and ACERO, A. (1994). "Signal Processing for Robust Speech Recognition," *Proc. of the International Conference on Spoken Language Processing*, Yokohama, Japan, September, 1994.

HANAI, N., and STERN, R. M. (1994). "Robust Speech Recognition in the Automobile," *Proc. of the International Conference on Spoken Language Processing*, Yokohama, Japan, September, 1994.

OHSHIMA, Y., and STERN, R. M. (1994). "Environmental Robustness in Automatic Speech Recognition Using Physiologically-Motivated Signal Processing," *Proc. of the International Conference on Spoken Language Processing*, Yokohama, Japan, September, 1994.

LIU, F.-H., STERN, R. M., ACERO, A., and MORENO, P. J. (1994). "Environment Normalization for Robust Speech Recognition using Direct Cepstral Comparison," *Proc. of the IEEE International Conference on Acoustics, Speech, and Signal Processing*, Adelaide, Australia, pp. II-61 II-64.

MORENO, P. J., and STERN, R. M. (1994). “Sources of Degradation of Speech Recognition in the Telephone Network,” *Proc. of the IEEE International Conference on Acoustics, Speech, and Signal Processing*, Adelaide, Australia, pp. I-109 - I-112.

LIU, F.-H., MORENO, P. J., STERN, R. M., and ACERO, A. (1994). “Signal Processing For Robust Speech Recognition,” *Proceedings of the Seventh ARPA Workshop on Human Language Technology*, Princeton, New Jersey, Morgan Kaufmann, C. J. Weinstein, Ed.

LIU, F.-H., MORENO, P. J., STERN, R. M., and ACERO, A. (1994). “Signal Processing For Robust Speech Recognition,” *Proceedings of the ARPA Workshop on Spoken Language Technology*, Princeton, New Jersey, R. M. Stern, Ed.

LIU, F.-H., STERN, R. M., HUANG, X., and ACERO, A. (1993). “Efficient Cepstral Normalization For Robust Speech Recognition,” *Proceedings of the Sixth ARPA Workshop on Human Language Technology*, Princeton, New Jersey, Morgan Kaufmann, M. Bates, Ed., pp. 69-74.

SULLIVAN, T. M., and STERN, R. M. (1993). “Multi-Microphone Correlation- Based Processing for Robust Speech Recognition,” *Proc. of the IEEE International Conference on Acoustics, Speech, and Signal Processing*, Minneapolis, Minnesota, **2**: 91-94.

STERN, R. M., LIU, F.-H., OHSHIMA, Y., SULLIVAN, T. M., and ACERO, A. (1992a). “Multiple Approaches to Robust Speech Recognition”, *Proc. of the Fifth DARPA Speech and Natural Language Workshop*, Harriman, New York, February, 1992.

STERN, R. M., LIU, F.-H., OHSHIMA, Y., SULLIVAN, T. M., and ACERO, A. (1992b). “Multiple Approaches to Robust Speech Recognition,” *Proc. of the Second International Conference on Spoken Language Processing*, Banff, Alberta, Canada, pp. 695-698, October, 1992.

LIU, F.-H., ACERO, A., and STERN, R. M. (1992). “Efficient Joint Compensation of Speech for the Effects of Additive Noise and Linear Filtering,” *Proc. of the IEEE International Conference on Acoustics, Speech, and Signal Processing*, San Francisco, California, pp. 865-868.

WARD, W., ISSAR, S., HUANG, X., HON, H.-W., HWANG, M.-Y., YOUNG, S., MATESSA, M., LIU, F.-H., and STERN, R. (1992). “Speech Understanding in Open Tasks”, *Proc. of the Fifth DARPA Speech and Natural Language Workshop, Harriman*, New York, February, 1992.

STERN, R. M., XU, X., and TAO, S. (1991). “A Coincidence-Based Model that Describes “Straightness” Weighting in Binaural Perception,” *Abstracts of the Fourteenth Midwinter Research Meeting of the Association for Research in Otolaryngology*, St. Petersburg Beach, Florida, p. 33(A).

STERN, R. M., and TRAHIOTIS, C. (1991). “The Role of Consistency of Interaural Timing over Frequency in Binaural Lateralization,” *Proc. of the Ninth International Symposium on Auditory Physiology and Perception*, Carcans, France.

ACERO, A. and STERN, R. M. (1991). “Robust Speech Recognition by Normalization of the Acoustic Space,” *Proc. of the IEEE International Conference on Acoustics, Speech, and Signal Processing*, Toronto, Ontario, pp. 893-896.

ROZZI, W. A. and STERN, R. M. (1991). “Fast Estimation of Mean Vectors using Adaptive Filtering,” *Proc. of the IEEE International Conference on Acoustics, Speech, and Signal Processing*, Toronto, Ontario, pp. 865-868.

ACERO, A. and STERN, R. M. (1990a). "Environmental Robustness in Automatic Speech Recognition," *Proc. of the IEEE International Conference on Acoustics, Speech, and Signal Processing*, Albuquerque, New Mexico, pp. 849-852.

ACERO, A. and STERN, R. M. (1990b). "Toward Microphone-Independent Spoken Language Systems," *Proceedings of the DARPA Speech and Natural Language Workshop*, Hidden Valley, PA, R. M. Stern, Ed., Morgan Kaufmann Publishers, Inc., San Mateo, CA,

ACERO, A. and STERN, R. M. (1990c). "Acoustical Pre-Processing for Robust Spoken Language Systems," *Proc. First International Conference on Spoken Language Processing*, pp. 1121-1124, Kobe, Japan, November, 1990.

STERN, R. M., ZEPPENFELD, T., and SHEAR, G. D. (1990). "Lateralization of RectangularlyModulated Noise: An Explanation for "Illusory" Reversals," *Abstracts of the Thirteenth Midwinter Research Meeting of the Association for Research in Otolaryngology*, St. Petersburg Beach, Florida, pp. 163-164(A).

STERN, R. M. and ACERO, A. (1989). "Acoustical Pre-Processing for Robust Speech Recognition," presented at the October, 1989, DARPA Workshop on Speech and Natural Language.

WARD, W. H., HAUPTMANN, A. G., STERN, R. M., and CHANAK, T. (1988). "Parsing Spoken Phrases Despite Missing Words," *Proc. of the IEEE International Conference on Acoustics, Speech, and Signal Processing*, pp. 275-278.

COAST, D. A., STERN, R. M., CANO, G. C., and BRILLER, S. A. (1987). "Cardiac Arrhythmia Analysis Using Hidden Markov Models," presented at the IEEE Engineering in Medicine Society Conference, November, 1987.

STERN, R. M. (1988). "Nuevos Enfoques en Reconocimiento Automatico de Habla (New Directions in Automatic Speech Recognition)", invited lecture at the XIII Simposium Internacional de Sistemas Computacionales, Monterrey, Mexico.

STERN, R. M., WARD, W. H., HAUPTMANN, A. G., and LEON, J. (1987). "Sentence Parsing with Weak Grammatical Constraints," *Proc. of the IEEE International Conference on Acoustics, Speech, and Signal Processing*, pp. 380-383.

LASRY, M.J., AND STERN, R.M. (1984). "Unsupervised Adaptation to New Speakers in Feature-Based Letter Recognition," *Proc. of the IEEE International Conference on Acoustics, Speech, and Signal Processing*, pp. 17.6.1-17.6.4.

STERN, R. M., and BACHORSKI, S. J. (1983). "Dynamic Cues in Binaural Perception," in *Hearing-Physiological Bases and Psychophysics*, R. Klinke and R. Hartmann, Eds., Springer Verlag Press, Heidelberg.

STERN, R. M., and LASRY, M. J. (1983). "Dynamic Speaker Adaptation for Isolated Letter Recognition Using MAP Estimation," *Proc. of the IEEE International Conference on Acoustics, Speech, and Signal Processing*, pp. 734-737.

COLE, R. A., STERN, R. M., PHILLIPS, M. S., BRILL, S. M., PILANT, A. P., and SPECKER, P. (1983). "Feature-based Speaker-Independent Recognition of Isolated English Letters," *Proc. of the IEEE International Conference on Acoustics, Speech, and Signal Processing*, pp. 731-734.

STERN, R. M., Jr. and RUBINOV, E. M. (1980). “Subjective Laterality of Noise-Masked Binaural Targets,” in *Psychological, Physiological, and Behavioural Studies of Hearing*, G. v.d. Brink and F. A. Bilsen, Eds., Delft University Press, Delft.

Non-Reviewed Submitted Conference Papers and Other

STECKER, G. C., DIETZ, M., and STERN, R. M. (2019). "Straightness" versus "Briefness" in Binaural Cue Extraction, talk at the 177th meeting of the Acoustical Society of America, Louisville, Kentucky, May, 2019.

STERN, R. M. (2017). “Predicting binaural lateralization and discrimination using the position variable model,” talk at the 173rd meeting of the Acoustical Society of America, Boston, Massachusetts, June, 2017.

DIETZ, M., MARQUARDT, T., MAJDAK, P., STERN, R. M., HARTMANN, W. M., GOODMAN, D. F., and EWERT, S. D. (2017). “An initiative for testability and comparability of binaural models, talk at the 173rd meeting of the Acoustical Society of America, Boston, Massachusetts, June, 2017.

Stern, R. M. (2016). The position-variable model at 40. talk at the 172nd meeting of the Acoustical Society of America, Honolulu, Hawaii, December, 2016.

PARK, H.-M., MACIEJEWSKI, M., KIM, C., and STERN, R. M. (2015). “Robust automatic speech recognition in reverberation: onset enhancement versus binaural source separation,” J. Acoust. Soc. Amer., **137**:2303 (A).

ROMIGH, G. D., STERN, R. M., BRUNGART, D. S., and SIMPSON, B. D. (2015). “A Bayesian framework for the estimation of head-related transfer functions,” J. Acoust. Soc. Amer., **137**:2323 (A).

STERN, R. M. (2003). “Signal Processing for Robust Recognition,” Speech Information Technology Research Center Seminar Series, Seoul, Korea, May, 2003.

NEDEL, J. P., and STERN, R. M. (2002). “Duration Normalization for Improved Automatic Speech Recognition,” *J. Acoust. Soc. Amer.*, **112**: 2321 (A).

HUERTA, J. M., CHEN, S. J., and STERN, R. M. (1999). “The 1998 Carnegie Mellon University Sphinx-3 Spanish Broadcast News Transcription System,” *Proc. of the DARPA Broadcast News Transcription and Understanding Workshop*, March, 1999, Herndon, Virginia.

HUERTA, J. M., THAYER, E., RAVISHANKAR, M., and STERN, R. M. (1998). “The Development of the 1997 CMU Spanish Broadcast News Transcription System,” *Proc. of the DARPA Broadcast News Transcription and Understanding Workshop*, February, 1998, Landsdowne, Virginia.

SEYMORE, K., CHEN, S., DOH, S.-J., ESKENAZI, M., GOUVEA, E., RAJ, B., RAVISHANKAR, M. ROSENFELD, R., SIEGLER, M. A., STERN, R. M., AND THAYER, E. (1998). “The 1997 CMU Sphinx-3 English Broadcast News Transcription System,” *Proc. of the DARPA Broadcast News Transcription and Understanding Workshop*, February, 1998, Landsdowne, Virginia.

SIEGLER, M. A., JAIN, U., RAJ, B., and STERN, R. M. (1997). “Automatic Segmentation, Classification, and Clustering of Broadcast News Audio,” *Proc. DARPA Speech Recognition Workshop*, February, 1997, Chantilly, Virginia.

PARIKH, V. N., RAJ, B., and STERN, R. M. (1997). “Speaker Adaptation and Environmental Compensation for the 1996 Broadcast News Task,” *Proc. DARPA Speech Recognition Workshop*, February, 1997, Chantilly, Virginia.

PLACEWAY, P., CHEN, S., ESKENAZI, M., JAIN, U., PARIKH, V., RAJ, B., RAVISHANKAR, M., ROSENFELD, R., SEYMORE, K., SIEGLER, M., STERN, R., and THAYER, E. (1997). "The 1996 Hub-4 SPHINX-3 System," *Proc. DARPA Speech Recognition Workshop*, February, 1997, Chantilly, Virginia.

STERN, R. M. (1997). "Specification of the 1996 Hub 4 Broadcast News Evaluation," *Proc. DARPA Speech Recognition Workshop*, February, 1997, Chantilly, Virginia.

GOUVEA, E. B., MORENO, P. J., RAJ, B., SULLIVAN, T. M., and STERN, R. M. (1996). "Adaptation and Compensation: Approaches To Microphone And Speaker Independence In Automatic Speech Recognition," *Proceedings of the ARPA Workshop on Speech Recognition Technology*, Harriman, NY, Morgan Kaufmann, D. Pallett, Ed.

JAIN, U., SIEGLER, M. A., DOH, S.-J., GOUVEA, E., MORENO, P. J., RAJ, B. and STERN, R. M. (1996). "Recognition Of Continuous Broadcast News With Multiple Unknown Speakers And Environments," *Proceedings of the ARPA Workshop on Speech Recognition Technology*, Harriman, NY, Morgan Kaufmann, D. Pallett, Ed.

STERN, R. M. (1996). "Specification of the 1995 ARPA Hub 3 Evaluation: Unlimited Vocabulary NAB News Baseline," *Proceedings of the ARPA Workshop on Speech Recognition Technology*, Harriman, NY, Morgan Kaufmann, D. Pallett, Ed.

MORENO, P. J., SIEGLER, M. A., JAIN, U., and STERN, R. M. (1995). "Continuous Recognition of Large-Vocabulary Telephone-Quality Speech," *Proceedings of the ARPA Workshop on Spoken Language Technology*, Austin, TX, Morgan Kaufmann, J. Cohen, Ed.

MORENO, P. J., JAIN, U., RAJ, B., and STERN, R. M. (1995). "Approaches to Microphone Independence in Automatic Speech Recognition," *Proceedings of the ARPA Workshop on Spoken Language Technology*, Austin, TX, Morgan Kaufmann, J. Cohen, Ed.

LEE, W., and STERN, R. M. (1994). "Consistency Over Frequency in High- Frequency Binaural Lateralization," *J. Acoust. Soc. Amer.*, **95**: 2896 (A).

STERN, R. M., and SULLIVAN, T. M. (1993). "Microphone-Array Algorithms for Robust Speech Recognition," talk at the Speech Research Symposium-XIII, Baltimore, Maryland, June, 1993.

TAO, S. H., and STERN, R. M. (1992). "Additive versus Multiplicative Combination of Differences of Interaural Time and Intensity," *J. Acoust. Soc. Amer.*, 91: 2414(A).

STERN, R. M., LIU, F.-H., OHSHIMA, Y., SULLIVAN, T. M., and ACERO, A. (1992c). "Alternative Approaches to Acoustical Pre-Processing for Robust Speech Recognition," presented at the Speech Research Symposium-XII, June, 1992.

ACERO, A., and STERN R. M. (1992). "Cepstral Normalization for Robust Speech Recognition," *Proc. of the ESCA Workshop on Speech Processing in Adverse Conditions*, Cannes-Mandelieu, France, pp 89-92.

STERN, R. M., and ACERO, A. (1990a). "Acoustic Pre-Processing for Robust Spoken Language Systems," invited presentation at the Speech Research Symposium- X, Baltimore, MD, October 16, 1990.

STERN, R. M., and ACERO, A. (1990b). "Signal Processing for Robust Spoken Language Systems," invited tutorial lecture at the Satellite Workshop on Statistical Approaches to Spoken Language Processing, Tokyo, Japan, November 17, 1990.

STERN, R. M., and RUDNICKY, A. I. (1990). “Spoken-Language Workstations in the Office Environment,” invited lecture at the SpeechTech90 Conference, New York City.

STERN, R. M., SHEAR, G. D., and ZEPPENFELD, T. (1988). “Lateralization Predictions for High-Frequency Binaural Stimuli,” *J. Acoust.Soc. Amer*. **84**: S80 (A).

SHEAR, G. D., and STERN, R. M. (1987). “Extending the Position-Variable Model: Dependence of Lateralization on Various Spectral Parameters,” *J. Acoust. Soc. Amer*. **81**: S27 (A).

STERN, R. M., and ZEIBERG, A. S. (1986). “A Weighted Image Model for Binaural Lateralization,” *J. Acoust. Soc. Amer.* **80**: S107 (A).

BEECHER, L., and STERN, R. M. (1986). “Perception of Modulations in Pitch and Lateralization,” *J. Acoust. Soc. Amer*. **79**: S22 (A).

STERN, R.M., ELSNER, A.E., and SCHIANO, J.L. (1984). “Interaural Time Discrimination in Tonal Maskers,” *J. Acoust. Soc. Amer*. **76**: S91(A).

LASRY, M. J., and STERN, R. M. (1984). “Unsupervised Speaker Adaptation in Feature-Based Isolated Letter Recognition,” *J. Acoust. Soc. Amer*. **74**: S16(A).

BACHORSKI, S. J., and STERN, R. M. (1983). “Dynamic Cues in Binaural Perception,” *J. Acoust. Soc. Amer*. **73**: S42(A).

STERN, R. M. and LASRY, M. J. (1982). “Tuning to the Speaker: Dynamic Adaptation of Statistical Parameters in Isolated Letter Recognition,” *J. Acoust. Soc. Amer*. 72 S31(A).

BRILL, S. M., PHILLIPS, M. S., LASRY, M. J., and STERN, R. M. (1982). “Decisions about Features,” *J. Acoust. Soc. Amer*. 72 S32(A).

KAISER, D. L., and STERN, R. M. (1981). “Interaural Time Discrimination in Tonal Maskers,” *J. Acoust. Soc. Amer*. 70, S88 (A).

SLOCUM, J. E., and STERN, R. M., Jr., (1980). “Interaural Time and Amplitude Discrimination in Noise,” *J. Acoust. Soc. Amer*. 68, S60(A).

STERN, R. M., Jr. and WAIBEL, A. H. (1980). “Audibility of Phase Changes in Vowel Sounds and Complex Tones,” *J. Acoust. Soc. Amer*. 68, S50(A).

DUMOND, G. J. and STERN, R. M., Jr. (1979). “A Forced-Choice Paradigm for PulsationThreshold Measurements,” *J. Acoust. Soc. Amer*. 65, S58(A).

RUBINOV, E. M. and STERN, R. M., Jr. (1979). “Effects of Binaural Maskers on the Subjective Laterality of Diotic Targets,” *J. Acoust. Soc. Amer*., 65, S121(A).

STERN, R. M., Jr. (1979). “On the Use of Multiple Perceptual Images in Binaural Discrimination Experiments,” *J. Acoust. Soc. Amer*. 65, S122(A).

RUOTOLO, B. R., STERN, R. M., Jr., and COLBURN, H. S. (1977). “Discrimination of Symmetric, Time-Intensity Traded Stimuli,” *J. Acoust. Soc. Amer*., 61, S60(A).

STERN, R. M., Jr. (1977). “Lateralization of the MLD: Detection-Threshold Performance of an Auditory-Nerve-Based Model for Lateral Position,” *J. Acoust. Soc. Amer*., 61, S60(A).

COLBURN, H. S., DOMNITZ, R. H., STERN, R. M., Jr., and DURLACH, N. I. (1976). "Current Problems in Binaural Hearing Research," *J. Acoust. Soc. Amer*., 59, S16(A).

STERN, R. M., Jr. (1976a). "Lateral Position, Interaural Discrimination, and Binaural Detection: Model Based on Auditory-Nerve Activity," *J. Acoust. Soc. Amer*., 59, S23(A).

STERN, R. M., Jr. (1972). "Perception of Simultaneously-Presented Musical Timbres," *Quart.*

Prog. Rep. No. 106, Res. Lab. of Electronics, MIT, Cambridge, MA.

Ph.D. THESES SUPERVISED

D. A. Coast, *Cardiac Arrhythmia Analysis Using Hidden Markov Models*, September, 1988.

A. Acero, *Acoustical and Environmental Robustness for Automatic Speech Recognition*, September, 1990.

W. A. Rozzi, *Speaker Adaptation in Automatic Speech Recognition via Estimation of Correlated Mean Vectors*, May, 1991.

Y. Ohshima, *Environmental Robustness in Speech Recognition using Physiologically-Motivated Signal Processing*, December, 1993.

F.-H. Liu, *Environmental Adaptation for Robust Speech Recognition*, June, 1994.

P. J. Moreno, *Speech Recognition in Noisy Environments*, May, 1996.

T. M. Sullivan, *Multi-Microphone Correlation-Based Processing for Robust Automatic Speech Recognition*, June, 1996.

E. Gouvea, *Acoustic-Feature-Based Frequency Warping for Speaker Normalization,* February, 1999.

M. Siegler, *Integration of Continuous Speech Recognition and Information Retrieval for Mutually Optimal Performance*, December, 1999.

B. Raj, *Reconstruction of Incomplete Spectrograms for Robust Speech Recognition,* April, 2000.

J. Huerta, *Robust Speech Recognition in GSM Codec Environments*, April, 2000.

S.-J. Doh, *Enhancements to Transformation-Based Speaker Adaptation: Principal Component and Inter-Class Maximum Likelihood Linear Regression*, July, 2000.

M. Seltzer, *Microphone Arrays for Robust Speech Recognition,* July, 2003.

J, Nedel, *Duration Normalization for Robust Recognition of Spontaneous Speech via Missing Feature Methods,* April, 2004.

X. Li, *Combination and Generation of Parallel Feature Streams for Improved Speech Recognition*, February, 2005.

Z. Al Bawab, *An Analysis-by-Synthesis Approach to Vocal Tract Modeling for Robust Speech Recognition,* September, 2009.

Y.-H. Chiu, *Learning-Based Auditory Encoding for Robust Speech Recognition,* April, 2010.

L. Gu, *Single-Channel Speech Separation Based on Instantaneous Frequency,* May, 2010.

C. Kim, *Signal Processing for Robust Speech Recognition Motivated by Auditory Processing*, September, 2010.

K. Kumar, *A Spectro-Temporal Framework for Compensation of Reverberation for Speech Recognition,* January, 2011.

G. Romigh, *Individualized Head-Related Transfer Functions: Efficient Modeling and Estimation from Small Sets of Spatial Samples*, December, 2012.

A. Moghimi, *Array-Based Spectro-Temporal Masking for Automatic Speech Recognition,* April, 2014.

M. J. Harvilla, *Compensentation for Nonlinear Distortion in Noise for Robust Speech Recognition,* October, 2014.

A. Menon, *Robust Recognition Of Binaural Speech Signals Using Techniques Based On Human Auditory Processing*, February, 2019.

T. Vuong, *Incorporating Modulation Information into Deep Neural Networks for Robust Speech Processing,* April, 2023.

Y. Xia, *Speech Enhancement Using Deep Neural Networks,* April, 2024.

M. Lindsey, *Online Active Learning with Corrective Feedback,* August 2024.

M.S. THESES and PROJECTS SUPERVISED

K. G. Murti, *Sonometric Evaluation of Eustacian-Tube Function*, May, 1979.

G. J. DuMond, *A Forced-Choice Paradigm for Pulsation-Threshold Measurement of Monaural and Binaural Phenomena*, August, 1979.

E. M. Rubinov, *Auditory Lateralization in Noise*, September, 1979.

J. E. Slocum, *Discrimination of Interaural Time and Intensity in Noise*, November, 1980.

D. L. Kaiser, *Interaural Time Discrimination in Tonal Maskers*, June, 1982.

A. H. Ferguson, *Effects of Frequency Modulation on Complex Pitch Perception*, December, 1982.

M. J. Lasry, *Dynamic Adaptation of Statistical Parameters in a Feature-Based Isolated Letter Recognition System*, December, 1982.

Y. L. Chu, *Zero Crossings and Ear Modelling in Speech*, February, 1983.

S. J. Bachorski, *Dynamic Cues in Binaural Perception*, December, 1983.

W. J. Lue, *Identification of Transient Features in Speech*, January, 1984.

L. E. Beecher, *Perception of Modulations in Pitch, Pitch Strength, and Subjective Lateral Position*, December, 1986.

S. J. Bartschat, *Bayesian Combination of Knowledge Sources in Automatic Speech Recognition*, August, 1987.

M. P. Vea, *Multisensor Signal Enhancement for Speech Recognition*, September, 1987.

G. D. Shear, *Modelling the Dependence of Auditory Lateralization on Frequency and Bandwidth*, September, 1987.

B. Chigier, *Classification of Stop Consonants in Natural Continuous Speech*, May, 1988.

P. H. Dietz, *Simulation of Trumpet Tones Via Physical Modeling*, December, 1988.

S. Palm, *Enhancement of Reverberated Speech using Models of the Human Binaural System*, April, 1989.

A. Liss, *Lateralization of Complex Binaural Sounds*, July, 1990.

S. H. Tao, *Additive versus Multiplicative Combination of Differences of Interaural Time and Intensity*, May, 1992.

P. J. Moreno, *Speech Recognition in Telephone Environments*, January, 1993.

J. Nehme, *Robust Recognition of Speech Input to Computer Workstations*, March, 1993.

N. Hanai, *Speech Recognition in the Automobile*, May, 1993.

W. Lee, *Lateralization of High-Frequency Bandpass Noise*, May, 1994.

M. Siegler, *Effects of Speech Rate on Speech Recognition Accuracy*, December, 1995.

U. Jain, *Speech Recognition over the Telephone*, May, 1995.

J. Nedel, *Integration of Speech and Video. Applications for Lip Synch: Lip Movement Synthesis and Time Warping,* May, 1999.

A. Ripepi, *Lateralization and Identification of Simultaneously-Presented Whispered Vowels and Speech Sounds,* May, 1999.

M. Seltzer, *Automatic Detection of Corrupted Speech Features for Robust Speech Recognition*, May, 2000.

M. Balakrishnan, *Extracting Additional Information from Gaussian Mixture Probabilities for Unsupervised Text-Independent Speaker recognition,* May 2005.

D. Zurale, *TaanSyn - A Synthesizer to add critical Indian Classical Singing elements to Western Vocals,* May 2017.

J. Michelson, *Unsupervised Guitar String Classification for Tablature Transcription*, August 2017.

C. Liang, *A Comparative Evaluation of Statistical Models for Score Following,* October 2017.

M. Yu. *How Important are the Headphone and Ear Canal Frequency Responses in Music Perception?* October 2018.

Y. Zhang. *Design of Matching Criteria for Audio-Based Polyphonic Score-following Systems Using Harmonic Product Spectra,* December 2018.

Y. Ma. *Learnable Front Ends Based on Temporal Modulation for Music Tagging,* August 2022.

R. Jin. *Controlling the Breathiness in Singing Voice Synthesis,* September 2022.

D.-H. Lee, *Polyphonic Piano Transcription Using a Simple RNN Encoder and Decoder Model*, October 2022.

S. Ganesh Thopaiah. *Estimation of Personalized HRTFs from 3D Point Clouds,* May 2023.

A. Pillay, *A Neural Score Follower for Computer Accompaniment of Polyphonic Musical Instruments,* May 2024.

COURSES TAUGHT AT CARNEGIE MELLON UNIVERSITY

18-101	Linear Circuit Analysis (Sophomore)
18-290	Signals and Systems (Sophomore and Junior)
18-396	Signals and Systems (Junior)

18-301	Fundamentals of Control (Senior)
18-308	Introduction to Digital Communication (Senior)
18-491	Introduction to Signal Processing (Junior/Senior)
18-493	Electroacoustics (Senior)
18-550	Fundamentals of Communication (Senior)
18-551	Communication and Signal Processing Project Laboratory (Senior)
18-340	Senior Seminar on Speech Processing
18-762	Signal Transmission and Detection (Graduate)
18-791	Digital Signal Processing (Graduate)
18-792	Advanced Digital Signal Processing (Graduate)
18-799	Digital Signal Processing for Robust Speech Recognition (Graduate)
18/42-795	Sensory Processes: Perception and Psychophysics (Graduate//Senior)
42-726	Hearing: Physiology and Perception (Graduate Seminar)

(18-308, 18-493, 18-791, 18-792, 18-795, and 42-726 were all new courses when first taught)

SERVICE ON DEPARTMENT AND UNIVERSITY COMMITTEES

Major contributions and initiatives:

Developed and obtained university approval for the CMU Programs in Music and Technology. Service on the CMU Music and Technology Executive Committee, 2009 - present.

As Associate Director of the CMU Information Networking Institute managed the MSIN master's program that spanned the engineering, computer science, management, and public policy colleges, and had primary responsibility for admissions and recruiting, 1994 - 2003.

Chair, Faculty Senate Intellectual Property Policy Committee, 1983 - 1986. Drafted and negortiated the approval of the intellectual property policy now in use by CMU. Chairing committee to review and recommend modifications to the Policy, 2017 - present.

Faculty Senate International Affairs Forum, 1989 - 1991. Initiated cooperative educational programs with Latin American universities.

Chair and Senate Delegate, Faculty Senate Student Affairs Council, 1981 - 1984. Drafted CMU alcohol policy and set faculty policy on a range of other student-related issues.

Other activities (partial listing):

Faculty Senate Executive Committee

University Trustees Student Affairs Council

University Education Council

University Research Council

University Review Committee for Non-Tenured Appointments Senate

Delegate, Faculty Senate Student Affairs Council

Senate Delegate, Faculty Senate Educational Affairs Council

International Committee of the Dean's Council

CMU Faculty Review Committee (Member and Co-Chair)

Carnegie Institute of Technology Faculty Chair

Music and Technology Program Executive Committee

MISCELLANEOUS SKILLS AND INTERESTS

Solo and chamber music performance on harpsichord with ensembles playing historical instruments, and with members of the Pittsburgh Symphony Orchestra. Participation in mountain biking, running, backpacking, racquet sports.

Extensive travel experience throughout Europe, Latin America, and Asia. s Fluent in Spanish; some French, Italian, and German.