

Carol Yvonne Espy-Wilson  
Department of Electrical and Computer Engineering  
Institute for Systems Research,  
University of Maryland, College Park  
[espy@umd.edu](mailto:espy@umd.edu)

**Education:**

MIT, EECS, MS (1981) Eng. (1984) PhD (1987)

Stanford University, BSEE (1979)

**Awards, Honors, and Special Recognition**

- Distinguished University Professor
- Keynote Speaker, International Speech Communication Association (ISCA) Conference Interspeech 2025 (August)
- Vice-President Elect/Vice President/Vice President Outgoing, Acoustical Society of America, 2025-2028
- Institute of Electrical and Electronics Engineers (IEEE) Fellow, 2021
- Keynote Speaker, 180<sup>th</sup> Meeting of the Acoustical Society of American, June 2021
- University of Maryland Campus Woman of Influence (2020)
- Keynote Speaker, 12<sup>th</sup> International Seminar on Speech Production, Dec. 2020
- Member, National Institute on Deafness and Communication Disorders Advisory Council, 2020-2023
- Member, National Institute on Biomed. Imaging and Bioengineering Advisory Council, 2015-2018
- Member, National Institute on Child Health and Human Development Advisory Board , 2010-2013
- First African American woman, and first African American, in ECE to achieve tenure and be promoted to the rank of full professor (University of Maryland First to ADVANCE Program, 2019)
- Jimmy Lin Award for Innovation in Electrical and Computer Engineering, University of Maryland, 2018

- Tutorial, Interspeech 2018, “Articulatory Representations: Measurement, Estimation and Application to State-of-the-Art Automatic Speech Recognition”
- Tutorial, 149<sup>th</sup> Meeting of the Acoustical Society of America, “Automatic Speech Recognition” May 2005.
- Fellow, International Speech Community Association (ISCA), 2018
- Fellow, Radcliffe Institute for Advanced Study, Harvard University, 2008-2009
- Fellow of the Acoustical Society of America, Nov. 2005
- IEEE Senior Member
- Elected to the Speech and Language Technical Committee of IEEE, 2010-2012
- Elected Chair, Speech Technical Committee of ASA, Spring 2007-Fall 2009
- Member of the Advisory Council for the National Institutes on Deafness and Other Communication Disorders 2020-2023
- Member of the Advisory Council for the National Institutes on Biomedical Imaging and Bioengineering 2015-2018
- 2012-2013 Distinguished Scholar-Teacher Award, University of Maryland
- 2015-2017 Institute for Systems Research Senior Faculty Fellow Award
- Advance Professor 2011-2012
- Clare Boothe Luce Professor, August 1990 to August 1995
- Invention of the Year Award for “Multi-pitch tracking in adverse environments,” University of Maryland, 2010.
- 2010 Maryland Innovator of the Year Award for “Multi-pitch tracking in adverse environments”, Baltimore Daily Record.
- Featured in “*Stem Gems: How 44 Women Shine in Science, Technology, Engineering and Mathematics, And How You Can Too!*” published by Stephanie Espy in 2016

- First Place, University of Maryland \$75K Business Plan Competition, High Technology Category (\$25,000), 2010
- First Place, University of Maryland \$75K Business Plan Competition, Social Impact Category (\$15,000), 2010
- StartRight! Women's Business Plan Competition Award
- SAIC Award to a Venture Accelerator Company at UMD (OmniSpeech LLC)
- Honda Initiation Award, 2003 and 2004
- NIH Career Development Award, 1998-2003
- NIH Language and Communication Study Section, 2001-2004.
- NSF Research Initiation Award, 1990-1992
- Xerox Doctoral Fellow, 1981-1987

**Editorial Service:**

- Editorial Board, Computer Speech Language, 2021- present
- Associate Editor, Journal of the Acoustical Society of America, 2010-2018
- Editorial Board, Acoustics Today, Acoustical Society of America, 2007-2009

**Academic Service:**

**University of Maryland College Park, Department of Electrical and Computer Engineering and Institute for Systems Research**

2001-2007: Associate Professor, 2007-present: Full Professor

**Boston University, Department of Electrical and Computer Engineering**

1990-1999: Assistant Professor, 1999-2001: Associate Professor

**Massachusetts Institute of Technology, Electrical Engineering and Computer Science**

1988-1990: Research Scientist and Lecturer for 2 undergraduate courses; 1990-2000: Visiting Scientist; 1987-1988: Postdoctoral Fellow, Speech Communication Group; 1988-2001: Associate Housemaster in MacGregor Dormitory; 1985-1988: Graduate Tutor in MacGregor Dormitory

## **Industrial Experience**

2011-present: Founder and CEO of OmniSpeech LLC

1987-1988: Part-time member of technical staff, Linguistics Research, Bell Telephone Laboratories, Murray Hill, New Jersey

## **Selected Invited Talks and Participations**

- Invited lectures in academic and industrial research labs including MIT, Ga. Tech, NYU, Harvard, Stanford, UC Berkeley, University of Maryland Medical Division of Otolaryngology, National Institute of Mental Health, Brown University, Cirrus Logic, Nuance, Apple, John Hopkins, Northwestern University, Ohio State, Rutgers University, NIDCD workshop, MIT Lincoln Laboratories, Honda, Dartmouth College, Boston University, UCLA, Motorola
- Invited Speaker, Workshop on Computational Linguistics and Clinical Psychology, CLPsych 2021, “Speech Acoustics for Mental Health Assessment”.
- Invited Speaker, From Sound to Sense Conference: 50+ Years of Discoveries in Speech Communication Conference, MIT, Cambridge, MA, June 11-13, 2004.
- Invited Speaker, National Institute of Mental Health’s Innovation Speaker Series, “Multimodal Signal Integration for Predictive Mental Health Analytics”, 2015.
- Invited Speaker, Symposia “Inspiring Women in Science”, Brown University, Celebration of Women’s History Month, a special one-day event with six women in science and engineering, 2017.
- Invited Speaker, NIH NIDCD workshop on Brain-Computer Interface Technologies, “Speech and Low Bandwidth Communication Technologies”, May 2006.
- Organized an NSF workshop on Speech for Robotics at University of Maryland, October 2019. 45 Participants from the US and abroad. Report written for program officers and is current in revision for Computer Speech and Language.
- Invited speaker on OmniSpeech LLC to the Committee on Economic Development and Technology Commercialization, Maryland board of Regents, Nov. 2013.
- Invited panelist, Science Coalition Panel for a Senate Staff Briefing, House Energy and Commerce Committee on Sparking Economic Growth, Capital Hill, Nov. 2013.
- External Evaluator, The Beckman Institute at the University of Illinois Urbana Champaign, Oct. 2015.
- Invited participant at several NSF panels and workshops (Directorates of Computer and Information Science and Engineering, and Education and Human Resources), and an NIH panel, 1995-present

- Keynote speaker for many events for students of color including some in the past 10 years: UMBC Meyerhoff Scholars Seminar (April 2019); Closing Reception for the Fall Regional NSBE Conference (Nov. 2018), Features Speaker in the STEM GEMs Virtual Summer Book Club (June, 2017), Keynote Speaker for Black Girls Who Code Event put on by Women in Engineering, Center for Minorities in Science and Engineering and DC Women of Color in STEM (Dec. 2016); College Board Conference (2015); MIT Summer Research Program (August 2012); Inaugural Willie Hobbs Moore Lecture Series, at U. Mich. Physics Department (March 2012); WEB DuBois Honor Society Induction Ceremony at UMD (2010); Center for Minorities in Science and Engineering/s Annual Student Recognition and Alumni Banquet (April 2008)

### Select Refereed Journal Publications

30. C. Espy-Wilson, K. Loganathan, N. Seneveratne and N. Benway (2026) “Depression, Speech Intelligibility and Articulatory Coordination”, *J. Acoustical Society of America*, *May 1;159(5):3988-3998*. *PubMed PMID: 42112822*.
29. N.R., Benway, S. Tabatabaee, Wang, D., B. Munson, J.L. Preston, C. Espy-Wilson (2025) “Perceptual Ratings Predict Speech Inversion Articulatory Kinematics in Childhood Speech Sound Disorders” *Journal of Speech, Language and Hearing Research*, 69 (2): 541:561. DOI: 10.1044/2025\_JSLHR-25-00515.
28. Y. Siriwardena, S. Boyce, M. Tiede, L. Oren, B. Fletcher, M. Stern and C. Espy-Wilson, (2024) “Speaker-Independent Speech Inversion for Recovery of Velopharyngeal Port Constriction Degree,” *Journal of the Acoustical Society of America*, vol.156, no. 2, August 2024, doi: 10.1121/10.0028124
27. M. Marge, C. Espy-Wilson, N. Ward, A. Alwan, Y. Artzi, M. Bansal, G. Blankenship, J. Chai, H. Daumé III, D. Dey , M. Harper, T.Howard, C. Kennington I. Kruijff-Korbayová, D. Manocha, C. Matuszek, R. Mead, R.Mooney, R. K. Moore, M. Ostendorf, H. Pon-Barry, A. Rudnicky, M. Scheutz, R. St. Amant, T. Sun, S. Tellex, D. Traum, Z. Yu , “Spoken Language Interaction for Robotics: Research Issues and Recommendations”, *Computer Speech Language*, Volume 71, January 2022.
26. **S. Sahu**, R. Gupta and C. Espy-Wilson, “Modeling Feature Representations for Affective Speech using Generative Adversarial Networks”, *IEEE Transactions on Affective Computing*, May 2020.
25. **G. Sivaraman**, V. Mitra, H. Nam, M. Tiede, and C. Espy-Wilson, “Unsupervised speaker adaptation for speaker independent acoustic to articulatory speech inversion”, *The Journal of the Acoustical Society of America*, vol. 146, no. 1, 2019, pp. 316-329, doi: 10.1121/1.5116130.
24. V. Mitra, **G. Sivaraman**, H. Nam, C. Espy-Wilson, E. Saltzman, M. Tiede, (2017) “Hybrid Convolutional Neural Networks For Articulatory And Acoustic Information Based Speech Recognition”, *Speech Communication*, Volume 89, Issue C, Pages 103-112, May 2017.
23. S. Gordon-Salant, D. Zion and C. Espy-Wilson (2014) “Recognition of time-compressed speech does not predict recognition of natural fast-rate speech by older listeners” *Journal of the Acoustical Society*

of America, Express Letters, vol. 136, pp. 268-274.

22. X. Zhou, J. Zhou, M. Stone, J. Prince and C. Espy-Wilson (2013), “Improve vocal tract reconstruction and modeling using an image super-resolution technique”, Journal of the Acoustical Society of America Express Letters, vol. 133, no. 6, pp. 439-445.
21. H. Nam, **V. Mitra**, M. Tiede, M. Hasegawa-Johnson, C. Espy-Wilson, E. Saltzman, L. Goldstein, (2012), “A procedure for estimating gestural scores from speech acoustics”, Journal of the Acoustical Society of America”, vol. 132, no. 6, 3980-3989.
20. **V. Mitra**, H. Nam, C. Espy-Wilson, E. Saltzman, and L. Goldstein (2012) “Recognizing articulatory gestures from speech for robust speech recognition”, Journal of the Acoustical Society of America, vol. 131, no. 3, pp. 2270-2287.
19. **V. Mitra**, H. Nam, C. Espy-Wilson, E. Saltzman, L. Goldstein (2011), “Articulatory Information for Noise Robust Speech Recognition“, IEEE Transactions on Audio, Speech and Language Processing, vol. 19, no. 7, pp. 1913-1924.
18. **V. Mitra**, H. Nam, C. Espy-Wilson, E. Saltzman, L. Goldstein (2010) “Retrieving Tract Variables from Acoustics: a comparison of different Machine Learning strategies” IEEE Journal of Selected Topics, vol. 4, issue 6, pp. 1027-1045.
17. **X. Zhou**, C. Espy-Wilson, S. Boyce, M. Tiede, Christy Holland and Ann Choe (2008) “A magnetic resonance imaging-based articulatory and acoustic study of “retroflex” and “bunched” American English /r/ sounds”, Journal of the Acoustical Society of America, Vol. 123, no. 6, pp. 4466-4481.
16. **A. Juneja** and Carol Espy-Wilson (2008) “Probabilistic landmark detection for automatic speech recognition using acoustic-phonetic information”, Journal of the Acoustical Society of America, vol. 123, no. 2, pp. 1154-1168.
15. **T. Pruthi**, C. Espy-Wilson and Brad Story (2007) “Simulation and analysis of nasalized vowels based on MRI data”, Journal of the Acoustical Society of America, vol. 121, no. 6, pp. 3858-3873.
14. **O. Deshmukh**, C. Espy-Wilson, L. Carney, (2007) “Speech Enhancement using the Modified Phase Opponency Model”, Journal of the Acoustical Society of America, vol. 121, no. 6, pp. 3886-3898.
13. **O. Deshmukh**, C. Espy-Wilson, **A. Salomon**, and **Jawahar Singh** (2005) “Use of Temporal Information: Detection of the Periodicity, Aperiodicity and Pitch in Speech”, IEEE Transactions on Speech and Audio Processing, vol. 13, pp. 776-786.
12. **T. Pruthi** and C. Espy-Wilson (2004) “Acoustic Parameters for Automatic Detection of Nasal Manner”, Speech Communication, vol. 43, no. 3, pp.225-239.
11. Z. Zhang and Carol Espy-Wilson (2004) “A vocal tract model for American English /l/”, Journal of the Acoustical Society of America, vol. 115, no. 3, pp. 1274-1280.

10. A. Salomon, C. Espy-Wilson and Om Deshmukh (2004) “Detection of Speech Landmarks: Use of Temporal Information”, *Journal of the Acoustical Society of America*, vol. 115, no. 3, pp. 1296-1305.
9. M. Jackson, C. Espy-Wilson and S. Boyce (2001) “Verifying a vocal tract model with a closed side branch”, *Journal of the Acoustical Society of America*, vol. 109, no. 6, pp. 2983-2987.
8. C. Y. Espy-Wilson, S. Boyce, M. Jackson, S. Narayanan and A. Alwan (2000) “Acoustic Modeling of American English /r/”, *Journal of the Acoustical Society of America*, pp. 343-356.
7. F. Guenther, C. Espy-Wilson, S. Boyce, M. Matthies, M. Zandipour and J. Perkell (1999) “Articulatory tradeoffs reduce acoustic variability during American English /r/ production”, *Journal of the Acoustical Society of America*, vol. 105, no. 5, pp. 2854-2865.
6. C. Y. Espy-Wilson, V. R. Chari, J. M. MacAuslan, C. B. Huang and M. J. Walsh (1998) “Enhancement of Electrolaryngeal Speech by Adaptive Filtering”, *Journal of Speech, Language and Hearing Research*, vol. 41, no. 6, December, pp. 1253-1264.
5. S. Boyce and C. Y. Espy-Wilson (1997) “Coarticulatory Stability in American English /r/s”, *Journal of the Acoustical Society of America*, vol. 101, no. 6, pp. 3741-3753.
4. V. Chari and C. Espy-Wilson (1995) “Extraction of Formant Frequencies by Adaptive Enhancement of Fourier Spectra,” *IEEE Transactions on Speech and Audio Processing*, vol. 3, no. 1, pp. 35-39.
3. C. Y. Espy-Wilson (1994) “A Feature-Based Semivowel Recognition System,” *Journal of the Acoustical Society of America*, vol. 96, no. 1, pp. 65-72.
2. C. Y. Espy-Wilson (1992) “Acoustic Measures for Linguistic Features Distinguishing the Semivowels /wjr/ in American English,” *Journal of the Acoustical Society of America*, vol. 92, no. 2, pp. 736-751.
1. C. Y. Espy and J. Lim (1983) “Effects of Additive Noise on Signal Reconstruction from Fourier Transform Phase,” *IEEE Transactions on Acoustics, Speech and Signal Processing*, pp. 894-898.

### **Book Chapters**

1. C. Espy-Wilson, G. Sivaraman, M. Tiede, V. Mitra, E. Saltzmann, L. Goldstein, H. Nam, “Modeling of Articulatory Gestures to Control Effects of Production Variability on Speech Technologies”. In Cangemi, Clayards, Niebuhr, Schupler & Zellers (eds). *Rethinking Reduction*, Berlin: Mouton de Gruyter, 2018.
2. M. Tiede, S. Boyce and C. Espy-Wilson, “Variability of North-American /r/ production in response to palatal perturbation”. In Maassen, Ben and van Lieshout, Pascal (eds.), *Speech Motor Control: New developments in basic and applied research*, Oxford University Press, 2010.
3. Carol Espy-Wilson, “Phonological Models of Variation in Computer Speech Processing: Commentary on the papers by Nam, Son et al., and Hirschberg”. In Cole, Jennifer and Hualde, José I. (eds.), *Laboratory Phonology 9*. Berlin: Mouton de Gruyter. pp. 535-546, 2007.

4. S. Nawab, C. Espy-Wilson, R. Mani, and N. Bitar, "Knowledge-Based Analysis of Speech Mixed with Sporadic Environmental Sounds," Computational Auditory Scene Analysis edited by Rosenthal and Okuno Lawrence Erlbaum Associates Inc. Publishers, 1998.

### **Selected Peer-Reviewed Conference Proceedings (last 10 years)**

1. Tabatabaee, S, Tiede, M, Boyce SE, Oren L, and Espy-Wilson, C, (2026) Cross-language stop and nasal timings quantified using speech inversion, Proc. LabPhon20, 2026.
2. Tabatabaee, S, Espy-Wilson, C, (2026) Towards Noise-Robust Speech Inversion Through Multi-Task Learning with Speech Enhancement, ICASSP 2026, 1520-6149
3. Withanage, T, Espy-Wilson, C, (2026) Global Articulatory Entrainment and Conversational Success in Spontaneous Dyadic Speech, ASA Spring 2026 Annual Meeting, May 2026.
4. Withanage, T, Espy-Wilson, C, Redcay, E. (2026) Acoustic and facial markers of perceived conversational success in spontaneous speech, ICASSP 2026, 1520-6149
5. Withanage, T, Jones, D, Espy-Wilson, C, Redcay, E, Sasson, N. (2026) Differences in Facial Synchrony Across Autistic and Non-Autistic Dyads During Naturalistic Conversations, INSAR 2026 Annual Meeting, April 2026.
6. Tabatabaee, S, Boyce SE, Oren L, Tiede, M and Espy-Wilson, CY. (2025) Acoustic to Articulatory Speech Inversion for Children with Velopharyngeal Insufficiency, Proceedings of IEEE Automatic Speech Recognition and Understanding (ASRU) Workshop, Dec. 2025
7. Attia, A., Demszky, D., Liu, J., Espy-Wilson, C. (2025) From Weak Labels to Strong Results: Utilizing 5,000 Hours of Noisy Classroom Transcripts with Minimal Accurate Data. *Proc. Interspeech 2025*, 3678-3682
8. Benway, N.R., Tabatabaee, S., Munson, B., Preston, J., Espy-Wilson, C. (2025) Subtyping Speech Errors in Childhood Speech Sound Disorders with Acoustic-to-Articulatory Speech Inversion. *Proc. Interspeech 2025*, 2800-2804
9. Premananth, G., Kugathasan, V., Espy-Wilson, C. (2025) Analyzing the Impact of Accent on English Speech: Acoustic and Articulatory Perspectives. *Proc. Interspeech 2025*, 1493-1497
10. Premananth, G., Resnik, P., Bansal, S., Kelly, D.L., Espy-Wilson, C. (2025) Multimodal Biomarkers for Schizophrenia: Towards Individual Symptom Severity Estimation. *Proc. Interspeech 2025*, 3065-3069
11. Tabatabaee, S., Liu, J., Espy-Wilson, C. (2025) FT-Boosted SV: Towards Noise Robust Speaker Verification for English Speaking Classroom Environments . *Proc. Interspeech 2025*, 2815-2819
12. Tabatabaee, S., Boyce, S., Oren, L., Tiede, M., Espy-Wilson, C. (2025) Enhancing Acoustic-to-Articulatory Speech Inversion by Incorporating Nasality. *Proc. Interspeech 2025*, 325-329

13. Attia, A. A., Demszky, D., Ógúnremí, T., Liu, J., & Espy-Wilson, C. (2025). CPT-Boosted Wav2vec2.0: Towards Noise Robust Speech Recognition for Classroom Environments. *ICASSP 2025 – 2025 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*.
14. Premananth, G., & Espy-Wilson, C. (2025). Self-supervised Multimodal Speech Representations for the Assessment of Schizophrenia Symptoms. *ICASSP 2025 – 2025 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*.
15. Premananth, G., & Espy-Wilson, C. (2025). Speech-Based Estimation of Schizophrenia Severity Using Feature Fusion. *2025 IEEE International Conference on Acoustics, Speech, and Signal Processing Workshops (ICASSPW)*.
16. Ojha, S., Gervits, F., & Espy-Wilson, C. (2025). Speaking with Robots in Noisy Environments. *Proceedings of the 2025 ACM/IEEE International Conference on Human-Robot Interaction*, 1057–1061. Melbourne, Australia: IEEE Press.
17. Attia, A. A., Liu, J., Ai, W., Demszky, D., & Espy-Wilson, C. (2024). Kid-whisper: Towards bridging the performance gap in automatic speech recognition for children vs. adults. *Proceedings of the AAAI/ACM Conference on AI, Ethics, and Society*, 7, 74–80.
18. Benway, N.R., Preston, J.L., Espy-Wilson, C. (2024) Examining Vocal Tract Coordination in Childhood Apraxia of Speech with Acoustic-to-Articulatory Speech Inversion Feature Sets. *Proc. Interspeech 2024*, 5138-5142.
19. Premananth, G., Siriwardena, Y.M., Resnik, P., Bansal, S., L.Kelly, D., Espy-Wilson, C. (2024) A Multimodal Framework for the Assessment of the Schizophrenia Spectrum. *Proc. Interspeech 2024*, 1470-1474.
20. Siriwardena, Y. M., Swedlow, N., Howard, A., Gitterman, E., Darcy, D., Espy-Wilson, C., & Fanelli, A. (2024). Accent Conversion with Articulatory Representations. *Proc. Interspeech 2024*, 4383-4387
21. Attia, A. A., Siriwardena, Y. M., & Espy-Wilson, C. (2024). Improving speech inversion through self-supervised embeddings and enhanced tract variables. *2024 32nd European Signal Processing Conference (EUSIPCO)*, 306–310.
22. Premananth, G., Siriwardena, Y. M., Resnik, P., & Espy-Wilson, C. (2024). A multi-modal approach for identifying schizophrenia using cross-modal attention. *2024 46th Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC)*.
23. Siriwardena, Y.M., Espy-Wilson, C., Shamma, S. (2023) Learning to Compute the Articulatory Representations of Speech with the MIRRORNET. *Proc. INTERSPEECH 2023*
24. Benway, N.R., Siriwardena, Y.M., Preston, J.L., Hitchcock, E., McAllister, T., Espy-Wilson, C. (2023) Acoustic-to-Articulatory Speech Inversion Features for Mispronunciation Detection of /ɪ/ in Child Speech Sound Disorders. *Proc. INTERSPEECH 2023*
25. Siriwardena, Y.M., Espy-Wilson, C., Boyce, S., Tiede, M., Oren, L. (2023) Speaker-independent Speech Inversion for Estimation of Nasalance. *Proc. INTERSPEECH 2023*

26. Attia, A.A., Tiede, M., Espy-Wilson, C. (2023) Enhancing Speech Articulation Analysis Using A Geometric Transformation of the X-ray Microbeam Dataset. Proc. INTERSPEECH 2023
27. Siriwardena, Y. M., Adel Attia, A., Sivaraman, G., and Espy-Wilson, C. “Audio Data Augmentation for Acoustic-to-Articulatory Speech Inversion,” *2023 31st European Signal Processing Conference (EUSIPCO)*
28. Adel Attia, A., and Espy-Wilson, C., “Masked Autoencoders are Articulatory Learners,” ICASSP 2023
29. Seneviratne, N., Espy-Wilson, C., “Multimodal Depression Severity Score Prediction Using Articulatory Coordination Features and Hierarchical Attention Based Text Embeddings.” Proc. Interspeech 2022, 3353-3357
30. Siriwardena, Y. M., and Espy-Wilson, C., “The Secret Source: Incorporating Source Features to Improve Acoustic-To-Articulatory Speech Inversion,” *ICASSP 2023*
31. Siriwardena, Y.M., Sivaraman, G., Espy-Wilson, C. “Acoustic-to-articulatory Speech Inversion with Multi-task Learning”. Proc. Interspeech 2022
32. Parikh, R., Seneviratne, N., Sivaraman, G., Shamma, S., Espy-Wilson, C. “Acoustic To Articulatory Speech Inversion Using Multi-Resolution Spectro-Temporal Representations Of Speech Signals”. Proc. Interspeech 2022
33. Parikh, Rahil, Gaspar Rochette, Carol Espy-Wilson, and Shihab Shamma. “An Empirical Analysis on the Vulnerabilities of End-to-End Speech Segregation Models.” Proc. Interspeech 2022
34. Parikh, Rahil, Ilya Kavalerov, Carol Espy-Wilson, and Shihab Shamma. “Harmonicity Plays a Critical Role in DNN Based Versus in Biologically-Inspired Monaural Speech Segregation Systems.” In *ICASSP 2022-2022 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pp. 536-540. IEEE, 2022
35. N. Seneviratne and C. Espy-Wilson, “Multimodal Depression Classification using Articulatory Coordination Features and Hierarchical Attention Based text Embeddings,” ICASSP 2022 – 2022 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), 2022, pp. 6252-6256.
36. Siriwardena, Y.M., Kitchen, C., Kelly, D.L. & Espy-Wilson, C. “Multimodal Approach for Assessing Neuromotor Coordination in Schizophrenia using Convolutional Neural Networks”, In Proceedings of the 2021 International Conference on Multimodal Interaction (ICMI '21)
37. N. Seneviratne, C. Espy-Wilson, "Generalized Dilated CNN Models for Depression Detection Using Inverted Vocal Tract Variables", accepted for Interspeech 2021, Aug. 30-Sept. 3, 2021 (arxiv: 2011.06739)
38. N. Seneviratne, C. Espy-Wilson, "Speech based Depression Severity Level Classification Using a Multi-Stage Dilated CNN-LSTM Model", accepted for Interspeech 2021, Aug. 30-Sept. 3, 2021 (arxiv: 2104.04195)

39. Sirwadana, Y., Kitchen, C., Kelly, D.L. & Espy-Wilson, C. (2021). Inverted vocal tract variables and facial action units to quantify neuromotor coordination in schizophrenia. In Proceedings of the 12th International Seminar on Speech Production (ISSP 2020), 174-177.
40. N. Seneviratne, A. Lammert, Jim Williamson, T. Quatieri and C. Espy-Wilson “Extended Study on the Use of Vocal Tract Variables to Quantify Neuromotor Coordination in Depression”, Proc. Interspeech 2020.
41. S. Sahu, V. Mitra, N. Seneviratne and C. Espy-Wilson. “Multi-Modal Learning for Speech Emotion Recognition: An Analysis and Comparison of ASR Outputs with Ground Truth Transcription”, Proc. of Interspeech, Sept. 2019.
42. N. Seveviratne, G. Sivaraman, V. Mitra and C. Espy-Wilson, “Noise Robust Articulatory to Acoustic Speech Inversion, Proc. of Interspeech, Sept. 2019.
43. C. Espy-Wilson, A. Lammert, N. Seneviratne, T. Quatieri, “Assessing Neuromotor Coordination in Depression Using Inverted Vocal Tract Variables”, Proc. Interspeech 2019, 1448-1452, DOI: 10.21437/Interspeech.2019-1815.
44. S. Sahu and C. Espy-Wilson, “On Enhancing Speech Emotion Recognition Using Generative Adversarial Networks”, Proc. of Interspeech 2018.
45. N. Seneviratne, G. Sivaraman, V. Mitra and C. Espy-Wilson, “Noise Robust Articulatory to Acoustic Speech Inversion, Proc. of Interspeech 2018.
46. S. Sahu, R. Gupta, G. Sivaraman, and C. Espy-Wilson, “Smoothing Model predictions Using Adversarial Training Procedures for Speech based Emotion Recognition,” Proc. of ICASSP 2018.
47. R. Gupta, S. Sahu, C. Espy-Wilson and S. Naryanan, “Semi-supervised and Transfer learning approaches for low resource sentiment classification,” Proc. of ICASSP 2018.
48. Mitra, V., Sivaraman, G., Bartels, C., Nam, H., Wang, W., Espy-Wilson, C., Vergyri, D. and Franco, H., 2017, March. Joint modeling of articulatory and acoustic spaces for continuous speech recognition tasks. In 2017 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP) (pp. 5205-5209). IEEE.
49. R. Gupta, S. Sahu, C. Espy-Wilson, and S. Narayanan. An Affect Prediction Approach through Depression Severity Parameter Incorporation in Neural Networks. In Proceedings of Interspeech, August 2017.
50. S. Sahu, R. Gupta, G. Sivaraman, W. AbdAlmageed, and C. Espy-Wilson, “Adversarial auto-encoders for speech based emotion recognition,” in Interspeech, 2017, pp. 1243–1247.
51. G. Sivaraman, V.Mitra, H. Nam, M.K. Tiede, C. Espy-Wilson (2016) Vocal tract length normalization for speaker independent acoustic-to-articulatory speech inversion, Proc. of INTERSPEECH 2016.
52. S.Sahu, C.Espy-Wilson, “Speech features for depression detection”, In Proceedings of Interspeech Sept. 2016, pp. 1928-1932.

**Supervised: 16 PhD, 31 Master’s, and 40 undergraduate students**

## **Academic Research Program**

- **Grants and Contracts:** More than 16 million dollars from NSF, NIH, DoD, industry.
- **Patents:** 7

## **OmniSpeech LLC**

- **Grants:** More than 1 million dollars from NSF (SBIR Phase, 1, Phase II and Phase IIB) and NIH (STTR).
- **Angel Investments:** More than 5 million dollars.
- **Patents:** 6