

# William Hartmann

Raytheon BBN Technologies  
10 Moulton Street Cambridge, MA 02138  
✉ hartmannw@gmail.com  
🌐 www.betweenzeroandone.com

## Research Interests

deep learning, automatic speech recognition, machine translation, semi-supervised and unsupervised learning, reinforcement learning

## Education

- June 2012 **Ph.D. Computer Science (Artificial Intelligence)**, *The Ohio State University*, Columbus, OH  
Linguistics, Algorithms/Theory of Computation minors  
Advisor: Eric Fosler-Lussier
- June, 2010 **M.S. Computer Science (Artificial Intelligence)**, *The Ohio State University*, Columbus, OH  
Advisor: Eric Fosler-Lussier
- May, 2006 **B.S. Computer Science**, *Northern Kentucky University*, Highland Heights, KY  
Mathematics minor

## Professional and Academic Experience

- 2019–present **Senior Scientist**, *Owen Kimball (Manager)*, Raytheon BBN Technologies  
Led the speech recognition effort for the IARPA MATERIAL program and multiple government and commercial customers. Trained and developed a team of junior researchers. Led the transition from hybrid ASR systems to sequence-to-sequence models.
- 2014–2019 **Scientist**, *Stavros Tsakalidis (Manager)*, Raytheon BBN Technologies  
Oversaw research to improve ASR across multiple programs. Improved feature extraction and acoustic models for the Sage speech processing platform. Adapted speech processing methods for sonar-based ship classification.
- 2012–2014 **Post-Doctoral Researcher**, *Lori Lamel (Supervisor)*, LIMSI-CNRS  
Researched methods for improving pronunciation lexicons and automatically discovering acoustic units for under resourced languages. Explored sub-word lexical units for OOV term detection. Investigated text-based speaker recognition without acoustic information.
- 2007–2012 **Graduate Research Assistant**, *Eric Fosler-Lussier (Advisor)*, The Ohio State University  
Investigated lexical access and the effects of underspecified dictionaries on ASR systems. Researched methods for utilizing the Ideal Binary Mask in ASR. Developed a speech separation system that incorporated high-level linguistic information.
- 2011 **Graduate Teaching Assistant**, *Kathryn Reeves (Supervisor)*, The Ohio State University  
Taught a senior level course on Artificial Intelligence. Course topics included search, planning, logic, and uncertainty.

- 2008–2010 **Graduate Research Assistant**, *Donna Byron (Supervisor)*, The Ohio State University  
 Developed a system using Experiment Builder and Eyelink to perform eye-tracking experiments on human subjects. Created necessary audio and video materials for the experiments. Performed data collection with human subjects.
- 2007–2008 **NSF GK-12 Fellow**, *Susan Olesik (Supervisor)*, The Ohio State University  
 Worked with two 5th grade teachers in the Columbus Public school district to improve science education. Responsibilities included developing lesson plans, teaching classes, and serving as scientific advisor for elementary school teachers. Example lessons include water-powered rockets, cryptography, and robotics.
- 2006–2007 **Graduate Research Assistant**, *James W. Davis (Advisor)*, The Ohio State University  
 Installed a motion sensor network on seven floors of Dreese Laboratories. Implemented a real-time system in C++ for creating activity maps for camera with pan and tilt control. Researched methods for activity pattern recognition.
- 2006 **Computer Programmer**, *Matthew Zacate (Supervisor)*, Northern Kentucky University  
 Developed a physics application in C for simulating compounds with large concentrations of defects using a combined energy minimization-Monte Carlo technique.
- 2005–2006 **Undergraduate Research Assistant**, *Richard Fox (Advisor)*, Northern Kentucky University  
 Worked on a hand-written character recognition system that incorporated abduction.
- 2005 **Running Start Leader**, *Barbara Hamilton (Supervisor)*, Northern Kentucky University  
 Taught a supplemental course in developmental mathematics to assist students in meeting university requirements.

---

## Refereed Publications

- [1] S. Wotherspoon, W. Hartmann, M. Snover, and O. Kimball, “Improved data selection for domain adaptation in asr,” in *Proceedings of IEEE ICASSP*, 2021.
- [2] R. Meermeier, L. Zhang, F. Keith, W. Hartmann, S. Tsakalidis, and A. Tabarez, “Godec: An open-source data processing framework for deploying ml data flows in edge-computing environments,” in *International Conference on Edge Computing*, pp. 84–93, Springer, 2020.
- [3] H. Gish, J. Silovsky, M.-L. Sung, M.-H. Siu, W. Hartmann, and Z. Jiang, “Towards a new understanding of the training of neural networks with mislabeled training data,” in *Proceedings of IEEE ICASSP*, pp. 8394–8398, IEEE, 2020.
- [4] D. Karakos, R. Zbib, W. Hartmann, R. Schwartz, and J. Makhoul, “Reformulating information retrieval from speech and text as a detection problem,” in *Proceedings of the workshop on Cross-Language Search and Summarization of Text and Speech (CLSSTS)*, pp. 38–43, 2020.
- [5] L. Zhang, D. Karakos, W. Hartmann, M. Srivastava, L. Tarlin, D. Akodes, S. K. Gouda, N. Bathool, L. Zhao, Z. Jiang, R. Schwartz, and J. Makhoul, “The 2019 bbn cross-lingual information retrieval system,” in *Proceedings of the workshop on Cross-Language Search and Summarization of Text and Speech (CLSSTS)*, pp. 44–51, 2020.
- [6] D. Bagchi and W. Hartmann, “Learning from the best: A teacher-student multilingual

- framework for low-resource languages,” in *Proceedings of IEEE ICASSP*, pp. 6051–6055, IEEE, 2019.
- [7] R. Zbib, L. Zhao, D. Karakos, W. Hartmann, J. DeYoung, Z. Huang, Z. Jiang, N. Rivkin, L. Zhang, R. Schwartz, and J. Makhoul, “Neural-network lexical translation for cross-lingual ir from text and speech,” in *Proceedings of ACM SIGIR*, pp. 645–654, 2019.
- [8] F. Keith, W. Hartmann, M.-H. Siu, J. Ma, and O. Kimball, “Optimizing multilingual knowledge transfer for time-delay neural networks with low-rank factorization,” in *Proceedings of IEEE ICASSP*, pp. 4924–4928, IEEE, 2018.
- [9] D. Karakos, J. Silovsky, R. Schwartz, W. Hartmann, and J. Makhoul, “Individual ship detection using underwater acoustics,” in *Proceedings of IEEE ICASSP*, pp. 2121–2125, IEEE, 2018.
- [10] W. Hartmann, D. Karakos, R. Hsiao, L. Zhang, T. Alumäe, S. Tsakalidis, and R. Schwartz, “Analysis of keyword spotting performance across iarpa babel languages,” in *Proceedings of IEEE ICASSP*, pp. 5765–5769, IEEE, 2017.
- [11] T. Alumäe, D. Karakos, W. Hartmann, R. Hsiao, L. Zhang, L. Nguyen, S. Tsakalidis, and R. Schwartz, “The 2016 bbn georgian telephone speech keyword spotting system,” in *Proceedings of IEEE ICASSP*, pp. 5755–5759, IEEE, 2017.
- [12] W. Hartmann, R. Hsiao, T. Ng, J. Z. Ma, F. Keith, and M.-H. Siu, “Improved single system conversational telephone speech recognition with vgg bottleneck features.,” in *Proceedings of Interspeech*, pp. 112–116, 2017.
- [13] D. Karakos, W. Hartmann, R. Schwartz, J. Makhoul, S. Tsakalidis, E. Insanic, and G. Shepard, “Applying speech technology to the ship-type classification problem,” in *Proceedings of IEEE OCEANS*, pp. 1–8, IEEE, 2017.
- [14] W. Hartmann, R. Hsiao, and S. Tsakalidis, “Alternative networks for monolingual bottleneck features,” in *Proceedings of IEEE ICASSP*, pp. 5290–5294, IEEE, 2017.
- [15] R. Hsiao, R. Meermeier, T. Ng, Z. Huang, M. Jordan, E. Kan, T. Alumäe, J. Silovsky, W. Hartmann, F. Keith, *et al.*, “Sage: The new bbn speech processing platform.,” in *Proceedings of Interspeech*, pp. 3022–3026, 2016.
- [16] W. Hartmann, T. Ng, R. Hsiao, S. Tsakalidis, and R. M. Schwartz, “Two-stage data augmentation for low-resourced speech recognition.,” in *Proceedings of Interspeech*, pp. 2378–2382, 2016.
- [17] W. Hartmann, L. Zhang, K. Barnes, R. Hsiao, S. Tsakalidis, and R. M. Schwartz, “Comparison of multiple system combination techniques for keyword spotting.,” in *Proceedings of Interspeech*, pp. 1913–1917, 2016.
- [18] N. Kleynhans, W. Hartmann, D. Van Niekerk, C. Van Heerden, R. Schwartz, S. Tsakalidis, and M. Davel, “Code-switched english pronunciation modeling for swahili spoken term detection,” *Procedia Computer Science*, vol. 81, pp. 128–135, 2016.
- [19] L. Zhang, D. Karakos, W. Hartmann, R. Hsiao, R. Schwartz, and S. Tsakalidis, “Enhancing low resource keyword spotting with automatically retrieved web documents,” in *Proceedings of Interspeech*, 2015.

- [20] R. Hsiao, J. Ma, W. Hartmann, M. Karafiát, F. Grézl, L. Burget, I. Szöke, J. H. Černocký, S. Watanabe, Z. Chen, *et al.*, “Robust speech recognition in unknown reverberant and noisy conditions,” in *Proceedings of IEEE ASRU*, pp. 533–538, IEEE, 2015.
- [21] A. Roy, H. Bredin, W. Hartmann, V. B. Le, C. Barras, and J.-L. Gauvain, “Lexical speaker identification in tv shows,” *Multimedia Tools and Applications*, vol. 74, no. 4, pp. 1377–1396, 2015.
- [22] M. Davel, E. Barnard, C. v. Heerden, W. Hartmann, D. Karakos, R. Schwartz, and S. Tsakalidis, “Exploring minimal pronunciation modeling for low resource languages,” in *Proceedings of Interspeech*, 2015.
- [23] A. Laurent, W. Hartmann, and L. Lamel, “Unsupervised acoustic model training for the korean language,” in *The 9th International Symposium on Chinese Spoken Language Processing*, pp. 469–473, IEEE, 2014.
- [24] W. Hartmann, L. Lamel, and J.-L. Gauvain, “Efficient rule scoring for improved grapheme-based lexicons,” in *Proceedings of IEEE EUSIPCO*, pp. 1477–1481, IEEE, 2014.
- [25] W. Hartmann, L. Lamel, and J.-L. Gauvain, “Cross-word sub-word units for low-resource keyword spotting.,” in *Proceedings of SLTU*, pp. 112–117, Citeseer, 2014.
- [26] W. Hartmann, V.-B. Le, A. Messaoudi, L. Lamel, and J.-L. Gauvain, “Comparing decoding strategies for subword-based keyword spotting in low-resourced languages,” in *Proceedings of Interspeech*, 2014.
- [27] V.-B. Le, L. Lamel, A. Messaoudi, W. Hartmann, J.-L. Gauvain, C. Woehrling, J. Despres, and A. Roy, “Developing stt and kws systems using limited language resources,” in *Proceedings of Interspeech*, 2014.
- [28] W. Hartmann, A. Roy, L. Lamel, and J.-L. Gauvain, “Acoustic unit discovery and pronunciation generation from a grapheme-based lexicon,” in *Proceedings of IEEE ASRU*, pp. 380–385, IEEE, 2013.
- [29] W. Hartmann, A. Narayanan, E. Fosler-Lussier, and D. Wang, “A direct masking approach to robust asr,” *IEEE Transactions on Audio, Speech, and Language Processing*, vol. 21, no. 10, pp. 1993–2005, 2013.
- [30] W. Hartmann and E. Fosler-Lussier, “Improved model selection for the asr-driven binary mask,” in *Proceedings of Interspeech*, 2012.
- [31] W. Hartmann and E. Fosler-Lussier, “Asr-driven top-down binary mask estimation using spectral priors,” in *Proceedings of IEEE ICASSP*, pp. 4685–4688, IEEE, 2012.
- [32] W. Hartmann and E. Fosler-Lussier, “Investigations into the incorporation of the ideal binary mask in asr,” in *Proceedings of IEEE ICASSP*, pp. 4804–4807, IEEE, 2011.
- [33] R. Prabhavalkar, P. Jyothi, W. Hartmann, J. Morris, and E. Fosler-Lussier, “Investigations into the crandem approach to word recognition,” in *Proceedings of NAACL-HLT*, pp. 725–728, 2010.
- [34] W. Hartmann and E. Fosler-Lussier, “Investigating phonetic information reduction and lexical confusability,” in *Proceedings of Interspeech*, 2009.

- [35] R. Fox and W. Hartmann, “An abductive approach to hand-written character recognition for multiple domains,” in *Proceedings of ICAI*, pp. 349–355, Citeseer, 2006.
- [36] R. Fox and W. Hartmann, “Hand-written character recognition using layered abduction,” in *Advances in Systems, Computing Sciences and Software Engineering*, pp. 141–147, Springer, 2006.

---

## Honors and Awards

Ed Starr Award for Technical Leadership, 2019  
OSU CSE Research Exhibition Best Poster Award, 2012  
NSF GK-12 Fellowship, 2007  
The Ohio State University Fellowship, 2006  
COMAP MCM, Meritorious Winner, 2005, 2006 (Consortium for Mathematics and Applications, Mathematical Contest in Modeling)  
William H. Greaves Scholarship, 2004, 2005  
Dean’s Scholarship, 2003  
COMAP MCM, Honorable Mention, 2003

---

## Professional Activities

IEEE Speech and Language Technical Committee 2022-2024  
Co-organizer for Interspeech 2021 Special Session for OpenASR20 and Low Resource ASR Development  
North American Computational Linguistics Olympiad (NACLO) Local Organizer, 2011-2012  
Buckeye Language Network Steering Committee, The Ohio State University, 2010-2012  
Reviewer for IEEE Transactions on Audio, Speech, and Language Processing  
Reviewer for Neurocomputing  
Reviewer for EURASIP Journal on Audio, Speech, and Music Processing  
Reviewer for Computer, Speech, and Language  
Reviewer for Interspeech, ICASSP, ASRU, SLT

---

## Computer skills

Languages: C/C++, Python, Perl, Bash  
Applications: Kaldi, PyTorch, Tensorflow, HTK,  
OpenFST, Latex  
OS: Linux, Windows, Mac OS X