Evandro Gouvêa

Please email me for current address.

`evandro.gouvea@alumni.cmu.edu`

Expertise

- Information retrieval
- Speech recognition
- Natural language processing

Education

- **Ph.D.** – Electrical and Computer Engineering — **Carnegie Mellon University**, *USA, 1999*
  Thesis: *Acoustic-Feature-Based Frequency Warping for Speaker Normalization.*

- **M.S.** – Electrical Engineering — **Universidade de São Paulo**, *Brazil, 1993*
  Thesis: *Speech Synthesis in Portuguese.*

- **B.S.** – Electrical Engineering — **Instituto Tecnológico de Aeronáutica**, *Brazil, 1990*

Experience

- **Fortune 500 Company** — **NJ, USA** remotely from **Heidelberg, Germany**
  Full Time Speech Consultant — *Aug 2012 - current*

  - Created Python tools to process and collect linguistic data in English, Spanish, and Portuguese.
  - Improved accuracy in production deployments by personalizing applications via construction of language models and frontend plugins.

- **Independent Projects** — **Heidelberg, Germany**
  Speech Consultant — *Mar 2012 - current*

  - Built interactive voice response (IVR) system based on CMU Sphinx, FreeSwitch, and UniMCRP
  - Built confidence score tool and optimized speech recognition in a virtual assistant research project.

- **European Media Laboratory GmbH (EML)** — **Heidelberg, Germany**
  Software Developer — *July 2010 - Feb 2012*

  - Optimized the company's automatic speech recognizer (ASR), resulting in large gain in speed.
  - Implemented speaker adaptation (cMLLR) and segment clustering API to ASR.
  - Created the company's regression test infrastructure based on Jenkins.
  - Co-implemented statistical language modeling (SLM) tool.

- **Mitsubishi Electric Research Labs (MERL)** — **Cambridge, MA, USA**
  Principal Technical Staff — *Jan 2008 - Jun 2010*

  - Fixed long-standing issues with scoring methods and handling of large document collections in MERL's SpokenQuery voice search (document retrieval) system.
  - Determined system's operating point by evaluating recall rate and recognition error rate across different languages and audio conditions.
  - Constructed subword unit-based spoken document retrieval system, achieving performance invariant to out-of-vocabulary (OOV) rates.
  - Continued contribution to CMU Sphinx, an open source speech recognition system.

- **Carnegie Mellon University** — **Pittsburgh, PA, USA**
  Research Associate — *Feb. 2002 - Dec 2007*

  - Developed Sphinx-4, a Java open source speech recognizer, jointly with Sun Microsystems, and MERL.
  - Maintained the open source CMU Sphinx project, creating code releases, fixing bugs, maintaining regression tests and managing CVS and SVN repositories.
  - Created the open source CMU Sphinx tutorial and scripts.
  - Released the AN4 audio database, and acoustic models in English, Spanish, and French. These were the first free resources widely available to the speech recognition community.
  - Built acoustic models for children's speech used by the Project LISTEN's reading tutor.

- **Vocollect, Inc.**   **Pittsburgh, PA, USA**
  Research Scientist   *Jan 1999 - July 2001*
  - Created Vocollect's Hidden Markov Model-based speech recognition system. System designed to run on an embedded platform in noisy environments, with language independent models trained on small amounts of data.
  - Developed the acoustic model trainer in full; rewrote the front end; designed the collection of audio databases to benchmark speech recognition performance.

- **Carnegie Mellon University**   **Pittsburgh, PA, USA**
  Research Assistant   *Jan 1994 - Dec 1998*
  - Invented vocal tract length normalization (VTLN) algorithms towards the Ph.D.

Courses, Tutorials

- "Machine Learning", Andrew Ng, Coursera, 2015.

- "Natural Language Processing", Dan Jurafsky and Christopher Manning, Stanford Online Course, 2012.

- "Speech Mashups: A Framework For Multimodal Mobile Services And Voice Search", Giuseppe Di Fabbrizio, ICASSP Tutorial, 2010.

- "Making You A Leader Fast Track", IEEE Boston Section Course, 2010.

Computer Skills

Speech technology tools : CMU Sphinx, AT&T Watson, Nuance VoCon, Aachen RASR, SRI LM toolkit.
Operating systems : Unix/Linux, Mac OSX, Windows.
System administration : PBS, Torque, Maui, Jenkins, TWiki.
Languages : C/C++, Java, Python, Perl, Unix/Linux Shell (bash, tcsh), JNI.
Version control : Subversion, CVS, MS SourceSafe.
Compilers and IDE : ant, gcc, Eclipse, Microsoft Visual Studio.
Databases : MySql, PostgreSQL, SOLR.
Productivity tools : Matlab, LaTeX, Emacs, Microsoft Office (Excel, PowerPoint, Word).

Patents

- Ezzat, T., **Gouvêa, E.**, "Method for Retrieving Items Represented by Particles from an Information Database", U.S. Pat. 8,055,693, Granted November 2011.

- Raj, B., **Gouvêa, E.**, "Method for Determining Distributions of Unobserved Classes of a Classifier", U.S. Pat. 8,219,510, Granted July 2012.

- Raj, B., **Gouvêa, E.**, *et al.*, "Method for Indexing for Retrieving Documents Using Particles", U.S. Pat. 8,229,921, Granted July 2012.

Published Research

- Radeck-Arneth, S., Milde, B., Lange, A., **Gouvêa, E.**, Radomski, S., Mühlhäuser, M., Biemann, C., "Open-Source German Distant Speech Recognition: Corpus and Acoustic Model", *Proc. TSD*, Plzen, 2015.

- **Gouvêa, E.**, Moreno-Daniel, A, Reddy, A., Chengalvarayan, R., Thomson, D., Ljolje, A., "The AT&T Speech API: A Study on Practical Challenges for Customized Speech to Text Service", *Proc. Interspeech*, Lyon, 2013.

- **Gouvêa, E.**, "Hybrid Speech Recognition for Voice Search: a Comparative Study", *Proc. Interspeech*, Florence, 2011.

- **Gouvêa, E.**, Davel, M. H., "Kullback-Leibler divergence-based ASR training data selection", *Proc. Interspeech*, Florence, 2011.

- Reddy, S., **Gouvêa, E.**, "Learning from Mistakes: Expanding Pronunciation Lexicons using Word Recognition Errors", *Proc. Interspeech*, Florence, 2011.

- Mostow, J., Beck, J., Cuneo, A., **Gouvêa, E.**, Heiner, C., Juarez, O., "Lessons from Project LISTEN's Session Browser" in C. Romero, S. Ventura, S. R. Viola, M. Pechenizkiy, & R. S. J. d. Baker (Eds.), *Handbook of Educational Data Mining*, Taylor & Francis Group, 2010.

- **Gouvêa, E.**, Ezzat, T., "Vocabulary Independent Spoken Query: a Case for Subword Units", *Proc. Interspeech*, Tokyo, 2010.

- **Gouvêa, E.**, Ezzat, T., Raj, B., "Subword Unit Approaches For Retrieval By Voice", *SpokenQuery 2010 Workshop on Voice Search*, Dallas, 2010.

- **Gouvêa, E.**, Raj, B., "Word Particles Applied to Information Retrieval", *European Conference on Information Retrieval*, Toulouse, 2009.

- Stern, R.M., **Gouvêa, E.**, Kumar, K., "Polyaural Array Processing for Robust Automatic Speech Recognition in Noisy and Reverberant Environments", *Journal of the Acoustical Society of America*, v. 123, n. 5, 2008.

- Stern, R.M., **Gouvêa, E.**, Kim, C., Kumar, K., Park, H.-M., "Binaural and Multiple-Microphone Signal Processing Motivated by Auditory Perception", *HSCMA Joint Workshop on Hands-free Speech Communication and Microphone Arrays*, Trento, 2008.

- Singh, R., **Gouvêa, E.**, Raj, B., "Probabilistic Deduction of Symbol Mappings for Extension of Lexicons", *Proc. Interspeech*, Antwerp, 2007.

- Stern, R.M., **Gouvêa, E.**, Thattai, G., "Polyaural Array Processing for Automatic Speech Recognition in Degraded Environments", *Proc. Interspeech*, Antwerp, 2007.

- Mostow, J., Beck, J., Cen, H., Cuneo, A., **Gouvêa, E.**, Heiner, C, "An Educational Data Mining Tool to Browse Tutor-Student Interactions: Time Will Tell!", *Proc. Workshop on Educational Data Mining*, Pittsburgh, 2005.

- Mostow, J., Beck, J., Cuneo, A., **Gouvêa, E.**, Heiner, C, "A Generic Tool to Browse Tutor-Student Interactions: Time Will Tell!", *Proc. AIED*, Amsterdam, 2005.

- Walker, W., Lamere, P., Kwok, P., Raj, B., Singh, R., **Gouvêa, E.**, Wolf, P., Woelfel, J., "Sphinx-4: A flexible open source framework for speech recognition", *Technical report*, Sun Microsystems, 2004.

- Lamere, P., Kwok, P., Walker, W., **Gouvêa, E.**, Singh, R., Raj, B., Wolf, P., "Design of the CMU Sphinx-4 decoder", *Proc. Eurospeech*, Genebra, 2003.

- **Gouvêa, E.**, "Acoustic-Feature-Based Frequency Warping for Speaker Normalization", Ph.D. Thesis, Carnegie Mellon University, Pittsburgh, 1999.

- Seymore, K., Chen, S., Doh, S., Eskenazi, M., **Gouvêa, E.**, Raj, B., Ravishankar, M., Rosenfeld, R., Siegler, M., Stern, R., Thayer, E., "The 1997 CMU Sphinx-3 English Broadcast News Transcription System", *Proc. DARPA Speech Recognition Workshop*, Chantilly, 1998.

- **Gouvêa, E.**, Stern, R.M., "Speaker Normalization Through Formant-Based Warping of the Frequency Scale", *Proc. Eurospeech*, Rhodes, 1997.

- Raj, B., **Gouvêa, E.**, Stern, R.M., "Cepstral Compensation Using Statistical Linearization", *Proc. ETRW -RSR*, Pont-a-Mousson, 1997.

- Campos, G.L., **Gouvêa, E.**, "Speech Synthesis using the CELP Algorithm", *Proc. ICSLP*, Philadelphia, 1996.

- Raj, B., **Gouvêa, E.**, Stern, R.M., "Cepstral Compensation by Polynomial Approximation for Environment-Independent Speech Recognition", *Proc. ICSLP*, Philadelphia, 1996.

- **Gouvêa, E.**, Moreno, P.J., Raj, B., Sullivan, T.M., Stern, R.M., "Adaptation and Compensation: Approaches to Microphone and Speaker Independence in Automatic Speech Recognition", *Proc. DARPA Speech Recognition Workshop*, Harriman, 1996.

- Jain, U., Siegler, M.A., Doh, S.-J., **Gouvêa, E.**, Moreno, P.J., Raj, B., Stern, R.M., "Recognition of Continuous Broadcast News With Multiple Unknown Speakers And Environments", *Proc. DARPA Speech Recognition Workshop*, Harriman, 1996.

- Moreno, P.J., Raj, B., **Gouvêa, E.**, Stern, R.M., "Multivariate-Gaussian-Based Cepstral Normalization for Robust Speech Recognition", *Proc. ICASSP*, Detroit, 1995.

- **Gouvêa, E.**, "Speech Synthesis in Portuguese", Master's Thesis, Universidade de São Paulo, Brazil, 1993.

Professional Activities

- Committee member, M.S. Thesis, for Rafael Oliveira, Universidade Federal do Pará, Brazil, 2012.

- Co-organizer of the SpokenQuery 2010 Workshop on Voice Search.

- Co-organizer of the CMU Sphinx Users and Developers Workshop 2010.

- Internship mentor for S. Reddy, Summer 2009.

- Committee member, Ph.D. Thesis, for Carlos Lima, Universidade do Minho, Portugal, 2002.

Technical Reviewer

- Panel Reviewer for the U.S. Government's National Science Foundation (NSF).
- Book Reviewer for Elsevier Publishing Company and Springer International Publisher.
- Journal Reviewer for IEEE Transactions on Audio, Speech, and Language Processing.
- Journal Reviewer for Computers in Biology and Medicine.
- Reviewer for ICASSP 2008 – current.
- Reviewer for InterSpeech 2011 – current.
- Reviewer for ASRU 2005 – current.
- Reviewer for SAPA 2004 – 2008.
- Reviewer for ICMI 2008 – 2009.

Honors and Awards

- Courtesy appointment, Affiliate Member of the Machine Learning for Signal Processing Group, Language Technologies Institute, Carnegie Mellon University, May 2011 – present
- MERL President's Award, in recognition of contributions to Voice Interface Support and SpokenQuery in Car Navigation, 2008.
- Scholarship for the Ph.D. program by CNPq (National Council for Scientific and Technological Development, the Brazilian government's research funding agency), January 1994 – July 1998.
- Scholarship for the M.S. program by FAPESP (São Paulo Research Foundation, the research funding agency in the Brazilian state of São Paulo), August 1992 – December 1993.
- Scholarship for scientific initiation by CNPq, Studies in Mathematics, on *topology, functional analysis, and numerical analysis.*, March 1986 – December 1988.
- Participation in the XXII International Mathematics Olympiad, Poland, 1986.
- First Prize in the VIII Brazilian Mathematics Olympiad, 1985.
- Fifth Prize in the I Physics Olympiad of the State of São Paulo, 1985.
- Third Prize in the VI Mathematics Olympiad of the State of São Paulo, 1982.
- First Prize in the IV Mathematics Olympiad of the State of São Paulo, 1980.

Languages

- Native: Portuguese.
- Fluent: English, Spanish.
- Other: French (CEFR Level B1), German (CEFR Level C1), and Italian.

Non-Professional Activities

- Co-hosted the "Brazilian Radio Hour" on WRCT-FM, 1999 to 2007.
- Co-produced video presentation for event celebrating Brazil's 500 years.
- Co-organizer of the 1990 Graduating Class Commencement Ceremony.