

# Mahsa Elyasi (Mahsa Sadat Elyasi Langarani)

---

E-mail: [ms.elyasi.1@gmail.com](mailto:ms.elyasi.1@gmail.com)

Website: [mahsae.github.io](http://mahsae.github.io)

## RESEARCH INTERESTS

Machine learning and its applications in language and speech processing.

## EDUCATION

**Ph.D.**, Computer Science and Engineering

Center for Spoken Language Understanding, OHSU, Portland, OR, Aug 2020

**M.Sc.**, Computer Engineering, Artificial Intelligence

Sharif University of Technology, Tehran, IRAN, September 2012

**B.Sc.**, Computer Engineering, Software Engineering

University College of Nabi Akram, Tabriz, IRAN, September 2010

## POSITIONS

**BioSpeech Inc.**, Portland, OR

March 2020 - Present

Speech Research contractor

- Word boundary alignment for three types of speech: normal, fast, and slow. Performed a word boundary detection using automatic pause detection, speech to text method followed by automatic fine tuning of word boundaries. Compared to baseline, our method improves  $F_1$  measure from 63 to 95.

**ObEN Inc.**, Pasadena, CA

Spring 2017

Speech Research Intern

- TTS adaptation: Developed an intonation adaptation model to transform the perceived identity of a DNN-based TTS (Merlin) system to that of a target speaker. Resulting in higher similarity compared to baseline.
- TTS: developed a discrete cosine transform intonation model for generating fundamental frequency ( $F_0$ ) contour in Mandarin. Achieving higher naturalness compared to rule-based approach.
- Data preparation: Developed a script to speed up the labeling process and error checking which saved one week for the team.

**ObEN Inc.**, Pasadena, CA

Summer 2016

Speech Research Intern

- Speech-To-Music: developed a bidirectional LSTM syllabification method to generate syllable boundaries given speech signal and phonetic labels. Achieving  $F_1$  measure of 95.
- TTS: developed a superpositional intonation model for English. Achieving higher naturalness compared to intonation generated with HTS.

**Sensory, Inc.**, Portland, OR

Fall 2015

Speech Research Intern

- Implemented intent classification to determine the intent expressed in a given query for an NLU system. Trained logistic regression in one-vs-rest setting. Achieved on average  $F_1$  measure of 85.

**Center for Spoken Language Processing**, OHSU, Portland, OR 2012 - 2020

Graduate Research Assistant

- Experimenting effect of fundamental frequency ( $F_0$ ) contour in speaker group classification using combination of CNN and RNN. 1D CNN was used in lower layers in a deep architecture to capture supra-segmental patterns of the  $F_0$  contour. Applying a Recurrent layer on top of the Convolutional helped the

model to capture these temporal pattern. In a balanced data set, achieved an average accuracy of 91

- Proposed a quantitative model for analysis and synthesis of English intonation. Developed and invented various Machine learning method to examine the performance and potential of the proposed intonation model in a variety of speech processing applications, including TTS, TTS adaptation, and speaker group classification.

**PUBLICATIONS** **M.S. Elyasi Langarani**, J. van Santen, Investigating prosodic unit effects of fundamental frequency dynamics in clear and conversational speech (under submission).  
**M.S. Elyasi Langarani**, J. van Santen, Prosody based dialect classification using NMF and sparsity criteria (under submission).  
**M.S. Elyasi Langarani**, J. van Santen, Recurrent Convolutional Neural Network for Classification of Speaker Groups based on Prosodic Information, 12th Women in Machine Learning Workshop (WiML), 2017.  
**M.S. Elyasi Langarani**, J. van Santen, Automatic, model-based detection of pauseless phrase boundaries from fundamental frequency and duration features, 9th ISCA Speech Synthesis Workshop, 2016.  
**M.S. Elyasi Langarani**, J. van Santen, Foot-based Intonation for Text-to-Speech Synthesis using Neural Networks, Speech Prosody 2016.  
**M.S. Elyasi Langarani**, J. van Santen, Speaker Intonation Adaptation for Transforming Text-To-Speech Synthesis Speaker Identity, ASRU 2015.  
**M.S. Elyasi Langarani**, J. van Santen, S.H. Mohammadi, A. Kain, Data-driven Foot-based Intonation Generator for Text-to-Speech Synthesis, Interspeech 2015.  
**M.S. Elyasi Langarani**, J. van Santen, Modeling fundamental frequency dynamic in hypokinetic dysarthria, SLT 2014.  
**M.S. Elyasi Langarani**, E. Klabbers, J. van Santen, A Novel Pitch Decomposition method for the Generalized Linear Alignment Model, ICASSP 2014.  
**M.S. Elyasi Langarani**, H. Veisi, H. Sameti: The effect of phase information in speech enhancement and speech recognition. ISSPA 2012.  
S.H. Mohammadi, H. Sameti, **M.S. Elyasi Langarani**, , A. Tavanaei, KNNDIST: A Non-parametric distance measure for speaker segmentation, Interspeech 2012.

**ACTIVITIES** **Reviewer**, Interspeech, WiML, Speech prosody  
**Organizing Committee**, Volunteer at Interspeech 2012 conference.  
**Student Member**, ISCA, IEEE Signal Processing Society  
**Award**  
Top student in bachelor of Science class (GPA: 3.83)  
Nominated by OHSU for a HHMI fellowship 2015

**COMPUTER SKILLS** *Languages:* Python, R, C++.  
*ML Toolkits:* Keras, TensorFlow.  
*Speech Toolkits:* Merlin, Festival, HTS, Praat, TextGrid.  
*Academic skills:* Data cleaning, Amazon mechanical turk, A/B testing, Error analysis, Statistical reliability.

**REFERENCES** **Jan van Santen**, Professor, Center for Spoken Language Understanding, Oregon Health and Science University, vansantj@ohsu.edu.  
**Alexander Kain**, Associate Professor, Center for Spoken Language Understanding, Oregon Health and Science University, kaina@ohsu.edu.  
**Abeer Alwan**, Professor and Vice Chair, Electrical and Computer Engineering, University of California, Los Angeles (UCLA) alwan@ee.ucla.edu .