

# Shruti Palaskar

✉ [spalaska@cs.cmu.edu](mailto:spalaska@cs.cmu.edu)  
🌐 <https://shrutijpalaskar.github.io>  
🐙 <https://github.com/shrutijpalaskar>  
in <https://www.linkedin.com/in/shrutipalaskar/>

---

## RESEARCH INTERESTS

Multimodal Machine Learning, Speech Recognition and Natural Language Processing

---

## EDUCATION

2016-Current **Carnegie Mellon University, Pittsburgh, USA**  
Ph.D., LTI, School of Computer Science  
Advisor: Prof. Florian Metze

2016-2018 **Carnegie Mellon University, Pittsburgh, USA**  
Master of Language Technologies, LTI, School of Computer Science  
Advisor: Prof. Florian Metze

2012-2016 **Pune Institute of Computer Technology, University of Pune, India**  
Bachelor of Engineering, Computer Engineering  
Grade: First Class with Distinction

---

## HONORS & AWARDS

PhD Fellowship Received the **Facebook Fellowship** for years 2019-2021

PhD Fellowship Received the **Center for Machine Learning and Health Fellowship** for 2018-2019

AAAI'19 **Winner** of the 7<sup>th</sup> Dialog System Technology Challenges: Audio-Visual Scene Aware Dialog Track

Interspeech'18 **PhD student panelist** at the Young Female Researchers in Speech Workshop at Interspeech '18

Grad. Fellowship Recipient of the **CMU LTI Graduate Fellowship** for academic years 2016-2018

---

## PUBLICATIONS

- AAAI'19 DSTC7 "CMU Sinbad's Submission to the DSTC7 AVSD Track",  
**Oral** Ramon Sanabria\*, **Shruti Palaskar\***, and Florian Metze
- ICASSP'19 "Learned in Speech Recognition: Contextual Acoustic Word Embeddings",  
**Shruti Palaskar\***, Vikas Raunak\* and Florian Metze
- ICASSP'19 "Learning from Multiview Correlations in Open-Domain Videos",  
Nils Holzenberger\*, **Shruti Palaskar\***, Pranava Madhyastha, Raman Arora and Florian Metze
- ICASSP'19 "Multimodal Grounding for Sequence-to-Sequence Speech Recognition",  
Ozan Caglayan, Ramon Sanabria, **Shruti Palaskar**, Loic Barrault and Florian Metze
- NIPS ViGIL'18 "Multimodal Abstractive Summarization for Open-Domain Videos",  
**Spotlight** Jindrich Libovicky, **Shruti Palaskar**, Spandana Gella and Florian Metze
- NIPS ViGIL '18 "How2: A Large-scale Dataset for Multimodal Language Understanding", Ramon Sanabria, Ozan Caglayan, **Shruti Palaskar**, Desmond Elliot, Loic Barrault, Lucia Specia and Florian Metze
- SLT'18 "Acoustic-to-Word Recognition with Sequence-to-Sequence Models",  
**Shruti Palaskar** and Florian Metze
- ICASSP'18 "End-to-End Multimodal Speech Recognition",  
**Shruti Palaskar\***, Ramon Sanabria\*, Florian Metze
- ICASSP'18 "Linguistic Unit Discovery from Multimodal Inputs in Unwritten Languages",  
Odette Scharenborg,..., **Shruti Palaskar**,..., Emmanuel Dupoux

- ICNLSSP'17 "Building an ASR System for a Low-resource Language Through the Adaptation of a High-resource Language ASR System: Preliminary Results", Odette Scharenborg, Francesco Ciannella, **Shruti Palaskar**, Alan Black, Florian Metze, Lucas Ondel, Mark Hasegawa-Johnson
- Interspeech'17 "End-to-End Audio Visual Speech Recognition and Summarization", **Shruti Palaskar** and Florian Metze (Poster at YFRS Workshop)
- arXiv "Combining LSTM and Latent Topic Modeling for Mortality Prediction", Yohan Jo, Lisa Lee, **Shruti Palaskar**

---

## ONGOING & PAST RESEARCH

- Sep'18-Current **Information Extraction from Speech Dialogs**, *Graduate Research Assistant, CMU*  
*Advisor: Prof. Florian Metze*
- Developing direct speech-to-information extraction systems for medical conversations
- May'17-Current **Multimodal Speech Recognition and Summarization**, *Graduate Research Assistant, CMU*  
*Advisor: Prof. Florian Metze*
- Building multimodal models to jointly learn to recognize and summarize speech
- Jan'18-Aug'18 **DARPA AIDA**, *Graduate Research Assistant, CMU*  
*Advisors: Prof. Florian Metze, Prof. Eduard Hovy*
- Building multimodal speech recognition models for online inference with streaming data

---

## PAST INTERNSHIPS

- Jun'18-Aug'18 **Grounded Sequence-to-Sequence Transduction**, *Graduate Student Participant*  
JSALT, JHU *Advisors: Prof. Lucia Specia, Prof. Florian Metze*
- Building better multimodal learning representations for speech recognition, summarization
- Jun'17-Aug'17 **Multimodal Linguistic Unit Discovery**, *Graduate Student Participant*  
JSALT, CMU *Advisors: Prof. Emmanuel Dupoux, Prof. Odette Scharenborg, Prof. Florian Metze*
- Worked on sub-unit adaptation from high to low resource languages for ASR
- May'15-Jul'15 **Real-Time Motion Tracking**, *Innovation Engineer*  
MIT Media Lab *Advisors: Dr. Hyunsung Park, Prof. Ramesh Raskar*
- ReDx • Collected chewing videos data for real-time detection of unilateral chewing and TMJ syndrome
- Sep'15-Dec'15 **Statistical Machine Translation**, *Research Intern*  
IIT Bombay *Advisor: Prof. Ganesh Ramakrishnan*
- Focused on generating correct constituent structure and word sense disambiguation

---

## PROGRAMMING EXPERIENCE

Skills Python, PyTorch, TensorFlow, MATLAB, Lua, Java, C, C++, Bash

---

## RELEVANT COURSES

- CMU Topics in Deep Learning, Probabilistic Graphical Models, Speech Recognition and Understanding, Introduction to Machine Learning, Algorithms for NLP, Large Scale Multimedia Analysis
- PICT Data Structures, Advanced Algorithms, Advanced Operating Systems, Databases, Computer Networks, Compilers, Parallel and Distributed Computing, NLP, Discrete Math, Signal Processing

---

## TEACHING ASSISTANT & PROFESSIONAL SERVICE

- TA Fall'18 TA for Speech Recognition and Understanding course taught by Prof. Florian Metze
- Reviewer Reviewer for ICASSP 2019, Dialog State Tracking Challenge 7 AVSD Track 2019