

Elizabeth E. Salesky

esalesky.github.io • elizabeth.salesky@gmail.com

Education

- **Johns Hopkins University** **Baltimore, MD**
Ph.D. in Computer Science 08/2019 – Present
 - **Carnegie Mellon University** **Pittsburgh, PA**
M.S. in Language Technologies 08/2017 – 05/2019
 - **Dartmouth College** **Hanover, NH**
B.A. in Mathematics and Linguistics 09/2008 – 06/2012
- Additional Education**
- | | |
|--|-------------|
| Massachusetts Institute of Technology, Speech Recognition | Spring 2017 |
| Harvard Extension School, Chinese I | Fall 2015 |
| Linguistic Society of America, Linguistic Summer Institute | July 2015 |
| Massachusetts Institute of Technology, Cognitive Science | Spring 2014 |

Awards

- **Women in Defense HORIZONS Scholarship** 2018-2019
- **Interpretability Prize, SIGMORPHON** 2019
- **IEEE SPS Travel Grant** 2018
- **Honors in Linguistics** 2012
- **Citation in Intro to Computational Linguistics** 2011

Experience

- **ETH Zürich, Media Tech Center** **Zürich, Switzerland**
Graduate Student Intern 10/2020 – 03/2021
Researched translation and evaluation of models for dialectal speech and text
- **JSALT Summer Workshop** **Montreal, Canada**
Graduate Student Intern 06/2019 – 08/2019
Researched machine translation of noisy text from social media
- **Karlsruhe Institute of Technology** **Karlsruhe, Germany**
Mitarbeiterin (Research Staff) 05/2019 – 08/2018
Researched implicit removal of disfluencies in end-to-end speech translation;
Created new dataset to facilitate research and evaluation of this task
- **Karlsruhe Institute of Technology** **Karlsruhe, Germany**
Mitarbeiterin (Research Staff) 06/2017 – 08/2017
Researched end-to-end models for speech-to-text translation;
Researched neural models for re-punctuation and casing between ASR and MT steps in SLT
- **MIT Lincoln Laboratory** **Lexington, MA**
Associate Staff 07/2015 – 06/2017
Researched novel alignment and inflection generation techniques for machine translation,
trained competitive speech and text translation systems for international evaluation campaigns;
Researched adaptation techniques for translation with limited parallel data;
Wrote webservice for online text difficulty assessment application;
Supervised summer intern on topic classification project

- **MIT Lincoln Laboratory** **Lexington, MA**
Assistant Staff *05/2012 – 07/2015*
 Incorporated linguistic features into machine translation models;
 Defined features for a language-independent text readability assessment system;
 Created an interface for an audio keyword searching system; ran Mechanical Turk experiments to determine user efficacy with different amounts of training data
 - **Rosetta Project** **San Francisco, CA**
Summer Intern *06/2012 – 08/2012*
 Created an application to upload files from Record-a-thons to RosettaProject's Internet Archive collection, using the Internet Archive API
 - **Logic and Language Course** **Hanover, NH**
Teaching Assistant *Winter 2012, Spring 2012*
 Tutored undergrads in first-order logic, held tutorials 3 times weekly, and graded student work
-

Skills

Programming Languages: Python, Scala, Julia, Matlab, HTML/CSS, Javascript

Natural Languages: English (native), Spanish (proficient), German (limited)

Software: L^AT_EX, PyTorch, Moses, Praat

Service and Professional Activities

Organizing

- **Tutorials:** End-to-End Speech Translation at EACL 2021
- **Workshops:** IWSLT 2021; SIGTYP 2021
- **Evaluation Campaigns:** IWSLT 2021, 2020, 2019; SIGTYP 2021, 2020; SIGMORPHON 2020
- **Seminars Series:** NLP with Friends

Reviewing

- **Journals:** IEEE/ACM Transactions on Audio, Speech, and Language Processing
 - **Conferences:** ACL 2021, 2020*, 2018; Interspeech 2021; CoNLL 2021, 2020; EMNLP 2021, 2019, 2018; IWSLT 2021, 2020, 2019, 2018; WMT 2021, 2020, 2019, 2018
 - **Workshops:** ACL SRW 2021; SIGTYP 2021, 2020; SIGMORPHON 2020; NLPTEA 2018, 2017; BEA 2017, 2015
- * *Outstanding Reviewer*

Volunteering

- NACLO, CMU TechNights, Cambridge Literacy Project, Girls Who Code, Girls Who Build
-

Invited Talks

- **Heidelberg University**, "Source Language Representations for Speech Translation" 05/2021
- **Cambridge University**, "Improving Speech Translation with Linguistically-Informed Representations" 11/2020
- **MIT Lincoln Laboratory**, "Cross-Language Information Retrieval" 11/2015
- **Brigham Young University**, "Natural Language Learning Research and Development" 10/2014

Publications

- **E. Salesky**, D. Etter, M. Post. "Robust Open-Vocabulary Translation from Visual Text Representations." *EMNLP*. 2021.
- T. Pimentel, **E. Salesky**, C. Meister, S. Teufel, D. Blasi, R. Cotterell. "Phone-level Uniform Information Density across and within Languages." *EMNLP*. 2021
- **E. Salesky**, M. Wiesner, J. Bremerman, R. Cattoni, M. Negri, M. Turchi, D.W. Oard, M. Post. "The Multilingual TEDx Corpus for Speech Recognition and Translation." *Interspeech*. 2021.
- N.Q. Pham, T.L. Ha, T.N. Nguyen, T.S. Nguyen, **E. Salesky**, S. Stueker, J. Niehues, A. Waibel. "Relative Positional Encoding for Speech Recognition and Direct Translation." *Interspeech*. 2020.
- **E. Salesky**, A.W. Black. "Phone Features Improve Speech Translation." *ACL*. 2020.
- **E. Salesky**, E. Chodroff, T. Pimentel, M. Wiesner, R. Cotterell, A.W. Black, J. Eisner. "A Corpus for Large-Scale Phonetic Typology." *ACL*. 2020.
- C. Meister, **E. Salesky**, R. Cotterell. "Generalized Entropy Regularization or: There's Nothing Special about Label Smoothing." *ACL*. 2020.
- **E. Salesky**, M. Sperber, A.W. Black. "Exploring Phoneme-Level Speech Representations for End-to-End Speech Translation." *ACL*. 2019.
- **E. Salesky**, M. Sperber, A. Waibel. "Fluent Translations from Disfluent Speech in End-to-End Speech Translation." *NAACL*. 2019.
- **E. Salesky**, A. Runge, A. Coda, J. Niehues, G. Neubig. "Optimizing Segmentation Granularity for Neural Machine Translation." *Machine Translation*. 2020. arXiv:1810.08641 Oct. 2018.
- **E. Salesky**, S. Burger, J. Niehues, A. Waibel. "Towards Fluent Translations from Disfluent Speech." *SLT*. 2018.
- F. Dessloch, T.L. Ha, M. Müller, J. Niehues, T.S. Nguyen, N.Q. Pham, **E. Salesky**, M. Sperber, S. Stüker, T. Zenkel, A. Waibel. "KIT Lecture Translator: Multilingual Speech Translation with One-Shot Learning." *COLING*. 2018.
- **E. Salesky**, J. Ray, W. Shen. "Operational Assessment of Keyword Search on Oral History." *LREC*. 2016.
- **E. Salesky**, W. Shen, "Exploiting Morphological, Grammatical, and Semantic Correlates for Improved Text Difficulty Assessment." *BEA Workshop at ACL*. 2014.
- W. Shen, J. Williams, T. Marius, **E. Salesky**, "A Language-Independent Approach to Automatic Text Difficulty Assessment for Second-Language Learners." *PITR Workshop at ACL*. 2013.

Shared Tasks

- A. Anastasopoulos, O. Bojar, J. Bremerman, R. Cattoni, M. Elbayad, M. Federico, X. Ma, S. Nakamura, M. Negri, J. Niehues, J. Pino, **E. Salesky**, S. Stüker, K. Sudoh, M. Turchi, A. Waibel, C. Wang, M. Wiesner. "Findings of the IWSLT 2021 Evaluation Campaign." *IWSLT*. 2021.
- **E. Salesky**, B.M. Abdullah, S.J. Mielke, E. Klyachko, O. Serikov, E.M. Ponti, R. Kumar, R. Cotterell, E. Vylomova. "SIGTYP 2021 Shared Task: Robust Spoken Language Identification." *SIGTYP*. 2021.
- J. Bjerva, **E. Salesky**, S.J. Mielke, A. Chaudhary, G.G.A. Celano, E.M. Ponti, E. Vylomova, R. Cotterell, I. Augenstein. "SIGTYP 2020 Shared Task: Prediction of Typological Features." *SIGTYP*. 2020.
- E. Ansari, N. Bach, O. Bojar, R. Cattoni, F. Dalvi, N. Durrani, M. Federico, C. Federmann, J. Gu, F. Huang, K. Knight, X. Ma, A. Nagesh, M. Negri, J. Niehues, J. Pino, **E. Salesky**, S. Shi, S. Stüker, M. Turchi, A. Waibel, C. Wang. "Findings of the IWSLT 2020 Evaluation Campaign." *IWSLT*. 2020.
- E. Vylomova, J. White, **E. Salesky**, S. J. Mielke, S. Wu, E. Ponti, R. Hall Maudslay, R. Zmigrod, J. Valvoda, S. Toldova, F. Tyers, E. Klyachko, I. Yegorov, N. Krizhanovsky, P. Czarnowska, I. Nikkarinen, A. Krizhanovsky, T. Pimentel, L. Torroba Hennigen, C. Kirov, G. Nicolai, A. Williams, A. Anastasopoulos, H. Cruz, E. Chodroff, R. Cotterell, M. Silfverberg, M. Hulden. "SIGMORPHON 2020 Shared Task 0: Typologically Diverse Morphological Inflection." *SIGMORPHON*. 2020.

- J. Niehues, R. Cattoni, S. Stüker, M. Negri, M. Turchi, T.L. Ha, **E. Salesky**, R. Sanabria, L. Barrault, L. Specia, M. Federico. "The IWSLT 2019 Evaluation Campaign." *IWSLT*. 2019.
- A. Chaudhary, **E. Salesky**, G. Bhat, D. Mortenson, J. Carbonell, Y. Tsvetkov. "CMU-01 at the SIGMORPHON 2019 Shared Task on Crosslinguality and Context in Morphology." *SIGMORPHON*. 2019.
- N.Q. Pham, M. Sperber, **E. Salesky**, T.L. Ha, J. Niehues, A. Waibel. "KIT's Multilingual Neural Machine Translation systems for IWSLT 2017." *IWSLT*. 2017.
- J. Gwinnup, T. Anderson, G. Erdmann, K. Young, M. Kazi, **E. Salesky** B. Thompson, J. Taylor. "The AFRL-MITLL WMT17 Systems: Old, New, Borrowed, BLEU." *WMT*. 2017.
- M. Kazi, **E. Salesky**, B. Thompson, J. Taylor, J. Gwinnup, T. Anderson, G. Erdmann, E. Hansen, B. Ore, K. Young, M. Hutt. "The MITLL-AFRL IWSLT 2016 Systems." *IWSLT*. 2016.
- J. Gwinnup, T. Anderson, G. Erdmann, K. Young, M. Kazi, **E. Salesky**, B. Thompson. The AFRL-MITLL WMT16 News-Translation Task Systems. *WMT*. 2016.
- M. Kazi, B. Thompson, **E. Salesky**, T. Anderson, G. Erdmann, E. Hansen, B. Ore, J. Gwinnup, K. Young, M. Hutt, C. May. "The MITLL-AFRL IWSLT 2015 Systems." *IWSLT*. 2015.
- J. Gwinnup, T. Anderson, M. Kazi, **E. Salesky**, G. Erdmann, K. Young, B. Thompson, C. May. "The AFRL-MITLL WMT15 System." *WMT*. 2015.
- M. Kazi, **E. Salesky**, B. Thompson, J. Ray, M. Coury, W. Shen, T. Anderson, G. Erdmann, J. Gwinnup, K. Young, B. Ore, M. Hutt. "The MIT-LL/AFRL IWSLT-2014 MT system." *IWSLT*. 2014.
- M. Kazi, M. Coury, **E. Salesky**, J. Ray, W. Shen, T. Gleason, T. Anderson, G. Erdmann, L. Schwartz, B. Ore, R. Slyh, J. Gwinnup, K. Young, M. Hutt. "The MIT-LL/AFRL IWSLT-2013 MT system." *IWSLT*. 2013.