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.Sc. in Language Technologies — *08/2017 – 05/2019*

- **Dartmouth College** — **Hanover, NH**
  B.A. in Linguistics and Mathematics — *09/2008 – 06/2012*

  Additional Education
  Massachusetts Institute of Technology, Speech Recognition — *Spring 2017*
  Linguistic Society of America, Linguistic Summer Institute — *July 2015*
  Massachusetts Institute of Technology, Cognitive Science — *Spring 2014*

Awards

- **Apple Scholars in AI/ML PhD Fellowship** — **2022**

- **Interpretability Prize, SIGMORPHON** — **2019**

- **Women in Defense HORIZONS Scholarship** — **2018**

- **IEEE SPS Travel Award** — **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 of 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) — *05/2018 – 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**
  Assistant → Associate Staff — *11/2012 – 06/2017*
  Researched novel alignment and inflection generation techniques for machine translation;
  Researched adaptation techniques for translation with limited parallel data;

Incorporated linguistic features into translation and text difficulty assessment applications;
Trained competitive speech and text translation systems for international evaluation campaigns;
Created an interface for an audio keyword searching system and ran Mechanical Turk experiments
to determine user efficacy with different amounts of training data

Teaching

- **HEART: Modeling Spoken Language Translation**   **Johns Hopkins University**
  Instructor   *Fall 2021*
  Developed and instructed seminar on the interdisciplinary research area of speech translation,
  with introductions to speech processing, phonetics, and machine translation

- **Logic and Language Course**   **Dartmouth College**
  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: LaTeX, PyTorch, Moses, Praat

Service and Professional Activities

Organizing

- **Tutorials:** End-to-End Speech Translation @ EACL 2021

- **Workshops:** IWSLT 2022, 2021; SIGTYP 2021

- **Steering Committee:** IWSLT (2021–*present*)

- **Evaluation Campaigns:** IWSLT 2022, 2021, 2020, 2019; SIGTYP 2021, 2020; SIGMORPHON 2020

- **Seminar Series:** NLP with Friends; SIGSLT Lectures

Reviewing

- **Journals:** IEEE/ACM Transactions on Audio, Speech, and Language Processing

- **Conferences:** ACL Rolling Review 2022, 2021; ICASSP 2022; Interspeech 2021; ACL 2021, 2020*, 2018; CoNLL 2021, 2020; EMNLP 2021, 2019, 2018; IWSLT 2021, 2020, 2019, 2018; WMT 2021, 2020, 2019, 2018

- **Workshops:** ComputEL 2022; AfricaNLP 2022; ACL SRW 2022, 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

- **Spectrum Labs**, "Robust Open-Vocabulary Translation from Visual Text Representations"   04/2022

- **IST–Unbabel Seminar**, "Robust Open-Vocabulary Translation from Visual Text Representations" 01/2022

- **Saarland University**, "Looking beyond Unicode for Open-Vocabulary Text Representations" 01/2022

- **SIGTYP Lecture Series**, "Looking beyond Unicode for Open-Vocabulary Text Representations" 12/2021

- **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

- S.J. Mielke, Z. Alyafeai, **E. Salesky**, C. Raffel, M. Dey, M. Gallé, A. Raja, C. Si, W.Y. Lee, B. Sagot, S. Tan. (2021). "Between words and characters: A Brief History of Open-Vocabulary Modeling and Tokenization in NLP." *preprint*.

- **E. Salesky**, J. Mäder, S. Klinger. (2021). "Assessing Evaluation Metrics for Speech-to-Speech Translation." *ASRU*.

- **E. Salesky**, D. Etter, M. Post. (2021). "Robust Open-Vocabulary Translation from Visual Text Representations." *EMNLP*.

- T. Pimentel, C. Meister, **E. Salesky**, S. Teufel, D. Blasi, R. Cotterell. (2021). "A surprisal–duration trade-off across and within the world's languages." *EMNLP*.

- **E. Salesky**, M. Wiesner, J. Bremerman, R. Cattoni, M. Negri, M. Turchi, D.W. Oard, M. Post. (2021). "The Multilingual TEDx Corpus for Speech Recognition and Translation." *Interspeech*.

- N.Q. Pham, T.L. Ha, T.N. Nguyen, T.S. Nguyen, **E. Salesky**, S. Stueker, J. Niehues, A. Waibel. (2020). "Relative Positional Encoding for Speech Recognition and Direct Translation." *Interspeech*.

- **E. Salesky**, A.W. Black. (2020). "Phone Features Improve Speech Translation." *ACL*.

- **E. Salesky**, E. Chodroff, T. Pimentel, M. Wiesner, R. Cotterell, A.W. Black, J. Eisner. (2020). "A Corpus for Large-Scale Phonetic Typology." *ACL*.

- C. Meister, **E. Salesky**, R. Cotterell. (2020). "Generalized Entropy Regularization or: There's Nothing Special about Label Smoothing." *ACL*.

- **E. Salesky**, A. Runge, A. Coda, J. Niehues, G. Neubig. (2020). "Optimizing Segmentation Granularity for Neural Machine Translation." *Machine Translation*.

- **E. Salesky**, M. Sperber, A.W. Black. (2019). "Exploring Phoneme-Level Speech Representations for End-to-End Speech Translation." *ACL*.

- **E. Salesky**, M. Sperber, A. Waibel. (2019). "Fluent Translations from Disfluent Speech in End-to-End Speech Translation." *NAACL*.

- **E. Salesky**, S. Burger, J. Niehues, A. Waibel. (2018). "Towards Fluent Translations from Disfluent Speech." *SLT*.

- 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. (2018). "KIT Lecture Translator: Multilingual Speech Translation with One-Shot Learning." *COLING*.

- **E. Salesky**, J. Ray, W. Shen. (2016). "Operational Assessment of Keyword Search on Oral History." *LREC*.

- **E. Salesky**, W. Shen. (2014). "Exploiting Morphological, Grammatical, and Semantic Correlates for Improved Text Difficulty Assessment." *BEA Workshop @ ACL*.

- W. Shen, J. Williams, T. Marius, **E. Salesky**. (2013). "A Language-Independent Approach to Automatic Text Difficulty Assessment for Second-Language Learners." *PITR Workshop @ ACL*.

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. (2021). "Findings of the IWSLT 2021 Evaluation Campaign." *IWSLT*.

- **E. Salesky**,* B.M. Abdullah,* S.J. Mielke,* E. Klyachko, O. Serikov, E.M. Ponti, R. Kumar, R. Cotterell, E. Vylomova. (2021). "SIGTYP 2021 Shared Task: Robust Spoken Language Identification." *SIGTYP*.

- J. Bjerva, **E. Salesky**, S.J. Mielke, A. Chaudhary, G.G.A. Celano, E.M. Ponti, E. Vylomova, R. Cotterell, I. Augenstein. (2020). "SIGTYP 2020 Shared Task: Prediction of Typological Features." *SIGTYP*.

- 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. (2020). "Findings of the IWSLT 2020 Evaluation Campaign." *IWSLT*.

- 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. (2020). "SIGMORPHON 2020 Shared Task 0: Typologically Diverse Morphological Inflection." *SIGMORPHON*.

- J. Niehues, R. Cattoni, S. Stüker, M. Negri, M. Turchi, T.L. Ha, **E. Salesky**, R. Sanabria, L. Barrault, L. Specia, M. Federico. (2019). "The IWSLT 2019 Evaluation Campaign." *IWSLT*.

- A. Chaudhary, **E. Salesky**, G. Bhat, D. Mortenson, J. Carbonell, Y. Tsvetkov. (2019). "CMU-01 at the SIGMORPHON 2019 Shared Task on Crosslinguality and Context in Morphology." *SIGMORPHON*.

- N.Q. Pham, M. Sperber, **E. Salesky**, T.L. Ha, J. Niehues, A. Waibel. (2017). "KIT's Multilingual Neural Machine Translation systems for IWSLT 2017." *IWSLT*.

- J. Gwinnup, T. Anderson, G. Erdmann, K. Young, M. Kazi, **E. Salesky** B. Thompson, J. Taylor. (2017). "The AFRL-MITLL WMT17 Systems: Old, New, Borrowed, BLEU." *WMT*.

- M. Kazi, **E. Salesky**, B. Thompson, J. Taylor, J. Gwinnup, T. Anderson, G. Erdmann, E. Hansen, B. Ore, K. Young, M. Hutt. (2016). "The MITLL-AFRL IWSLT 2016 Systems." *IWSLT*.

- J. Gwinnup, T. Anderson, G. Erdmann, K. Young, M. Kazi, **E. Salesky**, B. Thompson. (2016). "The AFRL-MITLL WMT16 News-Translation Task Systems." *WMT*.

- M. Kazi, B. Thompson, **E. Salesky**, T. Anderson, G. Erdmann, E. Hansen, B. Ore, J. Gwinnup, K. Young, M. Hutt, C. May. (2015). "The MITLL-AFRL IWSLT 2015 Systems." *IWSLT*.

- J. Gwinnup, T. Anderson, M. Kazi, **E. Salesky**, G. Erdmann, K. Young, B. Thompson, C. May. (2015). "The AFRL-MITLL WMT15 System." *WMT*.

- M. Kazi, **E. Salesky**, B. Thompson, J. Ray, M. Coury, W. Shen, T. Anderson, G. Erdmann, J. Gwinnup, K. Young, B. Ore, M. Hutt. (2014). "The MIT-LL/AFRL IWSLT-2014 MT system." *IWSLT*.

- 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. (2013). "The MIT-LL/AFRL IWSLT-2013 MT system." *IWSLT*.

  ★ indicates equal contribution