

IWSLT 2021

**The 18th International Conference on
Spoken Language Translation**

Proceedings of the Conference

August 5 – 6, 2021
Bangkok, Thailand (online)

©2021 The Association for Computational Linguistics
and The Asian Federation of Natural Language Processing

Order copies of this and other ACL proceedings from:

Association for Computational Linguistics (ACL)
209 N. Eighth Street
Stroudsburg, PA 18360
USA
Tel: +1-570-476-8006
Fax: +1-570-476-0860
acl@aclweb.org

ISBN 978-1-954085-74-9

Preface

The International Conference on Spoken Language Translation (IWSLT) is the premiere annual scientific conference for the study, development and evaluation of spoken language translation technology. Launched in 2004 and spun out from the C-STAR speech translation consortium before it (1992-2003), IWSLT is the main venue for scientific exchange on all topics related to speech-to-text translation, speech-to-speech translation, simultaneous and consecutive translation, speech dubbing, cross-lingual communication including all multimodal, emotional, paralinguistic, and stylistic aspects and their applications in the field. The conference organizes evaluations around challenge areas, and presents scientific papers and system descriptions.

This year, IWSLT features four shared tasks: (i) Simultaneous Speech Translation; (ii) Offline Speech Translation; (iii) Multilingual Speech Translation; and (iv) Low-Resource Speech Translation. These topics represent open problems toward effective cross-lingual communication and we expect the community effort and discussion will greatly advance the state of the field. Each shared task was coordinated by a chair. The resulting evaluation campaigns attracted a total of 22 teams, from academy, research centers and industry. System submissions resulted in system papers that will be presented at the conference. Following our call for papers, this year 40 submissions were received. In a blind review process, 11 research papers were selected out of 19 for oral presentation (58%) in addition to 21 system papers. The program this year will also host 4 so-called ACL *findings* papers (not included in this proceedings), that expressed interest in being presented at IWSLT 2021. The program committee is excited about the quality of the accepted papers and expects lively discussion and exchange at the conference.

The conference chairs and organizers would like to express their gratitude to everyone who contributed and supported IWSLT. We thank the shared tasks chairs, organizers, and participants, the program chair and committee members, as well as all the authors that went the extra mile to submit system and research papers to IWSLT, and make this year's conference a most vibrant event. We also wish to express our sincere gratitude to ACL for hosting our conference and for arranging the logistics and infrastructure that allow us to hold IWSLT 2021 as a virtual online conference.

Welcome to IWSLT 2021 wherever you are joining from!

Marta R. Costa-jussà, Program Chair

Marcello Federico and Alex Waibel, Conference co-Chairs

Organizers:

Marcello Federico, Amazon, USA (Conference Chair)
Alex Waibel, CMU, USA (Conference Chair)
Marta R. Costa-jussà, UPC, Spain (Program Chair)
Jan Niehues, Maastricht U., Netherlands (Evaluation Chair)
Sebastian Stüker, KIT, Germany (Evaluation Chair)
Elizabeth Salesky, JHU, USA (Website Chair)

Shared Task Organizers:

Simultaneous Speech Translation

Juan Pino Facebook, USA, (Chair)
Katsuhito Sudoh, NAIST, Japan (Chair)
Satoshi Nakamura, NAIST, Japan
Ondřej Bojar, Charles University, Czech Republic
Xutai Ma, JHU/Facebook, USA
Maha Elbayad, Facebook, USA
Changhan Wang, Facebook, USA

Offline Speech Translation

Marco Turchi, FBK, Italy (Chair)
Sebastian Stüker, KIT, Germany
Jan Niehues, Maastricht University, Netherlands
Matteo Negri, FBK, Italy
Roldano Cattoni, FBK, Italy

Multilingual Speech Translation

Elizabeth Salesky, JHU, USA (Chair)
Jake Bremerman, UMD, USA
Jan Niehues, Maastricht University, Netherlands
Matt Post, JHU, USA
Matthew Wiesner, JHU, USA

Multilingual Speech Translation

Antonios Anastasopoulos, George Mason University, USA (Chair)
Grace Tang, Translators without Borders, Canada
Will Lewis, University of Washington, USA
Sylwia Tur, Appen, USA
Rosie Lazar, Appen, USA
Marcello Federico, Amazon, USA
Alex Waibel, CMU, USA

Program Committee:

Hirofumi Inaguma, Kyoto University, Japan
Surafel Melaku Lakew, Amazon AI, USA
Carlos Escolano, Universitat Politècnica de Catalunya, Spain
Jiatao Gu, Facebook AI Research, USA
Changhan Wang, Facebook AI Research, USA
Yun Tang, Facebook AI Research, USA
Akiko Eriguchi, Microsoft, USA
Xian Li, Facebook AI Research, USA
Qianqian Dong, Chinese Academy of Sciences, China
Yuchen Liu, Princeton, USA
Matthias Sperber, Apple, USA
Antonio Toral, U. Groningen, Netherlands
Thanh-le Ha, Karlsruhe Institute of Technology, Germany
Krzysztof Wolk, Polish-Japanese Academy of Information Technology, Poland
Gerard I. Gallego, Universitat Politècnica de Catalunya, Spain
Julian Salazar, Amazon AWS AI, USA
Chengyi Wang, Nankai University, China
Christian Hardmeier, Uppsala U., Sweden
David Vilar, Amazon, Germany
Elizabeth Salesky, JHU, USA
Evgeny Matusov, AppTek, Germany
Juan Pino, Facebook, USA
Katsuhito Sudoh, NAIST, Japan
Laurent Besacier, IMAG, France
Marco Turchi, FBK, Italy
Markus Freitag, Google, USA
Matteo Negri, FBK, Italy
Matthias Huck, LMU, Germany
Mattia di Gangi, AppTek, Germany
Cuong Hoang, Amazon AI, USA
Nguyen Bach, Alibaba, USA
Preslav Nakov, QCRI, Qatar
Yves Lepage, U. Waseda, Japan
Jiajun Zhang, Institute of Automation Chinese Academy of Sciences, China
Duygu Ataman, University of Zurich, Switzerland
Xutai Ma, Johns Hopkins University, USA
Melvin Johnson, Google, USA
Cuong Hoang, Amazon, USA
José Guilherme Camargo de Souza, Unbabel, Portugal

Table of Contents

FINDINGS OF THE IWSLT 2021 EVALUATION CAMPAIGN

Antonios Anastasopoulos, Ondřej Bojar, Jacob Bremerman, Roldano Cattoni, Maha Elbayad, Marcello Federico, Xutai Ma, Satoshi Nakamura, Matteo Negri, Jan Niehues, Juan Pino, Elizabeth Salesky, Sebastian Stüker, Katsuhito Sudoh, Marco Turchi, Alexander Waibel, Changhan Wang and Matthew Wiesner 1

The USTC-NELSLIP Systems for Simultaneous Speech Translation Task at IWSLT 2021

Dan Liu, Mengge Du, Xiaoxi Li, Yuchen Hu and Lirong Dai 30

NAIST English-to-Japanese Simultaneous Translation System for IWSLT 2021 Simultaneous Text-to-text Task

Ryo Fukuda, Yui Oka, Yasumasa Kano, Yuki Yano, Yuka Ko, Hirotaka Tokuyama, Kosuke Doi, Sakriani Sakti, Katsuhito Sudoh and Satoshi Nakamura 39

The University of Edinburgh's Submission to the IWSLT21 Simultaneous Translation Task

Sukanta Sen, Ulrich Germann and Barry Haddow 46

Without Further ADO: Direct and Simultaneous Speech Translation by AppTek in 2021

Parnia Bahar, Patrick Wilken, Mattia A. Di Gangi and Evgeny Matusov 52

The Volctrans Neural Speech Translation System for IWSLT 2021

Chengqi Zhao, zhicheng liu, Jian Tong, Tao Wang, Mingxuan Wang, Rong Ye, Qianqian Dong, Jun Cao and Lei Li 64

THE IWSLT 2021 BUT SPEECH TRANSLATION SYSTEMS

hari Krishna vydana, Martin Karafiat, Lukas Burget and Jan Černocký 75

Dealing with training and test segmentation mismatch: FBK@IWSLT2021

Sara Papi, Marco Gaido, Matteo Negri and Marco Turchi 84

The NiuTrans End-to-End Speech Translation System for IWSLT 2021 Offline Task

Chen Xu, Xiaoqian Liu, Xiaowen Liu, Tiger Wang, Canan Huang, Tong Xiao and Jingbo Zhu .. 92

ESPnet-ST IWSLT 2021 Offline Speech Translation System

Hirofumi Inaguma, Brian Yan, Siddharth Dalmia, Pengcheng Guo, Jiatong Shi, Kevin Duh and Shinji Watanabe 100

End-to-End Speech Translation with Pre-trained Models and Adapters: UPC at IWSLT 2021

Gerard I. Gállego, Ioannis Tsiamas, Carlos Escolano, José A. R. Fonollosa and Marta R. Costa-jussà 110

VUS at IWSLT 2021: A Finetuned Pipeline for Offline Speech Translation

Yong Rae Jo, Youngki Moon, Minji Jung, Jungyoon Choi, Jihyung Moon and Won Ik Cho 120

KIT's IWSLT 2021 Offline Speech Translation System

Tuan Nam Nguyen, Thai Son Nguyen, Christian Huber, Ngoc-Quan Pham, Thanh-Le Ha, Felix Schneider and Sebastian Stüker 125

FST: the FAIR Speech Translation System for the IWSLT21 Multilingual Shared Task

Yun Tang, Hongyu Gong, Xian Li, Changhan Wang, Juan Pino, Holger Schwenk and Naman Goyal 131

<i>Maastricht University’s Multilingual Speech Translation System for IWSLT 2021</i>	
Danni Liu and Jan Niehues	138
<i>ZJU’s IWSLT 2021 Speech Translation System</i>	
Linlin Zhang	144
<i>Multilingual Speech Translation with Unified Transformer: Huawei Noah’s Ark Lab at IWSLT 2021</i>	
Xingshan Zeng, Liangyou Li and Qun Liu	149
<i>Multilingual Speech Translation KIT @ IWSLT2021</i>	
Ngoc-Quan Pham, Tuan Nam Nguyen, Thanh-Le Ha, Sebastian Stüker, Alexander Waibel and Dan He	154
<i>Edinburgh’s End-to-End Multilingual Speech Translation System for IWSLT 2021</i>	
Biao Zhang and Rico Sennrich	160
<i>ON-TRAC’ systems for the IWSLT 2021 low-resource speech translation and multilingual speech translation shared tasks</i>	
Hang Le, Florentin Barbier, Ha Nguyen, Natalia Tomashenko, salima mdhaffar, Souhir Gabiche Gahbiche, Benjamin Lecouteux, Didier Schwab and Yannick Estève	169
<i>IMS’ Systems for the IWSLT 2021 Low-Resource Speech Translation Task</i>	
Pavel Denisov, Manuel Mager and Ngoc Thang Vu	175
<i>The USYD-JD Speech Translation System for IWSLT2021</i>	
Liang Ding and Dacheng Tao	182
<i>mixSeq: A Simple Data Augmentation Method for Neural Machine Translation</i>	
Xueqing Wu, Yingce Xia, Jinhua Zhu, Lijun Wu, Shufang Xie, Yang Fan and Tao Qin	192
<i>On Knowledge Distillation for Translating Erroneous Speech Transcriptions</i>	
Ryo Fukuda, Katsuhito Sudoh and Satoshi Nakamura	198
<i>Self-Guided Curriculum Learning for Neural Machine Translation</i>	
Lei Zhou, Liang Ding, Kevin Duh, Shinji Watanabe, Ryohei Sasano and Koichi Takeda	206
<i>Between Flexibility and Consistency: Joint Generation of Captions and Subtitles</i>	
Alina Karakanta, Marco Gaido, Matteo Negri and Marco Turchi	215
<i>Large-Scale English-Japanese Simultaneous Interpretation Corpus: Construction and Analyses with Sentence-Aligned Data</i>	
Kosuke Doi, Katsuhito Sudoh and Satoshi Nakamura	226
<i>Inverted Projection for Robust Speech Translation</i>	
Dirk Padfield and Colin Cherry	236
<i>Towards the evaluation of automatic simultaneous speech translation from a communicative perspective</i>	
Claudio Fantinuoli and Bianca Prandi	245
<i>Tag Assisted Neural Machine Translation of Film Subtitles</i>	
Aren Siekmeier, WonKee Lee, Hongseok Kwon and Jong-Hyeok Lee	255
<i>A Statistical Extension of Byte-Pair Encoding</i>	
David Vilar and Marcello Federico	263

Integrated Training for Sequence-to-Sequence Models Using Non-Autoregressive Transformer
Evgeniia Tokarchuk, Jan Rosendahl, Weiyue Wang, Pavel Petrushkov, Tomer Lancewicki, Shahram
Khadivi and Hermann Ney 276

Data Augmentation by Concatenation for Low-Resource Translation: A Mystery and a Solution
Toan Q. Nguyen, Kenton Murray and David Chiang 287

Conference Program

Day 1

Evaluation Overview

FINDINGS OF THE IWSLT 2021 EVALUATION CAMPAIGN

Antonios Anastasopoulos, Ondřej Bojar, Jacob Bremerman, Roldano Cattoni, Maha Elbayad, Marcello Federico, Xutai Ma, Satoshi Nakamura, Matteo Negri, Jan Niehues, Juan Pino, Elizabeth Salesky, Sebastian Stüker, Katsuhito Sudoh, Marco Turchi, Alexander Waibel, Changan Wang and Matthew Wiesner

Day 1-2

System Papers

The USTC-NELSLIP Systems for Simultaneous Speech Translation Task at IWSLT 2021

Dan Liu, Mengge Du, Xiaoxi Li, Yuchen Hu and Lirong Dai

NAIST English-to-Japanese Simultaneous Translation System for IWSLT 2021 Simultaneous Text-to-text Task

Ryo Fukuda, Yui Oka, Yasumasa Kano, Yuki Yano, Yuka Ko, Hirotaka Tokuyama, Kosuke Doi, Sakriani Sakti, Katsuhito Sudoh and Satoshi Nakamura

The University of Edinburgh's Submission to the IWSLT21 Simultaneous Translation Task

Sukanta Sen, Ulrich Germann and Barry Haddow

Without Further Ado: Direct and Simultaneous Speech Translation by AppTek in 2021

Parnia Bahar, Patrick Wilken, Mattia A. Di Gangi and Evgeny Matusov

The Volctrans Neural Speech Translation System for IWSLT 2021

Chengqi Zhao, zhicheng liu, Jian Tong, Tao Wang, Mingxuan Wang, Rong Ye, Qianqian Dong, Jun Cao and Lei Li

THE IWSLT 2021 BUT SPEECH TRANSLATION SYSTEMS

hari Krishna vydana, Martin Karafiat, Lukas Burget and Jan Černocký

Dealing with training and test segmentation mismatch: FBK@IWSLT2021

Sara Papi, Marco Gaido, Matteo Negri and Marco Turchi

The NiuTrans End-to-End Speech Translation System for IWSLT 2021 Offline Task

Chen Xu, Xiaoqian Liu, Xiaowen Liu, Tiger Wang, Canan Huang, Tong Xiao and Jingbo Zhu

Day 1-2 (continued)

ESPnet-ST IWSLT 2021 Offline Speech Translation System

Hirofumi Inaguma, Brian Yan, Siddharth Dalmaia, Pengcheng Guo, Jiatong Shi, Kevin Duh and Shinji Watanabe

End-to-End Speech Translation with Pre-trained Models and Adapters: UPC at IWSLT 2021

Gerard I. Gállego, Ioannis Tsiamas, Carlos Escolano, José A. R. Fonollosa and Marta R. Costa-jussà

VUS at IWSLT 2021: A Finetuned Pipeline for Offline Speech Translation

Yong Rae Jo, Youngki Moon, Minji Jung, Jungyoon Choi, Jihyung Moon and Won Ik Cho

KIT's IWSLT 2021 Offline Speech Translation System

Tuan Nam Nguyen, Thai Son Nguyen, Christian Huber, Ngoc-Quan Pham, Thanh-Le Ha, Felix Schneider and Sebastian Stüker

FST: the FAIR Speech Translation System for the IWSLT21 Multilingual Shared Task

Yun Tang, Hongyu Gong, Xian Li, Changan Wang, Juan Pino, Holger Schwenk and Naman Goyal

Maastricht University's Multilingual Speech Translation System for IWSLT 2021

Danni Liu and Jan Niehues

ZJU's IWSLT 2021 Speech Translation System

Linlin Zhang

Multilingual Speech Translation with Unified Transformer: Huawei Noah's Ark Lab at IWSLT 2021

Xingshan Zeng, Liangyou Li and Qun Liu

Multilingual Speech Translation KIT @ IWSLT2021

Ngoc-Quan Pham, Tuan Nam Nguyen, Thanh-Le Ha, Sebastian Stüker, Alexander Waibel and Dan He

Edinburgh's End-to-End Multilingual Speech Translation System for IWSLT 2021

Biao Zhang and Rico Sennrich

ON-TRAC' systems for the IWSLT 2021 low-resource speech translation and multilingual speech translation shared tasks

Hang Le, Florentin Barbier, Ha Nguyen, Natalia Tomashenko, salima mdhaffar, Souhir Gabiche Gahbiche, Benjamin Lecouteux, Didier Schwab and Yannick Estève

IMS' Systems for the IWSLT 2021 Low-Resource Speech Translation Task

Pavel Denisov, Manuel Mager and Ngoc Thang Vu

Day 1-2 (continued)

The USYD-JD's Speech Translation System for IWSLT2021

Liang Ding and Dacheng Tao

Research Papers

mixSeq: A Simple Data Augmentation Method for Neural Machine Translation

Xueqing Wu, Yingce Xia, Jinhua Zhu, Lijun Wu, Shufang Xie, Yang Fan and Tao Qin

On Knowledge Distillation for Translating Erroneous Speech Transcriptions

Ryo Fukuda, Katsuhito Sudoh and Satoshi Nakamura

Self-Guided Curriculum Learning for Neural Machine Translation

Lei Zhou, Liang Ding, Kevin Duh, Shinji Watanabe, Ryohei Sasano and Koichi Takeda

Between Flexibility and Consistency: Joint Generation of Captions and Subtitles

Alina Karakanta, Marco Gaido, Matteo Negri and Marco Turchi

Large-Scale English-Japanese Simultaneous Interpretation Corpus: Construction and Analyses with Sentence-Aligned Data

Kosuke Doi, Katsuhito Sudoh and Satoshi Nakamura

Inverted Projection for Robust Speech Translation

Dirk Padfield and Colin Cherry

Towards the evaluation of automatic simultaneous speech translation from a communicative perspective

Claudio Fantinuoli and Bianca Prandi

Tag Assisted Neural Machine Translation of Film Subtitles

Aren Siekmeier, WonKee Lee, Hongseok Kwon and Jong-Hyeok Lee

A Statistical Extension of Byte-Pair Encoding

David Vilar and Marcello Federico

Integrated Training for Sequence-to-Sequence Models Using Non-Autoregressive Transformer

Evgeniia Tokarchuk, Jan Rosendahl, Weiyue Wang, Pavel Petrushkov, Tomer Lancewicki, Shahram Khadivi and Hermann Ney

Day 1-2 (continued)

Data Augmentation by Concatenation for Low-Resource Translation: A Mystery and a Solution

Toan Q. Nguyen, Kenton Murray and David Chiang