

IWSLT 2022

**The 19th International Conference on Spoken Language
Translation**

Proceedings of the Conference

May 26-27, 2022

The IWSLT organizers gratefully acknowledge the support from the following sponsors and donors:

Diamond



Platinum



Bronze



©2022 Association for Computational Linguistics

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 978-1-955917-41-4

Introduction

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 eight shared tasks: (i) Simultaneous speech translation, (ii) Offline speech translation, (iii) Speech to speech translation, (iv) Low-resource speech translation, (v) Multilingual speech translation, (vi) Dialect speech translation. (vii) Formality control for spoken language translation, (viii) Isometric spoken language 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 one or more chairs. The resulting evaluation campaigns attracted a total of 27 teams, from academia, research centers and industry. System submissions resulted in system papers that will be presented at the conference. Following our call for papers, this year 44 submissions were received. In a blind review process, 9 research papers were selected out of 18 for oral presentation (50%) in addition to 25 system papers.

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. In particular, we wish to thank our Diamond sponsors and donors Apple, AWS, Meta and Zoom, our Platinum sponsor Microsoft, and our Bronze sponsor AppTek. 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 2022, for the first time, as a hybrid conference.

Welcome to IWSLT 2022 wherever you are joining us in person, in Dublin, or remotely!

Marcello Federico and Alex Waibel, Conference Chairs

Organizing Committee

Conference Chairs

Marcello Federico, AWS AI Labs, USA
Alex Waibel, CMU, USA

Program Chair

Marta Costa-jussà, Meta AI, France

Evaluation Chairs

Sebastian Stüker, KIT, Germany
Jan Niehues, KIT, Germany

Website and Publication Chair

Elizabeth Salesky, JHU, USA

Program Committee

Program Committee

Duygu Ataman, University of Zurich, Switzerland
Nguyen Bach, Alibaba, USA
Laurent Besacier, IMAG, France
Anna Currey, AWS AI Labs, USA
Mattia Di Gangi, AppTek, Germany
Georgiana Dinu, AWS AI Labs, Germany
Akiko Eriguchi, Microsoft, USA
Carlos Escolano, Universitat Politècnica de Catalunya, Spain
Markus Freitag, Google, USA
Gerard I. Gallego, Universitat Politècnica de Catalunya, Spain
Cuong Hoang, AWS AI Labs, USA
Matthias Huck, LMU, Germany
Hirofumi Inaguma, Kyoto University, Japan
Takatomo Kano, Nara Institute of Science and Technology, Japan
Yves Lepage, U. Waseda, Japan
Yuchen Liu, Princeton, USA
Xutai Ma, Johns Hopkins University, USA
Evgeny Matusov, AppTek, Germany
Surafel Melaku Lakew, Amazon AI, USA
Maria Nadejde, AWS AI Labs, USA
Matteo Negri, FBK, Italy
Juan Pino, Meta AI, USA
Raghavendra Pappagari, AWS AI Labs, USA
Julian Salazar, Amazon AWS AI, USA
Elizabeth Salesky, Johns Hopkins University, USA
Matthias Sperber, Apple, USA
Sebastian Stüker, Karlsruhe Institute of Technology, Germany
Katsuhito Sudoh, NAIST, Japan
Yun Tang, Meta AI, USA
Brian Thompson, AWS AI Labs, USA
Ioannis Tsiamas, Universitat Politècnica de Catalunya, Spain
Marco Turchi, FBK, Italy
David Vilar, Google, Germany
Changhan Wang, Meta AI, USA
Chengyi Wang, Nankai University, China
Krzysztof Wolk, Polish-Japanese Academy of Information Technology, Poland

Invited Speakers

Frederic Chaume, Universitat Jaume I

Keynote Talk: Synchronization in translation for dubbing: implications for its automation

Frederic Chaume
Universitat Jaume I

Abstract: Synchronization (or lip-sync, also spelled lip-synch) is one of the key factors in audiovisual translation, especially in the context of dubbing. Although it is often considered as the distinguishing feature of dubbing, it is only one of several important aspects such as the 'natural' reproduction of a pre-fabricated oral discourse or the translation problems posed by the interaction between image and word. If we take a look at the research on lip-sync, it is regarded as an urgent, vital issue, as can be seen from the wide range of publications on the subject. Beyond doubt synchronization has a direct impact on the translation process and product, and as such, puts all the translator's creative skills to the test. Dubbing is a well-known example of the invisibility of translation, an artistic and technical exercise that intentionally replaces the original dialogue track with a new track on which target language (TL) dialogue exchanges are recorded. In contrast to voice-over for example, the emphasis in dubbing lies in matching the translation to the silent mouths of the original actors. The result is that viewers watch and hear foreign actors speaking in the viewers' own language, a paradox which has been naturally accepted in all dubbing countries. This talk will deal with the definition and scope of synchronization in the audiovisual translation field, will explain the three main synchronization types, will tackle issues related to different language pairs combinations and will present the last efforts carried out by some start-ups and research groups to automate this technical and artistic process. The talk will be illustrated with clips from films and TV series dubbed into six different languages.

Bio: Frederic Chaume is a Full Professor of Audiovisual Translation at Universitat Jaume I (Spain), where he teaches audiovisual translation theory and translation and adaptation for dubbing; and Honorary Professor at University College London (UK), where he teaches translation and adaptation for voice-over and dubbing, Universidad Ricardo Palma (Perú) and Universidad Peruana de Ciencias Aplicadas (Perú). He is author of eight books and has also coedited two books and three special journal issues (Textus, Perspectives, Prosopopeya). He is the director of the TRAMA book series (Publicacions de la Universitat Jaume I), the first collection of monographs on audiovisual translation and media localization. Prof. Chaume has published over 100 articles, book chapters and encyclopedic entries on audiovisual translation and has given numerous keynote lectures on this topic in international translation studies conferences and in several European and American universities. He also teaches regularly in some of them (University College London-UK, Universidad de Granada-Spain, Università di Torino-Italy, among others). He has supervised or co-supervised 20 PhD theses on the topic of audiovisual translation and some of them have received different Spanish and European awards. He is also in close contact with the industry, serves as a consultant for Netflix and has signed several research agreements with different stakeholders of the media localization sector. He coordinates the research group TRAMA (www.trama.uji.es) and is the recipient of the Berlanga Award (2010), the Xènia Martínez Award (2016) and the Jan Ivarsson's Award (2020) for his constant and enthusiastic support to media localization as well as his constant university training in this field.

Table of Contents

Scientific Papers:

<i>SubER - A Metric for Automatic Evaluation of Subtitle Quality</i> Patrick Wilken, Panayota Georgakopoulou and Evgeny Matusov	1
<i>Improving Arabic Diacritization by Learning to Diacritize and Translate</i> Brian Thompson and Ali Alshehri	11
<i>Simultaneous Neural Machine Translation with Prefix Alignment</i> Yasumasa Kano, Katsuhito Sudoh and Satoshi Nakamura	22
<i>Locality-Sensitive Hashing for Long Context Neural Machine Translation</i> Frithjof Petrick, Jan Rosendahl, Christian Herold and Hermann Ney	32
<i>Anticipation-Free Training for Simultaneous Machine Translation</i> Chih-Chiang Chang, Shun-Po Chuang and Hung-yi Lee	43
<i>Who Are We Talking About? Handling Person Names in Speech Translation</i> Marco Gaido, Matteo Negri and Marco Turchi	62
<i>Joint Generation of Captions and Subtitles with Dual Decoding</i> Jitao Xu, François Buet, Josep Crego, Elise Bertin-Lemée and François Yvon	74
<i>MirrorAlign: A Super Lightweight Unsupervised Word Alignment Model via Cross-Lingual Contrastive Learning</i> Di Wu, Liang Ding, Shuo Yang and Mingyang Li	83
<i>On the Impact of Noises in Crowd-Sourced Data for Speech Translation</i> Siqi Ouyang, Rong Ye and Lei Li	92

Evaluation Campaign:

<i>Findings of the IWSLT 2022 Evaluation Campaign</i> Antonios Anastasopoulos, Loïc Barrault, Luisa Bentivogli, Marceley Zanon Boito, Ondřej Bojar, Roldano Cattoni, Anna Currey, Georgiana Dinu, Kevin Duh, Maha Elbayad, Clara Emmanuel, Yannick Estève, Marcello Federico, Christian Federmann, Souhir Gahbiche, Hongyu Gong, Roman Grundkiewicz, Barry Haddow, Benjamin Hsu, Dávid Javorský, Věra Kloudová, Surafel M. Lakew, Xutai Ma, Prashant Mathur, Paul McNamee, Kenton Murray, Maria Nădejde, Satoshi Nakamura, Matteo Negri, Jan Niehues, Xing Niu, John Ortega, Juan Pino, Elizabeth Salesky, Jiatong Shi, Matthias Sperber, Sebastian Stüker, Katsuhito Sudoh, Marco Turchi, Yogesh Virkar, Alexander Waibel, Changan Wang and Shinji Watanabe	98
<i>The YiTrans Speech Translation System for IWSLT 2022 Offline Shared Task</i> Ziqiang Zhang and Junyi Ao	158
<i>Amazon Alexa AI's System for IWSLT 2022 Offline Speech Translation Shared Task</i> Akshaya Vishnu Kudlu Shanbhogue, Ran Xue, Ching-Yun Chang and Sarah Campbell	169
<i>Efficient yet Competitive Speech Translation: FBK@IWSLT2022</i> Marco Gaido, Sara Papi, Dennis Fucci, Giuseppe Fiameni, Matteo Negri and Marco Turchi	177

<i>Effective combination of pretrained models - KIT@IWSLT2022</i>	
Ngoc-Quan Pham, Tuan Nam Nguyen, Thai-Binh Nguyen, Danni Liu, Carlos Mullov, Jan Niehues and Alexander Waibel	190
<i>The USTC-NELSLIP Offline Speech Translation Systems for IWSLT 2022</i>	
Weitai Zhang, Zhongyi Ye, Haitao Tang, Xiaoxi Li, Xinyuan Zhou, Jing Yang, Jianwei Cui, Dan Liu, Junhua Liu and Lirong Dai	198
<i>The AISP-SJTU Simultaneous Translation System for IWSLT 2022</i>	
Qinpei Zhu, Renshou Wu, Guangfeng Liu, Xinyu Zhu, Xingyu Chen, Yang Zhou, Qingliang Miao, Rui Wang and Kai Yu	208
<i>The Xiaomi Text-to-Text Simultaneous Speech Translation System for IWSLT 2022</i>	
Bao Guo, Mengge Liu, Wen Zhang, Hexuan Chen, Chang Mu, Xiang Li, Jianwei Cui, Bin Wang and Yuhang Guo	216
<i>NVIDIA NeMo Offline Speech Translation Systems for IWSLT 2022</i>	
Oleksii Hrinchuk, Vahid Noroozi, Ashwinkumar Ganesan, Sarah Campbell, Sandeep Subramanian, Somshubra Majumdar and Oleksii Kuchaiev	225
<i>The NiuTrans’s Submission to the IWSLT22 English-to-Chinese Offline Speech Translation Task</i>	
Yuhao Zhang, Canan Huang, Chen Xu, Xiaoqian Liu, Bei Li, Anxiang Ma, Tong Xiao and Jingbo Zhu	232
<i>The HW-TSC’s Offline Speech Translation System for IWSLT 2022 Evaluation</i>	
Minghan Wang, Jiaxin Guo, Xiaosong Qiao, Yuxia Wang, Daimeng Wei, Chang Su, Yimeng Chen, Min Zhang, Shimin Tao, Hao Yang and Ying Qin	239
<i>The HW-TSC’s Simultaneous Speech Translation System for IWSLT 2022 Evaluation</i>	
Minghan Wang, Jiaxin Guo, Yinglu Li, Xiaosong Qiao, Yuxia Wang, Zongyao Li, Chang Su, Yimeng Chen, Min Zhang, Shimin Tao, Hao Yang and Ying Qin	247
<i>MLLP-VRain UPV systems for the IWSLT 2022 Simultaneous Speech Translation and Speech-to-Speech Translation tasks</i>	
Javier Iranzo-Sánchez, Javier Jorge Cano, Alejandro Pérez-González-de-Martos, Adrián Giménez Pastor, Gonçal V. Garcés Díaz-Munío, Pau Baquero-Arnal, Joan Albert Silvestre-Cerdà, Jorge Civera Saiz, Albert Sanchis and Alfons Juan	255
<i>Pretrained Speech Encoders and Efficient Fine-tuning Methods for Speech Translation: UPC at IWSLT 2022</i>	
Ioannis Tsiamas, Gerard I. Gállego, Carlos Escolano, José A. R. Fonollosa and Marta R. Costajussà	265
<i>CUNI-KIT System for Simultaneous Speech Translation Task at IWSLT 2022</i>	
Peter Polák, Ngoc-Quan Pham, Tuan Nam Nguyen, Danni Liu, Carlos Mullov, Jan Niehues, Ondřej Bojar and Alexander Waibel	277
<i>NAIST Simultaneous Speech-to-Text Translation System for IWSLT 2022</i>	
Ryo Fukuda, Yuka Ko, Yasumasa Kano, Kosuke Doi, Hirotaka Tokuyama, Sakriani Sakti, Katsuhito Sudoh and Satoshi Nakamura	286
<i>The HW-TSC’s Speech to Speech Translation System for IWSLT 2022 Evaluation</i>	
Jiaxin Guo, Yinglu Li, Minghan Wang, Xiaosong Qiao, Yuxia Wang, Hengchao Shang, Chang Su, Yimeng Chen, Min Zhang, Shimin Tao, Hao Yang and Ying Qin	293

<i>CMU's IWSLT 2022 Dialect Speech Translation System</i>	
Brian Yan, Patrick Fernandes, Siddharth Dalmia, Jiatong Shi, Yifan Peng, Dan Berrebbi, Xinyi Wang, Graham Neubig and Shinji Watanabe	298
<i>ON-TRAC Consortium Systems for the IWSLT 2022 Dialect and Low-resource Speech Translation Tasks</i>	
Marcely Zanon Boito, John Ortega, Hugo Riguidel, Antoine Laurent, Loïc Barrault, Fethi Bougares, Firas Chaabani, Ha Nguyen, Florentin Barbier, Souhir Gahbiche and Yannick Estève	308
<i>JHU IWSLT 2022 Dialect Speech Translation System Description</i>	
Jinyi Yang, Amir Hussein, Matthew Wiesner and Sanjeev Khudanpur	319
<i>Controlling Translation Formality Using Pre-trained Multilingual Language Models</i>	
Elijah Rippeth, Sweta Agrawal and Marine Carpuat	327
<i>Controlling Formality in Low-Resource NMT with Domain Adaptation and Re-Ranking: SLT-CDT-UoS at IWSLT2022</i>	
Sebastian T. Vincent, Loïc Barrault and Carolina Scarton	341
<i>Improving Machine Translation Formality Control with Weakly-Labelled Data Augmentation and Post Editing Strategies</i>	
Daniel Zhang, Jiang Yu, Pragati Verma, Ashwinkumar Ganesan and Sarah Campbell	351
<i>HW-TSC's Participation in the IWSLT 2022 Isometric Spoken Language Translation</i>	
Zongyao Li, Jiaxin Guo, Daimeng Wei, Hengchao Shang, Minghan Wang, Ting Zhu, Zhanglin Wu, Zhengzhe Yu, Xiaoyu Chen, Lizhi Lei, Hao Yang and Ying Qin	361
<i>AppTek's Submission to the IWSLT 2022 Isometric Spoken Language Translation Task</i>	
Patrick Wilken and Evgeny Matusov	369
<i>Hierarchical Multi-task learning framework for Isometric-Speech Language Translation</i>	
Aakash Bhatnagar, Nidhir Bhavsar, Muskaan Singh and Petr Motlicek	379