

Workshop on Computational Approaches to Linguistic Code Switching (CALCS 2023)

Singapore
7 December 2023

ISBN: 978-1-7138-8597-9

Printed from e-media with permission by:

Curran Associates, Inc.
57 Morehouse Lane
Red Hook, NY 12571



Some format issues inherent in the e-media version may also appear in this print version.

Copyright© (2023) by the Association for Computational Linguistics
All rights reserved.

Printed with permission by Curran Associates, Inc. (2024)

For permission requests, please contact the Association for Computational Linguistics
at the address below.

Association for Computational Linguistics
209 N. Eighth Street
Stroudsburg, Pennsylvania 18360

Phone: 1-570-476-8006
Fax: 1-570-476-0860

acl@aclweb.org

Additional copies of this publication are available from:

Curran Associates, Inc.
57 Morehouse Lane
Red Hook, NY 12571 USA
Phone: 845-758-0400
Fax: 845-758-2633
Email: curran@proceedings.com
Web: www.proceedings.com

Table of Contents

<i>TongueSwitcher: Fine-Grained Identification of German-English Code-Switching</i> Igor Sterner and Simone Teufel	1
<i>Towards Real-World Streaming Speech Translation for Code-Switched Speech</i> Belen Alastruey, Matthias Sperber, Christian Gollan, Dominic Telaar, Tim Ng and Aashish Agarwal	14
<i>Language Preference for Expression of Sentiment for Nepali-English Bilingual Speakers on Social Media</i> Niraj Pahari and Kazutaka Shimada	23
<i>Text-Derived Language Identity Incorporation for End-to-End Code-Switching Speech Recognition</i> Qinyi Wang and Haizhou Li	33
<i>Prompting Multilingual Large Language Models to Generate Code-Mixed Texts: The Case of South East Asian Languages</i> Zheng Xin Xin Yong, Ruochen Zhang, Jessica Zosa Forde, Skyler Wang, Arjun Subramonian, Holy Lovenia, Samuel Cahyawijaya, Genta Indra Winata, Lintang Sutawika, Jan Christian Blaise Cruz, Yin Lin Tan, Long Phan, Long Phan, Rowena Garcia, Thamar Solorio and Alham Fikri Aji	43
<i>CONFLATOR: Incorporating Switching Point based Rotatory Positional Encodings for Code-Mixed Language Modeling</i> Mohsin Ali Mohammed, Sai Teja Kandukuri, Neeharika Gupta, Parth Patwa, Anubhab Chatterjee, Vinija Jain, Aman Chadha and Amitava Das	64
<i>Unified Model for Code-Switching Speech Recognition and Language Identification Based on Concatenated Tokenizer</i> Kunal Dhawan, KDimating Reakesh and Boris Ginsburg	74
<i>Multilingual self-supervised speech representations improve the speech recognition of low-resource African languages with codeswitching</i> Tolulope Ogunremi, Christopher Manning and Dan Jurafsky	83