

4th Workshop on Computational Approaches to Code Switching (CALCS 2020)

Marseille, France
11 – 16 May 2020

Editors:

**Thamar Solorio
Monojit Choudhury
Kalika Bali**

**Sunayana Sitaram
Amitava Das
Mona Diab**

ISBN: 978-1-7138-1258-6

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© (2020) by the Association for Computational Linguistics
All rights reserved.

Copyright for individual papers remains with the authors and are licensed under a Creative Commons 4.0 license, CC-BY-ND. (<https://creativecommons.org/licenses/by-nd/4.0/>)

Printed with permission by Curran Associates, Inc. (2020)

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>An Annotated Corpus of Emerging Anglicisms in Spanish Newspaper Headlines</i> Elena Alvarez-Mellado	1
<i>A New Dataset for Natural Language Inference from Code-mixed Conversations</i> Simran Khanuja, Sandipan Dandapat, Sunayana Sitaram and Monojit Choudhury	9
<i>When is Multi-task Learning Beneficial for Low-Resource Noisy Code-switched User-generated Algerian Texts?</i> Wafia Adouane and Jean-Philippe Bernardy	17
<i>Evaluating Word Embeddings for Indonesian–English Code-Mixed Text Based on Synthetic Data</i> Arra’Di Nur Rizal and Sara Stymne	26
<i>Understanding Script-Mixing: A Case Study of Hindi-English Bilingual Twitter Users</i> Abhishek Srivastava, Kalika Bali and Monojit Choudhury	36
<i>Sentiment Analysis for Hinglish Code-mixed Tweets by means of Cross-lingual Word Embeddings</i> Pranaydeep Singh and Els Lefever	45
<i>Semi-supervised acoustic and language model training for English-isiZulu code-switched speech recognition</i> Astik Biswas, Febe De Wet, Ewald Van der westhuizen and Thomas Niesler	52
<i>Code-mixed parse trees and how to find them</i> Anirudh Srinivasan, Sandipan Dandapat and Monojit Choudhury	57
<i>Towards an Efficient Code-Mixed Grapheme-to-Phoneme Conversion in an Agglutinative Language: A Case Study on To-Korean Transliteration</i> Won Ik Cho, Seok Min Kim and Nam Soo Kim	65