

ACL 2020

**The 29th International Conference on Computational  
Linguistics**

**Proceedings of the First Workshop on Performance and  
Interpretability Evaluations of Multimodal, Multipurpose,  
Massive-Scale Models (MMMPIE 2022)**

October 12–17, 2022

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](mailto:acl@aclweb.org)

## **Organizing Committee**

- Maria Glenski, Pacific Northwest National Laboratory
- Vidhisha Balachandran, Carnegie Mellon University
- Megan Kohagen, Quantinuum
- Yulia Tsvetkov, University of Washington



## Table of Contents

<i>On the Effects of Video Grounding on Language Models</i> Ehsan Doostmohammadi and Marco Kuhlmann .....	1
<i>Rethinking Task Sampling for Few-shot Vision-Language Transfer Learning</i> Zhenhailong Wang, Hang Yu, Manling Li, Han Zhao and Heng Ji .....	7
<i>Analyzing BERT Cross-lingual Transfer Capabilities in Continual Sequence Labeling</i> Juan Manuel Coria, Mathilde Veron, Sahar Ghannay, Guillaume Bernard, Hervé Bredin, Olivier Galibert and Sophie Rosset .....	15
<i>Pixel-Level BPE for Auto-Regressive Image Generation</i> Anton Razzhigaev, Anton Voronov, Andrey Kaznacheev, Andrey Kuznetsov, Denis Dimitrov and Alexander Panchenko .....	26
<i>Cost-Effective Language Driven Image Editing with LX-DRIM</i> Rodrigo Santos, António Branco and João Ricardo Silva .....	31
<i>Shapes of Emotions: Multimodal Emotion Recognition in Conversations via Emotion Shifts</i> Keshav Bansal, Harsh Agarwal, Abhinav Joshi and Ashutosh Modi .....	44



## Conference Program

*On the Effects of Video Grounding on Language Models*

Ehsan Doostmohammadi and Marco Kuhlmann

*Rethinking Task Sampling for Few-shot Vision-Language Transfer Learning*

Zhenhailong Wang, Hang Yu, Manling Li, Han Zhao and Heng Ji

*Analyzing BERT Cross-lingual Transfer Capabilities in Continual Sequence Labeling*

Juan Manuel Coria, Mathilde Veron, Sahar Ghannay, Guillaume Bernard, Hervé Bredin, Olivier Galibert and Sophie Rosset

*Pixel-Level BPE for Auto-Regressive Image Generation*

Anton Razzhigaev, Anton Voronov, Andrey Kaznacheev, Andrey Kuznetsov, Denis Dimitrov and Alexander Panchenko

*Cost-Effective Language Driven Image Editing with LX-DRIM*

Rodrigo Santos, António Branco and João Ricardo Silva

*Shapes of Emotions: Multimodal Emotion Recognition in Conversations via Emotion Shifts*

Keshav Bansal, Harsh Agarwal, Abhinav Joshi and Ashutosh Modi

