



Advanced Language Processing Winter School — ALPS, 16-20
January 2023

[Home](#) [Speakers](#) [Organizers](#) [Application](#) [Schedule](#) [Sponsors](#)

Schedule

The speakers will provide pre-recorded lectures. Free slots allow you to watch those lectures before the Q&A sessions. The schedule will be provided as an ics calendar soon on this page.

CET	Monday 16/01	Tuesday 17/01	Wednesday 18/01	Thursday 19/01	Friday 20/01
8-9					
9-10					
10-11					
11-12	Gather town Poster Session 1		Slack Lab Session Neural Machine Translation		Gather Town Social Session 2
12-13					
13-14					
14-15	Zoom Q&A François Yvon	Slack Lab Session Prompt Engineering			Zoom Q&A Dirk Hovy
15-16				Gather town Poster Session 2	
16-17					
17-18	Zoom Q&A Kyunghyun Cho	Zoom Q&A Michael Auli	Zoom Q&A Collin Raffel		
18-19		Gather Town Social Session 1		Zoom Q&A Yejin Choi	

Social Sessions

Social Session 1

- **Titouan Parcollet** – Speech and open source toolkits
- **Maha Elbayad** – Large Scale Multilingual Machine Translation

- **Sara Hooker** – Tips and tricks for living your best life as a researcher (from attending conferences to producing research you are proud of)
- **Jesse Dodge** – Efficient and Reproducible NLP
- **Vlad Niculae** – Structure prediction
- **Ange Richard** – Applying NLP to social science

Social Session 2

- **William Havard** – Handling an interdisciplinary PhD topic
- **Shachar Mirkin** – Being a Remote Scientist
- **Diane Larlus** – Text and image multimodal models
- **Sebastian Ruder** – Collaboration in research
- **Fabien Ringeval** – Academic careers in France
- **Wilker Aziz** – Text generation (decoding) under uncertainty

Poster Sessions

Session 1

- A 1 **Seongmin Mun**

How do Transformer-Architecture Models Address Polysemy of Korean Adverbial Postpositions?

▼ [Abstract]

Several studies have used contextualized word-embedding models to reveal the functions of Korean postpositions. To add more interpretation, we devised a classification model by employing BERT and GPT-2 and introduces a computational simulation that interactively demonstrates how these transformer-architecture models simulate human interpretation of Korean adverbial postpositions -ey, -eyse, and -(u)lo. Results reveal that (i) the BERT model performs better than the GPT-2 model to classify the intended function of postpositions, (ii) there is an inverse relationship

between the classification performance and the number of functions that each postposition manifests, (iii) the models' performance gradually improves as the epoch proceeds, and (iv) the models are affected by the scarcity of input and/or semantic closeness between the items.

- A 2 **Kanishka Silva**

Self Attention Generative Adversarial Network based Authorship Attribution in Historical Texts

▶ [Abstract]

- A 3 **Thomas Stroehle**

How can pre-trained language models support idea evaluation?

▶ [Abstract]

- A 4 **Elise Lincker**

Multi-modal data extraction and enrichment: textbooks as a use case

▶ [Abstract]

- A 5 **Romain Meunier**

NLP for crisis management

▶ [Abstract]

- A 6 **Alkis Koudounas**

Transformer-based Non-Verbal Emotion Recognition

▶ [Abstract]

- A 7 **Aleksey Dorkin**

Comparison of Current Approaches to Lemmatization: A Case Study in Estonian

▶ [Abstract]

- A 8 **Till Überrück-Fries**

Linguse – Identifying Multi-Word Expressions for Language Learners

▶ [Abstract]

- A 9 **Emil Kalbaliyev**

Narrative Why-Question Answering: A Review of Challenges and Datasets

▶ [Abstract]

- **A 10 Deepak Kumar**
Parameter-efficient On-demand Bias Mitigation via AdapterFusion
▶ [Abstract]
- **A 11 Ronald Cardenas**
Cognitive Structures of Content for Controlled Summarization of Long Documents
▶ [Abstract]
- **A 12 Harshita Diddee**
Unsupervised Estimation of Quality of Data
▶ [Abstract]
- **A 13 Tanise Ceron**
Optimizing text representations to capture (dis)similarity between political parties
▶ [Abstract]
- **A 14 Amit Sah**
ADA: An Attention-Based Data Augmentation Approach to Handle Imbalanced Textual Datasets
▶ [Abstract]
- **A 15 Sumanth Doddapaneni**
Samanantar: The Largest Publicly Available Parallel Corpora Collection for 11 Indic Languages
▶ [Abstract]

Session 2

- **B 1 Florian Mai**
HyperMixer: An MLP-based Low Cost Alternative to Transformers
▶ [Abstract]
- **B 2 Shu Okabe**
Towards Automatic Gloss Generation for Computational Language Documentation

- ▶ [Abstract]
- **B 3 Febe de Wet**
Voices of Mzansi: Localising the Mozilla Common Voice platform for South Africa's official languages
▶ [Abstract]
- **B 4 Atilla Kaan Alkan**
TDAC: the First Time-Domain Astrophysics Corpus
▶ [Abstract]
- **B 5 Lila Kim**
Automatic classification of nasal vowels for a characterization of speakers' voice quality by convolutional neural network
▶ [Abstract]
- **B 6 Carolina Biliotti**
Breaking Down the Lockdown: The Causal Effect of Stay-At-Home Mandates on Uncertainty and Sentiments during the COVID-19 pandemic
▶ [Abstract]
- **B 7 Anna Aksenova**
Challenges of Information Extraction from Clinical Texts
▶ [Abstract]
- **B 8 Léa-Marie Lam-Yee-Mui**
Multilingual features for speech recognition for low-resource languages
▶ [Abstract]
- **B 9 Xu Yizhou**
Prompt Engineering-Based Text Anomaly Detection
▶ [Abstract]
- **B 10 Riccardo Pozzi**
Evaluation of Incremental Entity Extraction with Background Knowledge and Entity Linking
▶ [Abstract]
- **B 11 Nalin Kumar**
Exploring joint approaches to RDF triple parsing

- ▶ [Abstract]
- **B 12 Rakia Saidi**
BERT model for Arabic multi-tasks
▶ [Abstract]
- **B 13 Fabio Fehr**
A VAE for Transformers with Nonparametric Variational Information Bottleneck
▶ [Abstract]
- **B 14 Mukund Rungta**
Geographic Citation Gaps in NLP Research
▶ [Abstract]

[Advanced Language Processing Winter School](#), Proudly powered by [WordPress](#).