

# Lukas Muttenthaler

Email: [lukas-mt@live.at](mailto:lukas-mt@live.at) | Website: [lukasmut.github.io](http://lukasmut.github.io) | LinkedIn: [LukasMut](#) | Github: [LukasMut](#)

## EDUCATION

### University of Copenhagen (UCPH)

Copenhagen, DK

Master of Science in *IT and Cognition* (NLP specialisation); GPA: 11.33 / 12.0 (First Class equivalent)

Expected June 2020

**Thesis:** Subjective Question Answering (QA) - Supervisors: [Isabelle Augenstein](#) and [Johannes Bjerva](#)

**Relevant Coursework:** Advanced Topics in Natural Language Processing, Cognitive Science I, II & III, Data Mining, Introduction to Data Science (IDS), Language Processing I & II, Scientific Programming, Vision and Image Processing, Web Science

### University of Vienna (UV)

Vienna, AT

Bachelor of Science in Psychology; GPA: 1.5 / 5.0 (among top 5% of all students) (First Class equivalent)

October 2015 - June 2018

**Entrance Exam:** Performed among the top 2% of all participants

**Thesis I (theoretical):** [Effective enhancement of attentional functions in the amblyopic brain](#) (Grade: 1.0)

**Thesis II (empirical):** How language shapes motion event cognition (as part of [iCLAP](#)) (Grade: 1.0)

**Relevant Coursework:** Cognitive Psychology, Experimental Psychology, Psycholinguistics, Neuroscience, Statistical Methods

### University of Vienna (UV)

Vienna, AT

Bachelor of Arts in Political Science & Philosophy; GPA: 1.7 / 5.0 (First Class equivalent)

March 2012 - January 2015

**Thesis:** Policy networks in a postmodern, interdependent world (Grade: 1.0)

**Relevant Coursework:** Political Theory, Philosophy, International Politics, Economics, Quantitative Methods

## SKILLS

- **Natural Languages:** Spoken and written fluency in English (*IELTS academic 8.0*) and German (*native*)
- **Programming Languages:** Python (*advanced*), MATLAB, R, SQL
- **Technologies:** GitHub, GitLab, GitBash, LaTeX, Overleaf, Jupyter, Anaconda, Postman, Slurm
- **Libraries:** PyTorch, TensorFlow, Keras, Numpy, Scikit-Learn, Pandas, spaCy, NLTK
- **Neuroscientific Research Methods:** EEG (*studies conducted*), Eye-Tracking (*studies conducted*), fMRI (*only theory*)

## RELEVANT WORK EXPERIENCE

### University of Copenhagen (Department of Computer Science)

Copenhagen, DK

Student Research Assistant (*Natural Language Processing & Business Process Mining*)

May 2019 - Present

- **EcoKnow & DCR Graphs:** Research Assistant (RA) in the [EcoKnow](#) project and Research Programmer for [SDPS](#). Currently developing a tool (in collaboration with [DCR Solutions](#) and the ITU Copenhagen) to extract activities, roles and relations in business process documents using Python libraries for NLP such as spaCy and NLTK
- **Question Answering (QA):** Master's thesis student and RA at [copeNLU](#). Researching QA, Machine Translation and Sentiment Analysis and developing QA and Text2SQL algorithms using NLP and Deep Learning frameworks (e.g., PyTorch, Keras)

### Ernst & Young (EY)

Copenhagen, DK

Data Analyst

January 2019 - April 2019

- **Technology Advisory:** Assistant in the Data Analytics team. Worked in the domains of Business Intelligence, Data Analytics and ML. Supported the enhancement of utilization forecasting

### University of Vienna (Faculty of Psychology)

Vienna, AT

Student Research Assistant & Teaching Assistant (*Cognitive Psychology - Visual Perception lab*)

September 2016 - July 2018

- **iCLAP (Cognition, Language and Perception):** Investigated through cognitive experiments how language shapes perception and cognition. Compared motion event cognition in native German and native Korean speakers (collaboration with Konkuk University in Korea and funded by the Vienna Science and Technology Fund)
- **Drivers' Attention:** Assisted in the examination of light-dynamic effects induced by modern adaptive car lighting systems on visual attention and perception of drivers (collaboration with [ZKW group](#))
- **Perceptual Load:** Programmed a unique, one trial experiment and tested visual attention in over 100 subjects using Eye-Tracking (lead to a publication in [Frontiers](#))

### LMU Munich (Department of Neuropsychology)

Munich, GER

Erasmus+ Research Internship (*Research Intern*)

July 2017 - September 2017

- **Multiple Object Tracking:** Researched literature on Multiple Object Tracking (MOT), fMRI and EEG; recruited test subjects for the investigation of spatial visual attention; collected and analysed data

## RECENT PROJECTS

---

- **Relation extraction with ET & EEG:** Comparison between different reading tasks and their effects on Eye-Tracking and EEG data
  - Developed a data loader object for [ZuCo](#) that automatically transforms and analyses Eye-Tracking (ET) data into a form that is ready to be used for statistical analysis and data visualisation
  - Compared Normal Reading (NR) with Annotation Reading (AR), and their respective effects on ET and EEG features
  - Planning to augment neural networks with the previously analysed ET and EEG data, and investigate how this affects performance on NLP tasks such as Named Entity Recognition and Relation extraction
  - Supervisor: [Maria Barrett](#), Collaborator: [Nora Hollenstein](#)
- **Text-to-SQL:** Currently working with [SQLNet](#) and [TypeSQL](#) as part of Master's thesis project on subjective QA
  - Modified [TypeSQL's source code](#) to run on Python 3 versions
  - Developed a retokenizer to rejoin BERT byte-pair encoded tokens into WikiSQL tokens to plug max-pooled BERT context embeddings into TypeSQL. Created an ensemble network based on TypeSQL with BERT models (*Current stage: Ablation studies*)
  - Further augmented TypeSQL's encoder with *Part-of-Speech (POS)* embeddings (*Current stage: Ablation studies*)
  - Supervisors: [Isabelle Augenstein](#) and [Johannes Bjerva](#)
- **Cross-Domain Authorship Attribution:** Authorship attribution in fan-fictional texts given variable length character and word  $n$ -grams
  - Implemented a soft-voting ensemble classifier based on Support Vector Machines (SVMs) and  $n$ -gram features at different levels
  - Achieved a macro  $F1$ -score of 70.5% for the development and 69% for the test corpus of [PAN's 2019 shared task](#) (*Baseline: 54.5%*)
  - Algorithm was the [winning approach](#) at PAN's shared task for [CLEF 2019](#) and paper received two "strong accepts" (*Corresponding author*)
  - Supervisor: [Manex Agirrezabal](#)
- **Emoji Prediction:** Prediction of emojis in Twitter tweets given word  $n$ -grams
  - Implemented a tf-idf weighted word  $n$ -gram model, used as features for a feed-forward neural network, to achieve an  $F1$ -score of 45% on the UvA [Twemoji](#) dataset (*Random: ~ 10%, Baseline: ~ 20%*)
  - Predictions were made for the 10 most frequently used emojis in the corpus
  - Implemented our own Twitter tweet cleaning algorithm: [TweetCleaner](#)
  - Supervisor: [Patrizia Paggio](#)

## SCHOLARSHIPS & ACHIEVEMENTS

---

- [Fund of Excellence](#) - one of a few candidates who received funding to pursue a Master's in a STEM subject (**2018**)
- Scholarship for Academic Excellence awarded by UV for three consecutive years for excellent grades in Psychology (**2016, 2017, 2018**)
- Scholarship awarded by Julius-Raab-Foundation\* for three consecutive years due to high achievements in Psychology (**2016, 2017, 2018**) (*\*a performance scholarship that is awarded to students from low-income families such as single parent households*)
- Scholarship for Academic Excellence awarded by UV for three consecutive years for excellent grades in Political Science (**2013, 2014, 2015**)

## PUBLICATIONS

---

- Hugo A. López, Morten Marquard, Lukas Muttenthaler, and Rasmus Strømsted. 2019. Assisting Declarative Process Generation from Natural Language Descriptions. In 2019 IEEE 23rd International Enterprise Distributed Object Computing Conference (EDOC), IEEE, 2019.
- Lukas Muttenthaler, Gordon Lucas, and Janek Amann. Authorship Attribution in Fan-fictional Texts Given Variable Length Character and Word  $n$ -grams - Notebook for PAN at CLEF 2019. In Linda Cappellato, Nicola Ferro, David E. Losada and Henning Mueller, editors, [CLEF 2019 Labs and Workshops, Notebook Papers, September 2019](#). [CEUR-WS.org](#).
- Buesel, C., Ditye, T., Muttenthaler, L., Ansorge, U. (2019). A Novel Test of Pure Irrelevance-Induced Blindness. *Frontiers in Psychology*, 10:375. DOI: [10.3389/fpsyg.2019.00375](#)
- Muttenthaler, L. (2019). Effective enhancement of attentional functions in the amblyopic brain. *Journal of European Psychology Students*, 10(1), 1-10. DOI: [10.5334/jeps.435](#)

## OTHER ACTIVITIES

---

- Created a PyTorch tutorial about ResNets on GitHub: [ResNets with PyTorch](#) (2019)
- Volunteering for the NGO Green Kayak to collect litter in Copenhagen's waterways and the North Sea (2019)
- Community service for nine months in an institution for mentally ill refugees (Feb 2015 - Oct 2015)
- Worked as a journalist at a renowned Austrian newspaper on home affairs and economics alongside Political Science studies (2014)
- Professional Track & Field athlete for 5 years (2009 - 2014) in high school and alongside university for the Viennese core team with medals in 4x100m relay, 200m and 400m sprint