

Peter Schoener

EDUCATION

- 2018 – 2020
(expected) **University of Washington, Seattle**
Computational Linguistics (MS)
- 2015 – 2018
(expected) **Eberhard Karls Universität Tübingen**
Computational Linguistics (BA) (1.67 average* as of 10/2017)
Member of student association (Fachschaft), 2018
- 2008 – 2015 **German International School of Silicon Valley**
California High School Diploma (4.04 weighted GPA)
German International Abitur (1.5 average*)

EXPERIENCE

- November 2017 –
March 2018 **Teaching Assistant | Universität Tübingen | Tübingen, Baden-
Württemberg**
I was a teaching assistant for the Programming Course for the
Computational Linguistics/Java III course taught by Daniël de Kok.
- June 2017 –
April 2018 **Intern | Kamusi GOLD | Tübingen, Baden-Württemberg**
I led conceptual development and implementation of an abstractive
model to be used in the Global Online Living Dictionary machine
translation project.
- November 2016 –
May 2017 **Research Assistant | Universität Tübingen | Tübingen, Baden-
Württemberg**
I built LTI-compliant web applications (Java/GWT) for cognitive
assessment as part of a language learning study in the linguistics
department.
- June 2015 –
August 2015 **Independent Contractor | Apple via WeLocalize | Cupertino, CA**
I tagged English and German texts for a document classification and
data structuring project at Apple, Inc.
- April 2013 **Intern | SAP | Palo Alto, CA**
I worked as an intern for Dr. Butch Anton, Director of Customer
Engagement and Special Projects.

* Standard German 1–6 scale; 1 is best, 4 is lowest passing grade

RESEARCH

B.A. Thesis: A Neural Approach to Semantic Compatibility of Nouns and Adjectives on the Basis of Word Embeddings (Advisor: Dr. Daniël de Kok, 2018)

Original work on predicting semantic compatibility of arbitrary noun-adjective pairs for the SFB A3 embedding composition project at the University of Tübingen as well as the dependency parsing project. Uses a neural approach to reliably predict the semantic compatibility of a noun-adjective pair.

Identification of Semantic Shifts in English Using Word Embeddings (Course: Unsupervised Learning, Dr. Çağrı Çöltekin, 2018)

Reimplementation of (Kutuzov and Kuzmenko, 2018) with some adjustment as per (Leeuwenberg et al., 2016) and reapplication to shifts over time rather than domain, with qualified success.

SELECTED COURSEWORK

Efficient Linear Algebra for Machine Learning and Computational Linguistics (Dr. Daniël de Kok, summer 2018)

Rust course with a focus on efficient data structures and algorithms for linear algebra with a strong focus on parallelization, memory management, low-level underlying representations, the particular strengths of Rust's zero-cost abstractions, and practical use of libraries such as ndarray and Tensorflow.

Unsupervised Learning for Computational Linguistics (Dr. Çağrı Çöltekin, winter 2017)

Builds on previous SNLP course, discussing bleeding-edge methods of clustering, classification, distance measurement, embedding, and comparison for topic and sentiment classification, linguistic taxonomy, synonym detection, depression detection, and summarization

Artificial Intelligence (Prof. Dr. Andreas Zell, winter 2017)

Theory of graph search, reasoning, uncertain decision-making, heuristics, optimization tasks, and genetic algorithms, implementations in Lisp and Java

Deep Learning for Natural Language Processing (Dr. Daniël de Kok, summer 2017)

Basics of machine learning (regression, neural networks, feed-forward and recurrent adaptations), current use within the scope of NLP (existing implementations for classification, sentiment analysis, abstraction, etc.), introduction to Tensorflow, Twitter sentiment analysis project

Statistical Natural Language Processing (Dr. Çağrı Çöltekin, summer 2017)

Basics of SNLP (ngram/Markov models, naïve Bayes, basics of regression and neural networks, clustering), implementation of methods for authorship attribution, document classification, POS tagging, parsing, and sentiment analysis as exercises

Computational Analysis of Linguistic Complexity (Prof. Dr. Detmar Meurers, winter 2016)

Traditional complexity indices, current indices and use of ML models, esp. for second language acquisition, problem generation, and teaching aids

TECHNICAL SKILLS

Languages

Rust	Tensorflow linear algebra and other parallelizable operations
C/C++	Windows and POSIX APIs strict C89/C99 compliance
Java (7, 8)	Apache OpenNLP Google Web Toolkit
Python	NLTK Tensorflow and Scikit

Some exposure to Lisp, Prolog, R

Methods

- ❖ Regression, shallow and deep NNs
- ❖ Clustering
- ❖ Ngram and Markov models
- ❖ Vector space embeddings
- ❖ Efficient linear algebra application and implementation
- ❖ String comparison metrics
- ❖ Fuzzy search
- ❖ Rule-based and statistical parsing and meaning extraction

OTHER QUALIFICATIONS

Languages

- ❖ English (*native*)
- ❖ German (*functionally native*)
- ❖ French (*intermediate*)
- ❖ Hindi (*beginner*)

Other

- ❖ Experience with academic writing
- ❖ Can teach individuals quickly and effectively (see T.A. experience)
- ❖ Strong understanding of linear algebra, statistics, and mathematical logic
- ❖ Authorized to work in the United States (U.S. citizen)

CONTACT INFORMATION

Location

Tübingen, Germany

Website

peter.schoener.net

Phone

+49 1520 2815782

Telegram

[@pschoener](https://t.me/pschoener)

Email

peter.schoener@student.uni-tuebingen.de

GitHub

github.com/peterr-s