



# Shibingfeng Zhang

## WORK EXPERIENCE

[ 01/03/2024 – Current ]

### Research Fellow

*Department of Classical Philology and Italian Studies, Alma Mater Studiorum - Università di Bologna*

**City:** Bologna

**Country:** Italy

Development and testing of a pipeline for Digital Linguistic Biomarker extraction. Research activities carried out within the ReMind project - an ecological, cost-effective AI platform for early detection of prodromal stages of cognitive impairment.

## EDUCATION AND TRAINING

[ 30/09/2021 – 13/11/2023 ]

### Master of Science in Language Science and Technology

*Department of Language Science and Technology, Saarland University*

**City:** Saarbrücken

**Country:** Germany

**Thesis:** Language-driven Action Recognition and Localization

Main subjects covered:

- Sign Language Recognition and Translation
- Machine Translation
- Machine Learning for NLP
- Computational Semantics
- Data Science
- Statistics with R

[ 09/2019 – 13/07/2021 ]

### Master of Arts in Specialized Translation with a special focus on Computational Linguistics

*Department of Interpreting and Translation, University of Bologna*

**City:** Forlì

**Country:** Italy

**Thesis:** Emotion Identification in Italian Opera

Main subjects covered:

- Computational Linguistics
- Terminology
- Translation Technologies and Computer-Assisted Translation

[ 08/2014 – 07/2018 ]

### Bachelor of Arts in Italian Language

*Sichuan International Studies University*

**City:** Chongqing

**Country:** China

## PUBLICATIONS

[ 2023 ]

### [GPL at SemEval-2023 Task 1: WordNet and CLIP to Disambiguate Images](#)

This is a Workshop paper for task [Visual Word Sense Disambiguation](#) at the 17th International Workshop on Semantic Evaluation (SemEval-2023).

Given a word in context, the task of Visual Word Sense Disambiguation consists of selecting the correct image among a set of candidates. We used different text augmentation strategies and pre-trained language-vision models like [CLIP](#) and [ViLT](#) to address this task in a zero-shot manner.

[ 2022 ] [Unsupervised Cantonese-Mandarin Machine Translation](#)

This is a school work of software project. It's not a publication, but I'm super proud of it. Our [GitHub repository](#) of this project got 19 stars!

In this study, we investigated unsupervised machine translation between Mandarin Chinese and Cantonese. Our research involved the development of a Cantonese-Mandarin machine translation system, trained solely on Mandarin and Cantonese monolingual corpora obtained from various websites. We utilized RNN-based and transformer models and explored various methods for learning cross-lingual embeddings and tokenization techniques to improve system performance.

[ 2022 ] [AriEmozione 2.0: Identifying Emotions in Opera Verses and Arias](#)

Paper published on [Italian Journal of Computational Linguistics](#) vol. 8, n. 2 December 2022

We present the task of identifying the emotions conveyed by the lyrics of Italian opera arias. We produced an embedding space for 17th century Italian language and leveraged such representation to investigate the emotion transmitted by Italian opera at different granularities.

[ 2021 ] [AriEmozione: Identifying Emotions in Opera Verses](#)

Short paper published on proceedings of the Seventh [Italian Conference on Computational Linguistics](#).

In this study, we seek to identify the emotions transmitted in Italian opera arias at the verse level. This is a relevant problem for the organization of the vast repertoire of Italian Opera arias available and to enable further analyses by both musicologists and the lay public. We conducted experiments using various feature extraction methods and classification models to investigate the most effective approach.

## DIGITAL SKILLS

---

Python | R | Natural Languages Processing | Deep Learning | Machine Learning | PyTorch Lightning | PyTorch | Academic Writing | Version Controlling: Git, GitHub | Scikit-Learn | Linux (Terminal Commands, Bash/Shell) | Statistical Analysis

## LANGUAGE SKILLS

---

**Mother tongue(s):** Chinese

**Other language(s):**

**English**

**LISTENING C2 READING C2 WRITING C1**

**SPOKEN PRODUCTION C1 SPOKEN INTERACTION C2**

**Italian**

**LISTENING C2 READING C2 WRITING C1**

**SPOKEN PRODUCTION C1 SPOKEN INTERACTION C2**

*Levels: A1 and A2: Basic user; B1 and B2: Independent user; C1 and C2: Proficient user*

## HONOURS AND AWARDS

---

[ 2022 ] **Eramus+ Scholarship Awarding institution:** Saarland University

Scholarship for the exchange period abroad

- [ 2020 ] **Unibo Action 2 scholarship Awarding institution:** University of Bologna  
Merit-based scholarship offered by the University of Bologna.
- [ 2019 ] **Unibo Action 2 scholarship Awarding institution:** University of Bologna  
Merit-based scholarship offered by the University of Bologna.
- [ 2017 ] **State Scholarship Fund for Visiting study Awarding institution:** China  
Scholarship Council  
Scholarship awarded under the State Scholarship Fund of China to pursue study in Italy  
as a visiting student