

# ARUM PARK

luvalska1231@gmail.com | +49 (0)176 5576 8409

---

## EXPERIENCE & SKILL

- Over 5 years of experience in research projects optimizing Korean-to-English Machine Translation (MT) systems (see **RESEARCH PROJECTS** on page 4)
- Implementing a Korean null subject resolution system using heuristic rules and Machine Learning (ML) ([Rule-based null subject resolution system](#)) – Jupyter notebook & Python  
([ML for the classification problem between anaphoric and non-anaphoric](#)) – scikit-learn
- Data preparation and analysis of MT results for Multilingual Neural Machine Translation (NMT) models as a Machine Translation Specialist Trainee at Electronic Arts in Köln (see **PROFESSIONAL EXPERIENCE** on page 2)
  - Data cleaning with Regex
- NLP challenge: JSGF development ([JSGF development for English and Korean](#))
- Sharing knowledge about ML ([posting example](#))  
([posting example code](#)) – Markdown & Version controll with Git
- Sharing knowledge about the Korean language on YouTube ([Learn authentic Korean for 10 min.](#))
  - Using a self-built Korean transcriber ([Korean Transcriber based on McCune-Reischauer Romanization System](#))
  - Using Text-to-Speech AI
- Building annotated data
  - Korean anaphora with MMAX2 Tool
  - Usages of Pronouns in SNS corpus with Python & Excel
  - Usages of Emoticons in SNS corpus with Python & Excel
- Data preprocessing
  - converting json file to csv file ([code example](#))
  - generating csv file with conditions ([code example](#))
- Command Line & grep & sed
- SQL excercise for data analysis

## LANGUAGE SKILLS

- Korean (native)
- English (C1)
- German (C1)

## PROFESSIONAL EXPERIENCE

31.05.2022 – 04.11.2022

**Machine Translation Specialist Trainee** *Localization Team at Electronic Arts (EA), Köln*

- Goal: Customizing language models for NMT
- Responsibilities:
  - Supported the Machine Translation Team with executing EA's MT strategy
  - Analyzed MT output for quality improvements
  - Cleaned up TMs and fed glossaries into MT engines
  - Tested and documented the results of output of trained MT engines
  - Reviewed human post-edited output to produce high-quality training corpus for MT engines
  - Assisted with collecting feedback from vendors, preparing data for training engines, normalizing glossaries for automatic scoring and with conducting general research on technology

2019 – 2021

**Korean Teacher**

*Koreanische Schule, Düsseldorf*

- Role:

Taught German native speakers and developed a customized curriculum to explain specific Korean language phenomena that students struggled with due to language differences

## EDUCATION

2016.10 – 2024.07 (expected)

### Doctoral Candidate in Linguistics (Linguistic Data Science Lab)

Ruhr Universität Bochum, GERMANY

- Doctoral Thesis (submitted on May 29, 2024)
- „Nullsubjektresolutionssystem im Koreanischen: Eine computer- und korpusbasierte Untersuchung der Parameter zur Wiederauffindbarkeit von Nullsubjekten“
- Implementing heuristic rules based on morpho-syntactic clues in Korean to assign the referents of the Korean null subjects and integrating them as a null subject resolution system
- All prerequisite courses - bestanden
- WS 16/17 Seminar: *CL-Propädeutikum*
- WS 16/17 Proseminar: *Forschungsseminar: Computational Pragmatics*
- WS 16/17 Hauptseminar: *Probabilistische Pragmatik*
- WS 16/17 Hauptseminar: *Syntax in der generativen Grammatik*
- WS 19/20 Proseminar: *Tools für den linguistischen Alltag: Automatisierung, Reproduzierbarkeit und Kollaboration*
- DFG Project from 2023: „Praktiken der Personenreferenz“
- DFG Project in 2019: „Elektronische Kurznachrichtenkommunikation im Deutschen und Koreanischen - Konvergenzen und Divergenzen sprachlicher, stilistischer und interaktionaler Strukturen“
- Scholarship from DAAD in 2019
- A Third-party funded project: „PAJAKO: Partnerschaft mit Japan und Korea“

2013 – 2016

### Doctoral Course Completion in German Linguistics

Institute of German Language and Literature, Sungkyunkwan University, KOREA

- Cumulative GPA: 4.5/4.5
- Global PhD Fellowship for two years
- ESSLLI 2016 (European Summer School in Logic, Language and Information) with scholarship in Bolzano-Bozen, Italy
- ESSLLI 2015 in Barcelona, Spain

2010 – 2013

**M.A degree in German Linguistics**

Institute of German Language and Literature, Sungkyunkwan University, KOREA

- Master Thesis
  - „Untersuchung der mehrdeutig adnominalen Konstruktion bei der Übersetzung von technischen Dokumentationen von Smartphones: in Hinblick auf kontrollierte Sprache“
  - Proposing Korean controlled language to disambiguate the ambiguous adnominal construction found in smartphone manuals
- Cumulative GPA: 4.44/4.5

2006 – 2010

**B.A Degree in German Linguistics and Literature**

Institute of German Language and Literature, Sungkyunkwan University, Korea

- Cumulative GPA: 4.20/4.5
- Exchange Student Program for WS 07/08 at Heidelberg University, Germany
- Global Leader Fellowship in 2006, Korea

## RESEARCH PROJECTS

- |   |  |
|---|--|
| 06.2012 – 02.2017   | Electronics and Telecommunications Research Institute (ETRI) |
| <b>Developing Automatic Speech Translation System for Tourism/International Event</b>                       |  |
| 10.2014 – 12.2014   | SK Telecom   |
| <b>Establishing POS Tag Guidelines for Korean Spoken Corpus</b>   |  |
| 06.2011 – 01.2012   | ETRI   |
| <b>Studies on the Manual/Automatic Post-Editing methods for the Korean-English MT systems</b>               |  |
| 06.2010 – 01.2011   | ETRI   |
| <b>Studies on the Linguistic Properties of Korean Spoken Language for Korean-English Spoken Language MT</b> |  |
| 05.2009 – 01.2010   | ETRI   |
| <b>Context-based Disambiguation method for Connective Endings in Korean-English MT</b>                      |  |
| 05.2006 – 11.2007   | ETRI   |
| <b>Controlled Korean for Korean-English Academic Paper Machine Translation</b>                              |  |

## PUBLICATION

- Park Arum & Hong Munpyo (2014)

**Hybrid approach to zero subject resolution for multilingual MT: Spanish-to-Korean cases.** In proceedings of the 28th Pacific Asia conference on language, information and computation, 254–261. ([paper](#))

- Park Arum & Hong Munpyo (2015a)

**Subjektellipse im Koreanischen und deren Behandlung für die maschinelle Übersetzung ins deutsche.** Koreanische Gesellschaft für Germanistik 56(1). 197–223. ([paper](#))

- Park Arum & Hong Munpyo (2015b)

**Objektellipse im Koreanischen und deren Behandlung für die maschinelle Übersetzung ins Deutsche.** Koreanische Gesellschaft für deutsche Sprache und Literatur 70. 1–21. ([paper](#))

- Park Arum, Lim Seunghee, & Hong Munpyo (2015)

**Zero object resolution in Korean.** In proceedings of the 29th Pacific Asia conference on language, information and computation, 439–448. ([paper](#))