LaTex template of Université Paris Nanterre

par Seongmin Mun

Thèse présentée et soutenue publiquement le 18 juin 2021

en vue de l'obtention du grade de

docteur en Traitement Automatique des Langues

sous la direction de Guillaume Desagulier

Membres du jury:

Directeur: Prénom et nom	Université & UMR 7114, MoDyCo
Rapporteur: Prénom et nom	Université
Rapporteur: Prénom et nom	Université
Examinatrice: Prénom et nom	Université
Examinatrice: Prénom et nom	Université & UMR 7114, MoDyCo
Examinatrice: Prénom et nom	Université & UMR 7114, MoDyCo

Acknowledgements

There are many who helped me along the way on this journey.

Abstract

This dissertation reports computational accounts of resolving word-level polysemy in a lesser-studied language—Korean.

Keywords: polysemy, natural language processing, classification, word embedding models, data visualization, Korean

Contents

List of Tables

List of Figures

List of abbreviations

I follow the Leipzig glossing rules[1].

Abbreviation	Label
ACC	Accusative
AGT	Agent

---

[1]Available at: https://www.eva.mpg.de/lingua/pdf/Glossing-Rules.pdf

Chapter 1

Introduction

The project presented in this dissertation aims to address the possible ways and limitations in applying computational approaches to word-level polysemy in a lesser-studied language, Korean.

1.1 Background of beginning this project

I assume that a relationship of words (represented as probabilistic information) is one core construct in understanding how language works.

Appendix A

Code for the word-level embedding models

The following scripts are the code that I used for the training of *traditional word embedding models* (i.e., PPMI-SVD, SGNS) and *similarity-based estimation*.

Listing A.1: Python code for the word embedding by using the PPMI-SVD model

```
1
2 class PPMI_SVD_Algorithm:
3
4     def __init__ (self, fold, postposition, postposition_ko,
            window):
5         self.fold = fold
6         self.postposition = postposition
7         self.postposition_ko = postposition_ko
8         self.window = window
```

References