Contents

Title Page i
Abstract iii
Table of Contents iv
Citations to Previously Published Work vi
Acknowledgments vii
Dedication viii

1 Introduction **1**
1.1 Part-of-speech Tagger 1
1.2 Tag set 2
1.3 Formal definition 2
1.4 Evaluation 2
1.5 Challenges in POS tagger 2
1.6 Multilingual POS Tagger 3

2 Background and Literature Review **4**
2.1 Monolingual POS Tagger 4
2.1.1 Supervised 4
2.1.2 Unsupervised 12
2.1.3 Semi-supervised 15
2.2 Multilingual POS Tagger 19
2.2.1 Topologically related languages approach 19
2.2.2 Unsupervised simultaneously tagging approach 20
2.2.3 Tag projection approach 22

3 Universal Tagger **25**
3.1 Idea and Motivation 25
3.2 Universal Tagset 27
3.3 Methodology 29
3.3.1 Seed Model 29
3.3.2 Self training and revision 32
3.4 Experiment Detail 33
3.4.1 Source data 33

3.4.2 Word alignment 34
3.4.3 Direct tag projection 36
3.4.4 Seed model and final model 36
3.4.5 Evaluation 36
3.5 Experimental Results 37
3.6 Conclusion 39

4 Source Language Selection **41**

5 Conclusion and Future Work **42**

A Stuff that didn't belong in the thesis **47**