TABLE OF CONTENTS

APPROVAL PAGE ..... **ii**
SELF DECLARATION AGAINST PLAGIARISM ..... **iii**
ABSTRACT ..... **iv**
ABSTRAK ..... **v**
DEDICATION ..... **vi**
ACKNOWLEDGMENTS ..... **vii**
PREFACE ..... **viii**
TABLE OF CONTENTS ..... **ix**
LIST OF TABLES ..... **xi**
LIST OF FIGURES ..... **xii**
LIST OF TERMS ..... **xiii**
LIST OF NOTATIONS ..... **xiv**
CHAPTER 1 INTRODUCTION **1**
1.1 Rationale *1*
1.2 Theoretical Framework *3*
1.3 Conceptual Framework/Paradigm *4*
1.4 Statement of the Problem *6*
1.5 Objective and Hypotheses *6*
1.6 Assumption *7*
1.7 Scope and limitation *7*
1.8 Significance of the Study *8*
CHAPTER 2 REVIEW OF LITERATURE AND STUDIES **9**
2.1 Related Literatures *9*
2.1.1 Monolingual Alignment 9
2.1.2 Phrase-based approach for adaptive tokenization 10
2.1.3 Paraphrase Database (PPDB) 11
2.1.4 Word Similarity 11
2.1.5 POS Tagger 12
2.1.6 Syntax Parsing 13
2.2 Related Studies *15*
CHAPTER 3 RESEARCH METHODOLOGY **16**

3.1 Research Design *16*
3.1.1 System/product/method Implementation 16
3.1.2 Experiment Scenario 17
3.2 Population/Sampling *25*
3.3 Instrumentation and Data Collection *26*
3.4 Tools for Data Analysis *26*
CHAPTER 4 PRESENTATION, ANALYSIS AND INTERPRETATION OF DATA 28
4.1 Presentation of Data *28*
4.2 Analysis of the Data *32*
4.2.1 Correlation 32
4.2.2 Differences of Measures 33
4.2.3 Interpretation of Data 38
4.3 Summary of Findings *39*
CHAPTER 5 CONCLUSION AND RECOMMENDATIONS 40
5.1 Conclusions *40*
5.2 Recommendations *40*
BIBILOGRAPHY 41
APENDIX A Indonesian POS Tagset 43
APENDIX B Document Translation Surah Yasin 44
APENDIX C: Document of Gold Standard 48
APENDIX D Script Program 52
APENDIX E Treebank Surah Yasin (selected verses) 58
APENDIX F Alignment Result 65
APENDIX G Curriculum Vitae 73