

วัลย์วรา ไชยฤกษ์ : การพัฒนาโปรแกรมการถอดอักษรภาษาอังกฤษเป็นภาษาไทยโดยใช้คลังคำทับศัพท์ของราชบัณฑิตยสถาน. (THE DEVELOPMENT OF AN ENGLISH-THAI TRANSLITERATION PROGRAM BASED ON TRANSLITERATED-WORD CORPUS OF THE ROYAL INSTITUTE) อ. ที่ปรึกษา: ผู้ช่วยศาสตราจารย์ ดร.วิโรจน์ อรุณมานะกุล, 150หน้า. ISBN 974-17-6489-8.

169375

วิทยานิพนธ์ฉบับนี้มีวัตถุประสงค์เพื่อรวบรวมคำทับศัพท์ภาษาอังกฤษของราชบัณฑิตยสถาน และเปรียบเทียบกฎการทับศัพท์ของราชบัณฑิตยสถานกับข้อมูลจริงที่พบ พร้อมทั้งสร้างโปรแกรมการถอดอักษรภาษาอังกฤษเป็นภาษาไทยโดยใช้คลังคำทับศัพท์ดังกล่าวเพื่อสร้างตารางกฎการถอดอักษรภาษาอังกฤษเป็นภาษาไทยให้กับคอมพิวเตอร์

การวิจัยนี้ได้เก็บรวบรวมคำภาษาอังกฤษและคำทับศัพท์ภาษาไทยจำนวน 10,060 คำจากเอกสารที่เผยแพร่โดยราชบัณฑิตยสถานเพื่อนำมาใช้เป็นคลังคำทับศัพท์ ผลการวิจัยพบว่า กฎการถอดอักษรของราชบัณฑิตยสถานสอดคล้องกับข้อมูลในคลังทับศัพท์ดังนี้ กฎสำหรับสระมีความสอดคล้อง 92.68% สำหรับพยัญชนะต้นสอดคล้อง 97.35% สำหรับพยัญชนะท้ายสอดคล้อง 97.34% และเมื่อ ทดลองโปรแกรมที่พัฒนาขึ้นโดยใช้กฎที่สร้างขึ้นจากข้อมูล 80% ของคลังคำทับศัพท์ ผลการทดลองกับข้อมูลอีก 20% ปรากฏว่าโปรแกรมสามารถถอดอักษรในระดับคำได้ถูกต้อง 39.53% และถูกต้อง 79.8% ในระดับอักษร ซึ่งแสดงว่าประสิทธิภาพการถอดอักษรของโปรแกรมยังไม่ได้ดีเท่าที่ควร เนื่องจากสาเหตุหลายประการ เช่น โปรแกรมไม่ได้นำขอบเขตพยางค์มาพิจารณา ไม่ได้พิจารณาเสียงภาษาอังกฤษ รวมถึงความหลากหลายของคำทับศัพท์ของราชบัณฑิตยสถานที่รวบรวมไว้ในคลังข้อมูล และยังพบว่าในการถอดอักษรจริงนั้น มีการถอดอักษรแบบที่นอกเหนือไปจากที่ราชบัณฑิตยสถานให้ไว้ในตารางเทียบอักษรอยู่จำนวนหนึ่ง

ภาควิชา.....ภาษาศาสตร์.....ลายมือชื่อนิสิต.....
 สาขาวิชา.....ภาษาศาสตร์.....ลายมือชื่ออาจารย์ที่ปรึกษา.....
 ปีการศึกษา 2547

4480195322 : MAJOR LINGUISTICS

KEY WORD: TRANSLITERATION / CORPUS / ENGLISH / THAI / THE ROYAL INSTITUTE


WANWARA CHAIROEK : THE DEVELOPMENT OF AN ENGLISH-THAI
TRANSLITERATION PROGRAM BASED ON TRANSLITERATED-WORD CORPUS OF
THE ROYAL INSTITUTE. THESIS ADVISOR: WIROTE AROONMANAKUN, Ph.D., 150 pp.
ISBN 974-17-6489-8.

169375

The purpose of this thesis is to collect transliterated words published by the Royal Institute and compare the transliteration regulation of the Royal Institute with the collected data. Furthermore, an English-Thai transliteration program based on transliterated-word corpus of the Royal Institute was developed in order to generate tables of English-Thai transliteration rules for computers.

The transliterated-word corpus is composed of 10,060 pairs of English and transliterated Thai words published by the Royal Institute. The result shows that the transliteration regulation of the Royal institute conforms with the collected data in the corpus as follows: The transliteration rules for vowels have 92.68% accordance, the transliteration rules for initial consonants have 97.35% accordance and the transliteration rules for last consonants have 97.34% accordance. When the developed program was trained by the rules generated from 80% of the corpus, the result of testing on 20% remainder shows 39.53% accuracy in word level and 79.8% accuracy in alphabetical level. This suggests that the transliteration program yields quite a poor result due to considerable reasons, for instance syllable boundaries and phonetic transcriptions were not considered including various versions of transliterated words published by the Royal Institute collected in the corpus. It is also found that a number of transliteration rules are not yet included in the Royal Institute's table of transliteration.

Department.....Linguistics..... Student's signature..... 

Field of study...Linguistics..... Advisor's signature..... 

Academic year 2004