

จินตนา พลศรี : การพัฒนากรอบการทำงานในการแปลงคำปรัวรรตของอักษรไทยน้อยเป็นคำอ่านภาษาไทยถิ่นอีสาน (THE DEVELOPMENT OF A FRAMEWORK FOR TRANSLITERATING THAI NOI CHARACTERS INTO ISAN LANGUAGE PRONUNCIATION) อาจารย์ที่ปรึกษา : ผู้ช่วยศาสตราจารย์ ดร.นิสาชล จำนงศรี, รองศาสตราจารย์ ดร.จิรยุทธ ไชยจรรวม, 218 หน้า.

การวิจัยนี้ นำเสนอกรอบการทำงานในการแปลงคำปรัวรรตของอักษรไทยน้อยเป็นคำอ่านภาษาไทยถิ่นอีสาน วิจัยดำเนินการวิจัยแบ่งออกเป็น 3 ระยะ ได้แก่ ระยะที่ 1 ประกอบด้วยการศึกษาปัญหาในการปรัวรรตเอกสาร โบราณ การศึกษาโครงสร้างการบันทึกคำราชาธิปไตยโบราณที่บันทึกในใบลานด้วยอักษรไทยน้อยที่ส่งผลต่อประสิทธิภาพการปรัวรรต และลักษณะอักษร ภาษา และอักษรวิธอักษรไทยน้อยที่ใช้บันทึกคำราชาธิปไตยโบราณ วิจัยใช้วิธีวิจัยเชิงคุณภาพ รวบรวมข้อมูลโดยการสัมภาษณ์นักภาษาศาสตร์โบราณ จำนวน 10 คน ด้วยแบบสัมภาษณ์แบบกึ่งโครงสร้าง และการวิเคราะห์คำราชาธิปไตยโบราณที่ปรัวรรตเอาไว้แล้วพร้อมภาพคำราชาธิปไตยฉบับ จำนวน 11 ผูก 459 หน้าลาน ผลการศึกษาพบว่า ปัญหาการปรัวรรตเอกสาร โบราณและคำราชาธิปไตยโบราณมีลักษณะคล้ายคลึงกัน และที่เป็นปัญหาอยู่ในระดับมากมีสาเหตุจากลักษณะทางกายภาพของเอกสารโบราณ รูปอักษร ภาษา และอักษรวิธของอักษรไทยโบราณ การบันทึกคำราชาธิปไตยโบราณจะบันทึกด้วย 2 ภาษาปะปนกันระหว่างภาษาไทยถิ่นอีสาน และบาลี อักษรไทยน้อยมีจำนวนตัวอักษรน้อยกว่าอักษรไทยปัจจุบันและไม่มีวรรณยุกต์ การประสมอักษรเป็นคำส่วนมากคล้ายกับอักษรไทยปัจจุบัน และบางส่วนคล้ายกับอักษรธรรมอีสาน คำเดียวกันสะกดได้หลายแบบ และความสัมพันธ์ระหว่างคำปรัวรรตของอักษรไทยน้อยและคำอ่านภาษาไทยถิ่นอีสานมี 4 รูปแบบ คือ 1:1 1:M M:1 M:N ทำให้การปรัวรรตอักษรไทยน้อยต้องอาศัยบริบทและเข้าใจภาษาไทยถิ่นอีสาน ระยะที่ 2 การออกแบบและพัฒนากรอบการทำงานในการแปลงคำปรัวรรตของตัวอักษรไทยน้อยเป็นคำอ่านภาษาไทยถิ่นอีสาน ผลการวิจัยประกอบด้วย ได้แก่ (1) คลังศัพท์คำปรัวรรต ขนาด 4,645 คำ ใช้สำหรับตัดคำ และค้นหาคำอ่านภาษาไทยถิ่นอีสานที่เป็นไปได้ทั้งหมด และ (2) ออนโทโลยีคำศัพท์ใช้สำหรับการค้นหาคำอ่านภาษาไทยถิ่นอีสานที่มีความกำกวมและต้องใช้คำบริบทในการเจาะจงคำอ่านที่เหมาะสม โครงสร้างออนโทโลยีประกอบด้วย 20 คลาสหลัก และ (3) กระบวนการและขั้นตอนวิธีแปลงคำปรัวรรตของอักษรไทยน้อยเป็นคำอ่านภาษาไทยถิ่นอีสาน แบ่งเป็น 3 ขั้นตอน ได้แก่ การตัดคำ การค้นหาคำอ่านด้วยคลังศัพท์ และการค้นหาคำอ่านด้วยออนโทโลยี และระยะที่ 3 การประเมินประสิทธิภาพของกรอบการทำงาน ผลการประเมินในเชิงปริมาณมีความถูกต้องในการแปลงคำปรัวรรตของตำรับยาโบราณอีสาน จำนวน 807 ตำรับ (7,929 คำ) เป็นคำอ่าน

ภาษาไทยถิ่นอีสาน มีความถูกต้องร้อยละ 87.84 และผลการประเมินประสิทธิภาพเชิงคุณภาพโดยนักภาษาโบราณ จำนวน 12 คน และผู้ใช้ประโยชน์จากตำรายาโบราณอีสาน 27 คน มีค่าเฉลี่ยทุกรายการอยู่ในระดับดี ( $\bar{X} = 4.08$ ,  $SD = 0.73$ )



สาขาวิชาเทคโนโลยีสารสนเทศ

ปีการศึกษา 2564

ลายมือชื่อนักศึกษา Am ant.

ลายมือชื่ออาจารย์ที่ปรึกษา อ.อ.อ.

ลายมือชื่ออาจารย์ที่ปรึกษาร่วม อ.อ.อ.

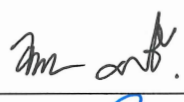

JINTANA POLSRI : THE DEVELOPMENT OF A FRAMEWORK FOR  
TRANSLITERATING THAI NOI CHARACTERS INTO ISAN LANGUAGE  
PRONUNCIATION. THESIS ADVISOR : ASST. PROF. NISACHOL  
CHAMNONGSRI, Ph.D., ASSOC. PROF. JEERAYUT CHAJARUWANICH,  
Ph.D., 218 PP.

#### ONTOLOGY/ TRANSLITERATING/ THAI NOI CHARACTER/ ISAN LANGUAGE

This research presented the framework to transliterate Thai Noi characters into Isan language pronunciations. This research was conducted into 3 phases summarized as follows: Phase 1 consisted of the studies of the problems of ancient document transliteration, the recording structures of ancient Isan medicine textbooks inscribed on palm leaves with Thai Noi characters affecting the efficiency of transliterating, and the character characteristics, language, and Thai Noi characters inscribed on the ancient Isan medicine textbooks. The research methodology used was the qualitative research method. Data were collected from interviewing 10 ancient language linguists using semi-structured interview forms, analyzing the ancient Isan medicine textbooks already transliterated, and 11 copies of the original digital pictures with 459 pages. The results showed that the problems of transliterating ancient documents and ancient Isan medicine textbooks were similar. The problems found were at a high level and from the physical characteristics of the ancient documents, the orthography of the ancient Thai scripts having no tonal markers, In terms of the recording structures and characteristics of Thai Noi characters of ancient Isan medicine textbooks, 2 languages, Isan language and Pali, were combinedly used. The Thai Noi characters are fewer number than standard Thai characters. Their orthography was similar to that of the present Thai characters, but some were derived from the orthography of Dhamma Isan scripts.

A Thai Noi word could be spelled in several ways. There were no both any marks placed over the final consonant of the word and tonal markers. The relationship between the transliterating of Thai Noi characters and words of Isan pronunciation language could be divided into 4 forms: 1:1, 1: M, M:1, and M: N. affected the context and understanding of Isan language was very important for transliterating Thai Noi words into Isan pronunciation language. Phase 2 was the design and development of a framework of Thai Noi characters into Isan pronunciation language. The components of the framework were described as follows: (1) The lexicon of 4,645 ancient Isan medicine words consisted of transliterating words, Isan pronunciation language words and Thai standard words. The lexicon was used for text segmentation and to find all Isan pronunciation language words. (2) The lexicon ontology of ancient Isan medicine textbooks for query the ambiguous Isan pronunciation language words and require context to identify appropriate word consisted of 20 main classes and (3) The process and algorithm for converting the transliteration of Thai Noi characters into Isan pronunciation language words was divided into 3 steps: text segmentation, searching all possible Isan pronunciation language words by lexicon and searching Isan pronunciation language words by ontology. Phase 3 were the evaluation of the efficacy of the framework. The result showed that the accuracy of the transliteration of 807 ancient Isan medicines recipes (7,929 words) into the Isan pronunciation language was 87.8 percent and the evaluate the framework efficacy have overall mean was good level in all items ( $\bar{x} = 4.08, SD = 0.73$ ).

School of Information Technology  
 Academic Year 2021

Student's Signature   
 Advisor's Signature   
 Co-Advisor's Signature 