**บทที่ 2 ทฤษฎีและงานวิจัยที่เกี่ยวข้อง**

เนื้อหาในส่วนนี้นำเสนอทฤษฎีและงานวิจัยที่เกี่ยวข้องกับ การปริวรรตอักษรด้วยเครื่องสำหรับภาษาสันสกฤตแบบโรมันเป็นไทย มีดังต่อไปนี้

**2.1 ทฤษฎีที่เกี่ยวข้อง**

**2.1.1 การปริวรรตอักษรหรือการเขียนคำทับศัพท์ (transliteration)**

การปริวรรตอักษร คือการดำเนินการแปลงอักษรหรืออักขรวิธีจากระบบการเขียนหรือภาษาหนึ่งไปสู่อีกระบบหนึ่งอย่างมีหลักการ เมื่อปริวรรตแล้วสามารถแปลงกลับเป็นอักษรหรืออักขรวิธีเดิมได้[16] เพื่อให้สามารถเขียนคำในภาษาต่างประเทศด้วยภาษาและอักษรในภาษานั้น ๆ ได้สะดวก และใกล้เคียงอักขรวิธีการเขียนเดิมให้ได้มากที่สุด เช่น การปริวรรตอักษรโรมันภาษาอังกฤษ มาเป็นอักษรไทยเพื่อใช้ในภาษาไทย หรือการปริวรรตอักษรภาษาไทย ไปเป็นอักษรโรมันเพื่อใช้ในภาษาอังกฤษ เป็น ปกติแล้วการปริวรรตอักษรคือการจับคู่จากระบบการเขียนหนึ่งไปยังอีกระบบหนึ่งแบบคำต่อคำหรืออักษรต่ออักษร การปริวรรตอักษรได้พยายามที่จะสร้างความสัมพันธ์หนึ่งต่อหนึ่งและทำให้เกิดความถูกต้องแม่นยำ เพื่อให้ผู้อ่านที่ได้รับรู้สามารถสะกดคำต้นฉบับจากคำปริวรรตอักษรได้ ดังนั้นจึงมีการกำหนดหลักการปริวรรตอักษรที่ซับซ้อนในการจัดการกับตัวอักษรบางตัวในภาษาต้นฉบับที่ไม่สัมพันธ์กับอักษรในภาษาเป้าหมาย ความหมายอย่างแคบของการปริวรรตอักษรคือ การปริวรรตอักษรแบบถอดอักษร (transliteration)นั้นเป็นการคงตัวอักษรและเครื่องหมายวรรคตอนทุกอย่างเอาไว้ ทั้งนี้การถอดอักษรไม่สนใจความแตกต่างของเสียงในภาษา เนื่องจากมีข้อจำกัดทางเทคนิค หรือการถอดอักษรโบราณเพื่อให้ยังคงรักษารูปแบบการเขียนเดิมเอาไว้มากที่สุด การปริวรรตอักษรเป็นการถอดอักษร ต่างจากการถอดเสียง (transcription) ซึ่งเป็นการจับคู่เสียงอ่านของภาษาหนึ่ง ๆ ไปยังรูปแบบการเขียนของอีกภาษาที่ใกล้เคียงที่สุด ถึงแม้ว่าระบบการถอดอักษรส่วนใหญ่จะยังคงจับคู่อักษรต้นฉบับกับอักษรในภาษาเป้าหมายที่ออกเสียงคล้ายกันในบางคู่ ถ้าหากความสัมพันธ์ระหว่างตัวอักษรกับเสียงเหมือนกันทั้งสองภาษา การถอดอักษรก็อาจแทบจะเหมือนกับการถอดเสียง

**2.1.2 ภาษาสันสกฤต**

ภาษาสันสกฤตเป็นภาษาที่เก่าแก่ที่สุดภาษาหนึ่งในภาษากลุ่มอินโด-ยูโรเปียน (Indo-European) สาขาย่อยอินโด-อิเรเนียน (Indo-Iranian) และอยู่ในกลุ่มย่อยอินโด-อารยัน (Indo-Aryan) โดยมีระดับวิวัฒนาการเก่าแก่ในระดับใกล้เคียงกับภาษาในกลุ่มอินโด-ยูโรเปียนด้วยกัน คือภาษาละตินและภาษากรีกโบราณ วรรณกรรมภาษาสันสกฤตพบการใช้ที่หลากหลาย ตั้งแต่เป็นวรรณคดี บทกวี บทละคร เป็นตำราทางวิชาการหลากหลายสาขา และเป็นใช้ทางศาสนา บันทึกบทสวด ปรัชญา หลักการทางศาสนา ทั้งในพบเอกสารทั้งใน ศาสนาพราหมณ์-ฮินดู พุทธ เชน และ ซิกข์ ซึ่งในส่วนของ ศาสนาพราหมณ์ – ฮินดูถือว่า ภาษาสันสกฤตเป็นภาษาศักดิ์สิทธิ์

คำว่า สันสกฤต แปลว่า “กลั่นกรองแล้ว” ซึ่งเป็นภาษาของชนชั้นพราหมณ์ ตรงข้ามกับภาษาพูดของชาวบ้านทั่วไปที่เรียกว่า “ปรากฤต” ภาษาสันสกฤตมีพัฒนาการในหลายยุคสมัย โดยมีหลักฐานเก่าแก่ที่สุด ภาษา ภาษาพระเวท (Vedic Sanskrit) ที่ปรากฏในคัมภีร์ฤคเวท เมื่อราว 1,200 ปีก่อน ค.ศ. อันเป็นบทสวดสรรเสริญพระเจ้าในลัทธิพราหมณ์ในยุคต้น ภาษาพระเวทดั้งเดิมยังมิได้มีการวางกฎเกณฑ์ให้เป็นระเบียบรัดกุมและสละสลวย และมีหลักทางไวยากรณ์อย่างกว้าง ๆ ราว 57 ปีก่อน พ.ศ. พราหมณ์ชื่อ "ปาณินิ" ชาวแคว้นคันธาระ ท่านเห็นว่าภาษาสันสกฤตแบบพระเวทนั้นมีภาษาถิ่นปนเข้ามา หากไม่เขียนไวยากรณ์ที่เป็นระเบียบแบบแผนไว้จะคละกับภาษาถิ่น ปาณินิได้ศึกษาและจัดเรียบเรียงตำราไวยากรณ์ขึ้น 8 บท ชื่อว่า “อัษฏาธยายี” ภาษาที่ปรับปรุงใหม่นี้เรียกว่า "ตันติสันสกฤต" หรือ สันสกฤตแบบแผน (Classical Sanskrit) วรรณคดีสันสกฤตแบบแผนที่สำคัญและเป็นที่รู้จักมากคือ มหาภารตะ และ รามายณะ ภาษาสันสกฤตอีกสาขาหนึ่ง เรียกว่า ภาษาสันสกฤตผสม หรือ ภาษาสันสกฤตผสมในพุทธศาสนา (Buddhist Hybrid Sanskrit or Mixed Sanskrit) เป็นภาษาสันสกฤตยุคหลังถัดจากภาษาสันสกฤตแบบแผน พบในคัมภีร์ทางพระพุทธศาสนาทั้งในนิกายสรรวาสติวาทและพระพุทธศาสนาฝ่ายมหายาน วรรณกรรมภาษาสันสกฤตพบการใช้ที่หลากหลาย มีพัฒนาการมาอย่างต่อเนื่องสามารถ แบ่งประเภทตามเนื้อหาได้เป็น 4 ประเภทใหญ่ๆคือ

1. อาคม(āgama) มีเนื้อหาเกี่ยวกับศาสนาและปรัชญา
2. อิติหาส(itihāsa) มีเนื้อหาเกี่ยวกับประวัติศาสตร์หรือวีรชนและประเพณีที่สืบทอดกันมา
3. ศาสตร์(śāstra) มีเนื้อหาเกี่ยวกับศิลปวิทยาการและงานวิชาการ
4. กาวยะ(kāvya) กวีนิพนธ์หรือบทประพันธ์ที่อยู่ในรูปของศิลปะ

ภาษาสันสกฤตไม่มีอักษรสำหรับเขียนชนิดใดชนิดหนึ่งโดยเฉพาะ เป็นภาษาที่ไม่มีระบบการเขียนเป็นของตนเอง แต่จะเขียนด้วยอักษรหลายชนิด อักษรเก่าแก่ที่ใช้เขียนภาษาสันสกฤตมีหลายชนิดด้วยกัน เช่น อักษรขโรษฐี อักษรพราหมี อย่างไรก็ตามในปัจจุบันโดยทั่วไปนิยมเขียนภาษาสันสกฤตด้วยอักษรเทวนาครี ส่วนอักษรอื่น ๆ แล้วแต่ความนิยมในแต่ละท้องถิ่นในอินเดีย ทั้งนี้เนื่องจากอักษรที่ใช้ในอินเดีย มักจะเป็นตระกูลเดียวกัน จึงสามารถดัดแปลงและถ่ายทอด (Transliteration) ระหว่างชุดอักษรได้ง่าย แม้กระทั่งในเอเชียตะวันออกเฉียงใต้มีชุดอักษรท้องถิ่นที่ใช้เขียนภาษาสันสกฤตได้ ยังมีจารึกโบราณภาษาสันสกฤตที่ใช้ อักษรปัลลวะ อักษรขอมโบราณ ในระดับสากลยังใช้อักษรโรมันเขียนภาษาสันสกฤตโดยมีมาตรฐานแตกต่างกันออกไปด้วย ตัวอย่างลักษณะการบันทึกแบบอักษรต่างๆ ดังที่แสดงในรูปที่ 2.1,2.2,2.3



รูปที่ 2.1 ตัวอย่างการเขียนภาษาสันสกฤตด้วยอักษรชนิดต่างๆ

(ที่มา: wikipedia.org, ปี พ.ศ. 2558)



รูปที่ 2.2 ตัวอย่างการเขียนภาษาสันสกฤตด้วยอักษรโรมันแบบต่างๆ



รูปที่ 2.3 ตัวอย่างจารึกภาษาสันสกฤตในประเทศไทย   
 จารึกวัดมเหยงค์ เป็นจารึกอักษรปัลลวะ อายุพุทธศตวรรษที่ ๑๒

(ที่มา: ศูนย์มานุษยวิทยาสิรินธร (องค์การมหาชน), ปี พ.ศ. 2551)

**2.1.3 การปริวรรตอักษรภาษาสันสกฤตสากล (IAST)**

การปริวรรตอักษรภาษาสันสกฤตสากล หรือ The International Alphabet of Sanskrit Transliteration (IAST) เป็นรูปแบบการปริวรรตอักษรภาษาสันสกฤตเป็นอักษรโรมัน(Romanization) รูปแบบหนึ่งที่ปราศจากการสูญเสีย(Lossless) คือสามารถจับคู่อักษรต้นทางและปลายทางได้ครบคู่ โดยมากอักษรต้นทางนั้นมักเป็นอักษรอินเดียตระกูลต่างๆ(Indic Script) นอกจากใช้ปริวรรตอักษรภาษาสันสกฤตให้เป็นอักษรโรมัน ยังนิยมใช้ปริวรรตอักษรภาษาปรากฤต อื่นๆเช่น ภาษาบาลี และภาษาอปภรัมศะ เป็นต้น

แต่เดิมในการการปริวรรตอักษรภาษาสันสกฤตเป็นอักษรโรมันยังไม่ได้มีมาตรฐานกลางที่ใช้ร่วมกันแต่ใช้วิธีการปริวรรตตามข้อเสนอของผู้เชี่ยวชาญหลายๆท่าน ดังที่แสดงในรูปที่ 2.4



รูปที่ 2.4 การปริวรรตอักษรภาษาสันสกฤตเป็นอักษรโรมันตามข้อเสนอของบุคคล   
 ต่างๆก่อนมีแบบการปริวรรตอักษรภาษาสันสกฤตสากล[13]

(ที่มา: shashir.autodidactus.org, ปี พ.ศ. 2555)

จนกระทั่งในปี ค.ศ. 1894 มีการประชุมทางวิชาการเกี่ยวกับด้านตะวันออกศึกษาหรือเอเชียศึกษาในปัจจุบัน การประชุมผู้เชี่ยวชาญด้านตะวันออกศึกษานานาชาติ ครั้งที่ 10 ที่นครเจนีวา ประเทศสวิสแลนด์ (10th International Congress of Orientalists, Held at Geneva) มีมติที่ประชุมให้รวมรูปแบบการปริวรรตอักษรภาษาสันสกฤตและบาลีเป็นอักษรโรมัน จากสองรูปแบบหลัก 2 รูปแบบคือ รูปแบบการปริวรรตของ ราชสมาคมเอเชียแห่งบริเตนใหญ่และไอร์แลนด์ (Royal Asiatic Society of Great Britain and Ireland) [14]และรูปแบบของสมาคมตะวันออกศึกษาแห่งเยอรมัน (German Oriental Society : Deutsche Morgenländische Gesellschaft) และตีพิมพ์สรุปรายงานการประชุมในปีเดียวกันเป็นภาษาฝรั่งเศสลงในหนังสือ Xme Congrès International des Orientalistes, Session de Genève. Rapport de la Commission de Transcription (1894) จากนั้นในปี ค.ศ.1895 ราชสมาคมเอเชียแห่งบริเตนใหญ่และไอร์แลนด์ได้แปลสรุปรายงานการประชุมเป็นภาษาอังกฤษลงในวารสารราชสมาคมเอเชีย [15] ต่อมารูปแบบการปริวรรตนี้ มีความสำคัญทางวิชาการภาษาสันสกฤตมาก  
 จึงได้เป็น “การปริวรรตอักษรภาษาสันสกฤตสากล” หรือ “The International Alphabet of Sanskrit Transliteration” (IAST) เพราะเป็นมาตรฐานหลักการปริวรรตอักษรภาษาสันสกฤตเป็นอักษรโรมัน (Romanization) จนถึงปัจจุบัน



รูปที่ 2.4 บางส่วนในรายงานมติที่ประชุม เรื่องการปริวรรตอักษรภาษาสันสกฤตและบาลี   
 สากล ในการประชุมผู้เชี่ยวชาญด้านตะวันออกศึกษานานาชาติ ครั้งที่ 10 [15]

(ที่มา: วารสารราชสมาคมเอเชีย, ปี ค.ศ.1985)

**2.1.4 การปริวรรตอักษรภาษาสันสกฤตแบบโรมัน IAST เป็นไทย**

การปริวรรตอักษรทั้งภาษาบาลีและภาษาสันสกฤตแบบโรมัน IAST เป็นไทย ราชบัณฑิตย- สถานยังไม่ได้วางหลักเกณฑ์การปริวรรตอักษร แต่ในบรรดาผู้ศึกษาจะยึดธรรมเนียมการปริวรรตอักษรที่สืบทอดต่อๆกันมา อย่างไรก็ตามหลักการยังไม่เป็นกฎเกณฑ์ที่แน่นอนตายตัวบางตำราบางอาจารย์อาจจะแตกต่างไป หรือมีการปรับเปลี่ยนกันเล็กน้อย ส่วนใหญ่จะปริวรรตอยู่ใน 2 รูปแบบ คือ

1. การปริวรรตอักษรแบบคงรูป หรือ แบบแผนคือการปริวรรตอักษรแบบตรงตามรูปศัพท์เดิมโดยใช้เครื่องหมายพินทุด้วย เพื่อให้ถอดกลับเป็นอักษรโรมันหรืออักษรเทวนาครีได้ถูกต้องและได้ลักษณะอักขรวิธีได้ใกล้เคียงกับคำเดิม การปริวรรตอักษรรูปแบบนี้เหมาะสำหรับใช้ในหนังสือที่ต้องการเขียนอย่างเป็นทางการหรือเอกสารทางวิชาการ

2. การปริวรรตอักษรแบบปรับรูป หรือ แบบทั่วไปคือหมายถึงการปริวรรตอักษรโดยปรับให้เข้ากับอักขรวิธีไทย เช่น โดยไม่ใช้พินทุ และเพิ่มวิสรรชนีย์หรือใส่เครื่องหมายทัณฑฆาต เพื่อให้อ่านได้ง่ายขึ้นและรูปคำกลมกลืนกับภาษาไทย จึงนิยมใช้การทับศัพท์อย่างง่าย เพื่อความสะดวกในการอ่านสำหรับบุคคลทั่วไป พบเห็นได้ตามหนังสือบทสวดมนต์ทั่วไป ลักษณะดังกล่าวทำให้การปริวรรตอักษรแบบปรับรูป นั้นใกล้เคียงกับการถอดเสียง (transcription) แต่การปริวรรตรูปแบบนี้ไม่ใช่การถอดเสียง เพราะพยัญชนะจากการปริวรรตบางตัวออกเสียงไม่ตรงมาตรา เช่น อักษร ค จากการปริวรรตออกเสียงคล้าย ก , อักษร ท ธ จากการปริวรรตออกเสียงคล้าย ด , อักษร พ ภ จากการปริวรรตออกเสียงคล้าย บ หรือเสียงของเครื่องหมาย วิสรรคะ

  
รูปที่ 2.5 การปริวรรตอักษรภาษาสันสกฤตแบบโรมัน IAST เป็นไทยรูปแบบต่างๆ

การปริวรรตอักษรภาษาสันสกฤตแบบโรมัน IAST เป็นไทยนั้นผู้ปริวรรตต้องมีความรู้ความเข้าใจเกี่ยวกับรูปแบบอักษรในแต่ละส่วนมีตำแหน่งละหน้าที่อย่างไรจึงจะปริวรรตอักษรได้อย่างถูกต้อง สำหรับการเปรียบเทียบรูปอักษรโรมัน IAST เป็นไทย มีส่วน ประกอบดังต่อไป 4 ส่วนดังนี้

1. รูปสระ(Vowels) แบ่งได้ 2 ประเภท คือ
   1. สระแท้ แสดงในตารางที่ 2.1
   2. สระประสม แสดงในตารางที่ 2.1
2. รูปพยัญชนะ(Consonants) แบ่งได้ 2 ประเภท คือ
   1. พยัญชนะวรรค แสดงในตารางที่ 2.2
   2. พยัญชนะอวรรค แสดงในตารางที่ 2.3
3. รูปเครื่องหมายและสัญลักษณ์พิเศษ(Diacritics) แสดงในตารางที่ 2.4
4. รูปตัวเลข(Numerals) รูปอักษรโรมัน IAST ใช้ตัวเลขฮินดู-อารบิก แสดงในตารางที่ 2.5

ตารางที่ 2.1 การเปรียบเทียบรูปสระ

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **ประเภท** | **ฐานที่เกิดเสียง** | **หรัสวะ**  **(สระเสียงสั้น)** | **ทีรฆะ (สระเสียงยาว)** | |
| **สระแท้**  **สมานากษะระ** | กัณฐยะ (คอ) | A a อะ | Ā ā อา | |
| ตาละวยะ(เพดาน) | I i อิ | Ī ī อี | |
| มูรธันยะ(ปุ่มเหงือก) | U u อุ | Ū ū อู | |
| ทันตยะ(ฟัน) | Ṛ ṛ ฤ | Ṝ ṝ ฤๅ | |
| โอษฐยะ(ริมผีปาก) | Ḷ ḷ ฦ | Ḹ ḹ ฦๅ | |
| **สระประสม**  **(สันธยักษะระ)** | กัณฐะตาละวยะ(คอและเพดาน) |  | E e เอ | AI ai ไอ |
| กัณโฐษฐยะ(คอและริมฝีปาก) |  | O o โอ | AU au เอา |

ตารางที่ 2.2 การเปรียบเทียบรูปพยัญชนะวรรค

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **วรรค** | **ฐานที่เกิดเสียง** | **สปรรศะ(เสียงกัก)** | | | | **อนุนาสิกะ (เสียงขึ้นจมูก)** |
| **อโฆษะ(เสียงไม่ก้อง)** | | | **โฆษะ(เสียงก้อง)** | |
| **ลมเบา\*** | **ลมหนัก\*\*** | **ลมเบา\*** | **ลมหนัก\*\*** | |
| ก วรรค | กัณฐยะ  (คอ) | K k  ก | KH kh  ข | G g  ค | gh  ฆ | Ṅ ṅ  ง |
| จ  วรรค | ตาละวยะ  (เพดาน) | C c  จ | CH ch  ฉ | J j  ช | JH jh  ฌ | Ñ ñ  ญ |
| ฎ วรรค | มูรธันยะ  (ปุ่มเหงือก) | Ṭ ṭ  ฏ | ṬH ṭh  ฐ | Ḍ ḍ  ฑ | ḌH ḍh  ฒ | Ṇ ṇ  ณ |
| ต วรรค | ทันตยะ  (ฟัน) | T t  ต | TH th  ถ | D d  ท | DH dh  ธ | N n  น |
| ป วรรค | โอษฐยะ  (ริมผีปาก) | P p  ป | PH ph  ผ | B b  พ | BH bh  ภ | M m  ม |

*\* ลมเบา คือ อัลปปราณะ \*\* ลมหนัก คือ มหาปราณะ*

ตารางที่ 2.3 การเปรียบเทียบรูปพยัญชนะอวรรค

|  |  |  |  |
| --- | --- | --- | --- |
| **ฐานที่เกิดเสียง** | **อันตะสถะ(เสียงเปิด)** | **อูษมะ(เสียงเสียดแทรก)** | |
| **โฆษะ(ก้อง)** | **อโฆษะ(ไม่ก้อง)** | **โฆษะ(ก้อง)** |
| **อัลปปราณะ(ลมเบา)** | **มหาปราณะ(ลมหนัก)** | |
| กัณฐยะ(คอ) |  |  | H h ห |
| ตาละวยะ(เพดาน) | Y y ย | Ś ś ศ |  |
| มูรธันยะ(ปุ่มเหงือก) | R r ร | Ṣ ṣ ษ |  |
| ทันตยะ(ฟัน) | L l ล | S s ส |  |
| โอษฐยะ(ริมผีปาก) | V v ว |  |  |
| พยัญชนะ อวรรคอื่นๆ | Ḻ ḻ ฬ \* | | |

*\* มีในภาษาสันสกฤตแบบพระเวท และภาษาบาลี*

ตารางที่ 2.4 การเปรียบเทียบรูปเครื่องหมายและสัญลักษณ์พิเศษ

|  |  |  |  |
| --- | --- | --- | --- |
| **ชื่อสัญลักษณ์** | **คำอธิบาย** | **สัญลักษณ์** | **ปริวรรตเป็นอักษรไทย** |
| อนุสวาระ | เครื่องหมายกำกับสระใช้กำกับเสียงขึ้นจมูก | Ṃ ṃ | ˚ (นิคหิต) |
| จันทรพินทุ/ อนุนาสิกะ | เครื่องหมายกำกับ บนเสียงอนุสวาระ ที่เน้นเสียงพิเศษคือ เน้นเสียงขึ้นจมูกกว่าปกติ | M̐ m̐ ,  Ṁ ṁ\* | ัํ , ̐ ( อัํ , อ̐ ) มีผู้ใช้ไม้หันอากาศและมีนิคหิตข้างบน และ ใช้จันทรพินทุสำหรับอักษรโรมัน(moon dot) |
| วิสรรคะ | เครื่องหมายกำกับสระ ใช้เพื่อแทนเสียงลมหายใจมาก | Ḥ ḥ | ะ (วิสรรชนีย์) ,  ห์ |
| อวครหะ | เครื่องหมายกำกับ ว่า สระอะมีการย่อหรือตัดออกในการสนธิในภาษาสันสกฤต | ' (apostrophe) | ’ (อัญประกาศเดียว) |
| ทัณฑะ | เครื่องหมายจบประโยคหรือจบบาทที่สองของโศลก | ।(danda) | ฯ (อังคั่นเดี่ยว) |
| ทวิทัณฑะ | จบประโยคเพื่อจะขึ้น  ย่อหน้าใหม่หรือจบโศลก | ॥ (double danda) | ๚ (อังคั่นคู่) |

*\* ประกาศสำนักนายกรัฐมนตรี เรื่อง หลักเกณฑ์การทับศัพท์ภาษาจีนและภาษาฮินดี ลงวันที่ 22 ธันวาคม 2549 ใช้จันทรพินทุเป็น Ṁ ṁ* [17]

*\*\*ปริวรรตเป็นอักษรไทยสำหรับ จันทรพินทุ มีการกำหนดเฉพาะ ภาษาฮินดี ตามประกาศสำนักนายกรัฐมนตรีข้างต้นให้ใช้ “ง”* [17] *ส่วนภาษาสันสฤตยังไม่มีหน่วยงานที่เกี่ยวข้องกำหนด*

ตารางที่ 2.5 การเปรียบเทียบรูปตัวเลข

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **ตัวเลขฮินดู-อารบิก** | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 |
| **ตัวเลขไทย** | ๐ | ๑ | ๒ | ๓ | ๔ | ๕ | ๖ | ๗ | ๘ | ๙ |

ยังมีรูปพยัญชนะอื่นๆในระบบอักษรที่ใช้ในการเขียนภาษาบาลีสันสกฤตและปรากฤต ใช้เขียนรูปพยัญชนะที่เรียกว่า พยัญชนะสังโยค หรือ พยัญชนะสังยุกต์ (Conjunct Consonants) คือเป็นพยัญชนะซ้อนกันหรือควบกันตั้งแต่ 2 ตัวขึ้นไป ในภาษาสันสกฤตนั้นพบการซ้อนกันมากถึง 5 อักษร โดยเมื่อพยัญชนะซ้อนกัน จะเขียนเป็นรูปพยัญชนะใหม่(Ligature) หรืออาจใช้รูปย่อของพยัญชนะนั้นซ้อนกันในแนวตั้ง ลักษณะของอักษรที่ใช้ระบบดังกล่าวหลายชนิด เช่น อักษรเทวนาครี หรือพบในประเทศไทยเช่น อักษรขอมไทย อักษรธรรมล้านช้าง อักษรธรรมล้านนา ส่วนอักษรโรมันและอักษรไทยไม่มีลักษณะการซ้อนพยัญชนะเช่นนี้ พยัญชนะสังโยคถ้าเขียนด้วยอักษรเทวนาครีจะเห็นชัดเจนเพราะอักษรสองตัวจะเชื่อมกันเป็นตัวเดียว แต่เมื่อปริวรรตเป็นอักษรโรมันก็จะได้เป็นพยัญชนะเรียงกัน



รูปที่ 2.6 ตัวอย่างลักษณะพยัญชนะสังโยค อักษรเทวนาครีเทียบกับอักษรโรมัน

การปริวรรตอักษรภาษาสันสกฤตแบบโรมัน IAST เป็นไทยทั้งสองรูปแบบ ให้ถือตามตารางเปรียบเทียบสระพยัญชนะเครื่องหมายและสัญลักษณ์พิเศษและตัวเลข ดังที่กล่าวมาเป็นหลัก ต่อไปนี้จะอธิบายถึงวิธีการปริวรรตอักษรที่ผู้ศึกษาได้พบทั้งสองรูปแบบหลัก อย่างไรก็ตามหลักการยังไม่เป็นกฎเกณฑ์ที่เป็นทางการ มีรูปแบบรายละเอียดแตกต่างกันบางประการ หรืออาจจะต้องอาศัยหลักการปริวรรตอักษรจากกลุ่มภาษาเดียวกันที่มีแบบแผนการปริวรรตที่เป็นแบบแผนและมาตรฐานอยู่แล้ว เช่น ภาษาบาลี ภาษาฮินดี

**วิธีการปริวรรตเป็นอักษรไทยแบบคงรูป หรือ แบบแผน**

1. การถอดรูปพยัญชนะ
   1. การถอดรูปพยัญชนะอักษรโรมันปกติ ให้ถอดตามตารางเปรียบเทียบพยัญชนะข้างต้น
2. การถอดรูปสระ
   1. ถอดรูปสระ a ā i ī u ū ṛ ṝ ḷ ḹ (รูปสระ อา อิ อี อุ อู ฤ ฤๅ ฦ ฦๅ) หากตามพยัญชนะต้นให้ถอดตามปกติ
   2. ถอดรูปสระ a (รูป สระอะ) ไม่ต้องถอดรูปหรือไม่เติม วิสรรชนีย์ (ะ) ตัวอย่างตาม ตารางที่ 2.6

ตารางที่ 2.6 ตัวอย่างการถอดรูปสระ a

|  |  |  |
| --- | --- | --- |
| bhagavadgītā | daśabala | śatakrata |
| ภควทฺคีตา | ทศพล | ศตกฺรต |

* 1. ถอดรูปสระ e ai o au (รูปสระ เอ ไอ โอ เอา) ที่ประสมอยู่พยัญชนะต้นอักษรโรมัน ต้องถอดรูปสระก่อนจึงค่อยถอดพยัญชนะต้น ยกเว้น รูปสระ au (รูปสระ เอา) นั้นให้ใส่ไม้หน้า( เ )ก่อน จึงค่อยถอดพยัญชนะต้นและปิดท้ายด้วย ลากข้าง   
      ( า ) ตัวอย่างตาม ตารางที่ 2.7

ตารางที่ 2.7 ตัวอย่างการถอดรูปสระ e ai o au ที่ประสมอยู่พยัญชนะต้นอักษรโรมัน

|  |  |  |  |
| --- | --- | --- | --- |
| teṣāṃ | maitrī | namo | sahitau |
| เตษํา | ไมตฺรี | นโม | สหิเตา |

* 1. หากสระอักษรโรมันขึ้นต้นคำ ให้ถอดโดยประสม อักษร อ กับสระนั้นๆเข้าไป และ a (รูปสระ อะ) ให้ใช้ อักษร อ ตัวเดียว เว้นไว้แต่สระอักษรโรมัน ṛ ṝ ḷ ḹ (รูปสระ ฤ ฤๅ ฦ ฦๅ) ไม่ต้องประสม อักษร อ ตัวอย่างตาม ตารางที่ 2.8

ตารางที่ 2.8 ตัวอย่างการถอดรูปสระอักษรโรมันขึ้นต้นคำ

|  |  |  |  |
| --- | --- | --- | --- |
| สระเสียงสั้น | | สระเสียงสั้น | |
| agniṃ | อคฺนิํ | ārya | อารฺย |
| indrā | อินฺทฺรา | īśāno | อีศาโน |
| usrā | อุสฺรา | ūtaya | อูตย |
| สระประสม | | | |
| eha | เอห | aibhiragne | ไอภิรคฺเน |
| ojasā | โอชสา | auśijaḥ | เอาศิชะ |

1. การใช้เครื่องหมายพินทุ
   1. โดยใส่ไว้ใต้พยัญชนะไทยที่ถอดมาจากพยัญชนะอักษรโรมันที่ไม่มีสระอักษรโรมันกำกับหรือตามมา ตัวอย่างตาม ตารางที่ 2.9

ตารางที่ 2.9 ตัวอย่างการถอดรูปพยัญชนะอักษรโรมันที่ไม่มีสระอักษรโรมันกำกับ

|  |  |  |
| --- | --- | --- |
| sarva | dharmacakra | sattvā |
| สรฺว | ธรฺมจกฺร | สตฺตฺวา |

* 1. ในกรณีถอดรูปพยัญชนะอักษรโรมันที่ประสมด้วยสระ ṛ ṝ ḷ ḹ (ฤ ฤๅ ฦ ฦๅ) มีสองแนวคิด คือ ใส่พินทุใต้พยัญชนะที่ประสมด้วยสระเหล่านี้ เช่น mṛga : มฺฤค  
     และไม่ใส่พินทุใต้พยัญชนะที่ประสมด้วยสระเหล่านี้ ตัวอย่างตาม ตารางที่ 2.10

ตารางที่ 2.10 ตัวอย่างการถอดรูปพยัญชนะอักษรโรมันที่ประสมด้วยสระ ṛ ṝ ḷ ḹ

|  |  |  |
| --- | --- | --- |
|  | mṛga | kṛṣṇa |
| แบบใส่พินทุ | มฺฤค | กฺฤษฺณ |
| แบบไม่ใส่พินทุ | มฤค | กฤษฺณ |

1. การถอดรูปเครื่องหมายและสัญลักษณ์พิเศษ
   1. รูปเครื่องหมายอนุสวาระ อักษรโรมันใช้ ṃ ถอดเป็นนิคหิตไว้ข้างบน( ˚ ) ลักษณะการถอดให้ถอดลำดับหลังสระอักษรโรมัน ยกเว้น รูป ā (รูปสระ อา) ให้ถอด  
      อนุสวาระก่อน ตัวอย่างตาม ตารางที่ 2.11

ตารางที่ 2.11 ตัวอย่างการถอดรูปเครื่องหมายอนุสวาระ

|  |  |  |
| --- | --- | --- |
| vihāraṃ | gṛhapatiṃ | mahāsāṃghikānāṃ |
| วิหารํ | คฤหปติํ | มหาสําฆิกานํา |

* 1. รูปเครื่องหมายจันทรพินทุหรืออนุนาสิกะ อักษรโรมันใช้ m̐ มี 3 แนวคิด คือ ตัดทิ้ง ถอดเป็นไม้หันอากาศและมีนิคหิตข้างบน( อัํ ) และ ใช้จันทรพินทุสำหรับอักษรโรมัน( ̐ : Moon dot) ลักษณะการถอดให้ถอดลำดับหลังสระอักษรโรมัน ยกเว้น รูป ā (รูปสระ อา) ให้ถอดจันทรพินทุก่อน ตัวอย่างตาม ตารางที่ 2.12

ตารางที่ 2.12 ตัวอย่างการถอดรูปรูปเครื่องหมายจันทรพินทุหรืออนุนาสิกะ

|  |  |  |
| --- | --- | --- |
|  | dasyūm̐radharām̐ | tālam̐lokān |
| แบบตัดทิ้ง | ทสฺยูรธรา | ตาลฺโลกานฺ |
| แบบไม้หันอากาศและ  มีนิคหิตข้างบน | ทสฺยูัํรธรัํา | ตาลัํโลกานฺ |
| แบบจันทรพินทุสำหรับ  อักษรโรมัน(Moon dot) | ทสฺยู̐รธร̐า | ตาล̐โลกานฺ |

* 1. รูปเครื่องหมายวิสรรคะ อักษรโรมันใช้ ḥ มี 2 แนวคิด คือ ถอดเป็น วิสรรชนีย์ (ะ) และ อักษร ห ใส่ทัณฑฆาต (ห์) เมื่อถอดแล้วไว้ตำแหน่งเดิม มักอยู่หลังสระอักษรโรมัน ตัวอย่างตาม ตารางที่ 2.13

ตารางที่ 2.13 ตัวอย่างการถอดรูปเครื่องหมายวิสรรคะ

|  |  |  |
| --- | --- | --- |
|  | mahāsattvāḥ | agniḥ |
| แบบวิสรรชนีย์ (ะ) | มหาสตฺตฺวาะ | อคฺนิะ |
| แบบ ห ใส่ทัณฑฆาต | มหาสตฺตฺวาห์ | อคฺนิห์ |

* 1. รูปเครื่องหมายอวครหะ อักษรโรมันใช้ ' (apostrophe) ถอดเป็น’ (อัญประกาศเดียว) เมื่อถอดแล้วไว้ตำแหน่งเดิม บ้างก็ไม่ถอดใช้รูปเดิม บ้างก็ตัดออกแล้วเปลี่ยนเป็นเว้นวรรค
  2. รูปเครื่องหมาย । (ทัณฑะ) ॥ (ทวิทัณฑะ) ถอดเป็น ฯ (อังคั่นเดี่ยว) ๚ (อังคั่นคู่) ตามลำดับ เมื่อถอดแล้วไว้ตำแหน่งเดิม บ้างก็ไม่ถอด ใช้รูปเดิม

1. การถอดรูปพยัญชนะสังโยค หรือ พยัญชนะสังยุกต์ ที่ประสมกับรูปสระ e ai o au (รูปสระ เอ ไอ โอ เอา) เนื่องจากการถอดพยัญชนะสังโยคจากอักษรเทวนาครีหรืออักษรที่ใช้ระบบสังโยคหรือการซ้อนของตัวพยัญชนะโดยตรง นั้นจะเห็นลักษณะการซ้อนของพยัญชนะได้ชัดเจนจะสามารถถอดรูปพยัญชนะสังโยคได้ง่าย แต่หากถอดจากอักษรโรมัน จะไม่สามารถเห็นลักษณะการซ้อนของตัวพยัญชนะ ต้องอาศัยความเชี่ยวชาญในภาษาสันสกฤต ดังนั้นการถอดรูปสระหน้าพยัญชนะสังโยค จึงมีวิธีปฏิบัติหลายแนวทางในการปริวรรตจึงแบ่งได้ลักษณะดังนี้
   1. กรณีพยัญชนะสังโยคซ้อนกันมากกว่า 2 ตัว ส่วนมากเหมือนกัน ไม่คำนึงถึงพยัญชนะสังโยค ถอดรูปสระและพยัญชนะปกติทั้งหมด ตัวอย่างพยัญชนะสังโยค ttv (ตฺตฺว) เช่น bodhisattvo ถอดเป็น โพธิสตฺตฺโว

**กรณีพยัญชนะสังโยคซ้อนกัน 2 ตัว**

* 1. หากทราบว่าชุดพยัญชนะอักษรโรมัน เป็นพยัญชนะสังโยคแล้ว ยกตัวอย่าง เช่น kṣ(กฺษ) , jñ(ชญ) หากผสมกับสระอักษรโรมันดังกล่าวเมื่อถอดเป็นสระอักษรไทยแล้ว จะเขียนสระหน้าไว้พยัญชนะตัวแรกของพยัญชนะสังโยค และใส่พินทุที่พยัญชนะตัวแรก และเว้นการใส่พินทุที่ตัวสุดท้าย ตัวอย่างตาม ตารางที่ 2.14

ตารางที่ 2.14 ตัวอย่างการถอดรูปชุดพยัญชนะอักษรโรมันที่เป็นพยัญชนะสังโยค

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| kṣetre | saṃjñe | tumhe | pṛṣṭhe | draupadī |
| เกฺษเตฺร | สํเชฺญ | ตุเมฺห | ปฺฤเษฺฐ | เทฺราปที |

* 1. ถอดเฉพาะชุดพยัญชนะอักษรโรมันที่เป็นพยัญชนะสังโยคซ้อนหรือควบกับพยัญชนะโรมันเฉพาะ y r l v ṣ (รูปพยัญชนะ ย ร ล ว และ ษ) หากผสมกับสระอักษรโรมันดังกล่าวเมื่อถอดเป็นสระอักษรไทยแล้ว จะเขียนสระหน้าไว้พยัญชนะตัวแรกของพยัญชนะสังโยค และใส่พินทุที่พยัญชนะตัวแรก และเว้นการใส่พินทุที่ตัวสุดท้าย นอกจากนั้นถอดรูปสระและพยัญชนะปกติ ตัวอย่างตาม ตารางที่ 2.15

ตารางที่ 2.15 ตัวอย่างการถอดรูปพยัญชนะสังโยคที่ซ้อนเฉพาะ y r l v ṣ

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| puṇyena | kṣetre | kleśa | aśvaiḥ | ādhyo | rātrau |
| ปุเณฺยน | เกฺษเตฺร | เกฺลศ | อไศฺวะ | อาโฒฺย | ราเตฺรา |

* 1. ถอดเฉพาะชุดพยัญชนะอักษรโรมัน ที่เป็นพยัญชนะสังโยคซ้อนหรือควบพยัญชนะโรมัน r (รูปพยัญชนะ ร) หากผสมกับสระอักษรโรมันดังกล่าวเมื่อถอดเป็นสระอักษรไทยแล้ว จะเขียนสระหน้าไว้พยัญชนะตัวแรกของพยัญชนะสังโยค และใส่พินทุที่พยัญชนะตัวแรก และเว้นการใส่พินทุที่ตัวสุดท้าย นอกจากนั้นถอดรูปสระและพยัญชนะปกติ ตัวอย่างตาม ตารางที่ 2.16

ตารางที่ 2.16 ตัวอย่างการถอดรูปพยัญชนะสังโยคที่ซ้อนเฉพาะ r

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| kṣetre | draupadī | rātrau | citro | traividyā | idrevato |
| กฺเษเตฺร | เทฺราปที | ราเตฺรา | จิโตฺร | ไตฺรวิทฺยา | อิเทฺรวโต |

* 1. ไม่คำนึงถึงพยัญชนะสังโยค ถอดรูปสระและพยัญชนะแบบปกติทั้งหมด ตัวอย่างตาม ตารางที่ 2.17

ตารางที่ 2.17 ตัวอย่างการถอดรูปพยัญชนะหรือสระที่ไม่คำนึงถึงพยัญชนะสังโยค

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| kṣetre | pṛṣṭhe | draupadī | traividyā | tumhe | saṃjñe |
| กฺเษตฺเร | ปฺฤษฺเฐ | ทฺเราปที | ตฺไรวิทฺยา | ตุมฺเห | สํชฺเญ |

ตารางที่ 2.18 สรุปตัวอย่างการ ตัวอย่างการถอดรูปพยัญชนะสังโยคแบบต่างๆ

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| คำศัพท์อักษรโรมัน | พยัญชนะสังโยค | แบบ 5.2 | แบบ 5.3 | แบบ 5.4 | แบบ 5.5 |
| kṣetre | kṣ (กฺษ),  tr (ตฺร) | เกฺษเตฺร | เกฺษเตฺร | กฺเษเตฺร | กฺเษตฺเร |
| saṃjñe | jñ(ชฺญ) | สํเชฺญ | สํชฺเญ | สํชฺเญ | สํชฺเญ |
| tumhe | mh(มฺห) | ตุเมฺห | ตุมฺเห | ตุมฺเห | ตุมฺเห |
| pṛṣṭhe | ṣṭh(ษฺฐ) | ปฺฤเษฺฐ | ปฺฤษฺเฐ | ปฺฤษฺเฐ | ปฺฤษฺเฐ |
| āḍhyo | ḍhy(ฒฺย) | อาโฒฺย | อาโฒฺย | อาฒฺโย | อาฒฺโย |
| puṇyena | ṇy(ณฺย) | ปุเณฺยน | ปุเณฺยน | ปุณฺเยน | ปุณฺเยน |
| kleśa | kl(กฺล) | เกฺลศ | เกฺลศ | กฺเลศ | กฺเลศ |
| aśvaiḥ | śv(ศฺว) | อไศฺวะ | อไศฺวะ | อศฺไวะ | อศฺไวะ |
| draupadī | dr (ทฺร) | เทฺราปที | เทฺราปที | เทฺราปที | ทฺเราปที |
| rātrau | tr (ตฺร) | ราเตฺรา | ราเตฺรา | ราเตฺรา | ราตฺเรา |
| citro | tr (ตฺร) | จิโตฺร | จิโตฺร | จิโตฺร | จิตฺโร |
| idrevato | tr (ตฺร) | อิเทฺรวโต | อิเทฺรวโต | อิเทฺรวโต | อิทฺเรวโต |
| traividyā | tr (ตฺร) | ไตฺรวิทฺยา | ไตฺรวิทฺยา | ไตฺรวิทฺยา | ตฺไรวิทฺยา |

**2.1.4 ทฤษฎีการปริวรรตอักษรด้วยเครื่อง**

การแปลภาษาด้วยเครื่องสามารถแบ่งตามลักษณะการทำงานของระบบออกเป็น 3 กลุ่ม คือ

2.2.1 การแปลภาษาด้วยเครื่องแบบใช้ฐานกฎ (Rule-based Machine Translation) เป็นการแปลภาษาด้วยเครื่องโดยใช้ความรู้ทางด้านภาษาศาสตร์มาใช้ในการกำหนดกฎเกณฑ์ของระบบ ซึ่งจะต้องมีการแยกคุณลักษณะและข้อมูลทางภาษาศาสตร์ของภาษาต้นฉบับจากนั้นจึงทำการวิเคราะห์ตามกฎไวยากรณ์ของภาษาต้นฉบับ แล้วส่งผ่านข้อมูลที่ได้จากการวิเคราะห์นั้นไปยังกระบวนการวิเคราะห์คุณลักษณะและข้อมูลทางภาษาศาสตร์ของภาษาเป้าหมายแล้วจึงทำการแปลจากพจนานุกรมคู่ภาษาและสร้างรูปประโยคของภาษาเป้าหมายออกมา ข้อจำกัดของการแปลภาษาด้วยเครื่องกลุ่มนี้คือ ผู้พัฒนาต้องมีความรู้ทางด้านภาษาศาสตร์ ของทั้งสองภาษาเป็นอย่างมาก จึงจะทำให้ได้การแปลที่มีประสิทธิภาพ

2.2.2 การแปลภาษาด้วยเครื่องแบบใช้สถิติ (Statistical Machine Translation)เป็นการแปลภาษาด้วยเครื่องโดยใช้วิธีการทางสถิติมาช่วยในการแปล ซึ่งการแปลกลุ่มนี้จำเป็นต้องมีฐานข้อมูลคู่ภาษาที่มีการจับคู่ประโยค เพื่อเป็นฐานความรู้ให้ระบบทำการเรียนรู้ และใช้ค่าทางสถิติ เอ็นแกรม (N-Gram Model) ซึ่งเป็นการคำนวณค่าของการที่คำเกิดขึ้นร่วมกัน ถ้าชุดคำชุดใดมีค่าเอ็นแกรมหรือค่าความน่าจะเป็น (Probability) สูง แสดงว่าชุดคำนี้มีโอกาสเกิดขึ้นร่วมกันบ่อยครั้ง การคำนวณหาค่าเอ็นแกรมของชุดคำที่มีอยู่ในฐานข้อมูลคู่ภาษา ทำให้ได้ค่าความน่าจะเป็นของชุดคำต่างๆ เพื่อนำไปเปรียบเทียบและใช้ในการแปลข้อความได้ โดยสามารถเลือกใช้ค่าเอ็นแกรมได้ตั้งแต่ 2 คำ (bigrams), 3 คำ (trigrams) เป็นต้น ยิ่งใช้ค่าเอ็นแกรมมาก ยิ่งทำให้ภาษาต้นฉบับมีการได้หลากหลายและมีความละเอียดมากขึ้นข้อดีของวิธีการนี้คือ ผู้พัฒนาไม่จำเป็นต้องมีความรู้ทางด้านภาษาศาสตร์ของทั้งสองภาษาจึงทำให้ไม่เกิดปัญหาเรื่องไวยากรณ์ต่างๆ แต่การแปลด้วยวิธีการนี้จำเป็นต้องมีฐานข้อมูลคู่ภาษาที่มีจำนวนข้อมูลมหาศาลเพื่อหาค่าทางสถิติที่จะนำไปใช้เปรียบเทียบกับประโยคที่จะนำมาแปลได้อย่างครอบคลุม

2.2.3 การแปลภาษาด้วยเครื่องแบบอ้างอิงตัวอย่าง (Example-based Machine Translation)เป็นการแปลภาษาด้วยเครื่องโดยมีฐานข้อมูลคู่ภาษาเป็นองค์ประกอบสำคัญ ซึ่งทำหน้าที่เก็บคำและประโยคตัวอย่างของคู่ภาษาที่ใช้จริงในชีวิตประจำวัน เพื่อนำไปคำนวณหาว่าประโยคที่รับเข้ามานั้น ควรทำการแปลออกมาเป็นประโยคในรูปแบบใด ระบบการทำงานของวิธีการนี้ แบ่งเป็น 2 ส่วนหลัก คือ การสร้างต้นแบบการแปล และการรวมประโยคคำแปลใหม่ โดยในส่วนแรกจะเป็นการจับคู่คำและประโยคตัวอย่างในฐานข้อมูลคู่ภาษา สร้างเป็นต้นแบบการแปลพื้นฐานเพื่อนำไปใช้เปรียบเทียบกับประโยคที่รับเข้ามา จากนั้นในส่วนที่ 2 จะทำการรวบรวมคำแปลที่ได้จากการเปรียบเทียบ มาสร้างเป็นประโยคผลลัพธ์ข้อดีของวิธีการนี้คือ ไม่ใช้กฎไวยากรณ์และพจนานุกรมคู่ภาษา จึงไม่เกิดปัญหาเรื่องไวยากรณ์ต่างๆ อีกทั้งยังช่วยประหยัดเวลาในการแก้ไขปรับปรุงระบบ เพราะใช้เวลาในการพัฒนาระบบที่จะทำงานเชื่อมต่อกับฐานข้อมูลคู่ภาษาเพียงครั้งเดียว ถ้าพัฒนาระบบให้ทำงานได้อย่างมีประสิทธิภาพแล้ว สามารถนำไปใช้กับฐานข้อมูลคู่ภาษาอื่นๆ ได้อีกด้วย แต่การแปลด้วยวิธีการนี้จำเป็นต้องมีฐานข้อมูลคู่ภาษาที่มีจำนวนคำและประโยคตัวอย่างเป็นจำนวนมาก และต้องทำการจับคู่คำและประโยคตัวอย่างให้มีความถูกต้องแม่นยำ เพื่อให้ได้การแปลที่มีประสิทธิภาพในงานวิจัยนี้จะเลือกใช้การแปลภาษาด้วยเครื่องแบบใช้ฐานกฎ ซึ่งเป็นวิธีเริ่มต้นที่มักจะใช้ในการแปลคู่ภาษาใหม่ๆ นอกจากนั้นการแปลภาษาด้วยเครื่องแบบใช้ฐานกฎไม่จำเป็นต้องใช้ฐานข้อมูลคู่ภาษาขนาดใหญ่ ซึ่งทำให้ง่ายต่อการพัฒนาเพื่อแปลประโยคพื้นฐานของคู่ภาษานั้นๆ

**2.2 งานวิจัยที่เกี่ยวข้อง**

ศรายุทธ ปลัดกอง ได้เสนอการพัฒนาโปรแกรมการถอดเสียงอักษรไทยเป็นอักษรโรมันตามวิธีการของราชบัณฑิตยสถาน โดยนำคำ วลี หรือประโยค มาแยกคำตามพจนานุกรมเพื่อแบ่งพยางค์ โดยใช้วิธีการจัดกลุ่มตัวอักขระ และนำกลุ่มอักขระไปแยกคำโดยวิธีเปรียบเทียบคำกับพจนานุกรมแบบฟอร์เวิร์ดและแบ็กเวิร์ด จะได้ผลลัพธ์เป็นหน่วยเสียงระดับพยางค์ แล้วจึงนำแต่ละพยางค์ไปถอดเสียงเป็นอักษรโรมันตามวิธีการของราชบัณฑิตยสถาน และเพิ่มสัญลักษณ์แทนระดับเสียงเพื่อใช้แทนการออกเสียงวรรณยุกต์เพื่อให้การออกเสียงมีความถูกต้องมากขึ้น ผลการทดลองพบว่ามีความถูกต้อง 99.95 %จากคำในพจนานุกรมจำนวน 21,399 คำ

ชุลีกร กิตติกูล ได้เสนอการถอดคำแบบถ่ายเสียงสำหรับชื่อบุคคลภาษาไทยที่เขียนด้วยอักษรโรมัน การถอดคำแบบถ่ายเสียงสำหรับแต่ละคำสามารถสร้างได้จากกฎ หรือใช้แบบจำลอง ทางสถิติ หรือค้นจากพจนานุกรม อย่างไรก็ตามการขาดมาตรฐานและความหลายหลากของการแปลงชื่อบุคคลไทยให้เป็นชื่อ ที่เขียนด้วยอักษรโรมันเป็นงานที่ท้าทาย และแม้ว่าวิธีที่ใช้พจนานุกรมเหมือนจะให้ผลที่ค่อนข้างถูกต้องที่สุด แต่ส่วนของการแปลงตัวอักษรเป็นเสียงก็ยังมีความจำเป็นสำหรับคำที่ไม่พบในพจนานุกรม งานวิจัยนี้เสนอวิธีการถอดคำแบบถ่ายเสียงสำหรับชื่อบุคคลภาษาไทยที่เขียนด้วยอักษรโรมันให้เป็นเสียงภาษาไทย โดยคำนึงถึงความนิยมในการใช้งาน ชื่อบุคคลภาษาไทยที่เขียนด้วยอักษรโรมันจะถูกแบ่งให้เป็นสายลำดับของแกรมโดยใช้พจนานุกรมแกรมสะสมซึ่งถูกสร้างจากชื่อมากกว่า 130,000 ชื่อ ผลการศึกษาพบว่าวิธีนี้ให้ความถูกต้องของ

อัจฉราภรณ์ ธาตุวิสัย ได้เสนองานการศึกษาวิเคราะห์คีตโควินทกาวยะ โดยมีวัตถุประสงค์ คือ 1. เพื่อปริวรรตและแปลคีตโควินทกาวยะเป็นภาษาไทย 2. เพื่อศึกษาวิเคราะห์เนื้อหาของคีตโควินทกาวยะ และ 3 .เพื่อศึกษาวิเคราะห์วรรณศิลป์ของคีตโควินทกาวยะ ต้นฉบับที่ใช้ในการศึกษาครั้งนี้คือต้นฉบับภาษาสันสกฤต ตัวอักษรเทวนาครีจากหนังสือ คีตโควินทะของชยเทพ ตีพิมพ์โดย สัตสันฆะ เสวะ สะมิถิ ไม่ระบุปี ที,พิมพ์ การวิจัยแบ่งเป็น 2 ส่วน คือ การวิเคราะห์เนื้อหาใน 3 หัวข้อ ได้แก่ การร้อยเรียงเรื่อง องค์ประกอบของเรื่องและแนวคิดของเรื่องและการวิเคราะห์วรรณศิลป์ โดยใช้ทฤษฎีอลังการศาสตร์ 2 ทฤษฎี ได้แก่ทฤษฎีอลังการ และทฤษฎีรส โดยมีขั้นตอนการศึกษา คือ ค้นคว้าและรวบรวมข้อมูลต่างๆ ที่เกี่ยวข้อง ปริวรรตคีตโควินทกาวยะจากอักษรเทวนาครีเป็นอักษรไทย และแปลจากภาษาสันสกฤตเป็นภาษาไทย จากนั้นจึงศึกษาวิเคราะห์ในด้านเนื้อหาและด้านวรรณศิลป์ แล้วอภิปรายและสรุปผลการศึกษา