**บทที่ 1**

**บทนำ**

**1.1 ความสำคัญและที่มาของงานวิจัย**

ภาษาสันสกฤตเป็นภาษาโบราณของอินเดีย และเป็นภาษาที่เก่าแก่ที่สุดภาษาหนึ่งในภาษากลุ่มอินโด-ยูโรเปียน (Indo-European) ภาษาสันสกฤตเป็นภาษาในกลุ่มภาษาเดียวกับภาษาบาลี ทั้งสองภาษาเป็นภาษาที่มาจากอินเดียที่มีอิทธิพลอย่างมากต่อภาษาพื้นเมืองในภูมิภาคเอเชียตะวันออกเฉียงใต้ตั้งแต่สมัยโบราณ โดยเฉพาะภาษาสันสกฤตนั้นดังจะเห็นได้จากบันทึกและหลักศิลาจารึกโบราณถูกบันทึกเป็นภาษาสันสกฤต ในบริบทของภาษาไทยเอง แม้จะเป็นภาษาคนละตระกูลกับภาษาบาลีและสันสกฤต แต่ภาษาไทยยืมคำภาษาทั้งสองนี้มาใช้เป็นจำนวนมาก และมักจะเรียกคู่กันเสมอว่าเป็น “ภาษาบาลีสันสกฤต” ส่วนคำในภาษาสันสกฤตนั้นถูกนำมาใช้อย่างกว้างขวาง ทั้งในมิติทางด้านศาสนา วัฒนธรรม ประเพณี งานวิชาการ อีกทั้งวรรณคดีสันสกฤตหลายเรื่องมีอิทธิพลต่อภูมิภาคแห่งนี้ อาทิ รามายณะ มหาภารตะ เป็นต้น การศึกษาภาษาสันสกฤตจึงมีคุณประโยชน์ในการเข้าใจภาษา วรรณคดี รวมทั้งพื้นฐานวัฒนธรรมไทยในด้านต่างๆที่มีความเกี่ยวพันกับอินเดียมาแต่โบราณ ส่วนระบบการเขียนภาษาบาลีและสันสกฤตมีความหลากหลายเนื่องจากเป็นภาษาที่ไม่มีระบบการเขียนเป็นของตนเอง แต่จะเขียนด้วยอักษรหลายชนิด เช่นในการตีพิมพ์งานภาษาสันสกฤตส่วนมากมักจะใช้อักษรเทวนาครี ซึ่งเป็นอักษรในกลุ่มอักษรอินเดียหรือไม่ก็ใช้อักษรโรมัน ส่วนภาษาบาลีก็มักใช้ตามเขียนด้วยอักษรหลายชนิดตามแต่ประเทศของผู้ที่นับถือพุทธศาสนาแบบเถรวาท แต่เดิมการบันทึกภาษาบาลีและสันสกฤตในอาณาเขตประเทศไทยปัจจุบันนั้น บันทึกด้วยตัวอักษรต่างๆกันไปตามยุคสมัย ใช้อักษรเช่น อักษรปัลลวะ อักษรมอญโบราณ อักษรขอมโบราณ อักษรขอมไทย และยังไม่มีการบันทึกภาษาบาลีและสันสกฤตเป็นอักษรไทย[1] จวบจนใน พ.ศ. 2436 สมัยพระบาทสมเด็จพระจุลจอมเกล้าเจ้าอยู่หัว ได้ทรงโปรดให้ จัดทำระบบปริวรรตอักษร จากอักษรขอมไทยเป็นอักษรไทยและยังได้นำอักษรโรมันที่ใช้เขียนมาเทียบกับอักษรไทยที่เป็นต้นแบบไว้ด้วย และทรงโปรดให้มีการปริวรรตอักษรจากพระไตรปิฎกบาลีอักษรขอมไทยเป็นอักษรไทยและตีพิมพ์เป็นครั้งแรก [2],[3] และต่อมาก็ได้มีการปรับปรุงการปริวรรตอีกเล็กน้อย การเขียนภาษาบาลีและสันสกฤตด้วยอักษรไทยจึงเป็นที่นิยมใช้จนถึงปัจจุบัน เพราะสามารถศึกษาเข้าใจได้ง่าย  
 ประเทศไทยในปี 2559 มีสถานศึกษาที่เปิดการเรียนการสอนภาษาสันสกฤต อยู่ 5 แห่งคือ 1.มหาวิทยาลัยศิลปากร 2.จุฬาลงกรณ์มหาวิทยาลัย 3.มหาวิทยาลัยเชียงใหม่ 4.มหามกุฏราชวิทยาลัย และ 5.มหาจุฬาลงกรณราชวิทยาลัย โดยเปิดการเรียนการสอนที่มีเนื้อหาเกี่ยวกับปรัชญา ศาสนา ภาษา วรรณคดี ประวัติศาสตร์และจารึกในงานโบราณคดี ส่วนสภาพการศึกษาภาษาสันสกฤตในประเทศไทยนั้น เมื่อเปรียบเทียบกับการศึกษาภาษาบาลี การศึกษาภาษาสันสกฤตจะมีผู้ศึกษาน้อยกว่าหลายเท่า เนื่องจากพุทธศาสนาในไทยนั้นเป็นแบบเถรวาทใช้ภาษาบาลีเป็นหลัก ทั้งมีหลักสูตรการเรียนการสอนภาษาบาลีอยู่หลายระดับและสถานศึกษาที่ทำการสอนอยู่ทั่วประเทศ อาทิ มหาวิทยาลัยสงฆ์ วัด สำนักเรียนบาลี เอกสารตำราบาลีก็ล้วนปริวรรตออกมาเป็นอักษรไทย แปลเป็นภาษาไทยจำนวนมาก ดังนั้นภาษาบาลีในไทยจึงมีผู้เชี่ยวชาญ อยู่เป็นจำนวนมาก ขณะเดียวกันในทางตรงข้ามภาษาสันสกฤตมีการศึกษาอยู่ในวงที่ไม่กว้างนัก สถานศึกษาที่ทำการสอนก็มีอยู่น้อย เอกสารและตำราโดยมากนั้นเป็นอักษรเทวนาครี หรือไม่ก็ปริวรรตเป็นอักษรโรมัน มีฉบับแปลเป็นภาษาอังกฤษ การปริวรรตออกเป็นอักษรไทย หรือแปลเป็นภาษาไทย มีน้อยมาก[4] อาจจะกล่าวได้ว่าปัจจัยเหล่านี้เป็นข้อจำกัดและอุปสรรคต่อการศึกษาภาษาสันสกฤต จึงทำให้มีผู้เชี่ยวชาญด้านภาษาสันสกฤตในประเทศไทยน้อยมาก และผู้เชี่ยวชาญด้านภาษาสันสกฤตที่มีชื่อเสียงของประเทศไทยในปัจจุบัน คือ สมเด็จพระเทพรัตนราชสุดาฯ สยามบรมราชกุมารี สนพระราชหฤทัยสันสกฤตศึกษา และทรงสำเร็จการศึกษาสาขาจารึกภาษาตะวันออก คณะโบราณคดี มหาวิทยาลัยศิลปากร ทั้งยังทรงส่งเสริมและสนับสนุนด้านสันสกฤตศึกษามาโดยตลอด[5][6] อีกทั้งสภาวัฒนธรรมสัมพันธ์แห่งอินเดีย (Indian Council for Cultural Relations: ICCR) ซึ่งเป็นองค์กรภายใต้กระทรวงต่างประเทศอินเดีย จะทูลเกล้าฯ ถวายรางวัล “World Sanskrit Award” ในช่วงเดือนพฤศจิกายน ปี พ.ศ. 2559 ในฐานะที่ทรงมีบทบาทส่งเสริมภาษาสันสกฤตในต่างประเทศ [7]

ในปัจจุบันความสนใจในการศึกษาภาษาสันสกฤตในประเทศไทยก็ยังคงน้อยอยู่ แต่ถ้าเปรียบเทียบกับอดีตนั้นก็นับว่า ภาษาสันสกฤตได้รับความสนใจมากขึ้นกว่าในอดีต เป็นผลมาจากเทคโนโลยีสื่อสังคมออนไลน์(Social Media) ที่สามารถมีพื้นที่ให้ผู้สนใจในภาษาสันสกฤตการจัดตั้งกลุ่มสนทนาได้โดยง่าย สมาชิกไม่เพียงเฉพาะแต่นักศึกษาและคณาจารย์ผู้ทำการเรียนการสอนภาษาสันสกฤตเท่านั้น ยังมีผู้ที่สนใจในศาสนา วรรณคดี ภาษาศาสตร์ ประวัติศาสตร์ และกลุ่มครูสอนภาษาไทย[8] จนไปถึงผู้เชี่ยวชาญภาษาอื่นๆ อีกทั้งข้อมูลสื่อภาษาสันสกฤตที่อยู่ในเครือข่ายอินเตอร์เน็ตนั้นมีอยู่เป็นจำนวนมาก โดยเฉพาะ เอกสารภาษาสันสกฤตออนไลน์ ทำให้เอื้อต่อการศึกษาของ นักศึกษา และผู้สนใจ ให้สะดวกยิ่งขึ้น แต่เอกสารภาษาสันสกฤตดังกล่าว นั้นมีหลากหลายตัวอักษรแต่โดยมากเป็น อักษรเทวนาครี ดังนั้นการศึกษาข้อมูลและเอกสารภาษาสันสกฤตนี้จะประกอบด้วย การปริวรรตอักษร(Transliteration) และการแปล(Translation) ซึ่งการปริวรรตอักษร คือ การทับศัพท์แบบถอดอักษร จะการดำเนินการแปลงระบบการเขียนหรือภาษาหนึ่งไปสู่อีกระบบหนึ่งอย่างมีหลักการ การปริวรรตอักษรต่างจากการถอดเสียง (Transcription) คือ การปริวรรตอักษรเคร่งครัดในการคงตัวอักษร ไม่สนใจความแตกต่างของเสียงในภาษา การปริวรรตตัวอักษรเทวนาครีเพื่อใช้ในการศึกษาภาษาสันสกฤตในภาษาไทยนั้นจะปริวรรตเป็น 2 ลักษณะ ได้แก่ ปริวรรตเป็นอักษรโรมันและอักษรไทย ซึ่งแต่ละลักษณะนั้นมีประโยชน์แตกต่างกันไป การปริวรรตตัวอักษรเทวนาครีเป็นตัวอักษรโรมันมีความเป็นสากลในการศึกษาและจำเป็นในการศึกษาภาษาสันสกฤตในระดับที่สูงขึ้นต่อไป ส่วนการปริวรรตตัวอักษรเทวนาครีเป็นตัวอักษรไทยนั้นจะทำให้ผู้ศึกษาที่เป็นคนไทยสามารถเรียนรู้และเข้าใจได้ง่าย[9] ในด้านงานค้นคว้าทางวิชาการภาษาสันสกฤต หากยังไม่รวมขั้นตอนการแปล การปริวรรตอักษรนี้ก็ยังถือว่าเป็นงานที่ใหญ่พอสมควร เอกสารภาษาสันสกฤต เช่น คัมภีร์ต่างๆนั้นมีเนื้อหาจำนวนมาก ผู้ศึกษานั้นจะต้องใช้เวลาและความวิริยะอุตสาหะในการปริวรรตอักษรและพิสูจน์อักษรเป็นอย่างมาก

จากที่มาและปัญหาดังกล่าว การศึกษาโครงการเฉพาะเรื่อง จึงได้เสนอแนวทางการประยุกต์ใช้เทคโนโลยีสารสนเทศ “การปริวรรตอักษรด้วยเครื่องสำหรับภาษาสันสกฤตแบบอักษรโรมันเป็นอักษรไทย” ขึ้นมาเพื่อช่วยแก้ปัญหานี้ ปริวรรตอักษรจะสามารถทำได้รวดเร็ว พร้อมกับตารางเปรียบเทียบอักษรให้สามารถตรวจสอบผลลัพธ์และพิสูจน์อักษรได้สะดวกขึ้น และสามารถปริวรรตชุดอักษรอินเดีย(Indian Script) และอักษรโรมันมาตรฐานต่างๆ(Romanization) ได้หลายชุด เนื่องจากมีซอฟแวร์ชุดไลบรารี่ที่เป็นโอเพนซอร์สอยู่หลายชุด โดยผู้ศึกษาได้เลือกชุดไลบรารี่โอเพนซอร์ส ชื่อว่า “SANSCRIPT” [10],[11] โดยจะทำการปริวรรตอักษรภาษาสันสกฤต อักษรอินเดียและอักษรโรมันมาตรฐานต่างๆ ให้เป็น อักษรโรมันตามการปริวรรตอักษรภาษาสันสกฤตสากล หรือ IAST (International Alphabet of Sanskrit Transliteration) จากนั้นผู้ศึกษาทำการศึกษาวิธีการปริวรรตอักษรโรมันมาตรฐาน IAST ให้เป็นอักษรไทย 2 รูปแบบ คือ 1. ปริวรรตอักษรแบบคงรูป คือการ ปริวรรตอักษรแบบตรงตามรูปศัพท์เดิมโดยใช้เครื่องหมายอื่นๆร่วมกำกับด้วย เพื่อให้ถอดกลับเป็นอักษรโรมันหรืออักษรเทวนาครีได้ถูกต้อง ส่วนมากใช้ในหนังสือหรือเอกสารทางวิชาการ 2.ปริวรรตอักษรแบบปรับรูป โดยปรับให้เข้ากับอักขรวิธีไทย เพื่อให้ออกเสียงได้ง่ายและรูปคำกลมกลืนกับภาษาไทย เพื่อความสะดวกในการอ่านสำหรับบุคคลทั่วไป พบในบทสวดมนต์ต่างๆ ทั้งนี้การศึกษาโครงการเฉพาะเรื่อง จะประเมินความสมบูรณ์ของการปริวรรตอักษร โดยผู้เชี่ยวชาญภาษาสันสกฤต ทดสอบและเปรียบเทียบผลลัพธ์ที่ได้จากระบบ

**1.2 วัตถุประสงค์ของการศึกษาโครงการเฉพาะเรื่อง**

การศึกษาโครงการเฉพาะเรื่อง การปริวรรตอักษรด้วยเครื่องสำหรับภาษาสันสกฤตแบบอักษรโรมันเป็นอักษรไทย มีวัตถุประสงค์ดังต่อไปนี้

1. เพื่อศึกษาวิธีการที่จะทำให้ผู้ใช้ที่มีความสนใจในการศึกษาภาษาสันสกฤตสามารถปริวรรตภาษาสันสกฤตแบบโรมันเป็นไทยได้อย่างรวดเร็ว
2. เพื่อศึกษาและนำเสนอกฎ (Rule-based) ขั้นตอนวิธี (Algorithm) ในการปริวรรตอักษรภาษาสันสกฤตแบบโรมันเป็นไทย
3. เพื่อพัฒนาการปริวรรตอักษรด้วยเครื่องสำหรับภาษาสันสกฤตแบบโรมันเป็นไทย

**1.3 ประโยชน์และผลที่ได้รับจากการศึกษาโครงการเฉพาะเรื่อง**

ประโยชน์และผลที่ได้รับจากการศึกษาโครงการเฉพาะเรื่อง การปริวรรตอักษรด้วยเครื่องสำหรับภาษาสันสกฤตแบบอักษรโรมันเป็นอักษรไทย มีดังต่อไปนี้

1. เอกสารภาษาสันสกฤตแบบต่าง มีอยู่อย่างมากมายในอินเตอร์เน็ต การปริวรรตอักษรด้วยเครื่องสำหรับภาษาสันสกฤตแบบโรมันเป็นไทย จะช่วยสร้างคุณค่าให้กับเอกสารเหล่านั้น สำหรับผู้ที่สนใจที่ไม่ถนัดการอ่านอักษรในรูปแบบต่างๆ ให้เป็นอักษรไทยได้
2. ปริวรรตเอกสารภาษาสันสกฤตจะทำได้อย่างรวดเร็ว จากเดิมที่ใช้เวลานานในการปริวรรต
3. การปริวรรตอักษรด้วยเครื่องสำหรับภาษาสันสกฤตแบบโรมันเป็นไทย จะเป็นเครื่องมือหนึ่งที่สนับสนุนการศึกษาภาษาสันสกฤตทั้งผู้สนใจและนักศึกษาในสาขาที่เกี่ยวข้อง

**1.4 ขอบเขตงานการศึกษาโครงการเฉพาะเรื่อง**

การศึกษาโครงการเฉพาะเรื่อง การปริวรรตอักษรด้วยเครื่องสำหรับภาษาสันสกฤตแบบอักษรโรมันเป็นอักษรไทย มีการกำหนดขอบเขตของงานวิจัยดังต่อไปนี้

1. การศึกษาโครงการเฉพาะเรื่อง ในส่วนพัฒนานี้ ในส่วนตัวปริวรรตอักษร(Transliteration Machine) มีขอบเขตทางเทคนิคดังต่อไปนี้
   1. พัฒนาเป็นแบบ Web Application
   2. ภาษาโปรแกรม ใช้ภาษา PHP รุ่น 5.6
   3. PHP Framework (MVC) ใช้ Codeigniter รุ่น 3.0.5
   4. Front-End Framework ใช้ Bootstrap v3.3.5
2. การศึกษาโครงการเฉพาะเรื่องนี้ จะทำการพัฒนาเพียงการปริวรรตภาษาสันสกฤตแบบอักษรโรมันมาตรฐาน International Alphabet of Sanskrit Transliteration (IAST) เท่านั้น เนืองจากมีชุดไลบรารี่ ฟังก์ชันบางส่วนอยู่บนพื้นฐานงาน SANSCRIPT ที่ http://learnsanskrit.org ในชุดไลบรารี่จะทำปริวรรตภาษาสันสกฤตแบบอักษรอินเดียชนิดต่างและอักษรโรมันมาตรฐานต่างๆ ดังนั้นประเมินความสมบูรณ์ของการปริวรรตอักษรจะไม่ประเมินในส่วนนี้
3. การศึกษาโครงการเฉพาะเรื่องนี้ มีลักษณะขั้นตอนการปริวรรตอยู่ 2 แบบ คือ
   1. ทำการปริวรรตอักษรภาษาสันสกฤตจาก แบบอักษรโรมันมาตรฐาน IAST เป็น อักษรไทยในรูปแบบทั่วไป(แบบปรับรูป)
   2. ทำการปริวรรตอักษรภาษาสันสกฤตจาก แบบอักษรโรมันมาตรฐาน IAST เป็น อักษรไทยในรูปแบบแบบแผน(แบบคงรูป)
4. ประเมินความสมบูรณ์ของการปริวรรตอักษร โดยวิธีการทดสอบและเปรียบเทียบผลลัพธ์ที่ได้จากระบบกับตัวอย่างการปริวรรต และผู้เชี่ยวชาญภาษาสันสกฤตประเมินผลการใช้งาน

**1.5 กระบวนการศึกษาโครงการเฉพาะเรื่อง**

กระบวนการศึกษาโครงการเฉพาะเรื่อง การปริวรรตอักษรด้วยเครื่องสำหรับภาษาสันสกฤตแบบอักษรโรมันเป็นอักษรไทย มีขั้นตอนดังต่อไปนี้

1. ศึกษาและทำความเข้าใจหลักการทฤษฎีและงานวิจัยรวมถึงเทคโนโลยีที่เกี่ยวข้อง ได้แก่ หลักการปริวรรตภาษาสันสกฤตแบบโรมันเป็นไทย ตัวอย่างงานวิจัยการปริวรรตอักษรด้วยเครื่อง และเทคโนโลยีที่ใช้ในการพัฒนาทั้งที่เป็น เครื่องมือพัฒนา ภาษาโปรแกรม ชุดไลบรารี่ ฟังก์ชัน เฟรมเวิร์ค เพื่อให้ทราบถึงการทำงาน และแนวทางในการพัฒนาการปริวรรตอักษรด้วยเครื่อง
2. จัดเตรียม จัดหาแหล่งข้อมูลและเก็บรวบรวมข้อมูลไว้สำหรับการศึกษาและการทดสอบระบบ ได้แก่ ตัวอย่างสำเนาเอกสารภาษาสันสกฤต ตัวอย่างการปริวรรตอักษรภาษาสันสกฤตแบบโรมันเป็นไทย
3. วิเคราะห์ข้อมูลที่เก็บรวบรวมมาได้และกำหนดภาพรวมของงาน นำข้อมูลที่วิเคราะห์ได้มา เลือกรูปแบบการพัฒนาและเทคโนโลยี ให้เหมาะสม
4. ออกแบบและกำหนด กฎ (Rule-based) ขั้นตอนวิธี (Algorithm) ในการปริวรรตอักษร
5. พัฒนาตัวต้นแบบการปริวรรตอักษร (Prototype) เพื่อตรวจสอบความเป็นไปได้ของการพัฒนาส่วนตัวปริวรรตอักษร
6. พัฒนาส่วนตัวปริวรรตอักษร(Transliteration Machine) ให้สามารถทำงานได้ตามกฎ (Rule-based) และขั้นตอนวิธี (Algorithm) ที่ได้ออกแบบไว้ อีกทั้งทำการทดสอบ ตรวจสอบผลลัพธ์ ภายใต้คำแนะนำของผู้เชี่ยวชาญภาษาสันสกฤต
7. พัฒนาส่วนต่อผู้ใช้งาน(User Interface) ที่ไว้สำหรับสำหรับในการป้อนอักษรต้นฉบับจากผู้ใช้ และส่วนการแสดงการเปรียบเทียบตรวจสอบจากอักษรต้นฉบับกับอักษรผลลัพธ์จากการปริวรรต
8. ทดสอบและประเมินผลการใช้งานจากผู้เชี่ยวชาญภาษาสันสกฤต
9. สรุปผลการวิจัยและข้อเสนอแนะ
10. จัดทำและเสนอรายงานการวิจัย

**บทที่ 2**

**ทฤษฎีและงานวิจัยที่เกี่ยวข้อง**

เนื้อหาในส่วนนี้นำเสนอทฤษฎีและงานวิจัยที่เกี่ยวข้องกับ การปริวรรตอักษรด้วยเครื่องสำหรับภาษาสันสกฤตแบบอักษรโรมันเป็นอักษรไทย มีดังต่อไปนี้

* 1. **ทฤษฎีที่เกี่ยวข้อง**

**2.1.1 การปริวรรตอักษรหรือการเขียนคำทับศัพท์ (transliteration)**

การปริวรรตอักษร หรือการเขียนคำทับศัพท์ [11][12] คือการดำเนินการแปลงอักษรหรืออักขรวิธีจากระบบการเขียนหรือภาษาหนึ่งไปสู่อีกระบบหนึ่งอย่างมีหลักการ เมื่อปริวรรตแล้วสามารถแปลงกลับเป็นอักษรหรืออักขรวิธีเดิมได้ 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[14]

(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) [15]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 [16] ต่อมารูปแบบการปริวรรตนี้ มีความสำคัญทางวิชาการภาษาสันสกฤตมาก จึงได้เป็น “การปริวรรตอักษรภาษาสันสกฤตสากล” หรือ “The International Alphabet of Sanskrit Transliteration” (IAST) เพราะเป็นมาตรฐานหลักการปริวรรตอักษรภาษาสันสกฤตเป็นอักษรโรมัน (Romanization) จนถึงปัจจุบัน



รูปที่ 2.4 บางส่วนในรายงานมติที่ประชุม เรื่องการปริวรรตอักษรภาษาสันสกฤตและบาลี   
 สากล ในการประชุมผู้เชี่ยวชาญด้านตะวันออกศึกษานานาชาติ ครั้งที่ 10 [16]

(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

  
รูปที่ 2.5 การปริวรรตอักษรภาษาสันสกฤตแบบโรมัน IAST เป็นไทยรูปแบบต่างๆ

การปริวรรตอักษรภาษาสันสกฤตแบบโรมัน IAST เป็นไทยนั้นผู้ปริวรรตต้องมีความรู้ความเข้าใจเกี่ยวกับรูปแบบอักษรในแต่ละส่วนมีตำแหน่งและหน้าที่อย่างไรจึงจะปริวรรตอักษรได้อย่างถูกต้อง 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**  **(ฐานกรณ์)** | **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**  **(ฐานกรณ์)** | **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**  **(ฐานกรณ์)** | **อันตัสถะ(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 ที่สะท้อนเสียงซ้ำของสระข้างหน้า | 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 หรืออาจจะต้องอาศัยหลักการปริวรรตอักษรจากกลุ่มภาษาเดียวกันมีการปริวรรตที่เป็นแบบแผนและมาตรฐานอยู่แล้ว เช่น ภาษาบาลี ภาษาฮินดี และจะสามารถจำแนกวิธีการการปริวรรตอักษรภาษาสันสกฤตแบบโรมัน เป็นอักษรไทยในรูปแบบต่างๆได้ดังนี้

**1. วิธีการปริวรรตเป็นอักษรไทยแบบคงรูป หรือ แบบแผน**

ลักษณะการปริวรรตเป็นอักษรไทยแบบคงรูป หรือ แบบแผน การปริวรรตอักษรแบบตรงตามรูปศัพท์เดิมโดยใช้เครื่องหมายพินทุด้วย เพื่อให้ถอดกลับเป็นอักษรโรมันหรืออักษรเทวนาครีได้ถูกต้องและได้ลักษณะอักขรวิธีได้ใกล้เคียงกับคำเดิม นั้นเมื่อเปรียบเทียบกับการปริวรรตแบบปรับรูปนั้น จะพบเอกสารที่ใช้แบบคงรูปมากกว่าเนื่องจากถูกใช้ในวงการวิชาการ และมีหลักการปริวรรตที่เป็นแบบแผนและมีหลักการแต่อย่างไรก็ดี หลักการที่ว่าก็ไปตามแต่ละตำราอาจารย์ ยังไม่เป็นมาตรฐานเดียวกันเสียทีเดียว หลักการยังไม่เป็นกฎเกณฑ์ที่แน่นอนตายตัวบางตำราบางอาจารย์อาจจะแตกต่างไปเล็กน้อย ดังสรุปได้ต่อไปนี้

1. การถอดรูปพยัญชนะ
   1. การถอดรูปพยัญชนะอักษรโรมันปกติ ให้ถอดตามตารางเปรียบเทียบพยัญชนะ
2. การถอดรูปสระ
   1. ถอดรูปสระ 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 |
| แบบไม้หันอากาศมีนิคหิตข้างบน | ทสฺยูัํรธรัํา | ตาลัํโลกานฺ |
| แบบ Moon dot | ทสฺยู̐รธร̐า | ตาล̐โลกานฺ |

* 1. รูปเครื่องหมายวิสรรคะ อักษรโรมันใช้ ḥ มี 2 แนวคิด คือ ถอดเป็น วิสรรชนีย์ (ะ) และ แบบอักษร ห ใส่ทัณฑฆาต (ห์) แบบนี้เป็นพระราชวินิจฉัยของพระบาทสมเด็จพระมงกุฎเกล้าเจ้าอยู่หัวที่ทรงใช้แทนวิสรรชนีย์ (ะ) [18] ลักษณะการถอด ให้ถอดลำดับหลังสระอักษรโรมัน 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. วิธีการปริวรรตเป็นอักษรไทยแบบปรับรูป หรือ แบบทั่วไป**

วิธีการปริวรรตเป็นอักษรไทยแบบปรับรูป หรือ แบบทั่วไป เป็นการปริวรรตอักษรโดยปรับให้เข้ากับอักขรวิธีไทย เช่น โดยไม่ใช้พินทุ และเพิ่มวิสรรชนีย์หรือใส่เครื่องหมายทัณฑฆาต เพื่อให้อ่านได้ง่ายขึ้นและรูปคำกลมกลืนกับภาษาไทย จึงนิยมใช้การทับศัพท์อย่างง่าย บางรูปแบบที่พบอาจมีรูปแบบการถอดเสียง (transcription) ผสมอยู่ด้วย วิธีการปริวรรตแบบนี้ยังไม่มีหลักการปริวรรตที่ชัดเจน แล้วแต่ผู้ใช้จะทำการปริวรรตตามแนวทางของตน หรือถ้าหากใช้ในลักษณะทางการอาจจะต้องอาศัยหลักการปริวรรตอักษรจากกลุ่มภาษาเดียวกันที่มีแบบแผนการปริวรรตที่เป็นแบบแผนและมาตรฐานอยู่แล้ว เช่น หลักการปริวรรตอักษรไทยแบบปรับรูปจากภาษาบาลี ยกตัวอย่างเช่น บทสวดมนต์สำหรับบุคคลทั่วไป หรือใช้หลักการเขียนคำทับศัพท์ภาษาฮินดี ตามหลักเกณฑ์การทับศัพท์ของราชบัณฑิตยสถาน[17] รูปแบบวิธีการปริวรรตเป็นอักษรไทยแบบปรับรูปจะกล่าวแบบกว้างๆ ดังสรุปได้ต่อไปนี้

1. การถอดรูปพยัญชนะ
   1. การถอดรูปพยัญชนะอักษรโรมันปกติ ให้ถอดตามตารางเปรียบเทียบพยัญชนะ
2. การถอดรูปสระ
   1. ถอดรูปสระ ให้ใช้ลักษณะเดียวกันกับแบบปริวรรตเป็นอักษรไทยแบบคงรูป
   2. ถอดรูปสระ a (รูป สระอะ) ถอดรูปโดยเติม วิสรรชนีย์ (ะ) ตัวอย่างตาม ตารางที่ 2.6

ตารางที่ 2.19 ตัวอย่างการถอดรูปสระ a

|  |  |  |
| --- | --- | --- |
| karaṇa | daśabala | śatakrata |
| กะระณะ | ทะศะพะละ | ศะตะกระตะ |

* 1. ถอดรูปสระ e ai o au (รูปสระ เอ ไอ โอ เอา) ที่ประสมอยู่พยัญชนะต้นอักษรโรมัน ให้ใช้ลักษณะเดียวกันกับแบบปริวรรตเป็นอักษรไทยแบบคงรูป
  2. หากสระอักษรโรมันขึ้นต้นคำ ให้ใช้ลักษณะเดียวกันกับแบบปริวรรตเป็นอักษรไทยแบบคงรูป และหากเป็นสระ a (รูปสระ อะ) ให้ใช้ อักษร อ เติมวิสรรชนีย์ (อะ) หากตามด้วยพยัญชนะอักษรโรมันที่ไม่มีสระอักษรโรมันกำกับหรือตามมาให้เป็นตัวสะกด ให้ใช้ อักษร อ เติมไม้หันอากาศ (อั)

ตารางที่ 2.20 ตัวอย่างการถอดรูปสระ a ขึ้นต้นคำ

|  |  |  |  |
| --- | --- | --- | --- |
| akaraṇa | อะกะระณะ | agni | อัคนี |

1. ถอดรูปสระ a กับพยัญชนะตัวสะกด (การใช้ วิสรรชนีย์,ไม้หันอากาศ, ร หัน)
   1. หากพยัญชนะอักษรโรมันประสมกับสระ a (รูปสระ อะ) ตามด้วยพยัญชนะอักษรโรมันที่ไม่มีสระอักษรโรมันกำกับหรือตามมาให้เป็นตัวสะกด ให้ถอดเป็นไม้หันอากาศ

ตารางที่ 2.21 ตัวอย่างการถอดรูปพยัญชนะที่ประสมสระ a ตามด้วยพยัญชนะ

|  |  |  |
| --- | --- | --- |
| sarva | cakra | sattvā |
| สรฺว | จักระ | สัตตวา |

* 1. กรณี ร-เรผะ (รฺ) คือพยัญชนะอักษรโรมันประสมกับสระ a (รูปสระ อะ) ตามด้วยพยัญชนะอักษรโรมันตัว r ที่ไม่มีสระอักษรโรมันกำกับหรือตามมา (ar) พบ 3 วิธี ให้ใช้ คือ ยังถอดเป็นไม้หันอากาศ กับวิธีที่โบราณจารย์ท่านนำมาภาษาสันสฤตมาใช้ในภาษาไทย ท่านถอดเป็น ร- หัน (รร) วิธีนี้ใช้ในหลักการเขียนคำทับศัพท์ภาษาฮินดีของราชบัณฑิตยสถาน[17] และใช้แบบเป็น ร ตัวเดียว

ตารางที่ 2.22 ตัวอย่างการถอดรูปพยัญชนะที่ประสมสระ a ตามด้วย r

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | sarva | dharma | karma | garbha |
| แบบ ร- หัน | สรรวะ | ธรรมะ | กรรมะ | ครรภะ |
| แบบไม้หันอากาศ | สัรวะ | ธัรมะ | กัรมะ | คัรภะ |
| แบบเป็น ร ตัวเดียว | สรวะ | ธรมะ | กรมะ | ครภะ |
| ตัวอย่างใช้ในภาษาไทย | สรรพ | ธรรม | กรรม | ครรภ์ |

* 1. กรณีพยัญชนะอักษรโรมันประสมกับสระ a (รูปสระ อะ) ตามด้วยพยัญชนะอักษรโรมันตัว v ที่ไม่มีสระอักษรโรมันกำกับหรือตามมา (av) เนื่องจากถอดเป็นไทยแล้วจะไปพ้องรูปสระอัว พบ 2 วิธี ให้ใช้ คือ ถอดเป็นวิสรรชนีย์ (ะ) กับ แบบตัดทิ้ง วิธีนี้ใช้ในหลักการเขียนคำทับศัพท์ภาษาฮินดีของราชบัณฑิตยสถาน[17]

ตารางที่ 2.23 ตัวอย่างการถอดรูปพยัญชนะที่ประสมสระ a ตามด้วย v

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | bhavyā | tālavya | yamavya | gavyū |
| แบบ วิสรรชนีย์ | ภะวยา | ตาละวยะ | ยะมะวยะ | คะวยู |
| แบบ ตัดทิ้ง | ภวยา | ตาลวยะ | ยะมวยะ | ควยู |

* 1. กรณีพยัญชนะอักษรโรมันประสมกับสระ a (รูปสระ อะ) ตามด้วยพยัญชนะอักษรโรมันตัว h ที่ไม่มีสระอักษรโรมันกำกับหรือตามมา (ah) พบ 2 วิธี คือ ยังถอดเป็นไม้หันอากาศ กับ แบบตัดทิ้ง วิธีนี้ใช้ในหลักการเขียนคำทับศัพท์ภาษาฮินดีของราชบัณฑิตยสถาน[17] พบการใช้ในภาษาบาลีทั้งสองแบบ

ตารางที่ 2.24 ตัวอย่างการถอดรูปพยัญชนะที่ประสมสระ a ตามด้วย h

|  |  |  |
| --- | --- | --- |
|  | brahma | mahyān |
| แบบ ไม้หันอากาศ | พรัหมะ | มัหยาน |
| แบบ ตัดทิ้ง | พรหมะ | มหยาน |

1. การถอดรูปเครื่องหมายและสัญลักษณ์พิเศษ
   1. รูปเครื่องหมายอนุสวาระ(ṃ)และ เครื่องหมายจันทรพินทุหรืออนุนาสิกะ(m̐) การถอดรูปนั้นมักจะถอดเหมือนกันทั้งสองเครื่องหมาย พบมีการใช้อักษร ง และ อักษร ม ในการถอดรูปเครื่องหมายดังกล่าว โดยเฉพาะการใช้อักษร ง หลักการเขียนคำทับศัพท์ภาษาฮินดีของราชบัณฑิตยสถาน ที่ให้ถอดรูปเครื่องหมายทั้งสอง เป็นอักษร ง [17] หรือหลักในภาษาบาลีการถอดรูปเครื่องหมายอนุสวาระ(ภาษาบาลีไม่มีเครื่องหมายจันทรพินทุ) กรณีใช้ปกตินั้นก็เป็นถอด อักษร ง แต่ในภาษาสันสกฤตเครื่องหมายทั้งสองนั้นออกเสียงในกรณีปกติคล้าย ม แต่เสียงจะขึ้นนาสิกเล็กน้อย ดังนั้นจึงใช้อักษร ม

การถอดรูปแบบทั้งหมดที่กล่าวมายังมีใช้อยู่โดยทั่วไป อย่างไรก็ตามภาษาบาลีและสันสันสกฤตยังมีกฎไวยากรณ์ในการออกเสียงเครื่องหมายดังกล่าวมีลักษณะที่ใกล้เคียงกัน เรียกว่า การสนธิ โดยในภาษาบาลี เรียกเฉพาะเจาะจงว่า นิคหิตสนธิ ต่อไปจะกล่าวถึงหลักการการสนธิถอดรูปเครื่องหมายอนุสวาระ(ṃ)และเครื่องหมายจันทรพินทุโดยจะกล่าวถึงเฉพาะเครื่องหมายอนุสวาระ ส่วนเครื่องหมายจันทรพินทุให้ใช้หลักการเดียวกัน ในภาษาสันสกฤตมีหลักการดังนี้

* + 1. กรณีใช้เครื่องหมายอนุสวาระ ไม่มีพยัญชนะโรมันอื่นตามติดมาหรือไม่มีการสนธิคำ เครื่องหมายอนุสวาระให้ถอดเป็น อักษร ม อยู่กับสระใดให้เป็นตัวสะกดสระนั้น

ตารางที่ 2.25 ตัวอย่างการถอดรูปอนุสวาระที่ไม่มีพยัญชนะโรมันอื่นตามติดมา

|  |  |  |
| --- | --- | --- |
| siddhaṃ | śaraṇaṃ | teṣāṃ |
| สิทธัม | ศะระณัม | เตษาม |

* + 1. กรณีใช้เครื่องหมายอนุสวาระและมีพยัญชนะวรรคโรมันตามติดมา คำนั้นมีการเข้าสนธิ เครื่องหมายอนุสวาระให้ถอดเป็นพยัญชนะนาสิก(อนุนาสิกะ)ในวรรคของตัวที่ตามมา อยู่กับสระใดให้เป็นตัวสะกดสระนั้น

(ดูรายละเอียดที่ อนุนาสิกะ ตารางที่ 2.2 การเปรียบเทียบรูปพยัญชนะวรรค)

ตารางที่ 2.26 ตัวอย่างการถอดรูปอนุสวาระแบบสนธิมีพยัญชนะวรรคตาม

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **คำศัพท์** | **พยัญชนะตาม** | | **วรรค** | **พยัญชนะนาสิก** | | **ถอดเป็น** |
| **โรมัน** | **ไทย** | **โรมัน** | **ไทย** |
| saṃga | g | ก | ก วรรค | ṅ | ง | สังคะ |
| paṃca | c | จ | จ วรรค | ñ | ญ | ปัญจะ |
| paṃḍita | ḍ | ฑ | ฏ วรรค | ṇ | ณ | ปัณฑิตะ |
| saṃtoṣa | t | ต | ต วรรค | n | น | สันโตษะ |
| saṃbhava | bh | ภ | ป วรรค | m | ม | สัมภะวะ |

* + 1. กรณีใช้เครื่องหมายอนุสวาระและมีพยัญชนะอวรรคโรมันที่เป็นอูษมะ ได้แก่ ś ṣ s h ตามติดมา ยังเป็นที่ถกเถียงกันอยู่แนวทางใหญ่ๆมี 2 แนวทาง ส่วนแนวทางอื่นๆนั้นจะใช้ 2 แนวทางนี้ปนกัน แนวทางทั้งสองมีรายละเอียดดังนี้
       1. แนวทางแรก หากมีพยัญชนะอวรรคโรมันที่เป็นแบบอูษมะ คำนั้นมีการเข้าสนธิ เครื่องหมายอนุสวาระให้ถอดเป็นพยัญชนะนาสิก(อนุนาสิกะ)ในฐานที่เกิดเสียงของตัวที่ตามมา อยู่กับสระใดให้เป็นตัวสะกดสระนั้น

ตารางที่ 2.27 ตัวอย่างการถอดรูปอนุสวาระแบบสนธิมีพยัญชนะตามแบบอูษมะ

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **คำศัพท์** | **พยัญชนะตาม** | | **ฐานกรณ์** | **พยัญชนะนาสิก** | | **ถอดเป็น** |
| **โรมัน** | **ไทย** | **โรมัน** | **ไทย** |
| siṃha | h | ห | กัณฐยะ | ṅ | ง | สิงหะ |
| saṃśaya | ś | ศ | ตาลวยะ | ñ | ญ | สัญศะยะ |
| saṃṣkā | ṣ | ษ | มูรธันยะ | ṇ | ณ | สัณษกา |
| Saṃskṛta | s | ส | ทันตยะ | n | น | สันสกฤตะ |

* + - 1. แนวทางที่สอง หากมีพยัญชนะอวรรคโรมันที่เป็นแบบอูษมะ ไม่มีการสนธิคำ เครื่องหมายอนุสวาระให้ถอดเป็นอักษร ม อยู่กับสระใดให้เป็นตัวสะกดสระนั้น

ตารางที่ 2.28 ตัวอย่างการถอดรูปอนุสวาระแบบไม่สนธิมีพยัญชนะตามแบบอูษมะ

|  |  |  |  |
| --- | --- | --- | --- |
| **คำศัพท์** | **พยัญชนะตาม** | | **ถอดเป็น** |
| **โรมัน** | **ไทย** |
| siṃha | h | ห | สิมหะ |
| saṃśaya | ś | ศ | สัมศะยะ |
| saṃṣkā | ṣ | ษ | สัมษกา |
| Saṃskṛta | s | ส | สัมสกฤตะ |

* + 1. กรณีใช้เครื่องหมายอนุสวาระและมีพยัญชนะอวรรคโรมันที่เป็นแบบอันตัสถะ ได้แก่ y r l v ตามติดมา ไม่มีการสนธิคำ เครื่องหมายอนุสวาระให้ถอดเป็นอักษร ม อยู่กับสระใดให้เป็นตัวสะกดสระนั้น

ตารางที่ 2.29 ตัวอย่างการถอดรูปอนุสวาระและมีพยัญชนะตามแบบอันตัสถะ

|  |  |  |  |
| --- | --- | --- | --- |
| **คำศัพท์** | **พยัญชนะตาม** | | **ถอดเป็น** |
| **โรมัน** | **ไทย** |
| saṃyoga | y | ย | สัมโยคะ |
| saṃrakṣa | r | ร | สัมรักษะ |
| saṃlāpa | l | ล | สัมลาปะ |
| saṃvatsara | v | ว | สัมวัตสะระ |

* 1. รูปเครื่องหมายวิสรรคะ ยังถกเถียงกันอยู่มาก ใช้หลากหลายกับตัวอักษร ห , ฮ เช่น ใช้พินทุ หฺ ฮฺ ใช้ทัณฑฆาต ห์ ฮ์ ใช้ วิสรรชนีย์ หะ ฮะ แต่แนวคิดหลังไม่ถูกต้องเท่าไรนัก เพราะเครื่องหมายวิสรรคะ ออกเสียงลมหายใจมาก ที่สะท้อนเสียงซ้ำของสระข้างหน้า แต่สรุปวิธีการที่ยอมรับกัน ได้มี 2 แนวคิด คือ แบบใช้อักษร ห ใส่ทัณฑฆาต (ห์) วิธีนี้เป็นพระราชวินิจฉัยของพระบาทสมเด็จพระมงกุฎเกล้าเจ้าอยู่หัว[19] และใช้ในหลักการเขียนคำทับศัพท์ภาษาฮินดีของราชบัณฑิตยสถาน[17] อีกแบบคือใช้อักษร ฮ หรือ ห แล้วถอดเสียง (transcription) เสียงซ้ำที่สะท้อนของสระข้างหน้า ตารางที่ 2.13

ตารางที่ 2.30 ตัวอย่างการถอดรูปเครื่องหมายวิสรรคะ

|  |  |  |  |
| --- | --- | --- | --- |
|  | แบบ ห ใส่ทัณฑฆาต | แบบถอดเสียง(ฮ) | แบบถอดเสียง(ห) |
| duḥka | ทุห์ขะ | ทุฮุขะ | ทุหุขะ |
| rāmaiḥ | ราไมห์ | ราไมฮิ | ราไมหิ |
| agniḥ | อัคนิห์ | อัคนิฮิ | อัคนิหิ |
| hetoḥ | เหโตห์ | เหโตโฮ | เหโตโห |
| namaḥ | นะมะห์ | นะมะฮะ | นะมะหะ |

* 1. รูปเครื่องหมายอวครหะ โดยมากตัดทิ้ง
  2. รูปเครื่องหมาย । (ทัณฑะ) ॥ (ทวิทัณฑะ) ให้ใช้ลักษณะเดียวกันกับแบบปริวรรตเป็นอักษรไทยแบบคงรูป

1. การถอดรูปพยัญชนะสังโยค หรือ พยัญชนะสังยุกต์ ที่ประสมกับรูปสระ e ai o au (รูปสระ เอ ไอ โอ เอา) ในแบบปรับรูปนั้นก็ยังไม่มีแบบแผนในการกำหนด แล้วแต่ผู้ปริวรรตเห็นว่าสมควร รูปแบบที่พบบ่อยเช่น หากมีการในภาษาไทยอยู่แล้วก็ใช้ลักษณะนั้น

ตารางที่ 2.31 ตัวอย่างการถอดรูปพยัญชนะสังโยคแบบคำที่ใช้ในภาษาไทย

|  |  |  |
| --- | --- | --- |
| śreṣṭha | śloka | kṣetra |
| เศรษฐะ | โศลก | เกษตระ |

หรือพยัญชนะสังโยคตัวนั้นมีลักษณะซ้อนคล้ายคำควบกล้ำในภาษาไทย(ซ้อนกับ อักษร ร ล ว) ให้ใช้แบบคำควบกล้ำ

ตารางที่ 2.32 ตัวอย่างการถอดรูปพยัญชนะสังโยคแบบคำควบกล้ำไทย

|  |  |  |
| --- | --- | --- |
| traividyā | draupadī | kleśa |
| ไตรวิทยา | เทราปะที | เกลศะ |

* 1. **ทฤษฎีการปริวรรตอักษรด้วยเครื่อง**
     1. **การปริวรรตอักษรด้วยเครื่อง**

แนวทางการปริวรรตอักษร(Machine Transliteration Approaches) สามารถแบ่งตามลักษณะการทำงานของระบบออกเป็น 2 กลุ่ม คือ

1. **การปริวรรตอักษรด้วยเครื่องแบบใช้ฐานกฎ (Rule-based Approach)**

เป็นปริวรรตอักษรด้วยเครื่องโดยใช้ความรู้ทางด้านกฎไวยากรณ์มาใช้ในการกำหนดกฎเกณฑ์ของระบบ ซึ่งจะต้องมีการแยกคุณลักษณะและข้อมูลทางกฎไวยากรณ์ของอักษรต้นฉบับจากนั้นจึงทำการวิเคราะห์ตามกฎไวยากรณ์ของอักษรเป้าหมาย ข้อจำกัดของการแปลภาษาด้วยเครื่องกลุ่มนี้คือ ผู้พัฒนาต้องมีความรู้ทางกฎไวยากรณ์ของทั้งสองตัวอักษรเป็นอย่างมาก

1. **การปริวรรตอักษรด้วยเครื่องแบบใช้สถิติ (Statistical Approach)**

เป็นการปริวรรตอักษรด้วยเครื่องโดยใช้วิธีการทางสถิติมาช่วยในการปริวรรต ซึ่งการปริวรรตกลุ่มนี้จำเป็นต้องมีฐานข้อมูลคู่ภาษาที่มีการจับคู่ประโยคในอักษรต้นฉบับและตัวอักษรเป้าหมาย เพื่อเป็นฐานความรู้ให้ระบบทำการเรียนรู้ และใช้ค่าทางสถิติ เอ็นแกรม (N-Gram Model) ซึ่งเป็นการคำนวณค่าของการที่คำเกิดขึ้นร่วมกัน ถ้าชุดคำชุดใดมีค่าเอ็นแกรมหรือค่าความน่าจะเป็น (Probability) สูง แสดงว่าชุดคำนี้มีโอกาสเกิดขึ้นร่วมกันบ่อยครั้ง การคำนวณหาค่าเอ็นแกรมของชุดคำที่มีอยู่ในฐานข้อมูลคู่อักษร ทำให้ได้ค่าความน่าจะเป็นของชุดคำต่างๆ เพื่อนำไปเปรียบเทียบและใช้ในการแปลข้อความได้ โดยสามารถเลือกใช้ค่าเอ็นแกรมได้ตั้งแต่ 2 คำ (bigrams), 3 คำ (trigrams) เป็นต้น ยิ่งใช้ค่าเอ็นแกรมมาก ยิ่งทำให้ภาษาต้นฉบับมีการได้หลากหลายและมีความละเอียดมากขึ้นข้อดีของวิธีการนี้คือ ผู้พัฒนาไม่จำเป็นต้องมีความรู้ทางด้านกฎไวยากรณ์ของทั้งสองตัวอักษรจึงทำให้ไม่เกิดปัญหาเรื่องไวยากรณ์ต่างๆ แต่ปริวรรตอักษรด้วยวิธีการนี้จำเป็นต้องมีฐานข้อมูลคู่อักษรที่มีจำนวนข้อมูลมหาศาลเพื่อหาค่าทางสถิติที่จะนำไปใช้เปรียบเทียบกับประโยคที่จะนำมาปริวรรตได้อย่างครอบคลุม

* + 1. **ประเภทของโมเดลปริวรรตอักษรด้วยเครื่อง**

ประเภทของโมเดลปริวรรตอักษรด้วยเครื่อง (Machine Transliteration Models) สามารถแบ่งตามลักษณะได้ 4 ประเภท คือ

1. **โมเดลแบบใช้ตัวอักษร (Grapheme Based Model)**

คือโมเดลการปริวรรตโดยจับคู่ตัวอักษรต้นฉบับกับตัวอักษรเป้าหมาย โดยไม่สนใจเกี่ยวกับการออกเสียงของคำที่ได้จากการปริวรรต มักใช้ในตัวอักษรในตระกูลเดียวกัน มีสระพยัญชนะใกล้เคียงกันหรือการปริวรรตที่ต้องการรักษารูปศัพท์เดิมไว้โดยไม่สนใจเสียงของตัวอักษรเป้าหมาย

1. **โมเดลแบบใช้เสียงตัวอักษร (Phoneme Based Models)**

คือโมเดลการปริวรรตโดยจับคู่เสียงของตัวอักษรต้นฉบับกับเสียงของตัวอักษรเป้าหมาย มักใช้ในกรณีที่ตัวอักษรต่างตระกูลกัน หรือมีสระพยัญชนะไม่เท่ากันอยู่มาก เป็นการปริวรรตที่ต้องการรักษาเสียงเดิมไว้ให้ใกล้เคียงที่สุดโดยไม่สนใจรูปศัพท์เดิมของตัวอักษรตัวอักษรต้นฉบับ

1. **โมเดลแบบผสม (Hybrid Model)**

คือโมเดลการปริวรรตที่ใช้โมเดลแบบใช้ตัวอักษรและโมเดลแบบใช้เสียงตัวอักษรรวมกันมักใช้ในกรณีที่ตัวอักษรต่างตระกูลกันแต่ต้องการรักษารูปศัพท์เดิมไว้ ถ้าส่วนใดมีสระพยัญชนะไม่ครบคู่ก็ใช้สระพยัญชนะที่เสียงใกล้เคียง หรือผสมให้เกิดเสียงคล้ายกันแทน

1. **โมเดลแบบสมนัย(Correspondence Based Model)**

คือโมเดลการปริวรรตที่คล้ายกับ โมเดลแบบผสม แต่ใช้ตัวอักษรต้นฉบับกับเสียงตัวอักษรต้นฉบับ ในการจับคู่เสียงสระพยัญชนะในอักษรเป้าหมาย ในกรณีที่ตัวอักษรเป้าหมายมีลักษณะเป็นคำหรือพยางค์ หรืออยู่ในลักษณะอักษรหนึ่งตัวประกอบไปด้วยสระพยัญชนะ เช่นการทับศัพท์หรือการปริวรรตอักษร อักษรโรมัน ภาษาอังกฤษ เป็น อักษรจีน ญี่ปุ่น เกาหลี เป็นต้น

* 1. **งานวิจัยที่เกี่ยวข้อง**

ศรายุทธ ปลัดกอง [20]ได้เสนอการพัฒนาโปรแกรมการถอดเสียงอักษรไทยเป็นอักษรโรมันตามวิธีการของราชบัณฑิตยสถาน โดยนำคำ วลี หรือประโยค มาแยกคำตามพจนานุกรมเพื่อแบ่งพยางค์ โดยใช้วิธีการจัดกลุ่มตัวอักขระ และนำกลุ่มอักขระไปแยกคำโดยวิธีเปรียบเทียบคำกับพจนานุกรมแบบฟอร์เวิร์ดและแบ็กเวิร์ด จะได้ผลลัพธ์เป็นหน่วยเสียงระดับพยางค์ แล้วจึงนำแต่ละพยางค์ไปถอดเสียงเป็นอักษรโรมันตามวิธีการของราชบัณฑิตยสถาน และเพิ่มสัญลักษณ์แทนระดับเสียงเพื่อใช้แทนการออกเสียงวรรณยุกต์เพื่อให้การออกเสียงมีความถูกต้องมากขึ้น ผลการทดลองพบว่ามีความถูกต้อง 99.95 %จากคำในพจนานุกรมจำนวน 21,399 คำ

ชุลีกร กิตติกูล [21]ได้เสนอการถอดคำแบบถ่ายเสียงสำหรับชื่อบุคคลภาษาไทยที่เขียนด้วยอักษรโรมัน การถอดคำแบบถ่ายเสียงสำหรับแต่ละคำสามารถสร้างได้จากกฎ หรือใช้แบบจำลอง ทางสถิติ หรือค้นจากพจนานุกรม อย่างไรก็ตามการขาดมาตรฐานและความหลายหลากของการแปลงชื่อบุคคลไทยให้เป็นชื่อ ที่เขียนด้วยอักษรโรมันเป็นงานที่ท้าทาย และแม้ว่าวิธีที่ใช้พจนานุกรมเหมือนจะให้ผลที่ค่อนข้างถูกต้องที่สุด แต่ส่วนของการแปลงตัวอักษรเป็นเสียงก็ยังมีความจำเป็นสำหรับคำที่ไม่พบในพจนานุกรม งานวิจัยนี้เสนอวิธีการถอดคำแบบถ่ายเสียงสำหรับชื่อบุคคลภาษาไทยที่เขียนด้วยอักษรโรมันให้เป็นเสียงภาษาไทย โดยคำนึงถึงความนิยมในการใช้งาน ชื่อบุคคลภาษาไทยที่เขียนด้วยอักษรโรมันจะถูกแบ่งให้เป็นสายลำดับของแกรมโดยใช้พจนานุกรมแกรมสะสมซึ่งถูกสร้างจากชื่อมากกว่า 130,000 ชื่อ ผลการศึกษาพบว่าวิธีนี้ให้ความถูกต้องของ

วินัย มาลีลัยและพุธษดี ศิริแสงตระกูล [22]ได้เสนอการแปลงภาษาบาลีอักษรไทยเป็นสัทอักษร โดยใช้พจนานุกรม (dictionary) ร่วมกับกฎบาลีไวยากรณ์ (Pali grammar) และกฎภาษาศาสตร์ (linguistic) และได้ใช้วิธีการตัดคำแบบเลือกคำที่ยาวที่สุด(longest matching algorithm) เพื่อใช้ในการคนหาความหมายของคำศัพท์ ในงานวิจัยนี้ผู้วิจัยได้สร้างพจนานุกรมภาษาบาลี-ไทย ประกอบด้วยคำศัพท์บาลี 14,323 คำ โดยอ้างอิงคำศัพท์บาลีจากหนังสือพจนานุกรมภาษาบาลี-ไทยของพระมหาไพโรจน พจนานุกรมที่สร้างขึ้นประกอบด้วยคำศัพท์ภาษาบาลีอักษรไทย ความหมายประเภทของคำศัพท์ คำในรูปภาษาไทย คำในรูปอักษรโรมันและคำในรูปสัทอักษร โดยจัดเก็บไวในระบบฐานข้อมูลและในงานวิจัยนี้ได้นำเสนอกฎในการแปลงภาษาบาลีเป็นสัทอักษรเพื่อใช้ในการแปลงคำที่ไม่ปรากฏในพจนานุกรมเพื่อทดสอบความถูกตองของระบบ ผู้วิจัยไดทำการทดสอบระบบที่พัฒนาโดยการแปลงเอกสารที่เขียนด้วยภาษาบาลีอักษรไทย ซึ่งประกอบด้วย บทสวดมนต์ 99 บท พุทธศาสนสุภาษิต 181 สุภาษิต และวินย. มหาวิภงฺโค 1 จากพระวินัยปิฎก เล่มที่ 1 ของพระไตรปิฎกภาษาบาลีฉบับสยามรัฐ ผลการทดสอบความถูกตองของการแปลงคิดเป็นร้อยละ 98.75

อัจฉราภรณ์ ธาตุวิสัย [23]ได้เสนองานการศึกษาวิเคราะห์คีตโควินทกาวยะ โดยมีวัตถุประสงค์ คือ 1. เพื่อปริวรรตอักษรและแปลคีตโควินทกาวยะเป็นภาษาไทย 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

**บทที่ 3**

**ขั้นตอนการดำเนินโครงการ**

ในบทนี้จะกล่าวถึงวิธีการออกแบบและขั้นตอนการพัฒนาการ การปริวรรตอักษรด้วยเครื่องสำหรับภาษาสันสกฤตแบบอักษรโรมันเป็นอักษรไทย โดยมีวิธีการดำเนินโครงการดังต่อไปนี้

1. ศึกษาหลักการทฤษฎีและงานวิจัยรวมถึงเทคโนโลยีที่เกี่ยวข้อง
2. จัดเตรียมแหล่งข้อมูลไว้สำหรับการศึกษาและการทดสอบระบบ
3. วิเคราะห์ข้อมูล กำหนดภาพรวมของงาน เลือกรูปแบบการพัฒนา
4. ออกแบบและกำหนด กฎ (Rule-based) ขั้นตอนวิธี (Algorithm)
5. พัฒนาตัวต้นแบบการปริวรรตอักษร (Prototype)
6. พัฒนาส่วนตัวปริวรรตอักษร(Transliteration Machine)
7. พัฒนาส่วนต่อผู้ใช้งาน(User Interface)
8. ทดสอบและประเมินผลการใช้งานจากผู้เชี่ยวชาญภาษาสันสกฤต
   1. **การศึกษาหลักการทฤษฎีและงานวิจัยรวมถึงเทคโนโลยีที่เกี่ยวข้อง**

ผู้จัดทำได้ทำการศึกษาและทำความเข้าใจหลักการทฤษฎีและงานวิจัยรวมถึงเทคโนโลยีที่เกี่ยวข้อง ได้แก่ หลักการปริวรรตภาษาสันสกฤตแบบอักษรโรมันเป็นอักษรไทย ตัวอย่างงานวิจัยการปริวรรตอักษรด้วยเครื่อง และเทคโนโลยีที่ใช้ในการพัฒนาทั้งที่เป็น เครื่องมือพัฒนา ภาษาโปรแกรม ชุดไลบรารี่ ฟังก์ชัน เฟรมเวิร์ค เพื่อให้ทราบถึงการทำงาน และแนวทางในการพัฒนาการปริวรรตอักษรด้วยเครื่อง โดยมีรายละเอียดต่อไปนี้

* + 1. **การศึกษาด้านทฤษฎีและงานวิจัยที่เกี่ยวข้อง**

1. **ศึกษาลักษณะการปริวรรตอักษรภาษาสันสกฤตแบบอักษรโรมันเป็นอักษรไทย**

เป็นการศึกษารวบรวมและสรุปแนวทางการปริวรรตอักษรภาษาสันสกฤตเป็นอักษรไทยในลักษณะต่างๆ จากตัวอย่างที่เกี่ยวกับการปริวรรตหรือภาษาสันสกฤตในงานวิจัยและเอกสารออนไลน์ อีกทั้งจากตัวอย่างที่ใช้กันในสังคมออนไลน์ที่มีเนื้อหาเกี่ยวกับภาษาสันสกฤต และจากการสอบถามผู้เชี่ยวชาญด้านภาษาสันสกฤตอย่างไรก็ตามหลักการยังไม่เป็นกฎเกณฑ์ที่แน่นอนตายตัวบางตำราบางอาจารย์อาจจะแตกต่างไป หรือมีการปรับเปลี่ยนกันเล็กน้อย หรืออาจจะต้องอาศัยหลักการปริวรรตอักษรจากกลุ่มภาษาเดียวกันมีการปริวรรตที่เป็นแบบแผนและมาตรฐานอยู่แล้ว เช่น ภาษาบาลี ภาษาฮินดี และสามารถสรุปลักษณะการปริวรรตจากการศึกษาได้ดังนี้

1. การปริวรรตอักษรแบบคงรูป หรือ แบบแผน คือการปริวรรตอักษรแบบตรงตามรูปศัพท์เดิมโดยใช้เครื่องหมายพินทุด้วย เพื่อให้ถอดกลับเป็นอักษรโรมันหรืออักษรเทวนาครีได้ถูกต้องและได้ลักษณะอักขรวิธีได้ใกล้เคียงกับคำเดิม

2. การปริวรรตอักษรแบบปรับรูป หรือ แบบทั่วไป คือหมายถึงการปริวรรตอักษรโดยปรับให้เข้ากับอักขรวิธีไทย เช่น โดยไม่ใช้พินทุ และเพิ่มวิสรรชนีย์หรือใส่เครื่องหมายทัณฑฆาต เพื่อให้อ่านได้ง่ายขึ้นและรูปคำกลมกลืนกับภาษาไทย จึงนิยมใช้การทับศัพท์อย่างง่าย

1. **ศึกษาด้านทฤษฎีแนวทางการพัฒนาการปริวรรตอักษร**

เป็นการศึกษาทฤษฎีทางเทคโนโลยีสารสนเทศที่เกี่ยวกับการปริวรรตอักษรว่ามีลักษณะการพัฒนาอย่างไรหลักการและแนวทางมีอะไรเพื่อใช้เป็นหลักในการพัฒนาที่เหมาะสมกับ การปริวรรตอักษรภาษาสันสกฤตแบบอักษรโรมันเป็นอักษรไทย ได้ทำการศึกษาแนวทางการปริวรรตอักษร(Machine Transliteration Approaches) และโมเดลปริวรรตอักษรด้วยเครื่อง (Machine Transliteration Models)

1. **ศึกษางานวิจัยอื่นที่เกี่ยวข้อง**

ศึกษาทางด้านเทคนิคการปริวรรต หรือกระบวนการที่ใกล้เคียงกันในงานวิจัยที่ผ่านมา เพื่อหาวิธีที่เหมาะสมต่อการพัฒนาและสามารถนำมาประยุกต์ใช้กับโครงการฉบับนี้

* + 1. **การศึกษาด้านเทคโนโลยีที่เกี่ยวข้อง**

ผู้จัดทำได้ทำการศึกษาและทำความเข้าใจเทคโนโลยีที่เกี่ยวข้อง ได้แก่ เครื่องมือพัฒนา ซอฟแวร์ ภาษาโปรแกรม ชุดไลบรารี่ ฟังก์ชัน เฟรมเวิร์ค ที่ช่วยพัฒนาส่วนเสริมการปริวรรตและสำหรับการทดสอบโดยมีรายละเอียดต่อไป

1. **ไลบรารี่การปริวรรตชุดอักษรอินเดียเป็นอักษรโรมันมาตรฐานต่างๆ**

เพื่อเพิ่มความสามารถให้ระบบเพื่อรองรับอักษรต้นฉบับภาษาสันสกฤตให้หลากหลายยิ่งขึ้น ผู้จัดทำได้ค้นหาไลบรารี่การปริวรรตชุดอักษรอินเดียเป็นอักษรโรมันมาตรฐานต่างๆ เนื่องจากมีซอฟแวร์และชุดไลบรารี่อยู่หลายชุด โดยผู้ศึกษาจะเลือกได้เลือกชุดไลบรารี่ ที่เป็นโอเพนซอร์ส โดยสามารถทำการปริวรรตอักษรภาษาสันสกฤต อักษรอินเดียให้เป็นอักษรโรมันมาตรฐานต่างๆ และสามารถแปลงไปกลับได้(Forward And Backward Transliteration) และทำการศึกษาวิธีใช้งานชุดไลบรารี่นั้น

1. **ซอฟแวร์สำหรับการเปรียบเทียบข้อความ**

ในการทดสอบความถูกต้องของการแปลนั้น มีส่วนการเปรียบเทียบผลลัพธ์จากระบบ นั้นต้องเปรียบเทียบกับตัวอย่างการปริวรรตซึ่งมีข้อความอยู่เป็นจำนวนมาก ดังนั้นผู้จัดทำจึงต้องค้นหา ซอฟแวร์สำหรับการเปรียบเทียบข้อความที่สามารถบอกความแตกต่างในระดับตัวอักษรได้

* 1. **การจัดเตรียมแหล่งข้อมูลไว้สำหรับการศึกษาและการทดสอบระบบ**

ผู้จัดทำได้ทำการจัดเตรียม จัดหาแหล่งข้อมูลและเก็บรวบรวมข้อมูลไว้สำหรับการศึกษาและการทดสอบระบบ ได้แก่ ตัวอย่างสำเนาเอกสารต้นฉบับภาษาสันสกฤต ตัวอย่างการปริวรรตอักษรภาษาสันสกฤตเป็นไทย โดยมีรายละเอียดต่อไป

* + 1. **แหล่งข้อมูลตัวอย่างสำเนาเอกสารต้นฉบับภาษาสันสกฤต**

ผู้จัดทำได้ทำการสืบค้นหาและเลือกแหล่งข้อมูลตัวอย่างสำเนาเอกสารต้นฉบับภาษาสันสกฤต เพื่อใช้ทดสอบการปริวรรต และใช้ประกอบการศึกษา โดยมีเงื่อนไขคือ เป็นข้อมูล ที่จัดเก็บเป็นข้อมูลอิเล็กทรอนิกส์ชนิด TEXT มีข้อมูลจัดเรียงกันเป็นหมวดหมู่และมีเอกสารอยู่หลากหลายชนิด โดยผู้จัดทำได้เลือกใช้แหล่งข้อมูลตัวอย่างสำเนาเอกสารต้นฉบับภาษาสันสกฤต จาก 2 แหล่ง เป็นแหล่งข้อมูลหลักในการศึกษาค้นคว้า ได้แก่

1. **Digital Sanskrit Buddhist Canon(DSBC)**

DSBC [24] ทำงานอยู่บนอินเตอร์เน็ตที่เว็บไซต์ www.dsbcproject.org โดยเป็นคลังเก็บข้อมูลเอกสารอิเล็กทรอนิกส์ที่เกี่ยวกับพระพุทธศาสนาในภาษาสันสกฤตหรือปรากฤต โดยที่ DSBC เป็นโครงการที่ทำการถอดความภาษาสันสกฤตหรือปรากฤตจากต้นฉบับตัวเขียน (Manuscript) และจัดเก็บเป็นข้อมูลอิเล็กทรอนิกส์ชนิด TEXT เป็นตัวอักษรเทวนา-ครีและตัวอักษรโรมันแบบ IAST โครงการ DSBC เป็นโครงการของมหาวิทยาลัยออฟ-เดอะเวสต์ (University of the West) รัฐแคลิฟอร์เนีย สหรัฐอเมริกา

1. **Göttingen Register of Electronic Texts in Indian Languages(GRETIL)**

GRETIL [25] ทำงานอยู่บนอินเตอร์เน็ตที่เว็บไซต์ gretil.sub.uni-goettingen.de โดยเป็นคลังเก็บข้อมูลเอกสารอิเล็กทรอนิกส์ในสาขาภารตวิทยา (Indology) โดยมีเอกสารทั้งภาษาสันสกฤต ปรากฤตและภาษาอินเดียอื่นๆ จากแหล่งข้อมูลจากในอินเดีย เอเชียกลาง และเอเชียตะออกเฉียงใต้ GRETIL เป็นโครงการของห้องสมุดกลางแห่งรัฐนีเดอร์ซัคเซินและมหาวิทยาลัยเกิททิงเงิน\* สหพันธ์สาธารณรัฐเยอรมนี

*\*หมายเหตุ ห้องสมุดกลางแห่งรัฐนีเดอร์ซัคเซินและมหาวิทยาลัยเกิททิงเงิน ใช้ชื่อภาษาอังกฤษ: Göttingen State and University Library ภาษาเยอรมัน: Niedersächsische Staats- und Universitätsbibliothek Göttingen หรือเรียกโดยย่อว่า SUB Göttingen[2]*

*มหาวิทยาลัยเกิททิงเงิน ใช้ชื่อในภาษาอังกฤษ : University of Göttingen ภาษาเยอรมัน :* *Georg-August-Universität Göttingen[3]*

* + 1. **ตัวอย่างการปริวรรตอักษรภาษาสันสกฤตเป็นไทย**

ผู้จัดทำได้รวบรวมตัวอย่างการปริวรรตอักษรภาษาสันสกฤตเป็นอักษรไทย เพื่อไว้ทำการศึกษาเปรียบเทียบ ถอดกฎการปริวรรต และทดสอบระบบเพื่อเปรียบเทียบผลลัพธ์ ตัวอย่างการปริวรรตส่วนหนึ่งมาจากภาคผนวกของงานวิจัยที่เกี่ยวข้องในด้านภาษาสันสกฤตและจากแหล่งข้อมูลออนไลน์ โดยมากเป็นงานที่ปริวรรตมาจากต้นฉบับอักษรเทวนาครี บางส่วนนั้นไม่แสดงข้อความต้นฉบับ ผู้จัดทำจำเป็นต้องหาต้นฉบับในแหล่งข้อมูลข้างต้น ตัวอย่างการปริวรรตอักษรภาษาสันสกฤตเป็นไทยที่ผู้จัดทำนำมาเป็นแนวทางหลักแบ่งเป็น 2 ประเภทได้แก่

1. **ตัวอย่างการปริวรรตอักษรภาษาสันสกฤตเป็นไทยแบบคงรูปหรือแบบแผน**

ผู้จัดทำใช้ตัวอย่างการปริวรรตอักษรภาษาสันสกฤตเป็นไทยแบบคงรูปหรือแบบแผน จาก 5 คัมภีร์เป็นแนวทางหลัก แต่ละคัมภีร์มีที่มาดังต่อไปนี้

* 1. **คัมภีร์มูลสรวาสติวาทวินยวัสตุ จีวรวัสตุและกรรมวัสตุ**

จากภาคผนวก ก ปริวรรตอักษรไทยของจีวรวสฺตุและกรฺมวสฺต ในวิทยานิพนธ์ของมหาวิทยาลัยศิลปากร ปี 2552 เรื่องการศึกษาเปรียบเทียบคัมภีร์มูลสรวาสติวาทวินยวัสตุกับพระวินัยปิฎก ศึกษาเฉพาะกรณีจีวรวัสตุและกรรมวัสตุ โดย พระมหาฉัตร ชัยมูลสาร[26]

* 1. **คัมภีร์ทิวยาวทาน**

จากภาคผนวก ข บทปริวรรตทิวยาวทาน ในวิทยานิพนธ์ของมหาวิทยาลัยศิลปากรปี 2556 เรื่องการศึกษาวิเคราะห์คัมภีร์ทิวยาวทานเรื่องที่ 1-19 โดย นางสาวปัทมา นาควรรณ[27]

* 1. **คัมภีร์โพธิจรรยาวตาร**

จากภาคผนวก ก คัมภีร์โพธิจรรยาวตาร ฉบับปริวรรต ในวิทยานิพนธ์ของมหาวิทยาลัยศิลปากร ปี 2549 เรื่อง การศึกษาเชิงวิเคราะห์คัมภีร์โพธิจรรยาวตาร โดย พระมหาวิชาญ กำเหนิดกลับ[28]

* 1. **คัมภีร์มูลมัธยมกการิกา**

จากภาคผนวก ก คัมภีร์มูลมัธยมกการิกา ฉบับปริวรรต ในวิทยานิพนธ์ของมหาวิทยาลัยธรรมศาสตร์ ปี 2554 เรื่อง การศึกษาเชิงวิเคราะห์การใช้เหตุผลแบบวิภาษวิธีในคัมภีร์มูลมัธยมกการิกา โดย นายกฤษฎา ภูมิศิริรักษ์[29]

* 1. **คัมภีร์ภควัทคีตา**

จากหนังสือภควัทคีตา(ฉบับโศลก )เกียรติขจร ชัยเธียร[30]

* 1. **คัมภีร์วัชรเฉทิกปรัชญาปารมิตาสูตร ฉบับอักษรไทย**

จาก Internet Archive: Digital Library โดย กรกิจ ดิษฐาน [31]

1. **ตัวอย่างการปริวรรตอักษรภาษาสันสกฤตเป็นไทยแบบปรับรูปหรือแบบทั่วไป**

เนื่องจากเอกสารประเภทนี้มีอยู่น้อยมาก และไม่ค่อยแพร่หลายเนื่องจากมีผู้ปริวรรตน้อยมาก ไม่ได้รับความนิยมเท่าภาษาบาลีที่มีผู้ปริวรรตเป็นบทสวดมนต์สำหรับบุคคลทั่วไปอยู่เป็นจำนวนมาก หรือหากพบก็พบเป็นข้อความหรือบทสั้นๆ ที่บุคคลทั่วที่สนใจปริวรรตกันขึ้นเองซึ่งปรากฏในสังคมออนไลน์และเว็บบอร์ดในกลุ่มที่สนใจในภาษาและวรรณคดีสันสกฤต กลุ่มที่สนใจหรือนับถือศาสนาพราหมณ์-ฮินดูและกลุ่มที่สนใจ ศาสนาพุทธแบบมหายาน ซึ่งส่วนมากก็เป็นในลักษณะการถอดเสียงไม่ใช่การปริวรรตอักษร ดังนั้นผู้จัดทำจึงใช้วิธีการ ใช้เอกสารที่เป็นบทสวดมนต์ภาษาบาลีที่มีอยู่ทั่วไปประกอบเข้าบางส่วนแต่ในที่นี้จะไม่กล่าวถึง

ส่วนเอกสารการปริวรรตประเภทนี้ที่ผู้จัดทำนำมาเป็นตัวอย่างเปรียบเทียบนั้นเป็นเอกสารที่เป็นผลงานการปริวรรตของคุณ กรกิจ ดิษฐาน ผู้ช่วยบรรณาธิการต่างประเทศและกีฬา หนังสือพิมพ์ M2F(2559) [32] ที่เผยแพร่ในอินเตอร์เน็ต ผู้จัดทำใช้ตัวอย่างการปริวรรตอักษรภาษาสันสกฤตเป็นไทยแบบปรับรูปหรือแบบทั่วไป 3 ตัวอย่างดังต่อไปนี้

* 1. **คัมภีร์ลลิตวิสตระ อัธยายที่ 26 ธรรมจักรประวรรตนสูตร**

จากห้องสมุดเสรีออนไลน์ วิกิซอร์ซ โดย กรกิจ ดิษฐาน [33]

* 1. **คัมภีร์วัชรเฉทิกปรัชญาปารมิตาสูตร ฉบับคำอ่านอักษรไทย**

จาก Internet Archive: Digital Library โดย กรกิจ ดิษฐาน [34]

* 1. **คัมภีร์ปรัชญาปารมิตาหฤทัยสูตร**

จากสารานุกรมเสรีออนไลน์ วิกิพีเดีย โดย กรกิจ ดิษฐาน [35]

* 1. **การวิเคราะห์ข้อมูล กำหนดภาพรวมของงาน เลือกรูปแบบการพัฒนา**

ผู้จัดทำได้ทำการวิเคราะห์ข้อมูลที่เก็บรวบรวมมาได้และกำหนดภาพรวมของงาน นำข้อมูลที่วิเคราะห์ได้มา เลือกรูปแบบการพัฒนาและเทคโนโลยี ให้เหมาะสม โดยมีรายละเอียดดังต่อไปนี้

* + 1. **สรุปวิธีการปริวรรตอักษรที่จะนำมาพัฒนา**

เนื่องจากกฎการปริวรรตอักษรภาษาสันสกฤตแบบอักษรโรมันเป็นอักษรไทยที่รวบรวมมานั้นบางส่วนมีความแตกต่างในการปริวรรตอยู่มาก อาจมีกฎอื่นๆที่ใช้กันแต่ผู้จัดทำยังรวบรวมมาไม่หมด หรือไม่กฎเหล่านั้นก็ยังคลุมเครืออยู่ยังไม่มีแนวทางหลักในการปริวรรตโดยเฉพาะการปริวรรตแบบปรับรูปมีปัญหามากที่สุด และเพื่อการพัฒนาให้เหมาะสมกับข้อจำกัดทางเวลา ผู้จัดทำจัดเป็นไปได้ยากที่จะจัดทำได้ครบทุกกฎ ดังนั้นจึงเลือกกฎจากทฤษฎีการปริวรรตอักษรภาษาสันสกฤตแบบอักษรโรมันเป็นอักษรไทยดังสรุปได้ต่อไปนี้

**วิธีการปริวรรตเป็นอักษรไทยแบบคงรูป หรือ แบบแผน**

* 1. การถอดรูปพยัญชนะ สระ เครื่องหมาย ตัวเลข หากมีวิธีเดียวให้ใช้ให้ใช้วิธีนั้น(อ้างอิงในบทที่ 2)
  2. ในกรณีถอดรูปพยัญชนะอักษรโรมันที่ประสมด้วยสระ ṛ ṝ ḷ ḹ (ฤ ฤๅ ฦ ฦๅ) มี 2 วิธี ให้ใช้ แบบไม่ใส่พินทุ
  3. รูปเครื่องหมายจันทรพินทุหรืออนุนาสิกะ มี 3 วิธี ให้ใช้ แบบไม้หันอากาศและมีนิคหิตข้างบน
  4. รูปเครื่องหมายวิสรรคะ มี 2 วิธี ให้ใช้ แบบวิสรรชนีย์ (ะ)
  5. รูปเครื่องหมายอวครหะ ไม่ถอดใช้ตามเดิม
  6. รูปเครื่องหมาย ทัณฑะและทวิทัณฑะ ถอดเป็น ฯ (อังคั่นเดี่ยว) ๚ (อังคั่นคู่)
  7. การถอดรูปพยัญชนะสังโยค หรือ พยัญชนะสังยุกต์ ถอดเฉพาะชุดพยัญชนะอักษรโรมัน ที่เป็นพยัญชนะสังโยคซ้อนหรือควบพยัญชนะโรมัน r (รูปพยัญชนะ ร)

**วิธีการปริวรรตเป็นอักษรไทยแบบปรับรูป หรือ แบบทั่วไป**

* 1. การถอดรูปพยัญชนะ สระ เครื่องหมาย ตัวเลข หากมีวิธีเดียวให้ใช้ให้ใช้วิธีนั้น(อ้างอิงในบทที่ 2)
  2. กรณี ร-เรผะ (รฺ) คือพยัญชนะอักษรโรมันประสมกับสระ a (รูปสระ อะ) ตามด้วยพยัญชนะอักษรโรมันตัว r ที่ไม่มีสระอักษรโรมันกำกับหรือตามมา (ar) มี 2 วิธี ให้ใช้ แบบ ร- หัน
  3. กรณีพยัญชนะอักษรโรมันประสมกับสระ a (รูปสระ อะ) ตามด้วยพยัญชนะอักษรโรมันตัว v ที่ไม่มีสระอักษรโรมันกำกับหรือตามมา (av) มี 2 วิธี ให้ใช้ แบบ วิสรรชนีย์
  4. กรณีพยัญชนะอักษรโรมันประสมกับสระ a (รูปสระ อะ) ตามด้วยพยัญชนะอักษรโรมันตัว h ที่ไม่มีสระอักษรโรมันกำกับหรือตามมา (ah) มี 2 วิธี ให้ใช้ แบบ ตัดทิ้ง
  5. ถอดรูปเครื่องหมายอนุสวาระและเครื่องหมายจันทรพินทุ พยัญชนะวรรคโรมันตามติดมา ให้ใช้ แบบมีการสนธิ
  6. ถอดรูปเครื่องหมายอนุสวาระและเครื่องหมายจันทรพินทุและมีพยัญชนะอวรรคโรมันที่เป็นอูษมะ ได้แก่ ś ṣ s h ตามติดมา ให้ใช้ แบบไม่มีการสนธิ
  7. รูปเครื่องหมายวิสรรคะ แบบ ห ใส่ทัณฑฆาต
  8. รูปเครื่องหมาย ทัณฑะและทวิทัณฑะ ถอดเป็น ฯ (อังคั่นเดี่ยว) ๚ (อังคั่นคู่)
  9. การถอดรูปพยัญชนะสังโยค หรือ พยัญชนะสังยุกต์ ที่ประสมกับรูปสระ e ai o au (รูปสระ เอ ไอ โอ เอา) ไม่มีแบบแน่นอน แต่ผู้จัดทำเห็นว่าสมควรว่าใช้แบบคงรูปคือ ถอดเฉพาะชุดพยัญชนะอักษรโรมัน ที่เป็นพยัญชนะสังโยคซ้อนหรือควบพยัญชนะโรมัน r (รูปพยัญชนะ ร)
     1. **สรุปแนวทางและโมเดลที่จะนำมาพัฒนา**

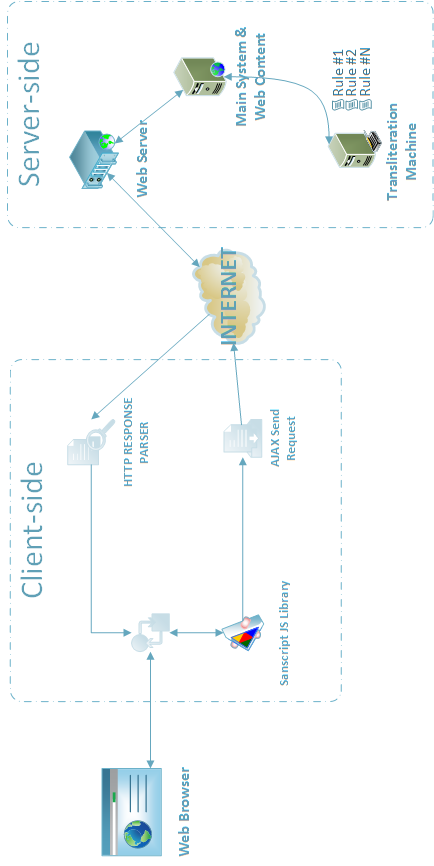
ในการเลือกแนวทางการปริวรรตอักษร(Machine Transliteration Approaches)เนื่องการศึกษาค้นคว้าจนได้รูปแบบกฎเกณฑ์การปริวรรต ผู้จัดทำจึงเลือกการปริวรรตอักษรด้วยเครื่องแบบใช้ฐานกฎ (Rule-based Approach)และในการเลือกโมเดลการปริวรรตอักษรด้วยเครื่อง (Machine Transliteration Models) ผู้จัดทำได้พิจารณาเห็นว่า ตัวอักษรโรมันแบบ IAST กับอักษรไทยนั้นครบคู่กัน อาจจะมาจากสาเหตุที่ไทยได้รับอิทธิพลภาษาสันสกฤตอย่างมาก ถึงแม้ว่าอักษรไทยบางตัวออกเสียงไม่เหมือนเสียทีเดียว แต่มองในรูปศัพท์แล้วจะให้ความหมายเดียวกัน ผู้จัดทำจึงเลือก โมเดลแบบใช้ตัวอักษร (Grapheme Based Model) ในการพัฒนา

* + 1. **สรุปเทคโนโลยีที่ช่วยพัฒนาส่วนเสริมการปริวรรตและการทดสอบ**

ไลบรารี่การปริวรรตชุดอักษรอินเดียเป็นอักษรโรมันมาตรฐานต่างๆ นั้นผู้จัดทำเลือกไลบรารี่ Sanscript [10],[11] เป็นไลบรารี่จาวาสคริปท์ที่สามารถทำการปริวรรตอักษรภาษาสันสกฤต อักษรอินเดียหลากหลายตัวอักษรให้เป็นอักษรโรมันมาตรฐานต่างๆ และสามารถแปลงไปกลับได้(Forward And Backward Transliteration) แม้ว่า Sanscript มุ่งเน้นการปริวรรตอักอักษรในภาษาสันสกฤต แต่ก็ยังสนับสนุนการปริวรรตสำหรับภาษาอินเดียอื่น ๆ อีกด้วยและยังสามารถติดตั้งและเรียกใช้งานโดยง่าย สำหรับซอฟแวร์สำหรับการเปรียบเทียบข้อความ ผู้จัดทำเลือก Pretty Diff v2.0.0 []ที่สามารถเปรียบเทียบข้อความเป็นจำนวนมากได้และสามารถเปรียบเทียบข้อความที่สามารถบอกความแตกต่างในระดับตัวอักษร พร้อมทั้งออกรายงานได้

* + 1. **สรุปภาพรวมของงาน**

สรุปภาพรวมของงานระบบจะทำงานในลักษณะ Server-Client โดยทำงานบนเว็บบราวเซอร์ โดยกระบวนการฝั่ง Client และการส่ง Request นั้น จะทำโดยผู้ใช้สามารถป้อนข้อความภาษาสันสกฤตโดยการเลือกชนิดตัวอักษรจากนั้น ไลบรารี่ Sanscript จะทำการปริวรรตอักษรชนิดต่าง ๆ และแสดงบนเว็บบราวเซอร์ได้ทันที่เนื่องจากเป็นไลบรารี่จาวาสคริปท์ที่จะทำงานในฝั่ง Client แต่หากจะต้องการปริวรรตอักษรชนิดอื่น ๆ เป็นภาษาไทยข้อความจะถูกส่งให้ไลบรารี่ Sanscript และปริวรรตชนิดตัวอักษรเหล่านั้นให้เป็นอักษรโรมันมาตรฐาน IAST แล้วส่งไปให้ Server โดยวิธีการ AJAX (Asynchronous JavaScript and XML) เมื่อระบบหลัก(Main System)รับคำขอดังกล่าวแล้วจะส่งข้อความให้กับ Transliteration Machine เพื่อทำการประมวลผลต่อไป ส่วนกระบวนการทางฝั่ง Server และ Response นั้น เมื่อ Transliteration Machine ได้ทำการประมวลผลจะนำข้อความไปผ่านกระบวนการตามกฎที่ได้กำหนดไว้ และคืนค่าเป็นผลการปริวรรตแบบคงรูปหรือแบบแผนและแบบปรับรูปหรือแบบทั่วไป โดยออกมาในรูปแบบ JSON (JavaScript Object Notation) จากนั้นก็ส่งต่อให้ระบบหลัก และส่งกลับไปยัง Client โดยมีตัวแปลง Response โดยจะทำหน้าที่แปลง JSON ที่ได้รับมาแปลงให้เป็นตารางเปรียบเทียบข้อความตามชนิดตัวอักษร และแสดงผลลัพธ์การปริวรรตบนเว็บบราวเซอร์



รูปที่ 3.1 สรุปภาพรวมของงาน

* 1. **การออกแบบและกำหนด กฎ (Rule-based) ขั้นตอนวิธี (Algorithm)**

ผู้จัดทำได้ทำการออกแบบและกำหนด กฎ (Rule-based) ขั้นตอนวิธี (Algorithm) ในการปริวรรตอักษร ดังมีรายละเอียดต่อไปนี้

* + 1. **สัญลักษณ์ประกอบขั้นตอนวิธีของกฎการปริวรรต**

โดยมีการกำหนดสัญลักษณ์เพื่อประกอบขั้นตอนวิธีของกฎการปริวรรต และในการกำหนดสัญลักษณ์ในที่นี้ให้ถือว่า อนุสวาระ จันทรพินทุ วิสรรคะ อวครหะ(ṁ,ṃ,ḥ,') เป็นพยัญชนะด้วย ดังมีรายละเอียดต่อไปนี้

ตารางที่ 3.1 สัญลักษณ์ประกอบขั้นตอนวิธีของกฎการปริวรรต

|  |  |
| --- | --- |
| **สัญลักษณ์** | **ชุดข้อมูลหรืออำอธิบายการทำงาน** |
| **TA** | พยัญชนะและสระอักษรไทยใด ๆ |
| **TS** | อักษรไทยใด ๆ ที่ไม่ใช้พยัญชนะและสระ |
| **TC** | พยัญชนะอักษรไทยใด ๆ |
| **TV** | สระอักษรไทยใด ๆ |
| **RNP** | ชุดตัวเลขและเครื่องหมายวรรคตอนอักษรโรมันได้แก่ {0,1 ,2 ,3 ,4 ,5 ,6 ,7 ,8 ,9 ,॥,।} |
| **TNP** | ชุดตัวเลขและเครื่องหมายวรรคตอนอักษรไทยได้แก่  {๐, ๑, ๒, ๓, ๔, ๕, ๖, ๗, ๘, ๙, ๚, ฯ} |
| **RMC** | ชุดพยัญชนะอักษรโรมันประกอบกัน 2 ตัวอักษรได้แก่  { kh,ch,ṭh,th,ph,gh,jh,ḍh,dh,bh } |
| **TMC** | ชุดพยัญชนะอักษรไทยที่แปลงมาจาก RMC ได้แก่  {ข,ฉ,ฐ,ถ,ผ,ฆ,ฌ,ฒ,ธ,ภ} |
| **RMV** | ชุดสระอักษรโรมันประกอบกัน 2 ตัวอักษรได้แก่ {ai,au} |
| **TMV** | ชุดสระอักษรไทยที่แปลงมาจาก RMV ได้แก่{ไ,เา} |
| **RSC** | ชุดพยัญชนะอักษรโรมันที่ตัวอักษรเดี่ยวได้แก่  {k,g,ṅ,c,j,ñ,ṭ,ḍ,ṇ,t,d,n,p,b,m,y,r,l,v,ḻ,ś,ṣ,s,h,ṁ,ṃ,ḥ,' } |
| **TSC** | ชุดพยัญชนะอักษรไทยที่แปลงมาจาก **RSC** ได้แก่  {ก,ค,ง,จ,ช,ญ,ฏ,ฑ,ณ,ต,ท,น,ป,พ,ม,ย,ร,ล,ว,ฬ,ศ,ษ,ส,ห,ँ,ํ,ะ,'} |
| **TSC2** | ชุดพยัญชนะอักษรไทยที่มาจาก **TSC** ใช้ในแบบปรับรูป แตกต่างที่ตำแหน่งวิสรรคะ(ห์){ก,ค,ง,จ,ช,ญ,ฏ,ฑ,ณ,ต,ท,น,ป,พ,ม,ย,ร,ล,ว,ฬ,ศ,ษ,ส,ห,ँ,ํ,ห์,'} |
| **RSV** | ชุดสระอักษรโรมันที่เป็นตัวอักษรเดี่ยวได้แก่  { a,ā,i,ī,u,ū,ṛ,ṝ,ḷ,ḹ,e,o } |
| **TSV** | ชุดสระอักษรไทยที่แปลงมาจาก **RSV** ใช้ในแบบคงรูป ได้แก่  { a,า,ิ,ี,ุ,ู,ฤ,ฤๅ,ฦ,ฦๅ,เ,โ } |
| **TSV2** | ชุดสระอักษรไทยที่มาจาก **TSV** ใช้ในแบบปรับรูป แตกต่างที่ตำแหน่งแรก ได้แก่{ะ,า,ิ,ี,ุ,ู,ฤ,ฤๅ,ฦ,ฦๅ,เ,โ } |
| **TIF** | ชุดสระอักษรไทยเมื่ออยู่ต้นคำต้องทำการเปลี่ยนแปลง ใช้ในแบบแบบคงรูป ได้แก่{ A ,า,ิ,ี,ุ,ู,เ,โ,ไ} |
| **TIFR** | ชุดสระอักษรไทยเมื่ออยู่ต้นคำเมื่อเปลี่ยนแปลงแล้ว ใช้ในแบบแบบคงรูปได้แก่ {อ, อา, อิ, อี, อุ, อู, เอ, โอ, ไอ} |
| **TIF2** | ชุดสระอักษรไทยชุดเดียวกันกับ **TIF-F** แต่ใช้ในแบบปรับรูป แตกต่างที่รูปตำแหน่งแรก ได้แก่ { ะ,า,ิ,ี,ุ,ู,เ,โ,ไ} |
| **TIFR2** | ชุดสระอักษรไทยชุดเดียวกันกับ **TIFR-F** แต่ใช้ในแบบปรับรูป แตกต่างที่รูปตำแหน่งแรก ได้แก่ {อะ, อา, อิ, อี, อุ, อู, เอ, โอ, ไอ} |
| **KVARGA** | ชุดพยัญชนะอักษรไทย วรรค ก{ก, ข, ค, ฆ, ง} |
| **CVARGA** | ชุดพยัญชนะอักษรไทย วรรค จ{จ, ฉ, ช, ฌ, ญ} |
| **THVARGA** | ชุดพยัญชนะอักษรไทย วรรค ฏ{ ฏ, ฐ, ฑ, ฒ, ณ} |
| **TVARGA** | ชุดพยัญชนะอักษรไทย วรรค ต{ต, ถ, ท, ธ, น} |
| **PVARGA** | ชุดพยัญชนะอักษรไทย วรรค ต{ป, ผ, พ, ภ, ม} |
| **F-RS**(X,Y,Z)  Replace String Function | X คือค่าพารามิเตอร์ที่เป็นตัวอักษรที่จะเปลี่ยน Y คือค่าพารามิเตอร์ที่เป็นตัวอักษรที่ไปแทนที่  Z คือค่าพารามิเตอร์ที่เป็นข้อความ  **F-RS**() คือฟังก์ชั่นการการแทนที่ตัวอักษร X ในข้อความ Zทั้งหมดด้วยตัวอักษร Y เช่น  X = C , Y = R , Z = CAT ดังนั้น RS(X,Y,Z) = RAT |
| **F-RSA** (X,Y,Z)  Replace String with Array Function | X คือค่าพารามิเตอร์ที่เป็นชุดตัวอักษรที่จะเปลี่ยน  Y คือค่าพารามิเตอร์ที่เป็นชุดตัวอักษรที่ไปแทนที่  X และ Y ต้องมีสมาชิกจำนวนเท่ากัน  Z คือค่าพารามิเตอร์ที่เป็นข้อความ  **F-RSA**() คือฟังก์ชั่นการการแทนที่ชุดตัวอักษร X ในข้อความ Z ทั้งหมดด้วยชุดตัวอักษร Y  X = {A,B} , Y = {APP,L} , Z = ABE เช่น ดังนั้น RSA(X,Y,Z) = APPLE |

* + 1. **กำหนดกฎการปริวรรตอักษรไทยแบบคงรูป**

ผู้จัดทำได้ทำการออกแบบและกำหนดกฎการปริวรรตอักษรไทยแบบคงรูป โดยแต่ละกฎจะทำงานตามลำดับเมื่อกฎสุดท้ายทำงานเสร็จจะได้ผลลัพธ์ ดังมีรายละเอียดกฎต่อไปนี้

ตารางที่ 3.2 กำหนดกฎการปริวรรตอักษรไทยแบบคงรูป

|  |  |  |
| --- | --- | --- |
| **ลำดับ** | **ชื่อกฎ** | **หน้าที่** |
| 1. | RomanChandrabinduToSingle | เปลี่ยนเครื่องหมายจันทรพินทุให้เป็นแบบตัวเดียว |
| 2. | RomanNumberAndPunctuate | เปลี่ยนตัวเลขและเครื่องหมายวรรคตอน |
| 3. | RomanMixConsonant | เปลี่ยนพยัญชนะโรมันแบบตัวผสม |
| 4. | RomanMixVowel | เปลี่ยนสระโรมันแบบตัวผสม |
| 5. | RomanSingleConsonant | เปลี่ยนพยัญชนะโรมันแบบตัวเดี่ยว |
| 6. | RomanSingleVowel | เปลี่ยนสระโรมันแบบตัวเดี่ยว |
| 7. | Bindu | ใส่จุดพินทุให้พยัญชนะไทย |
| 8. | ThaiVowelInFist | เปลี่ยนสระกรณีอยู่ต้นคำ โดยการเติมรูปพยัญชนะ อ ผสมเข้าไป |
| 9. | ThaiVowelPrefix | เลื่อนสระที่เป็นสระหน้าไปหน้าพยัญชนะ |
| 10. | ARemove | ลบสระ A |
| 11. | SwapAnusvaraAndChandrabindu | เลื่อนเครื่องหมายอนุสาวระอักษรไทยและจันทรพินทุอักษรโรมันไปข้างหน้ากรณีอยู่หลัง ลากข้าง(า) |
| 12. | Chandrabindu | แปลงเครื่องหมายจันทรพินทุ |
| 13. | ThaiAAInFist | เปลี่ยนลากข้าง(า) โดยผสมรูปพยัญชนะ อ กรณีอยู่ถัดรูปสระอื่น ซึ่งเป็นสระตัวสุดท้ายที่จะแปลง เนื่องจากต้องแปลงรูปสระเอา ก่อน |

* + 1. **กำหนดกฎการปริวรรตอักษรไทยแบบปรับรูป**

ผู้จัดทำได้ทำการออกแบบและกำหนดกฎการปริวรรตอักษรไทยแบบปรับรูป โดยแต่ละกฎจะทำงานตามลำดับเมื่อกฎสุดท้ายทำงานเสร็จจะได้ผลลัพธ์ ดังมีรายละเอียดกฎต่อไปนี้

ตารางที่ 3.3 กำหนดกฎการปริวรรตอักษรไทยแบบปรับรูป

|  |  |  |
| --- | --- | --- |
| **ลำดับ** | **ชื่อกฎ** | **หน้าที่** |
| 1. | AvagarahaRemove | ลบอวครหะ |
| 2. | RomanChandrabinduToSingle | เปลี่ยนเครื่องหมายจันทรพินทุให้เป็นแบบตัวเดียว |
| 3. | RomanNumberAndPunctuate | เปลี่ยนตัวเลขและเครื่องหมายวรรคตอน |
| 4. | RomanMixConsonant | เปลี่ยนพยัญชนะโรมันแบบตัวผสม |
| 5. | RomanMixVowel | เปลี่ยนสระโรมันแบบตัวผสม |
| 6. | RomanSingleConsonant | เปลี่ยนพยัญชนะโรมันแบบตัวเดี่ยว |
| 7. | RomanSingleVowel | เปลี่ยนสระโรมันแบบตัวเดี่ยว |
| 8. | AnusvaraAndChandrabindu | เข้าสนธิอนุสวาระและจันทรพินทุ |
| 9. | ThaiVowelInFist | เปลี่ยนสระกรณีอยู่ต้นคำ โดยการเติมรูปพยัญชนะ อ ผสมเข้าไป |
| 10. | ThaiVisarga | เปลี่ยนจัดการข้อความที่เป็นวิสรรชณีย์และไม้หันอากาศ |
| 11. | ThaiVowelPrefix | เลื่อนสระที่เป็นสระหน้าไปหน้าพยัญชนะ |
| 12. | ThaiAAInFist | เปลี่ยนลากข้าง(า) โดยผสมรูปพยัญชนะ อ กรณีอยู่ถัดรูปสระอื่น |
| 13. | convertAE | เลื่อนสระที่เป็นสระ เ ไปหน้าพยัญชนะหากตัวอักษรก่อนหน้าถูกแปลงเป็นไม้หันอากาศแล้วต้องแปลงกลับเป็นวิสรรชณีย์(ะ) |
| 14. | convertAO | เลื่อนสระที่เป็นสระ โ ไปหน้าพยัญชนะหากตัวอักษรก่อนหน้าถูกแปลงเป็นไม้หันอากาศแล้วต้องแปลงกลับเป็นวิสรรชณีย์(ะ) |

* + 1. **ขั้นตอนวิธีของกฎการปริวรรตอักษรไทยที่ใช้ร่วมกัน(กฎกลาง)**

ในขั้นตอนวิธีของกฎการปริวรรตอักษรไทยนั้นมีกฎที่ที่ใช้ร่วมกันอยู่ 9 กฎแต่อาจมีข้อมูลบางส่วนที่เป็นข้อมูลคนละชุดกัน เรียกว่า กฎกลาง ดังมีรายละเอียดต่อไปนี้

**กฎที่ 1. RomanChandrabinduToSingle**

กฎนี้ทำหน้าที่เปลี่ยนเครื่องหมายจันทรพินทุ จากการเขียนด้วยตัวอักษร m และเครื่องหมาย Dot Moonที่อยู่ด้านบน(m̐)ให้เป็นเครื่องหมายจันทรพินแบบตัวอักษรเดี่ยวคือ ṁ ออกแบบไว้เพื่อรองรับไว้สำหรับการใช้เครื่องหมายจันทรพินทุโรมันทั้งสองแบบ และเตรียมไว้สำหรับการแปลงต่อไปโดยมีขั้นตอนวิธีดังนี้

1. START
2. READ input
3. output = **F-RS**(‘m̐’ , ‘ṁ', input)
4. RETURN output
5. END

**กฎที่ 2. RomanNumberAndPunctuate**

กฎนี้ทำหน้าที่เปลี่ยนตัวเลขและเครื่องหมายวรรคตอนอักษรโรมันให้เป็นตัวเลขและเครื่องหมายวรรคตอนอักษรไทย โดยมีขั้นตอนวิธีดังนี้

1. START
2. READ input
3. output = **F-RSA**(**RNP** , **TNP**, input)
4. RETURN output
5. END

**กฎที่ 3. RomanMixConsonant**

กฎนี้ทำหน้าที่เปลี่ยนชุดพยัญชนะอักษรโรมันที่มีลักษณะอักษรที่เป็นคู่ ให้เป็นชุดพยัญชนะอักษรไทย โดยมีขั้นตอนวิธีดังนี้

1. START
2. READ input
3. output = **F-RSA**(**RMC**, **TMC**, input)
4. RETURN output
5. END

**กฎที่ 4. RomanMixVowel**

กฎนี้ทำหน้าที่เปลี่ยนชุดสระอักษรโรมันที่มีลักษณะอักษรที่เป็นคู่ให้เป็นชุดสระอักษรไทย โดยมีขั้นตอนวิธีดังนี้

1. START
2. READ input
3. output = **F-RSA**(**RMV**, **RMV**, input)
4. RETURN output
5. END

**กฎที่ 5. RomanSingleConsonant**

กฎนี้ทำหน้าที่เปลี่ยนชุดพยัญชนะอักษรโรมันที่เป็นอักษรเดี่ยว ให้เป็นชุดพยัญชนะอักษรไทย ในชุด **TSC** หากเป็นการปริวรรตแบบปรับรูปหรือแบบทั่วไปจะใช้ชุดข้อมูล **TSC2** และมีกรณีพิเศษคือ เครื่องหมายจันทรพิน แบบตัวอักษรเดี่ยวคือ ṁ จะถูกแปลงเป็น Moon dot (ँ )ก่อนเพื่อเตรียมประมวลผลในขั้นตอนต่อไป โดยมีขั้นตอนวิธีดังนี้

1. START
2. READ input
3. sc = **TSC**
4. IF thai general transliterate THEN
5. sc = **TSC2**
6. ENDIF
7. output = **F-RSA**(**RSC** , sc, input)
8. RETURN output
9. END

**กฎที่ 6. RomanSingleVowel**

กฎนี้ทำหน้าที่เปลี่ยนชุดสระอักษรโรมันที่เป็นอักษรเดี่ยว ให้เป็นชุดสระอักษรไทย ในชุด **TSV** หากเป็นการปริวรรตแบบปรับรูปหรือแบบทั่วไปจะใช้ชุดข้อมูล **TSV2** โดยมีขั้นตอนวิธีดังนี้

1. START
2. READ input
3. sv = **TSC**
4. IF thai general transliterate THEN
5. sv = **TSC2**
6. ENDIF
7. output = **F-RSA**(**RSV**, sv, input)
8. RETURN output
9. END

**กฎที่ 7. ThaiVowelInFist**

กฎนี้ทำหน้าที่เปลี่ยนเพิ่มรูปพยัญชนะ อ เข้าไปในกรณีสระอักษรไทยอยู่ต้นคำ แต่จะละเว้นไว้รูปสระอา ไว้ก่อน เพื่อประกอบรูปสระ เอา ก่อนแล้วจะพิจารณารูปสระ อา ในกฎสุดท้าย หากเป็นการปริวรรตแบบปรับรูปหรือแบบทั่วไปจะใช้ชุดข้อมูล **TIF2**, **TIFR2** โดยมีขั้นตอนวิธีดังนี้

1. START
2. READ input
3. IF =**TIF**, IFR = **TIFR**
4. IF thai general transliterate THEN
5. IF =**TIF2**, IFR = **TIFR2**
6. ENDIF
7. FOR each character of the input
8. current character = character
9. previous character = input[current index -1]
10. next character = input[current index +1]
11. main condition = [
12. Previous character NOT IN **TV** And
13. current character IN **IF**
14. ]
15. condition 1 = [
16. current character NOT IN **{ เ ,า}**
17. ]
18. condition 2 = [
19. current character ≠ ‘**เ**’ And next character ≠ ‘**า**’
20. ]
21. condition 3 = [
22. current character ≠ ‘**า**’ And previous character ≠ ‘**เ**’
23. ]
24. IF main condition AND condition 1 THEN
25. character = **F-RSA**(IF, IFR, input)
26. ELSE IF main condition AND condition 2 THEN
27. character = **F-RSA**(IF, IFR, input)
28. ELSE IF main condition AND condition 3 THEN
29. character = **F-RSA**(IF, IFR, input)
30. ENDIF
31. output = output Concatenated character
32. ENDFOR
33. RETURN output
34. END

**กฎที่ 8. ThaiVowelPrefix**

กฎนี้ทำหน้าที่ทำการสลับที่สระอักษรไทยที่เป็นสระโดยการเลื่อนจากหลังพยัญชนะ ไปไว้หน้าพยัญชนะหากพยัญชนะตัวหน้าเป็นรูปอักษร ร ให้ตรวจสอบดูว่าพยัญชนะที่ถัดจาก อักษร ร นั้นเป็นพยัญชนะหรือสระหากเป็นพยัญชนะ ให้เลื่อนสลักอักษรไทยที่เป็นสระหน้าไปไว้หน้าพยัญชนะตัวนั้น หากเป็นการปริวรรตแบบคงรูป จะมีการพิจารณาเครื่องหมายพินทุร่วมด้วย โดยมีขั้นตอนวิธีดังนี้

1. # bindu เครืองหมายพินทุ( ฺ )
2. START
3. READ input
4. ADD 4 space before input #Prevent Index Out of Bound
5. FOR each character of the input
6. current character = character
7. IF current character IN {'เ','โ', 'ไ'} THEN
8. previous character 1 = input[current index -1]
9. previous character 2= input[current index -2]
10. previous character 3= input[current index -3]
11. previous character 4 = input[current index-4 ]
12. conjunct r condition = [
13. previous character 1 = "ร" AND
14. previous character 2 = bindu AND
15. previous character 3 IN **TC** AND
16. previous character 4 ≠ "เ" AND
17. previous character 4 ≠ "โ" AND
18. previous character 4 ≠ "ไ"
19. ]
20. conjunct r thai general condition = [
21. previous character 1 = "ร" AND
22. previous character 2 IN **TC** AND
23. previous character 3 ≠ "เ" AND
24. previous character 3 ≠ "โ" AND
25. previous character 3 ≠ "ไ"
26. ]
27. normal condition = [
28. previous character 1 IN **TC**
29. ]
30. IF conjunct r condition THEN
31. SWAP value in current index TO current index -3
32. ELSE IF conjunct r thai general condition THEN
33. SWAP value in current index TO current index -2
34. ELSE IF normal condition THEN
35. SWAP value in current index TO current index -1
36. ENDIF
37. output = output Merge character
38. ENDFOR
39. RETURN output
40. END

**กฎที่ 9. ThaiAAInFist**

กฎนี้ทำหน้าที่เปลี่ยนลากข้าง(า) โดยผสมรูปพยัญชนะ อ กรณีอยู่ถัดรูปสระอื่น ซึ่งเป็นสระตัวสุดท้ายที่จะแปลง เนื่องจากต้องแปลงรูปสระเอา ก่อน โดยมีขั้นตอนวิธีดังนี้

1. START
2. READ input
3. FOR each character of the input
4. current character = character
5. previous character = input[current index -1]
6. IF previous character NOT IN **TV** AND character = "า" THEN
7. character = "อา"
8. ENDIF
9. output = output Concatenated character
10. ENDFOR
11. RETURN output
    * 1. **ขั้นตอนวิธีของกฎการปริวรรตอักษรไทยแบบคงรูป**

ขั้นตอนวิธีของกฎการปริวรรตอักษรไทยแบบคงรูป เมื่อถูกเรียกใช้งานจาก Transliteration Machine จะส่งข้อความลงไปประมวลผล ตามลำดับกฎดังต่อไปนี้

**กฎที่ 1. RomanChandrabinduToSingle**

เรียกใช้การทำงานที่กฎกลางเมื่อประมวลผลเสร็จส่งผลลัพธ์ไปที่กฎถัดไป

**กฎที่ 2. RomanNumberAndPunctuate**

เรียกใช้การทำงานที่กฎกลางเมื่อประมวลผลเสร็จส่งผลลัพธ์ไปที่กฎถัดไป

**กฎที่ 3. RomanMixConsonant**

เรียกใช้การทำงานที่กฎกลางเมื่อประมวลผลเสร็จส่งผลลัพธ์ไปที่กฎถัดไป

**กฎที่ 4. RomanMixVowel**

เรียกใช้การทำงานที่กฎกลางเมื่อประมวลผลเสร็จส่งผลลัพธ์ไปที่กฎถัดไป

**กฎที่ 5. RomanSingleConsonant**

เรียกใช้การทำงานที่กฎกลาง เมื่อประมวลผลเสร็จส่งผลลัพธ์ไปที่กฎถัดไป

**กฎที่ 6. RomanSingleVowel**

เรียกใช้การทำงานที่กฎกลาง เมื่อประมวลผลเสร็จส่งผลลัพธ์ไปที่กฎถัดไป

**กฎที่ 7. Bindu**

กฎนี้ทำหน้าที่เติมเครื่องหมายพินทุในพยัญชนะอักษรไทยที่ไม่มีสระกำกับ และยกเว้น 4 เครื่องหมายที่แสดงในกฎ โดยมีขั้นตอนวิธีดังนี้

1. # anusvara เครืองหมายนิคหิต
2. # chandrabindu เครืองหมาย Moon dot
3. # avagaraha เครืองหมาย apostrophe
4. # bindu เครืองหมายพินทุ
5. # ThaiVisarga เครืองหมายวิสรรชนีย์
6. START
7. READ input
8. FOR each character of the input
9. current character = character
10. next character = input[current index +1]
11. condition = [
12. current character NOT IN **TC** And
13. current character ≠ anusvara And
14. current character ≠ chandrabindu And
15. current character ≠ ThaiVisarga And
16. current character ≠ Avagaraha And
17. next character ≠ "a" And
18. next character NOT IN **TV**
19. ]
20. IF condition THEN
21. character = character Concatenated bindu
22. ENDIF
23. output = output Concatenated character
24. ENDFOR
25. RETURN output
26. END

**กฎที่ 8. ThaiVowelInFist**

เรียกใช้การทำงานที่กฎกลาง เมื่อประมวลผลเสร็จส่งผลลัพธ์ไปที่กฎถัดไป

**กฎที่ 9. ThaiVowelPrefix**

เรียกใช้การทำงานที่กฎกลาง เมื่อประมวลผลเสร็จส่งผลลัพธ์ไปที่กฎถัดไป

**กฎที่ 10. ARemove**

กฎนี้ทำหน้าที่ทำการลบสระอักษรโรมัน A ออกจากข้อความทั้งหมด โดยมีขั้นตอนวิธีดังนี้

1. START
2. READ input
3. output = **F-RS**(‘A’ , empty value, input)
4. RETURN output
5. END

**กฎที่ 11. SwapAnusvaraAndChandrabindu**

กฎนี้ทำหน้าที่เปลี่ยนเครื่องหมายอนุสาวระอักษรไทยและ Moon Dot ไปข้างหน้ากรณีอยู่หลัง ลากข้าง(า)โดยมีขั้นตอนวิธีดังนี้

1. #anusvara เครืองหมายนิคหิต(**ํ**)
2. #chandrabindu เครืองหมาย Moon dot( **ँ** )
3. START
4. READ input
5. output = **F-RS(" าํ** **" , " ํา** **" , input)**
6. output = **F-RS(" าँ** **" , " ँา** **", input)**
7. RETURN output
8. END

**กฎที่ 12. Chandrabindu**

กฎนี้ทำหน้าที่แปลงเครื่องหมาย Moon Dot โดยใช้ไม้หันอากาศและนิคหิตแทนโดยมีขั้นตอนวิธีดังนี้

1. START
2. READ input
3. output = **F-RS**(**Moon Dot**, **"ัํ"**, input)
4. RETURN output
5. END

**กฎที่ 13. ThaiAAInFist**

เรียกใช้การทำงานที่กฎกลาง เมื่อประมวลผลเสร็จส่งผลลัพธ์ไปที่ Transliteration Machine จบการทำงาน

* + 1. **ขั้นตอนวิธีของกฎการปริวรรตอักษรไทยแบบปรับรูป**

ขั้นตอนวิธีของกฎการปริวรรตอักษรไทยแบบปรับรูป เมื่อถูกเรียกใช้งานจาก Transliteration Machine จะส่งข้อความลงไปประมวลผล ตามลำดับกฎดังต่อไปนี้

**กฎที่ 1. AvagarahaRemove**

กฎนี้ทำหน้าที่ลบอวครหะออก (Apostrophe) โดยมีขั้นตอนวิธีดังนี้

1. #avagaraha เครื่องหมาย ' (apostrophe)
2. START
3. READ input
4. output = **F-RSA**(avagaraha , empty value, input)
5. RETURN output
6. END

**กฎที่ 2. RomanChandrabinduToSingle**

เรียกใช้การทำงานที่กฎกลางเมื่อประมวลผลเสร็จส่งผลลัพธ์ไปที่กฎถัดไป

**กฎที่ 3. RomanNumberAndPunctuate**

เรียกใช้การทำงานที่กฎกลางเมื่อประมวลผลเสร็จส่งผลลัพธ์ไปที่กฎถัดไป

**กฎที่ 4. RomanMixConsonant**

เรียกใช้การทำงานที่กฎกลางเมื่อประมวลผลเสร็จส่งผลลัพธ์ไปที่กฎถัดไป

**กฎที่ 5. RomanMixVowel**

เรียกใช้การทำงานที่กฎกลางเมื่อประมวลผลเสร็จส่งผลลัพธ์ไปที่กฎถัดไป

**กฎที่ 6. RomanSingleConsonant**

เรียกใช้การทำงานที่กฎกลางเมื่อประมวลผลเสร็จส่งผลลัพธ์ไปที่กฎถัดไป

**กฎที่ 7. RomanSingleVowel**

เรียกใช้การทำงานที่กฎกลางเมื่อประมวลผลเสร็จส่งผลลัพธ์ไปที่กฎถัดไป

**กฎที่ 8. AnusvaraAndChandrabindu**

กฎนี้ทำหน้าที่แปลงและเข้าสนธิ เครื่องหมายอนุสวาระและจันทรพินทุ มีพยัญชนะที่เป็นพยัญชนะวรรคตามติดมา เครื่องหมายอนุสวาระและจันทรพินทุ ให้ถอดเป็นพยัญชนะนาสิก(อนุนาสิกะ)ในวรรคของตัวที่ตามมา อยู่กับสระใดให้เป็นตัวสะกดสระนั้น หากเป็นพยัญชนะอวรรค ให้ใช้อักษร ม แทน โดยมีขั้นตอนวิธีดังนี้

1. # anusvara เครืองหมายนิคหิต
2. # chandrabindu เครืองหมาย Moon dot
3. START
4. READ input
5. FOR each character of the input
6. current character = character
7. next character = input[current index +1]
8. IF current character = anusvara OR chandrabindu THEN
9. IF next characterIN **KVARGA** THEN
10. character = ง
11. ELSE IF next characterIN **CVARGA** THEN
12. character = ญ
13. ELSE IF next characterIN **THVARGA** THEN
14. character = ณ
15. ELSE IF next characterIN **TVARGA** THEN
16. character = น
17. ELSE IF next character IN **PVARGA** THEN
18. character = ม
19. ELSE
20. character = ม
21. ENDIF
22. output = output Concatenated character
23. ENDIF
24. ENDFOR
25. RETURN output
26. END

**กฎที่ 9. ThaiVowelInFist**

เรียกใช้การทำงานที่กฎกลางเมื่อประมวลผลเสร็จส่งผลลัพธ์ไปที่กฎถัดไป

**กฎที่ 10. ThaiVisarga**

กฎนี้ทำหน้าที่จัดการกับเครื่องหมายวิสรรชนีย์และไม้หันอากาศ โดยจะละเว้นไม่พิจารณาเครื่องหมายวิสรรชนีย์กรณีที่มี อักษร ว ตามมา (av) เช่น ตะวยา จะคงรูปเดิมไม่แปลงเป็น ตัวยา โดยมีกฎย่อยอีก 3 กฎ โดยมีขั้นตอนวิธีดังนี้

1. START
2. READ input
3. ADD space 4 before input # reserve for condition
4. ADD space 6 after input # reserve for condition
5. FOR each character of the input
6. current character = character
7. next character = input[current index +1]
8. IF current character = ะ And next character ≠ ว THEN
9. output = **AH\_RULE**(input, current index)
10. output = **AR\_RULE**(output , current index )
11. output = **A\_AFTER\_RULE**(output, current index )
12. ENDIF
13. ENDFOR
14. RETURN output
15. END

**กฎย่อยที่ 10.1 AH\_RULE** กฎนี้ทำหน้าที่เปลี่ยนเครื่องหมายวิสรรชนีย์ที่อยู่หน้าอักษร ห จะตัด เครื่องหมายวิสรรชนีย์ออก เช่น พระหมะ เป็น พรหมะ แต่ยกเว้นเครื่องหมายวิสรรคะ (ห์ ) ไว้ โดยมีขั้นตอนวิธีดังนี้

1. START
2. READ input and index #input is array
3. character = input[index ]
4. next character 1 = input[current index +1]
5. next character 2= input[current index +2]
6. condition = [
7. character = ะ And
8. next character 1 = IN **TC**
9. next character 2 NOT IN **TV**
10. next character 1 = ห
11. next character 2 ≠ "์" #ไม้ทัณฑฆาต
12. ]
13. IF condition THEN
14. character = empty value
15. ENDIF
16. RETURN output
17. END

**กฎย่อยที่ 10.2 AR\_RULE** กฎนี้ทำหน้าที่เปลี่ยนเครื่องหมายวิสรรชนีย์ที่อยู่หน้าอักษร ร (กรณี ร-เรผะ)จะตัด เครื่องหมายวิสรรชนีย์ออก ใช้ ร-หัน แทน เช่น ธะรมะ เป็น ธรรมะ

1. START
2. READ input and index #input is array
3. previous character = input[current index -1]
4. character = input[index ]
5. next character 1 = input[current index +1]
6. next character 2= input[current index +2]
7. condition = [
8. character = ะ And
9. previous character IN **TC** And
10. next character 1 = ร
11. next character 2 NOT IN **TV**
12. ]
13. IF condition THEN
14. character = ร
15. ENDIF
16. RETURN output
17. END

**กฎย่อยที่ 10.1 A\_** **AFTER \_RULE** กฎนี้ทำหน้าที่เปลี่ยนเครื่องหมายวิสรรชนีย์ที่อยู่หลังพยัญชนะ โดยมีตัวสะกดตามมา ให้เป็นไม้หันอากาศ หากเครื่องหมายวิสรรชนีย์ที่อยู่หน้าสุดให้เปลี่ยนเป็น อ และ ไม้หันอากาศ เช่น ะนกะ เป็น อันกะ โดยมีขั้นตอนวิธีดังนี้

1. START
2. READ input and index #input is array
3. character = input[index ]
4. previous character 1 = input[current index -1]
5. previous character 2 = input[current index -2]
6. next character 1 = input[current index +1]
7. next character 2= input[current index +2]
8. condition = [
9. character = ะ And
10. previous character 1 IN **TC** And
11. previous character 2 = NOT IN **TV**
12. next character 1 = IN **TC** And
13. next character 2 NOT IN **TV**
14. ]
15. IF condition THEN
16. character = "ั" #ไม้หันอากาศ
17. ELSE IF condition And previous character 1 = empty value THEN
18. character = อั
19. ENDIF
20. RETURN output
21. END

**กฎที่ 11. ThaiVowelPrefix**

เรียกใช้การทำงานที่กฎกลางเมื่อประมวลผลเสร็จส่งผลลัพธ์ไปที่กฎถัดไป

**กฎที่ 12. ThaiAAInFist**

เรียกใช้การทำงานที่กฎกลางเมื่อประมวลผลเสร็จส่งผลลัพธ์ไปที่กฎถัดไป

**กฎที่ 13. convertAE**

เลื่อนสระที่เป็นสระ เ ไปหน้าพยัญชนะหากตัวอักษรก่อนหน้าถูกแปลงเป็นไม้หันอากาศแล้วต้องแปลงกลับเป็นวิสรรชณีย์(ะ)โดยมีขั้นตอนวิธีดังนี้

1. START
2. READ input
3. output = **F-RS**(**"** **ัเ" , "ะเ"**, input)
4. RETURN output
5. END

**กฎที่ 14. convertAO**

เลื่อนสระที่เป็นสระ โ ไปหน้าพยัญชนะหากตัวอักษรก่อนหน้าถูกแปลงเป็นไม้หันอากาศแล้วต้องแปลงกลับเป็นวิสรรชณีย์(ะ) เมื่อประมวลผลเสร็จส่งผลลัพธ์ไปที่ Transliteration Machine จบการทำงาน โดยมีขั้นตอนวิธีดังนี้

1. START
2. READ input
3. output = **F-RS**( **"**ัโ**"**, **"**ะโ**"**’,, input)
4. RETURN output
5. END
   * 1. **ขั้นตอนวิธี (Algorithm) ในการปริวรรตอักษร**

เนื่องจากการทำงานปริวรรตนั้นต้องทำกับข้อความจำนวนมากและโปรแกรมทำงานแบบ Server-Client หากการทำงานใช้เวลามากเกินไปจะเกิดปัญหา Time Out ใน Web Browser ได้ เพื่อให้การปริวรรตได้ราบรื่น และให้ได้ลักษณะการส่งข้อความไปประมวลผลให้เหมาะกับกฎที่ออกแบบไว้ ที่มี Loop อยู่เป็นจำนวนมาก ผู้จัดทำจึงได้ออกแบบ ขั้นตอนวิธีในการปริวรรตอักษรไว้ 3 วิธี

วิธีที่ 1. รับตัวอักษรต้นฉบับอักษรโรมันทำการปริวรรตทั้งข้อความ

1. START
2. READ roman text
3. thai text = TRANSLITERATION (roman text)
4. RETURN thai text
5. END

วิธีที่ 2. รับตัวอักษรต้นฉบับอักษรโรมันทำการปริวรรตที่ละบรรทัด

1. START
2. READ roman text
3. FOR each line of the roman text
4. thai text = thai text + TRANSLITERATION (line)
5. ENDFOR
6. RETURN thai text
7. END

วิธีที่ 3. รับตัวอักษรต้นฉบับอักษรโรมันทำการปริวรรตที่ละคำในแต่ละบรรทัด

1. START
2. READ roman text
3. FOR each line of the roman text
4. FOR each word of the line
5. thai text = thai text + TRANSLITERATION (word)
6. ENDFOR
7. ENDFOR
8. RETURN thai text
9. END
   1. **การพัฒนาตัวต้นแบบการปริวรรตอักษร (Prototype)**

ผู้จัดทำได้ทำการพัฒนาตัวต้นแบบการปริวรรตอักษร (Prototype) เพื่อตรวจสอบความเป็นไปได้และการพิสูจน์แนวคิดของการพัฒนาส่วนตัวปริวรรตอักษร(POC: Proof-of-Concept) ว่าระบบจะสามารถทำงานได้อย่างถูกต้องและตรงความต้องการหรือไม่ โดยผู้จัดทำได้ทำการพัฒนาโดยใช้ภาษา JAVA ทำการโปรแกรมให้ผลลัพธ์ออกมาในรูปแบบ Console Command โดยเลือกทำเฉพาะส่วนที่ประเมินว่ามีความยากในการพัฒนามากที่สุดคือ การปริวรรตอักษรแบบปรับรูป หรือ แบบทั่วไป โดยทดสอบใช้กฎ 6 กฎ ได้แก่ RomanMixConsonant, RomanMixVowel RomanSingleConsonant ,RomanSingleVowel, ThaiVowelInFist, ThaiVowelPrefix อีกทั้งการทำตัวต้นแบบนี้ยังเพื่อพิสูจน์ขั้นตอนวิธี (Algorithm) ในการปริวรรตอักษร 3 วิธีที่ออกแบบไว้ก่อนหน้าว่า ขั้นตอนวิธีใดมีประสิทธิภาพมากที่สุด เพื่อเลือกเอาไปพัฒนาต่อไป



รูปที่ 3.2 ผลลัพธ์หลังจากการพัฒนาตัวต้นแบบการปริวรรตอักษร (Prototype)

หลังจากพิสูจน์ขั้นตอนวิธี (Algorithm) ในการปริวรรตอักษร 3 วิธีที่ออกแบบไว้ โดยทดสอบกับตัวอักษรโรมันต้นฉบับ จำนวน 34,769 ตัวอักษร ขั้นตอนวิธีมีประสิทธิภาพมากที่สุด คือ วิธีที่ 2. การปริวรรตที่ละบรรทัด มีเวลาที่ใช้โดยเฉลี่ยจากการ execute 5 ครั้งที่น้อยที่สุด ผลลัพธ์ตามตารางที่

ตารางที่ 3.4 การพิสูจน์ขั้นตอนวิธี (Algorithm) ในการปริวรรตอักษร 3 วิธี

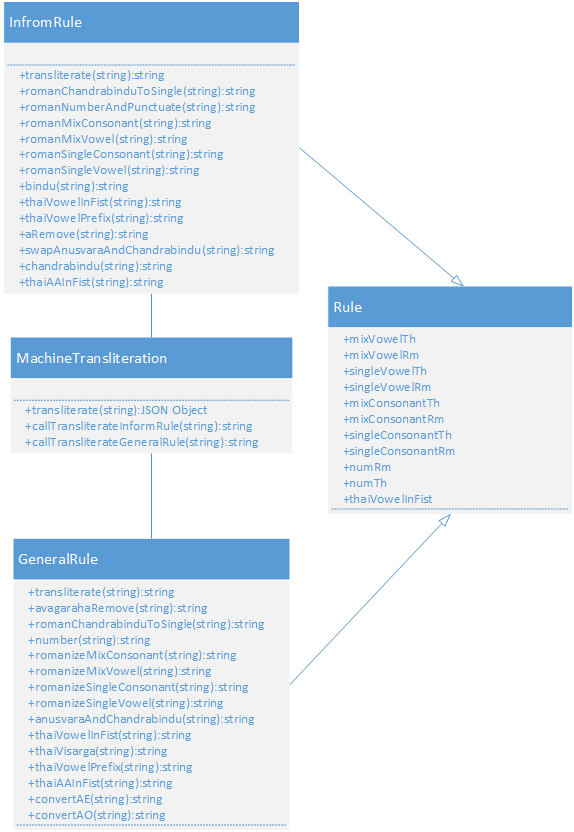
|  |  |
| --- | --- |
| **ขั้นตอนวิธีในการปริวรรตอักษร** | **เวลาที่ใช้โดยเฉลี่ย(วินาที)** |
| วิธีที่ 1. ทำการปริวรรตทั้งข้อความ | 9.86 |
| วิธีที่ 2. ทำการปริวรรตที่ละบรรทัด | 0.61 |
| วิธีที่ 3. ทำการปริวรรตที่ละคำในแต่ละบรรทัด | 0.76 |

* 1. **การพัฒนาส่วนตัวปริวรรตอักษร(Transliteration Machine)**

ผู้จัดทำได้ทำการพัฒนาส่วนตัวปริวรรตอักษร(Transliteration Machine) ให้สามารถทำงานได้ตามกฎ (Rule-based) และขั้นตอนวิธี (Algorithm) ที่ได้ออกแบบไว้ อีกทั้งทำการทดสอบ ตรวจสอบผลลัพธ์ ภายใต้คำแนะนำของผู้เชี่ยวชาญภาษาสันสกฤตโดยการพัฒนานั้นพัฒนาโดยใช้ภาษา PHP พัฒนาเป็นแบบ Object Oriented Programming ประกอบไปด้วยคลาส 3 Class ได้แก่ MachineTransliteration Class InfromRule Class และ GeneralRule Class

MachineTransliteration Class จะทำหน้าที่เป็น API สำหรับเชื่อมต่อกับระบบอื่นภายนอก

InfromRule Class จะทำหน้าที่ปริวรรตข้อความให้เป็นอักษรไทยแบบคงรูปหรือแบบแผน และGeneralRule Class จะทำหน้าที่ การปริวรรตข้อความให้เป็นอักษรไทยแบบปรับรูปหรือแบบทั่วไป ส่วนการทำงานนั้นหากมีการส่งข้อความเข้ามา MachineTransliteration Class จะรับเอาข้อความเข้ามา แล้วนำไปประมวลผลโดยจะแบ่งข้อความแบบทีละบรรทัด ตามที่ได้ออกแบบไว้ และส่งบรรทัดข้อความนั้น ๆ ไปให้ทั้งสอง Class โดยเรียกใช้ฟังก์ชัน transliterate() จากคลาส InfromRule และ GeneralRule ทำการประมวลผล ซึ่งจะได้ผลลัพธ์ทั้งแบบคงรูปและแบบปรับรูปเมื่อเสร็จจะส่งผลลัพธ์กลับไปยังที่ MachineTransliteration Class ทำการสะสมค่าผลลัพธ์เอาไว้ เมื่อเสร็จสิ้นกระบวนการทุกบรรทัดในข้อความแล้ว MachineTransliteration Class จะนำผลลัพธ์ที่ได้มาประมวลผลแล้วส่งออกเป็น JSON Object ให้กับระบบอื่นภายนอก ดังมีโครงสร้างใน class diagram ดังต่อไปนี้



รูปที่ 3.3 โครงสร้างใน class diagram

* 1. **การพัฒนาส่วนต่อผู้ใช้งาน(User Interface)**

ผู้จัดทำได้ทำการพัฒนาส่วนต่อผู้ใช้งาน(User Interface) ที่ไว้สำหรับสำหรับในการป้อนอักษรต้นฉบับจากผู้ใช้ เพื่อส่งไปยังส่วนตัวปริวรรตอักษร(Transliteration Machine) เพื่อทำประมวลให้ได้ผลลัพธ์ออกมาแสดงและส่วนการแสดงการเปรียบเทียบตรวจสอบจากอักษรต้นฉบับกับอักษรผลลัพธ์จากการปริวรรต การพัฒนาส่วนต่อผู้ใช้งาน(User Interface) โดยมีหน้าจอไว้สำหรับรับข้อมูลจากผู้ใช้ โดยลักษณะหน้าจอจะมี select box สำหรับเลือกตัวอักษรต้นฉบับ select box ตัวที่ 2 ใช้สำหรับเลือกอักษรชนิดอื่น ๆ ในการเปรียบเทียบหากผู้ใช้เลือกอักษรต้นฉบับและอักษรเปรียบเทียบตัวที่ 2 แล้ว จะปรากฏตารางเปรียบเทียบตัวอักษรที่ด้านล่างสุด ถัดลงมาด้านล่างจะเป็น text area ช่องแรกไว้สำหรับข้อความต้นฉบับช่องที่ 2 เป็นผลลัพธ์จากการปริวรรตอักษรอื่น ๆ ด้านล่างจะมีปุ่มไว้สำหรับทำการปริวรรตอักษรเป็นอักษรไทย

หากผู้ใช้กรอกตัวอักษรต้นฉบับและกดปุ่ม การปริวรรตเป็นอักษรไทยแล้ว จะมีผลลัพธ์ผลลัพธ์การปริวรรตเป็นอักษรไทยแบบคงรูป และแบบปรับรูป ออกแสดงที่ตารางเปรียบเทียบผลการปริวรรตโดยสามารถเลือกการเปรียบเทียบได้มากที่สุดถึง 4 ชนิดตัวอักษรและยังสามารถซ่อนหรือแสดงตารางเปรียบเทียบตัวใดตัวหนึ่งได้ อีกทั้งยังสามารถทำการเปรียบเทียบโดยการไฮไลท์ที่คำในแต่ละบรรทัดระบบจะทำการไฮไลท์ข้อความจากทุกๆตัวอักษร



รูปที่ 3.4 ส่วนต่อผู้ใช้งาน(User Interface) ส่วนบน



รูปที่ 3.5 ส่วนต่อผู้ใช้งาน(User Interface) ส่วนล่าง

* 1. **การทดสอบและประเมินผลการใช้งานจากผู้เชี่ยวชาญภาษาสันสกฤต**

เพื่อให้ผลการทดสอบสามารถประเมินผลการดำเนินโครงการ ให้ความถูกต้องตามวัตถุประสงค์ของการศึกษาโครงการเฉพาะเรื่อง จึงมีการกำหนดขอบเขตและออกแบบวิธีการทดสอบและประเมินผล ดังต่อไปนี้

* + 1. **การทดสอบการตัวปริวรรตอักษร**

การทดสอบการตัวปริวรรตอักษรเป็นขั้นตอนการทดสอบความถูกต้องของส่วนตัวปริวรรตอักษร ประกอบด้วย ชุดข้อมูลของการทดสอบ และวิธีการทดสอบ มีรายละเอียดดังต่อไปนี้

1. **ชุดข้อมูลของการทดสอบ**
   1. กรณีที่ใช้ในการทดสอบ(Test Case) โดยเป็นชุดข้อความผลลัพธ์ที่ถูกต้องและสามารถครอบคลุมกฎทั้งกรณีการปริวรรตแบบคงรูปและปรับรูป จำนวนอย่างละ 59 กรณี รวมเป็น 118 กรณี
   2. ตัวอย่างตัวอย่างการปริวรรต ผู้จัดทำได้เลือกหนึ่งในตัวอย่างการปริวรรตที่ กล่าวมาก่อนแล้ว โดยเลือกที่มีเนื้อหาไม่มากไม่น้อยจนเกินไป และมีต้นฉบับที่เก็บอยู่ในรูปแบบ TEXT โดยตัวอย่างการปริวรรต ผู้จัดทำได้เลือกคือ
      1. ตัวอย่างการปริวรรตแบบคงรูป ใช้ คัมภีร์วัชรเฉทิกปรัชญาปารมิตาสูตร ฉบับอักษรไทย ขนาด 31,436 ตัวอักษร
      2. ตัวอย่างการปริวรรตแบบปรับรูป ใช้ คัมภีร์วัชรเฉทิกปรัชญาปารมิตาสูตร ฉบับคำอ่านอักษรไทย ขนาด 34,117 ตัวอักษร
   3. ตัวอย่างต้นฉบับ คัมภีร์วัชรเฉทิกปรัชญาปารมิตา อักษรโรมัน จากโครงการ DSBC (www.dsbcproject.org )
2. **วิธีการทดสอบความถูกต้อง**

วิธีการทดสอบความถูกต้องของผลลัพธ์ที่ได้มาจาก ส่วนตัวปริวรรตอักษร มี 2 วิธีดังนี้

* 1. **การทดสอบแบบแยกส่วน (Unit Testing)**

การทดสอบแบบแยกส่วนนั้นจะใช้ PHPUnit (PHP Testing Framework) เป็นเครื่องมือในการตรวจสอบและทดสอบกฎ โดยใช้ชุดคำเป็นกรณีที่ใช้ทดสอบ(Test Case) 118 กรณี เพื่อทดสอบการทำงานของส่วนการปริวรรตว่า ผลลัพธ์ที่ออกมาตรงกับชุดคำดังกล่าวหรือไม่

* 1. **การทดสอบโดยการเปรียบเทียบกับตัวอย่างการปริวรรต**

การทดสอบโดยการเปรียบเทียบกับตัวอย่างการปริวรรต มีขั้นตอนดังต่อไปนี้

* + 1. ใช้ตัวอย่างต้นฉบับ คัมภีร์วัชรเฉทิกปรัชญาปารมิตา อักษรโรมัน ทำการปริวรรตเป็นอักษรไทยแบบคงรูปและปรับรูป
    2. นำผลลัพธ์การปริวรรตอักษรไทยแบบคงรูป ไปเปรียบเทียบกับ คัมภีร์วัชรเฉทิกปรัชญาปารมิตาสูตร ฉบับอักษรไทย และ ผลลัพธ์การปริวรรตอักษรไทยแบบปรับรูป ไปเปรียบเทียบกับ คัมภีร์วัชรเฉทิกปรัชญาปารมิตาสูตร ฉบับคำอ่านอักษรไทย โดยใช้โปรแกรม Pretty Diff v2.0.0 เป็นโปรแกรมเปรียบเทียบเพื่อคำนวณหาอัตราความถูกต้องของส่วนตัวปริวรรตอักษร
    3. **ประเมินผลการใช้งานจากผู้เชี่ยวชาญภาษาสันสกฤต**

การประเมินผลการใช้งานประเมินผลโดยผู้เชี่ยวชาญด้านภาษาสันสกฤตโดยมี ข้อคิดเห็นจากผู้ประเมินพร้อมหัวข้อแบบประเมิน การประเมินผลการใช้งานประเมินผล แบ่งหัวข้อได้ดังต่อไปนี้

1. **แบบประเมินผลการใช้งานระบบ ด้านการทำงานได้ตามฟังก์ชันงานของระบบ (Function)**

โดยจะประเมิน ความถูกต้องของการผลลัพธ์ ความรวดเร็วในการประมวลผล ความครอบคลุมของงาน ความสามารถป้องกันข้อผิดพลาดที่เกิดขึ้น และความน่าเชื่อถือได้ของระบบ โดยมีงานที่จะประเมินดังนี้

* 1. การปริวรรตจากอักษรโรมันเป็นอักษรไทยแบบคงรูป
  2. การปริวรรตจากอักษรโรมันเป็นอักษรไทยแบบปรับรูป
  3. ตารางการเปรียบเทียบผลการปริวรรต
  4. ตารางการเปรียบเทียบอักษร

1. **แบบประเมินผลการใช้งานระบบ ด้านความง่ายต่อการใช้งานระบบ (Usability)**

โดยจะประเมินจากส่วนต่อผู้ใช้งาน(User Interface)แบ่งหัวข้อได้ดังต่อไปนี้

1. ความง่ายต่อการใช้งานของระบบ
2. ความเหมาะสมในการเลือกใช้ชนิดตัวอักษรบนจอภาพ
3. ความเหมาะสมในการใช้สีของตัวอักษรและรูปภาพ
4. ความเหมาะสมในการใช้ข้อความเพื่ออธิบายสื่อความหมาย
5. ความเหมาะสมในการใช้สัญลักษณ์หรือรูปภาพในการสื่อความหมาย
6. ความเป็นมาตรฐานเดียวกันในการออกแบบหน้าจอภาพ
7. ความเหมาะสมในการปฏิสัมพันธ์โต้ตอบกับผู้ใช้
8. ความเหมาะสมในการวางตำแหน่งของส่วนประกอบบนจอภาพ
9. คำศัพท์ที่ประกอบใช้ผู้ใช้มีความคุ้นเคยและสามารถปฏิบัติตามได้โดยง่าย
   * 1. **เกณฑ์ประเมินผล**

การประเมินผลการใช้งานประเมินผลโดยผู้เชี่ยวชาญด้านภาษาสันสกฤตโดยมี ข้อคิดเห็นจากผู้ประเมินพร้อมหัวข้อแบบประเมิน

1. **ระดับความเหมาะสม/ความพึงพอใจ**

โดยในแบบการประเมิน จะมีตัวเลขของระดับความเหมาะสม/ความพึงพอใจในแต่ละด้านมีความหมายดังนี้

5 หมายถึง ระดับมากที่สุด

4 หมายถึง ระดับมาก

3 หมายถึง ระดับปานกลาง

2 หมายถึง ระดับน้อย

1 หมายถึง ระดับน้อยที่สุด

1. **การเกณฑ์การวัดระดับความเหมาะสม/ความพึงพอใจ**

สำหรับการเกณฑ์การวัดระดับความเหมาะสม/ความพึงพอใจในโดยรวมของแต่ละด้าน โดยคิดจากค่าเฉลี่ย มีความหมายดังนี้มีดังนี้

5.0 หมายถึง ระดับมากที่สุด

4.5 ถึง 4.9 หมายถึง ระดับมากค่อนไปทางมากที่สุด

4.0 ถึง 4.4 หมายถึง ระดับมาก

3.5 ถึง 3.9 หมายถึง ระดับปานกลางค่อนไปทางมาก

3.0 ถึง 3.4 หมายถึง ระดับปานกลาง

2.5 ถึง 2.9 หมายถึง ระดับน้อยค่อนไปทางปานกลาง

2.0 ถึง 2.4 หมายถึง ระดับน้อย

1.5 ถึง 1.9 หมายถึง ระดับน้อยค่อนไปทางน้อยที่สุด

1.0 ถึง 1.4 หมายถึง ระดับน้อยที่สุด

**บทที่ 4**

**ผลการดำเนินงาน**

ในบทนี้จะกล่าวถึงผลการดำเนินงาน การทดสอบ และการประเมินการใช้งานระบบโดยผู้เชียวชาญด้านภาษาสันสกฤต การปริวรรตอักษรด้วยเครื่องสำหรับภาษาสันสกฤตแบบอักษรโรมันเป็นอักษรไทย โดยรายละเอียดดังต่อไปนี้

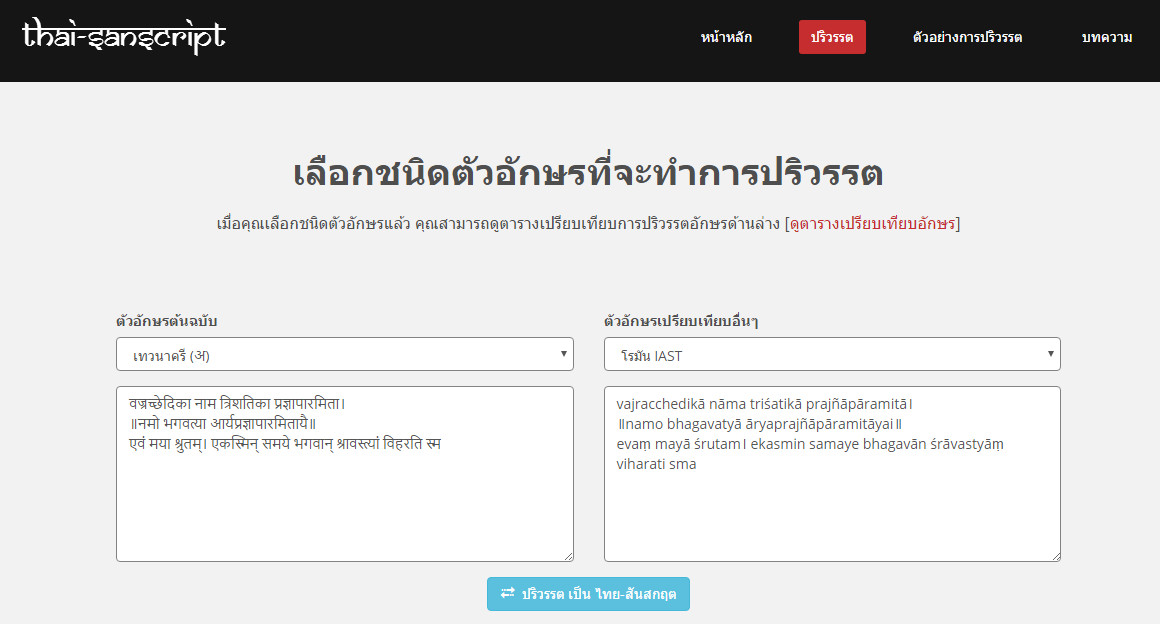
* 1. **ผลการดำเนินงาน**

หลังจากที่ได้ศึกษาวิเคราะห์ ออกแบบและพัฒนา การปริวรรตอักษรด้วยเครื่องสำหรับภาษาสันสกฤตแบบอักษรโรมันเป็นอักษรไทย จนเสร็จสมบูรณ์แล้ว ตัวระบบจะทำงานอยู่บนอินเตอร์เน็ตที่ www.thai-sanscript.com โดยมีหน้าหลัก ดังตัวอย่างในรูปที่ 4.1



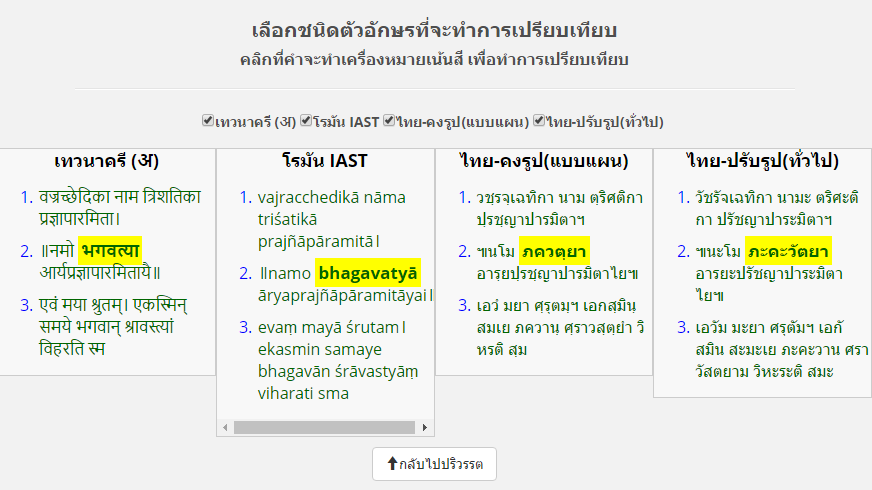
รูปที่ 4.1 หน้าแรกของตัวระบบ

ในเมนู ปริวรรต จะมีหน้าจอไว้สำหรับรับข้อมูลจากผู้ใช้ โดยหน้าจอจะมี Select Box สำหรับเลือกตัวอักษรต้นฉบับ Select Box ตัวที่ 2 ใช้สำหรับเลือกอักษรชนิดอื่น ๆ ถัดลงมาด้านล่างจะเป็น Text Area ช่องแรกไว้สำหรับข้อความต้นฉบับช่องที่ 2 เป็นผลลัพธ์จากการปริวรรตอักษรอื่น ๆ ด้านล่างจะมีปุ่มไว้สำหรับทำการปริวรรตอักษรเป็นอักษรไทย ตัวอย่างในรูปที่ 4.2



รูปที่ 4.2 หน้าปริวรรตอักษรส่วนฟอร์มรับข้อมูล

โดยผลลัพธ์การปริวรรตเป็นอักษรไทยแบบคงรูปและแบบปรับรูป จะออกแสดงที่ตารางเปรียบเทียบ โดยสามารถเลือกการเปรียบเทียบได้มากที่สุดถึง 4 ชนิดตัวอักษรและยังสามารถซ่อนหรือแสดงตารางเปรียบเทียบตัวใดตัวหนึ่งได้ อีกทั้งยังสามารถทำการเปรียบเทียบโดยการไฮไลท์ที่คำในแต่ละบรรทัดระบบจะทำการไฮไลท์ข้อความจากทุกๆตัวอักษร ตัวอย่างในรูปที่ 4.3



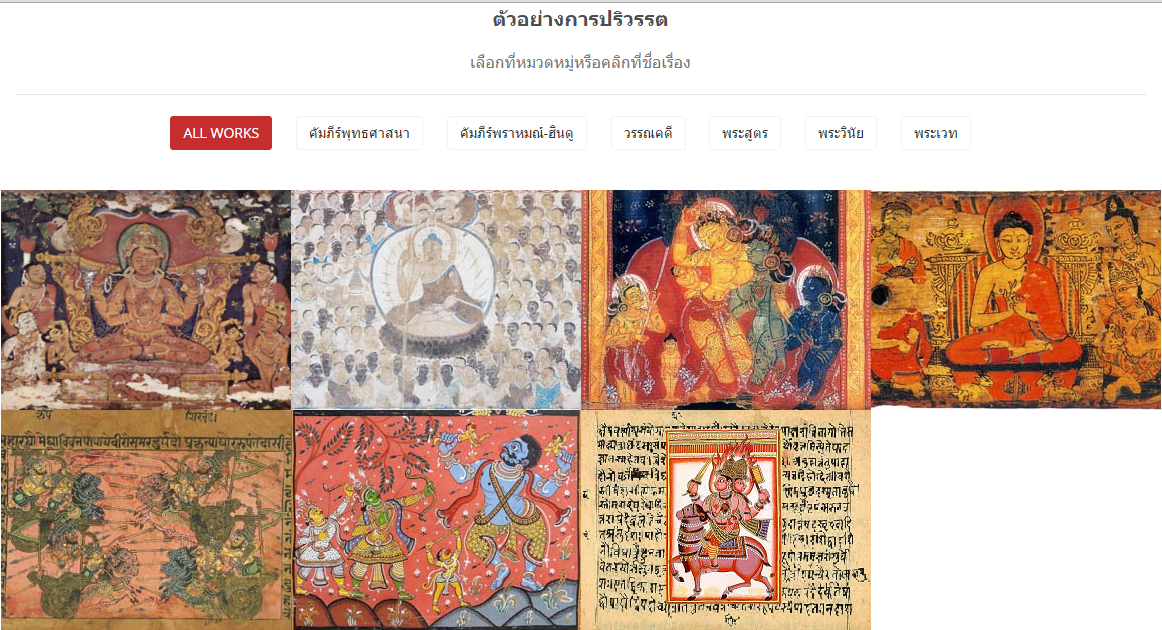
รูปที่ 4.3 หน้าปริวรรตอักษรส่วนตารางเปรียบเทียบผลการปริวรรต

ในการเปรียบเทียบหากผู้ใช้เลือกอักษรต้นฉบับและอักษรเปรียบเทียบตัวที่ 2 แล้ว จะปรากฏตารางเปรียบเทียบตัวอักษรที่ด้านล่างสุด ตัวอย่างในรูปที่ 4.3

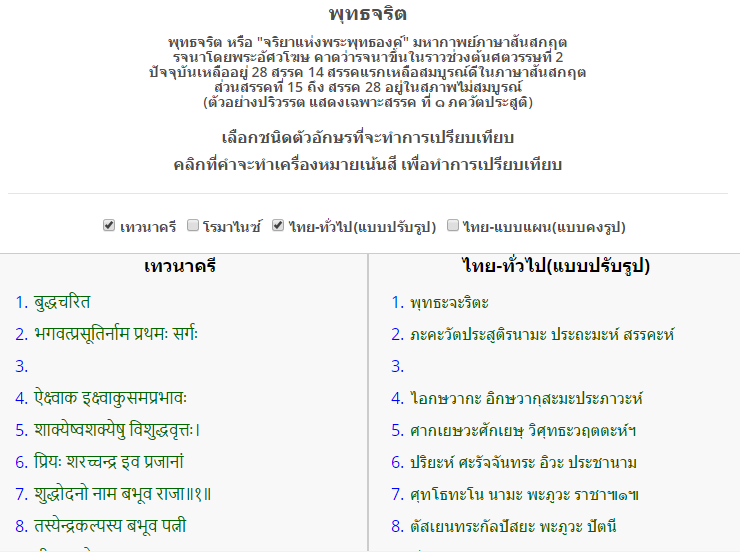


รูปที่ 4.4 หน้าปริวรรตอักษรส่วนตารางเปรียบเทียบตัวอักษร

และจะมีตัวอย่างการปริวรรตแยกไว้เป็นหมวดหมู่ ตัวอย่างในรูปที่ 4.5 , 4.6



รูปที่ 4.5 หน้าหลักตัวอย่างการปริวรรตอักษร



รูปที่ 4.6 หน้าตัวอย่างการปริวรรต

* 1. **ผลการทดสอบการตัวปริวรรตอักษร**

หลังจากทำการทดสอบการตัวปริวรรตอักษรเพื่อทดสอบความถูกต้องของส่วนตัวปริวรรตอักษร โดยทำการทดสอบกับชุดข้อมูลของการทดสอบ และวิธีการทดสอบ ที่กำหนดไว้ ผลการทดสอบมีรายละเอียดดังต่อไปนี้

* + 1. **ผลการทดสอบแบบแยกส่วน (Unit Testing)**

หลังจากทำการทดสอบ Unit Testing โดยใช้ PHPUnit (PHP Testing Framework) เป็นเครื่องมือในการตรวจสอบและทดสอบกฎ กับ Test Case โดยเป็นชุดข้อความผลลัพธ์ที่ถูกต้องและสามารถครอบคลุมกฎทั้งกรณีการปริวรรตแบบคงรูปและปรับรูป จำนวนอย่างละ 59 กรณี รวมเป็น 118 กรณี ได้ผลทดสอบดังต่อไปนี้ ตามตารางที่ 4.1

ตารางที่ 4.1 ผลการทดสอบแบบแยกส่วน (Unit Testing)

|  |  |  |  |
| --- | --- | --- | --- |
| **แบบการปริวรรต** | **จำนวน Test Case** | **Test Case ที่ผ่าน** | **ความถูกต้อง** |
| แบบคงรูป | 59 | 59 | 100% |
| แบบปรับรูป | 59 | 59 | 100% |

* + 1. **ผลการทดสอบโดยการเปรียบเทียบกับตัวอย่างการปริวรรต**

หลังจากทำการทดสอบ โดยการเปรียบเทียบกับตัวอย่างการปริวรรต ใช้ตัวอย่างต้นฉบับ คัมภีร์วัชรเฉทิกปรัชญาปารมิตา อักษรโรมัน ทำการปริวรรตเป็นอักษรไทยแบบคงรูปและปรับรูป นำผลลัพธ์การปริวรรตอักษรไทยแบบคงรูป ไปเปรียบเทียบกับ คัมภีร์วัชรเฉทิกปรัชญาปารมิตาสูตร ฉบับอักษรไทย และ ผลลัพธ์การปริวรรตอักษรไทยแบบปรับรูป ไปเปรียบเทียบกับ คัมภีร์วัชรเฉทิกปรัชญาปารมิตาสูตร ฉบับคำอ่านอักษรไทย โดยใช้โปรแกรม Pretty Diff v2.0.0 เป็นโปรแกรมเปรียบเทียบเพื่อคำนวณหาอัตราความถูกต้องของส่วนตัวปริวรรตอักษร ได้ผลทดสอบที่มีรายละเอียดต่อไปนี้

1. **การทดสอบโดยการเปรียบเทียบกับตัวอย่างการปริวรรตแบบคงรูป**

ต้นฉบับตัวอย่าง คือ คัมภีร์วัชรเฉทิกปรัชญาปารมิตาสูตร ฉบับอักษรไทย จำนวน มีจำนวนอักษร 31,436 ตัวอักษร หากไม่พิจารณาการเว้นวรรคและเครื่องวรรคตอนมี 27,440 ตัวอักษร ผลลัพธ์การปริวรรตจากเครื่อง คือการปริวรรตคัมภีร์วัชรเฉทิกปรัชญาปารมิตา ฉบับอักษรโรมัน เป็นอักษรไทยแบบคงรูป มีจำนวนอักษร 31,326 ตัวอักษร หากไม่พิจารณาการเว้นวรรคและเครื่องวรรคตอนมี 27,467 ตัวอักษร

เมื่อเปรียบเทียบความแตกต่างโดยไม่พิจารณาการเว้นวรรคและเครื่องวรรคตอนพบว่า มีจุดแตกต่าง 503 ตัวอักษร ดังสรุปได้ในตารางที่ 4.2

ตารางที่ 4.2 ผลการทดสอบโดยการเปรียบเทียบกับตัวอย่างการปริวรรตแบบคงรูป

|  |  |  |  |
| --- | --- | --- | --- |
| **ผลลัพธ์การปริวรรตจากเครื่อง** | **ต้นฉบับตัวอย่าง**  **(ตัวอักษร)** | **จุดแตกต่าง (ตัวอักษร)** | **ความถูกต้อง** |
| แบบคงรูป | 27,440 | 503 | 98.16 % |

ผลลัพธ์ จากตารางที่ 4.2 โดยภาพรวมมีความใกล้เคียงกันมากถึง 98.16 % และจากการสังเกตในส่วนที่มีความแตกต่าง พบว่าต้นฉบับใช้การปริวรรต แบบใส่พินทุใต้พยัญชนะที่ประสมด้วยสระ ฤ ฤๅ ฦ ฦๅ และการถอดพยัญชนะสังโยค กับ รูปสระ เอ ไอ โอ เอา นั้นต้นฉบับ ใช้แบบถอดรูปสระและพยัญชนะแบบปกติทั้งหมด ละเว้นพยัญชนะสังโยค

1. **การทดสอบโดยการเปรียบเทียบกับตัวอย่างการปริวรรตแบบปรับรูป**

ต้นฉบับตัวอย่าง คือ คัมภีร์วัชรเฉทิกปรัชญาปารมิตาสูตร ฉบับคำอ่านอักษรไทย มีจำนวนอักษร 34,117 ตัวอักษร หากไม่พิจารณาการเว้นวรรคและเครื่องวรรคตอนมี 30,049 ตัวอักษร ผลลัพธ์การปริวรรตจากเครื่อง คือการปริวรรตคัมภีร์วัชรเฉทิกปรัชญาปารมิตา ฉบับอักษรโรมัน เป็นอักษรไทยแบบปรับรูป มีจำนวนอักษร 34,064 ตัวอักษร หากไม่พิจารณาการเว้นวรรคและเครื่องวรรคตอนมี 30,268 ตัวอักษร เมื่อเปรียบเทียบความแตกต่างโดยไม่พิจารณาการเว้นวรรคและเครื่องวรรคตอนพบว่า มีจุดแตกต่าง 1,256 ตัวอักษร ดังสรุปได้ในตารางที่ 4.3

ตารางที่ 4.3 ผลการทดสอบโดยการเปรียบเทียบกับตัวอย่างการปริวรรตแบบปรับรูป

|  |  |  |  |
| --- | --- | --- | --- |
| **ผลลัพธ์การปริวรรตจากเครื่อง** | **ต้นฉบับตัวอย่าง**  **(ตัวอักษร)** | **จุดแตกต่าง (ตัวอักษร)** | **ความถูกต้อง** |
| แบบปรับรูป | 30,049 | 1,256 | 95.82 % |

ผลลัพธ์ จากตารางที่ 4.3 โดยภาพรวมมีความใกล้เคียงกันถึง 95.82 % ความถูกต้องลดลงเมื่อเปรียบเทียบกับการการปริวรรตแบบคงรูป เนื่องจากการปริวรรตแบบนี้มีหลักการอย่างกว้าง ๆ ยังไม่มีหลักที่แน่ชัด แต่อย่างไรก็ตามเปอร์เซ็นต์ความถูกต้องอยู่ในระดับสูงที่รับได้ และจากการสังเกตในส่วนที่มีความแตกต่าง พบว่าต้นฉบับใช้การปริวรรต ในการถอดอนุสวาระและจันทรพินทุ เป็นแบบใช้ ง โดยไม่เข้าสนธิ ซึ่งเป็นจุดที่แตกต่างส่วนมาก การถอด ร-เรผะ (รฺ) ต้นฉบับใช้การใช้แบบ ร-หันและ แบบ ร ตัวเดียว ผสมกันไป ซึ่งอาจจะเป็นแบบที่ให้ใกล้เคียงคำที่ใช้ในภาษาไทย เช่น อรหันตะ , ธรรมะ ผลลัพธ์จากระบบได้   
อรรหันตะ , ธรรมะ การถอดพยัญชนะสังโยค กับ รูปสระ เอ ไอ โอ เอา นั้นต้นฉบับ ใช้แบบถอดรูปสระและพยัญชนะแบบปกติทั้งหมด ละเว้นพยัญชนะสังโยค และการแปลง   
วิสรรชณีย์และไม้หันอากาศมีจุดแตกต่างอยู่หลายที่เช่นกัน

* + 1. **ผลการทดสอบประสิทธิภาพของระบบ**

หลังจากพิสูจน์ขั้นตอนวิธีในการปริวรรตอักษร 3 วิธีที่ออกแบบไว้ โดยทดสอบกับตัวอักษรโรมันต้นฉบับ จำนวน 34,769 ตัวอักษร ขั้นตอนวิธีมีประสิทธิภาพมากที่สุด คือ วิธีที่ 2. การปริวรรตที่ละบรรทัด มีเวลาที่ใช้โดยเฉลี่ยจากการ execute 5 ครั้งที่น้อยที่สุด เมื่อพัฒนาระบบจนมีกฎครบทุกกฎแล้ว ทำการทดสอบอีกครั้ง ได้ผลลัพธ์ตามตารางที่ 4.4

ตารางที่ 4.4 ผลการทดสอบประสิทธิภาพของระบบ

|  |  |  |
| --- | --- | --- |
| **ขั้นตอนวิธีในการปริวรรตอักษร** | **เวลาที่ใช้โดยเฉลี่ย(วินาที)** | **เวลาที่เพิ่มขึ้นจากเดิม** |
| วิธีที่ 1. ทำการปริวรรตทั้งข้อความ | 20.72 | 52.41 % |
| วิธีที่ 2. ทำการปริวรรตที่ละบรรทัด | 1.24 | 50.80 % |
| วิธีที่ 3. ทำการปริวรรตที่ละคำในแต่ละบรรทัด | 1.55 | 50.96 % |

* 1. **ผลการประเมินผลการใช้งานจากผู้เชี่ยวชาญภาษาสันสกฤต**

แบบประเมินผลการใช้งาน เพื่อใช้ประเมินผลการใช้งานและสอบถามความคิดเห็นเกี่ยวกับการใช้งานระบบ มีวัตถุประสงค์เพื่อประเมินหาระดับความพึงพอใจของ การปริวรรตอักษรด้วยเครื่องสำหรับภาษาสันสกฤตแบบอักษรโรมันเป็นอักษรไทย ที่ได้พัฒนาขึ้น ทำการประเมินโดย ท่านอาจารย์ธวัชชัย ดุลยสุจริต เป็นผู้เชี่ยวชาญด้านภาษาสันสกฤต ได้ทดสอบใช้ระบบระบบและประเมิน ซึ่งแบบประเมินผลการใช้งานได้แบ่งออกเป็น 3 ตอน ประกอบด้วย

1. แบบประเมินผล ด้านการทำงานได้ตามฟังก์ชันงานของระบบ (Function)
2. แบบประเมินผล ด้านความง่ายต่อการใช้งานระบบ (Usability)
3. แบบประเมินผล ส่วนความเห็นและข้อเสนอแนะจากผู้เชี่ยวชาญ

โดยสรุปผลการประเมินผลการใช้งานจากผู้เชี่ยวชาญภาษาสันสกฤต ได้ดังต่อไปนี้

* + 1. **ผลการประเมิน ด้านการทำงานได้ตามฟังก์ชันงานของระบบ (Function)**

โดยจะประเมิน ความถูกต้องของการผลลัพธ์ ความรวดเร็วในการประมวลผล ความครอบคลุมของงาน ความสามารถป้องกันข้อผิดพลาดที่เกิดขึ้น และความน่าเชื่อถือได้ของระบบ โดยมีงานที่จะประเมินดังนี้

1. การปริวรรตจากอักษรโรมันเป็นอักษรไทยแบบคงรูป
2. การปริวรรตจากอักษรโรมันเป็นอักษรไทยแบบปรับรูป
3. ตารางการเปรียบเทียบผลการปริวรรต
4. ตารางการเปรียบเทียบอักษร

โดยมีผลการประเมิน ตามตารางที่ 4.5, 4.6, 4.7, 4.8

ตารางที่ 4.5 ผลการประเมินการปริวรรตจากอักษรโรมันเป็นอักษรไทยแบบคงรูป

|  |  |  |
| --- | --- | --- |
| **รายการประเมิน** | **ระดับความพึงพอใจ** | **คะแนน** |
| ความถูกต้องของการผลลัพธ์ | ระดับมากที่สุด | 5 |
| ความรวดเร็วในการประมวลผล | ระดับมากที่สุด | 5 |
| ความครอบคลุมของงาน | ระดับมากที่สุด | 5 |
| ความสามารถป้องกันข้อผิดพลาดที่เกิดขึ้น | ระดับมากที่สุด | 5 |
| ความน่าเชื่อถือได้ของระบบ | ระดับมากที่สุด | 5 |
| ระดับความพึงพอใจโดยรวมอยู่ใน ระดับมากที่สุด (ค่าเฉลี่ย 5.00 ) คิดเป็น 100.00% | | |

ตารางที่ 4.6 ผลการประเมินการปริวรรตจากอักษรโรมันเป็นอักษรไทยแบบปรับรูป

|  |  |  |
| --- | --- | --- |
| **รายการประเมิน** | **ระดับความพึงพอใจ** | **คะแนน** |
| ความถูกต้องของการผลลัพธ์ | ระดับมากที่สุด | 5 |
| ความรวดเร็วในการประมวลผล | ระดับมากที่สุด | 5 |
| ความครอบคลุมของงาน | ระดับมาก | 4 |
| ความสามารถป้องกันข้อผิดพลาดที่เกิดขึ้น | ระดับมาก | 4 |
| ความน่าเชื่อถือได้ของระบบ | ระดับมากที่สุด | 5 |
| ระดับความพึงพอใจโดยรวมอยู่ใน ระดับมากค่อนไปทางมากที่สุด (ค่าเฉลี่ย 4.60 ) คิดเป็น 92.00% | | |

ตารางที่ 4.7 ผลการประเมินตารางการเปรียบเทียบผลการปริวรรต

|  |  |  |
| --- | --- | --- |
| **รายการประเมิน** | **ระดับความพึงพอใจ** | **คะแนน** |
| ความถูกต้องของการผลลัพธ์ | ระดับมากที่สุด | 5 |
| ความรวดเร็วในการประมวลผล | ระดับมากที่สุด | 5 |
| ความครอบคลุมของงาน | ระดับมากที่สุด | 5 |
| ความสามารถป้องกันข้อผิดพลาดที่เกิดขึ้น | ระดับมาก | 4 |
| ความน่าเชื่อถือได้ของระบบ | ระดับมากที่สุด | 5 |
| ระดับความพึงพอใจโดยรวมอยู่ใน ระดับมากค่อนไปทางมากที่สุด (ค่าเฉลี่ย 4.80 ) คิดเป็น 96.00% | | |

ตารางที่ 4.8 ผลการประเมินตารางการเปรียบเทียบอักษร

|  |  |  |
| --- | --- | --- |
| **รายการประเมิน** | **ระดับความพึงพอใจ** | **คะแนน** |
| ความถูกต้องของการผลลัพธ์ | ระดับมากที่สุด | 5 |
| ความรวดเร็วในการประมวลผล | ระดับมากที่สุด | 5 |
| ความครอบคลุมของงาน | ระดับมากที่สุด | 5 |
| ความสามารถป้องกันข้อผิดพลาดที่เกิดขึ้น | ระดับมาก | 4 |
| ความน่าเชื่อถือได้ของระบบ | ระดับมากที่สุด | 5 |
| ระดับความพึงพอใจโดยรวมอยู่ใน ระดับมากค่อนไปทางมากที่สุด (ค่าเฉลี่ย 4.80 ) คิดเป็น 96.00% | | |

* + 1. **แบบประเมินผล ด้านความง่ายต่อการใช้งานระบบ (Usability)**

โดยจะประเมินจากส่วนต่อผู้ใช้งาน(User Interface) ผลการประเมิน ตามตารางที่ 4.9

ตารางที่ 4.9 ผลการประเมินด้านความง่ายต่อการใช้งานระบบ (Usability)

|  |  |  |
| --- | --- | --- |
| **รายการประเมิน** | **ระดับความพึงพอใจ** | **คะแนน** |
| ความง่ายต่อการใช้งานของระบบ | ระดับมาก | 4 |
| ความเหมาะสมในการเลือกใช้ชนิดตัวอักษรบนจอภาพ | ระดับมากที่สุด | 5 |
| ความเหมาะสมในการใช้สีของตัวอักษรและรูปภาพ | ระดับมากที่สุด | 5 |
| ความเหมาะสมในการใช้ข้อความเพื่ออธิบายสื่อความหมาย | ระดับมากที่สุด | 5 |
| ความเหมาะสมในการใช้สัญลักษณ์หรือรูปภาพในการสื่อความหมาย | ระดับมากที่สุด | 5 |
| ความเป็นมาตรฐานเดียวกันในการออกแบบหน้าจอภาพ | ระดับมากที่สุด | 5 |
| ความเหมาะสมในการปฏิสัมพันธ์โต้ตอบกับผู้ใช้ | ระดับมากที่สุด | 5 |
| ความเหมาะสมในการวางตำแหน่งของส่วนประกอบบนจอภาพ | ระดับมาก | 4 |
| คำศัพท์ที่ประกอบใช้ผู้ใช้มีความคุ้นเคยและสามารถปฏิบัติตามได้โดยง่าย | ระดับมากที่สุด | 5 |
| ระดับความพึงพอใจโดยรวมอยู่ใน ระดับมากค่อนไปทางมากที่สุด (ค่าเฉลี่ย 4.77 ) คิดเป็น 96.55% | | |

* + 1. **แบบประเมินผลส่วน ส่วนความเห็นและข้อเสนอแนะจากผู้เชี่ยวชาญ**

ผู้เชี่ยวชาญภาษาสันสกฤตได้ให้ความเห็นและข้อเสนอแนะดังต่อไปนี้

โปรแกรมนี้สามารถใช้งานปริวรรตอักษรภาษาสันสกฤตได้เป็นอย่างดี ตามความจำเป็นในการศึกษาภาษาสันสกฤต ช่วยให้ประหยัดเวลาในการปริวรรตได้อย่างมาก การปรับปรุง อาจจัดเนื้อหาให้กะทัดรัดมากขึ้นในหน้าจอเดียว เพื่อความสะดวกในการอ่าน ในอนาคต หากปริวรรตอักษรกลุ่ม Indic Scripts ได้กว้างขวางขึ้น ครอบคลุมอักษรในเอเชียตะวันออกเฉียงใต้ (เช่น มอญ พม่า ลาว เขมร อักษรธรรม ฯลฯ) จะเป็นเครื่องมือช่วยในการศึกษาภาษาและวรรณกรรมได้สะดวกยิ่งขึ้น

**บทที่ 5**

**สรุปผลและข้อเสนอแนะ**

การศึกษาโครงการเฉพาะเรื่อง การปริวรรตอักษรด้วยเครื่องสำหรับภาษาสันสกฤตแบบอักษรโรมันเป็นอักษรไทย มีวัตถุประสงค์ ที่จะทำให้ผู้ใช้ที่มีความสนใจศึกษาค้นคว้าด้านภาษาสันสกฤตและสามารถปริวรรตภาษาสันสกฤตแบบโรมันเป็นไทยได้อย่างรวดเร็ว โดยการศึกษานี้นำเสนอกฎ,ขั้นตอนวิธีและวิธีการพัฒนา การปริวรรตอักษรด้วยเครื่องสำหรับภาษาสันสกฤตแบบอักษรโรมันมาตรฐาน IAST เป็นอักษรไทย ใน 2 รูปแบบ คือ อักษรไทยในรูปแบบคงรูปหรือแบบแผน และอักษรไทยในรูปแบบปรับรูปหรือแบบทั่วไป มีการประเมินความสมบูรณ์ของการปริวรรตอักษร โดยวิธีการทดสอบและเปรียบเทียบผลลัพธ์ที่ได้จากระบบกับตัวอย่างการปริวรรต และผู้เชี่ยวชาญภาษาสันสกฤตทำการประเมินผลการใช้งาน หลังจากได้ทำการศึกษาโครงการเฉพาะเรื่องเป็นที่เสร็จเรียบร้อยแล้ว จึงมีการสรุปผลและข้อเสนอแนะดังต่อไปนี้

* 1. **สรุปผล**

จากการศึกษาโครงการเฉพาะเรื่องสรุปได้ว่า ในส่วนของการพัฒนา การปริวรรตอักษรใน รูปแบบคงรูปหรือแบบแผนนั้นใช้กฎทั้งหมด จำนวน 13 กฎ ส่วนรูปแบบปรับรูปหรือแบบทั่วไป ใช้กฎทั้งหมด จำนวน 14 กฎ ขั้นตอนวิธีในการปริวรรตอักษรที่มีประสิทธิภาพมากที่สุด คือ การปริวรรตที่ละบรรทัด โดยวิธีนี้เป็นการประมวลผลที่ใช้เวลาน้อยกว่าวิธีอื่น

สรุปการทดสอบ Unit Testing การปริวรรตอักษรในรูปแบบคงรูปหรือแบบแผนใช้การทดสอบ 59 กรณีโดยผลการทดสอบปรากฏว่า ผ่านทุกกรณี ส่วนรูปแบบปรับรูปหรือแบบทั่วไป ทำการทดสอบ 59 กรณี โดยผลการทดสอบปรากฏว่า ผ่านทุกกรณีเช่นกัน

สรุปการทดสอบโดยการเปรียบเทียบกับตัวอย่างการปริวรรต การปริวรรตอักษรในรูปแบบคงรูปหรือแบบแผน ทำการเปรียบเทียบกับ ต้นฉบับตัวอย่างที่มีความยาวตัวอักษร 27,440 ตัวอักษร มีจุดแตกต่างกันจำนวน 503 ตัวอักษร คิดเป็นเปอร์เซ็นต์ความถูกต้องได้ 98.16 % ส่วนรูปแบบปรับรูปหรือแบบทั่วไป ทำการเปรียบเทียบกับ ต้นฉบับตัวอย่างที่มีความยาวตัวอักษร 30,049 ตัวอักษร มีจุดแตกต่างกันจำนวน 1,256 ตัวอักษร คิดเป็นเปอร์เซ็นต์ความถูกต้อง ได้ 95.82 %

สรุปผลการประเมินผลการใช้งานจากผู้เชี่ยวชาญภาษาสันสกฤต ด้านการทำงานได้ตามฟังก์ชันงานของระบบ (Function) การปริวรรตจากอักษรโรมันเป็นอักษรไทยแบบคงรูป มีระดับความพึงพอใจโดยรวมอยู่ใน ระดับมากที่สุด คิดเป็น 100.00% การปริวรรตจากอักษรโรมันเป็นอักษรไทยแบบปรับรูป ระดับความพึงพอใจโดยรวมอยู่ใน ระดับมากค่อนไปทางมากที่สุด คิดเป็น 92.00% ตารางการเปรียบเทียบผลการปริวรรต ระดับความพึงพอใจโดยรวมอยู่ใน ระดับมากค่อนไปทางมากที่สุด คิดเป็น 96.00% ตารางการเปรียบเทียบอักษร ระดับความพึงพอใจโดยรวมอยู่ใน ระดับมากค่อนไปทางมากที่สุด คิดเป็น 96.00% การประเมินด้านความง่ายต่อการใช้งานระบบ (Usability) ระดับความพึงพอใจโดยรวมอยู่ใน ระดับมากค่อนไปทางมากที่สุด คิดเป็น 96.55% และจากความเห็นและข้อเสนอแนะจากผู้เชี่ยวชาญเห็นว่า ระบบนี้สามารถใช้งานปริวรรตอักษรภาษาสันสกฤตได้เป็นอย่างดี ตามความจำเป็นในการศึกษาภาษาสันสกฤต ช่วยให้ประหยัดเวลาในการปริวรรตได้อย่างมาก

* 1. **ข้อเสนอแนะ**

โครงการศึกษาเฉพาะเรื่องในครั้งนี้เป็นเพียงการศึกษาการปริวรรตจากอักษรโรมันเป็นอักษรไทยเฉพาะบางกรณีที่เลือกมาเท่านั้น ยังไม่ได้ครอบคลุมลักษณะการปริวรรตทุกรูปแบบ เนื่องปัญหาหลักของการปริวรรตก็คือยังไม่มีหน่วยงานที่เกี่ยวของเข้ามาวางหลักเกณฑ์การปริวรรตอักษร เช่น ราชบัณฑิตยสถาน ดังนั้น จึงมีการปริวรรตอักษรที่แตกต่างกันไปตามตำราอาจารย์ แต่อย่างไรก็ตามผู้จัดทำได้จัดเตรียมตารางเปรียบเทียบเพื่อการพิสูจน์อักษรไว้ให้ อีกทั้งปัญหาการปริวรรตเป็นอักษรไทยแบบปรับรูป จะมีปัญหาเรื่องการตัดสินใจในการการวางตำแหน่งรูปสระหน้าหรือการเปลี่ยนรูปของสระอะเป็นไม้หันอากาศ กรณีพยัญชนะตัวควบหรือตัวซ้อน กับตัวสะกด เมื่อผสมกับสระ ยกตัวอย่าง เช่น vajrac ควรเป็น วัชรัจ หรือ วะชรัช หรือ dharmakṣetre ที่ถอดเสียง ธัน-ระ-มัก-เษต-เตร ควรเป็น ธรรมะเกษตเร , ธรรมักเษตเร , ธรรมะเกษเตร,ธรรมักเษเตร ซึ่งการเพิ่มกฎและตัวเลือกการปริวรรตในกรณีใหม่ๆ หรือตัวอย่าง(Example Base)เข้าไปช่วยปริวรรต หรือใช้การปริวรรตอักษรแบบใช้สถิติ (Statistical Approach) โดยมีพจนานุกรมศัพท์การปริวรรตเข้าทำงานร่วมกัน และอีกงานที่ท้าท้าย คือใช้แนวทางของ AI คือ Neural Network วิธีการเหล่านี้น่าจะเป็นแนวทางในการแก้ไขปัญหาที่เกิดขึ้นทั้งหมดอีกทางหนึ่ง