**บทที่ 1 บทนำ**

**1.1 ความสำคัญและที่มาของงานวิจัย**

ภาษาสันสกฤตเป็นภาษาโบราณของอินเดีย ที่มีอายุเก่าแก่ถึงประมาณ 3,500 ปี ภาษาสันสกฤต เป็นภาษาที่รับอิทธิพลมาจากอินเดียและส่งผลมาถึงอาณาจักรในแถบเอเชียตะวันออกเฉียงใต้ เป็นภาษาที่เก่าแก่ที่สุดภาษาหนึ่งในภาษากลุ่มอินโด-ยูโรเปียน (หรืออินเดีย-ยุโรป) สาขาย่อยอินโด-อิเรเนียน (อินเดีย-อิหร่าน) และอยู่ในกลุ่มย่อยอินโด-อารยัน (อินเดีย-อารยะ) มีวรรณคดีจำนวนมากที่ประพันธ์เป็นภาษาสันสกฤต อีกทั้งในภูมิภาคเอเชียตะวันออกเฉียงใต้ หลายเรื่องเป็นที่รู้จักในหมู่ชาวไทย อาทิ รามายณะ มหาภารตะ เป็นต้น ในบริบทของไทย มีศัพท์มากมายที่เป็นคำยืมภาษาสันสกฤต อีกทั้งวรรณคดีไทยมีความสัมพันธ์อย่างลึกซึ้งกับวรรณคดีสันสกฤต การศึกษาภาษาสันสกฤตจึงมีคุณประโยชน์อย่างยิ่งในการเข้าใจภาษาและวรรณคดี รวมทั้งวัฒนธรรมไทย ที่มีความเกี่ยวพันกับวัฒนธรรมอินเดียมาแต่โบราณ ในปัจจุบัน(ปี 2559 )ประเทศไทยมีมหาวิทยาลัยที่เปิดสอนภาษาสันสกฤตถึง 4 แห่งคือ 1.มหาวิทยาลัยศิลปากร 2.จุฬาลงกรณ์มหาวิทยาลัย 3.มหามกุฏราชวิทยาลัย และ 4.มหาจุฬาลงกรณราชวิทยาลัย ผุ้มีชื่อเสียงของประเทศไทย สมเด็จพระเทพรัตนราชสุดาฯ สยามบรมราชกุมารี สนพระราชหฤทัยสันสกฤตศึกษามาตั้งแต่ยังทรงพระเยาว์ และทรงสำเร็จการศึกษาสาขาจารึกภาษาตะวันออก คณะโบราณคดี มหาวิทยาลัยศิลปากร ทั้งยังทรงส่งเสริมและสนับสนุนด้านสันสกฤตศึกษามาโดยตลอด

ในกรณีการคำนวณมูลหนี้ คดีล้มละลาย ในระบบสารสนเทศของกรมบังคับคดี ทั้งมูลหนี้ตามคำพิพากษาและมูลหนี้อื่นซึ่งมิใช่หนี้ตามคำพิพากษา มีสูตรและวิธีคำนวณ ที่หลากหลาย เกิดจากสัญญาการทำธุรกรรมทางการเงินในแต่ละคดีมีความแตกต่างกัน ส่งผลให้มูลหนี้ในสำนวนคดีหรือคำพิพากษา มีการคำนวณต่างกันซึ่งเป็นไปตามข้อบังคับทางกฎหมาย ปกติการคำนวณมูลหนี้ จะเป็นหน้าที่ของเจ้าพนักงานบัญชีหรือเจ้าพนักงานปฏิบัติการ กองติดตามและเฉลี่ยทรัพย์ในคดีล้มละลาย กลุ่มงานบัญชี และจัดการติดตามทรัพย์สิน เป็นผู้คำนวณโดยเจ้าพนักงานปฏิบัติการจะรับสำนวนจากเจ้าพนักงานพิทักษ์ทรัพย์ จากนั้นเจ้าพนักงานบัญชีหรือเจ้าพนักงานปฏิบัติการทำการถอดสำนวนแล้วกรอกค่าต่างๆเข้าไปในระบบจากนั้นก็จะได้ผลลัพธ์ออกมา แต่ปัญหาที่เกิดขึ้นคือ การคำนวณมูลหนี้ คดีล้มละลาย ในระบบสารสนเทศของกรมบังคับคดีนั้น หากเป็นสำนวนที่ไม่ซับซ้อนมากนักระบบก็สามารถออกผลลัพธ์ได้อย่างถูกต้อง หากมีเงื่อนไขสูตรและวิธีคำนวณหลากหลายซับซ้อนมากเกิน เกินความสามารถของระบบ เจ้าพนักงานบัญชีก็จะใช้วิธีการคำนวณเอง ซึ่งการระบบสารสนเทศของกรมบังคับคดีนั้น หากมีการปรับเปลี่ยนวิธี หรือเพิ่มวิธีการคำนวณใหม่ ผู้ใช้ระบบไม่สามารถจะทำได้เอง หรือบางครั้งก็เกิดจากการทำงานผิดพลาดจากระบบเอง ต้องอาศัยนักพัฒนาระบบทำการแก้ไขที่ระบบสารสนเทศหลัก ส่งผลทำให้เกิดมูลค่าเพิ่มและความยุ่งยากในการแก้ไขสูตรต่างๆ

การศึกษาโครงการเฉพาะเรื่อง จึงได้เสนอแนวทางการประยุกต์ใช้ภาษาเฉพาะโดเมนมาขึ้นมารับรองการสร้างสูตรวิธีคำนวณมูลหนี้ คดีล้มละลาย ที่หลากหลาย เพื่อให้ผู้ใช้งานระบบ มีความยืดหยุ่นในการปรับแต่งหรือเพิ่มสูตรวิธีการคำนวณใหม่ได้อย่างครอบคลุม เพื่อให้ได้ผลลัพธ์ที่ต้องการ และสามารถตรวจสอบวิธีการคำนวณ โดยไม่ต้องทำการแก้ไขหรือการโปรแกรมใหม่ที่ระบบสารสนเทศหลัก เพื่อเป็นตัวช่วยหนึ่งในการคำนวณในกรณีระบบสารสนเทศของกรมบังคับคดีไม่สามารถทำการคำนวณในกรณีนั้นๆได้ อีกทั้งจะสามารถต่อประสานกับระบบที่จะพัฒนาในอนาคตได้

**1.2 วัตถุประสงค์ของการศึกษาโครงการเฉพาะเรื่อง**

การศึกษาโครงการเฉพาะเรื่อง การปริวรรตอักษรด้วยเครื่องสำหรับภาษาสันสกฤตแบบโรมันเป็นไทย มีวัตถุประสงค์ดังต่อไปนี้

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

**1.3 ประโยชน์และผลที่ได้รับจากการศึกษาโครงการเฉพาะเรื่อง**

ประโยชน์และผลที่ได้รับจากการศึกษาโครงการเฉพาะเรื่อง การปริวรรตอักษรด้วยเครื่องสำหรับภาษาสันสกฤตแบบโรมันเป็นไทย มีดังต่อไปนี้

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

**1.4 ขอบเขตงานการศึกษาโครงการเฉพาะเรื่อง**

การศึกษาโครงการเฉพาะเรื่อง การปริวรรตอักษรด้วยเครื่องสำหรับภาษาสันสกฤตแบบโรมันเป็นไทย 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 3.0.0
   4. Front-End Framework ใช้ Bootstrap v3.3.5
2. การศึกษาโครงการเฉพาะเรื่องนี้ จะทำการพัฒนาเพียงการปริวรรตภาษาสันสกฤตแบบอักษรโรมันมาตรฐาน International Alphabet of Sanskrit Transliteration (I.A.S.T) เท่านั้น
3. การศึกษาโครงการเฉพาะเรื่องนี้ มีลักษณะขั้นตอนการปริวรรตอยู่ 2 แบบ คือ
   1. การปริวรรตอักษรแบบไปข้างหน้า(Forward Transliteration) ทำการปริวรรตอักษรภาษาสันสกฤตจาก แบบอักษรโรมันมาตรฐาน I.A.S.T เป็น อักษรไทยในรูปแบบทั่วไป(แบบปรับรูป) และอักษรไทยในรูปแบบแบบแผน(แบบคงรูป)
   2. การปริวรรตอักษรแบบย้อนกลับ(backward Transliteration)ทำการปริวรรตอักษรภาษาสันสกฤตจาก อักษรไทยในรูปแบบแบบแผน(แบบคงรูป) เป็น แบบอักษรโรมันมาตรฐาน I.A.S.T แต่ไม่จะไม่ทำการปริวรรตอักษรภาษาสันสกฤตจาก อักษรไทยในรูปแบบทั่วไป(แบบปรับรูป) เป็น แบบอักษรโรมันมาตรฐาน I.A.S.T ได้เนื่องจาก
4. ประเมินความสมบูรณ์ของการปริวรรตอักษร โดยวิธีการใช้ผู้เชี่ยวชาญภาษาสันสกฤต ทดสอบและเปรียบเทียบผลลัพธ์ที่ได้จากระบบ

**1.5 กระบวนการศึกษาโครงการเฉพาะเรื่อง**

กระบวนการศึกษาโครงการเฉพาะเรื่อง การปริวรรตอักษรด้วยเครื่องสำหรับภาษาสันสกฤตแบบโรมันเป็นไทย 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. ทดสอบและประเมินผล จากผู้เชี่ยวชาญภาษาสันสกฤต
9. สรุปผลการวิจัยและข้อเสนอแนะ
10. จัดทำและเสนอรายงานการวิจัย

**บทที่ 2 ทฤษฎีและงานวิจัยที่เกี่ยวข้อง**

เนื้อหาในส่วนนี้นำเสนอทฤษฎีและงานวิจัยที่เกี่ยวข้องกับ การปริวรรตอักษรด้วยเครื่องสำหรับภาษาสันสกฤตแบบโรมันเป็นไทย มีดังต่อไปนี้

**2.1 ทฤษฎีที่เกี่ยวข้อง**

**2.1.1 การปริวรรตอักษรหรือการเขียนคำทับศัพท์ (transliteration)**

การทับศัพท์ คือการดำเนินการแปลงข้อความจากระบบการเขียนหรือภาษาหนึ่งไปสู่อีกระบบหนึ่งอย่างมีหลักการ เพื่อให้สามารถเขียนคำในภาษาต่างประเทศด้วยภาษาและอักษรในภาษานั้น ๆ ได้สะดวก เช่น การทับศัพท์ภาษาอังกฤษซึ่งเขียนด้วยอักษรโรมัน มาเป็นอักษรไทยเพื่อใช้ในภาษาไทย หรือการทับศัพท์ภาษาไทย ไปเป็นอักษรโรมันเพื่อใช้ในภาษาอังกฤษ เป็นต้น ส่วนมากใช้กับวิสามานยนาม อาทิ ชื่อบุคคล สถานที่ หรือชื่อเฉพาะที่ไม่สามารถแปลความหมายเป็นภาษาอื่นได้โดยสะดวก

ปกติแล้วการทับศัพท์คือการจับคู่จากระบบการเขียนหนึ่งไปยังอีกระบบหนึ่งแบบคำต่อคำ หรือตามทฤษฎีคืออักษรต่ออักษร การทับศัพท์ได้พยายามที่จะสร้างความสัมพันธ์หนึ่งต่อหนึ่งทั่วถึงและทำให้เกิดความถูกต้องแม่นยำ เพื่อให้ผู้อ่านที่ได้รับรู้สามารถสะกดคำต้นฉบับจากคำทับศัพท์ได้ และเพื่อที่จะบรรลุจุดประสงค์นี้ จึงมีการกำหนดหลักการทับศัพท์ที่ซับซ้อนในการจัดการกับตัวอักษรบางตัวในภาษาต้นฉบับที่ไม่สัมพันธ์กับอักษรในภาษาเป้าหมาย ความหมายอย่างแคบของการทับศัพท์คือ การทับศัพท์แบบถอดอักษร (transliteration) ซึ่งเคร่งครัดในการคงตัวอักษรและเครื่องหมายวรรคตอนทุกอย่างเอาไว้ ทั้งนี้การถอดอักษรไม่สนใจความแตกต่างของเสียงในภาษา ตัวอย่างหนึ่งของการถอดอักษรคือการใช้แป้นพิมพ์ภาษาอังกฤษ พิมพ์แทนภาษาอื่นที่ใช้ตัวอักษรต่างออกไปเช่นภาษารัสเซียเนื่องจากมีข้อจำกัดทางเทคนิค หรือการถอดอักษรโบราณเพื่อให้ยังคงรักษารูปแบบการเขียนเดิมเอาไว้

การทับศัพท์ (การถอดอักษร) ต่างจากการถอดเสียง (transcription) ซึ่งเป็นการจับคู่เสียงอ่านของภาษาหนึ่ง ๆ ไปยังรูปแบบการเขียนของอีกภาษาที่ใกล้เคียงที่สุด ถึงแม้ว่าระบบการถอดอักษรส่วนใหญ่จะยังคงจับคู่อักษรต้นฉบับกับอักษรในภาษาเป้าหมายที่ออกเสียงคล้ายกันในบางคู่ ถ้าหากความสัมพันธ์ระหว่างตัวอักษรกับเสียงเหมือนกันทั้งสองภาษา การถอดอักษรก็อาจแทบจะเหมือนกับการถอดเสียง ในทางปฏิบัติ ก็มีการทับศัพท์บางระบบที่ผสมกันระหว่างการถอดอักษรและการถอดเสียง โดยจะถอดอักษรต้นฉบับบางส่วนและถอดเสียงในส่วนที่เหลือ การถอดเสียงจะพยายามหาแนวทางที่ดีที่สุดในการเขียนภาษาต่างประเทศให้เป็นภาษาเฉพาะ เช่นการเขียนคำในภาษาอังกฤษเป็นภาษาไทยโดยไม่สลับภาษาบนแป้นพิมพ์ (บางครั้งผู้ใช้งานอาจพบว่าการกระทำเช่นนี้อาจทำให้อ่านเข้าใจยากกว่าการพิมพ์โดยสลับภาษาตามปกติ) การถอดเสียงขณะพิมพ์จึงถือเป็นกระบวนการประยุกต์โดยแท้จริงเพื่อการป้อนข้อความในภาษาเฉพาะเช่นนั้น แต่ก็มีข้อเสียคืออาจทำให้ไม่สามารถคาดเดาเพื่อถอดกลับไปเป็นภาษาเดิมได้ เพราะมีอักษรที่เพิ่มเข้ามาหรือถูกตัดออกไป หรือเปลี่ยนแปลงรูปแบบจนไม่เหลือเค้าโครงเดิม ความหมายอย่างกว้างของการทับศัพท์จึงหมายรวมทั้งการทับศัพท์แบบถอดอักษรและการทับศัพท์แบบถอดเสียง

นอกจากนี้ ไม่ควรสับสนระหว่างการทับศัพท์กับการแปล ซึ่งเกี่ยวข้องกับการเปลี่ยนหรือเลือกคำในภาษาเพื่อสงวนความหมายดั้งเดิมเอาไว้ ในขณะที่การทับศัพท์เป็นการแปลงตัวอักษรเท่านั้น

**2.1.2 ภาษาสันสกฤต**

ภาษาสันสกฤต เป็นภาษาที่รับอิทธิพลมาจากอินเดียและส่งผลมาถึงอาณาจักรในแถบเอเชียตะวันออกเฉียงใต้ เป็นภาษาที่เก่าแก่ที่สุดภาษาหนึ่งในภาษากลุ่มอินโด-ยูโรเปียน (หรืออินเดีย-ยุโรป) สาขาย่อยอินโด-อิเรเนียน (อินเดีย-อิหร่าน) และอยู่ในกลุ่มย่อยอินโด-อารยัน (อินเดีย-อารยะ) โดยมีระดับวิวัฒนาการในระดับใกล้เคียงกับภาษาละตินและภาษากรีก เป็นต้น โดยทั่วไปถือว่าเป็นภาษาที่ตายแล้ว ทว่ายังมีผู้ใช้ภาษาสันสกฤตอยู่บ้างในแวดวงที่จำกัดในประเทศอินเดีย โดยมีการคิดคำศัพท์ใหม่ ๆ ขึ้นมาด้วย ในศาสนาฮินดูเชื่อว่า ภาษาสันสกฤตเป็นภาษาสื่อที่เทพเจ้าใช้สื่อสารกับมวลมนุษย์ เพื่อถ่ายทอดความรู้แจ้งและปัญญาญาณแก่เหล่าฤษีทั้งหลายแต่ครั้งดึกดำบรรพ์

คำว่า สํสฺกฤต (संस्कृत) แปลว่า "กลั่นกรองแล้ว" ส่วนคำว่า สํสฺกฤตา วากฺ (संस्कृता वाक्) จะใช้เพื่อเรียก "ภาษาที่กลั่นกรองแล้ว" ซึ่งเป็นภาษาของชนชั้นพราหมณ์ ตรงข้ามกับภาษาพูดของชาวบ้านทั่วไปที่เรียกว่าปรากฤต ภาษาสันสกฤตมีพัฒนาการในหลายยุคสมัย โดยมีหลักฐานเก่าแก่ที่สุด คือภาษาที่ปรากฏในคัมภีร์ฤคเวท (เมื่อราว 1,200 ปีก่อนคริสตกาล) อันเป็นบทสวดสรรเสริญพระเจ้าในลัทธิพราหมณ์ในยุคต้น ๆ อย่างไรก็ตาม ในการจำแนกภาษาสันสกฤตโดยละเอียด นักวิชาการอาจถือว่าภาษาในคัมภีร์ฤคเวทเป็นภาษาหนึ่งที่ต่างจากภาษาสันสกฤตแบบแผน (Classical language) และเรียกว่า ภาษาพระเวท (Vedic language) ภาษาพระเวทดั้งเดิมยังมิได้มีการวางกฎเกณฑ์ให้เป็นระเบียบรัดกุมและสละสลวย และมีหลักทางไวยากรณ์อย่างกว้าง ๆ ปรากฏอยู่ในบทสวดในคัมภีร์พระเวทของศาสนาฮินดู เนื้อหาคือบทสวดสรรเสริญเทพเจ้า เอกลักษณ์ที่ปรากฏอยู่เฉพาะในภาษาพระเวทคือระดับเสียง (Accent) ซึ่งกำหนดไว้อย่างเคร่งครัด และถือเป็นสิ่งสำคัญของการสวดพระเวทเพื่อให้สัมฤทธิผล

ภาษาสันสกฤตมีวิวัฒนาการมาจากภาษาชนเผ่าอารยัน หรืออินโดยูโรเปียน (Indo-European) บรรพบุรุษของพวกอินโด-อารยัน ตั้งรกรากอยู่เหนือเอเซียตะวันออก (ตอนกลางของทวีปเอเชีย - Central Asia) โดยไม่มีที่อยู่เป็นหลักแหล่ง กลุ่มอารยันต้องเร่ร่อนทำมาหากินเหมือนกันชนเผ่าอื่น ๆ ในจุดนี้เองที่ทำให้เกิดการแยกย้ายถิ่นฐาน การเกิดประเพณี และภาษาที่แตกต่างกันออกไป ชนเผ่าอารยันได้แยกตัวกันออกไปเป็น 3 กลุ่มใหญ่ กลุ่มที่ 1 แยกไปทางตะวันตกเข้าสู่ทวีปยุโรป กลุ่มที่ 2 ลงมาทางตะวันออกเฉียงใต้ อนุมานได้ว่าน่าจะเป็นชนชาติอิหร่านในเปอร์เซีย และกลุ่มที่ 3 เป็นกลุ่มที่สำคัญที่สุด กลุ่มนี้แยกลงมาทางใต้ตามลุ่มแม่น้ำสินธุ (Indus) ชาวอารยันกลุ่มนี้เมื่อรุกเข้าในแถบลุ่มแม่น้ำสินธุแล้ว ก็ได้ไปพบกับชนพื้นเมืองที่เรียกว่า ดราวิเดียน (Dravidian) และเกิดการผสมผสานทางวัฒนธรรมและภาษา โดยชนเผ่าอารยันได้นำภาษาพระเวทยุคโบราณเข้าสู่อินเดียพร้อม ๆ กับความเชื่อทางศาสนา ซึ่งในยุคต่อมาได้เกิดตำราไวยากรณ์ภาษาสันสกฤตคือ อษฺฏาธฺยายี (अष्टाध्यायी "ไวยากรณ์ 8 บท") ของปาณินิ เชื่อกันว่ารจนาขึ้นในช่วงพุทธกาล ปาณินิเห็นว่าภาษาสันสกฤตแบบพระเวทนั้นมีภาษาถิ่นปนเข้ามามากพอสมควรแล้ว หากไม่เขียนไวยากรณ์ที่เป็นระเบียบแบบแผนไว้ ภาษาสันสกฤตแบบพระเวทที่เคยใช้มาตั้งแต่ยุคพระเวทจะคละกับภาษาท้องถิ่นต่าง ๆ ทำให้การประกอบพิธีกรรมไม่มีความศักดิ์สิทธิ์ ดังนั้น จึงแต่งอัษฏาธยายีขึ้น ความจริงตำราแบบแผนไวยากรณ์ก่อนหน้าปาณินิได้มีอยู่ก่อนแล้ว แต่เมื่อเกิดอัษฏาธยายีตำราเหล่านั้นก็ได้หมดความนิยมลงและสูญไปในที่สุด ผลของไวยากรณ์ปาณินิก็คือภาษาเกิดการจำกัดกรอบมากเกินไป ทำให้ภาษาไม่พัฒนา ในที่สุด ภาษาสันสกฤตแบบปาณินิ หรือภาษาสันสกฤตแบบฉบับ จึงกลายเป็นภาษาเขียนในวรรณกรรม ซึ่งผู้ที่สามารถจะอ่าน เขียนและแปลได้จะต้องใช้เวลามากพอสมควร ภาษาสันสกฤตแบ่งได้เป็น 2 กลุ่มกว้าง ๆ ได้แก่ ภาษาสันสกฤตแบบแผน และภาษาสันสกฤตผสม

1. **ภาษาสันสกฤตแบบแผน**

เกิดขึ้นจากการวางกฎเกณฑ์ของภาษาสันสกฤตให้มีแบบแผนที่แน่นอนในสมัยต่อมา โดยนักปราชญ์ชื่อ ปาณินิ ตามประวัติเล่าว่าเป็นผู้เกิดในตระกูลพราหมณ์ แคว้นคันธาระราว 57 ปีก่อนพุทธปรินิพพาน บางกระแสว่าเกิดราว พ.ศ. 143 ปาณินิได้ศึกษาภาษาในคัมภีร์พระเวทจนสามารถหาหลักเกณฑ์ของภาษานั้นได้ จึงจัดรวบรวมขึ้นเป็นหมวดหมู่ เรียบเรียงเป็นตำราไวยากรณ์ขึ้น 8 บทให้ชื่อว่า อัษฏาธยายี มีสูตรเป็นกฎเกณฑ์อธิบายโครงสร้างของคำอย่างชัดเจน นักวิชาการสมัยใหม่มีความเห็นว่า วิธีการศึกษาและอธิบายภาษาของปาณินิเป็นวิธีวรรณนา คือศึกษาและอธิบายตามที่ได้สังเกตเห็นจริง มิได้เรียบเรียงขึ้นตามความเชื่อส่วนตัว มิได้เรียบเรียงขึ้นตามหลักปรัชญา คัมภีร์อัษฏาธยายีจึงได้รับการยกย่องว่าเป็นตำราไวยากรณ์เล่มแรกที่ศึกษาภาษาในแนววิทยาศาสตร์และวิเคราะห์ภาษาได้สมบูรณ์ที่สุด[ต้องการอ้างอิง] ความสมบูรณ์ของตำราเล่มนี้ทำให้เกิดความเชื่อในหมู่พราหมณ์ว่า ตำราไวยากรณ์สันสกฤตหรือปาณินิรจนานี้ สำเร็จได้ด้วยอำนาจพระศิวะ อย่างไรก็ตาม นักภาษาศาสตร์เชื่อว่าการวางแบบแผนอย่างเคร่งครัดของปาณินิ ถือเป็นสาเหตุหนึ่งที่ทำให้ภาษาสันสกฤตต้องกลายเป็นภาษาตายอย่างรวดเร็วก่อนเวลาอันควร[ต้องการอ้างอิง] เพราะทำให้สันสกฤตกลายเป็นภาษาที่ถูกจำกัดขอบเขต (a fettered language) ด้วยกฎเกณฑ์ทางไวยากรณ์ที่เคร่งครัดและสลับซับซ้อน ภาษาสันสกฤตที่ได้ร้บการปรับปรุงแก้ไขหลักเกณฑ์ให้ดีขึ้นโดยปาณินินี้เรียกอีกชื่อหนึ่งว่า "เลากิกภาษา" หมายถึงภาษาที่ใช้กับสิ่งที่เป็นไปในทางโลก

1. **ภาษาสันสกฤตผสม**

ภาษาสันสกฤตผสม (Buddhist Hybrid Sanskrit or Mixed Sanskrit) เป็นภาษาสันสกฤตที่นักวิชาการบางกลุ่มได้จัดไว้เป็นพิเศษ เนื่องจากมีความแตกต่างจากภาษาพระเวทและภาษาสันสกฤตแบบแผน (ตันติสันสกฤต) ภาษาสันสกฤตแบบผสมนี้คือภาษาที่ใช้บันทึกวรรณคดีสันสกฤตทางพระพุทธศาสนา ทั้งในนิกาย สรรวาสติวาท และ มหายาน ภาษาสันสกฤตชนิดนี้คาดว่าเกิดขึ้นในราวพุทธศตวรรษที่ 3-4 นักปราชญ์บางท่านถือว่าเกิดขึ้นร่วมสมัยกับตันติสันสกฤต คือในปลายสมัยพระเวทและต้นของยุคตันติสันสกฤต โดยปรากฏอยู่โดยส่วนมากในวรรณกรรมของพระพุทธศาสนามหายาน อาทิ พระสูตร เช่น ลลิตวิสฺตร ลงฺกาวตารสูตฺร ปฺรชฺญาปารมิตา สทฺธรฺมปุณฺฑรีกสูตฺร และศาสตร์อันเป็นคำอธิบายหลักพุทธปรัชญาและตรรกวิทยา เช่น มธฺยมิกการิกา อภิธรฺมโกศ มหาปฺรชฺญาปารมิตาศาสฺตฺร มธฺยานฺตานุคมศาสฺตฺร เป็นต้น

ภาษาสันสกฤตไม่มีอักษรสำหรับเขียนชนิดใดชนิดหนึ่งโดยเฉพาะ และก็คล้ายกับภาษาอื่นหลายภาษา นั่นคือสามารถเขียนได้ด้วยอักษรหลายชนิด อักษรเก่าแก่ที่ใช้เขียนภาษาสันสกฤตมีหลายชนิดด้วยกัน เช่น อักษรขโรษฐี (Kharosthī) หรืออักษรคานธารี (Gāndhārī) นอกจากนี้ยังมีอักษรพราหมี (อักษรทั้งสองแบบพบได้ที่จารึกบนเสาอโศก) อักษรรัญชนา ซึ่งนิยมใช้จารึกคัมภีร์ทางพระพุทธศาสนาในอินเดียเหนือและเนปาล รวมถึง อักษรสิทธัม ซึ่งใช้บันทึกคัมภีร์พุทธศาสนารวมถึงบทสวดภาษาสันสกฤตในประเทศจีนและญี่ปุ่นโดยเฉพาะในนิกายมนตรยาน อย่างไรก็ตาม โดยทั่วไปนิยมเขียนภาษาสันสกฤตด้วยอักษรเทวนาครี (Devanāgarī) ส่วนอักษรอื่น ๆ เป็นความนิยมในแต่ละท้องถิ่น ทั้งนี้เนื่องจากอักษรที่ใช้ในอินเดีย มักจะเป็นตระกูลเดียวกัน จึงสามารถดัดแปลงและถ่ายทอด (transliteration) ระหว่างชุดอักษรได้ง่ายแม้กระทั่งในเอเชียตะวันออกเฉียงใต้ ยังมีจารึกภาษาสันสกฤตที่ใช้ อักษรปัลลวะ อักษรขอม นอกจากนี้ชาวยุโรปยังใช้อักษรโรมันเขียนภาษาสันสกฤต โดยเพิ่มเติมจุดและเครื่องหมายเล็กน้อยเท่านั้น



**2.1.3 หลักการปริวรรตอักษรภาษาสันสกฤตแบบโรมัน (IAST) เป็นไทย**

ตารางที่ การเปรียบเทียบสระ

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **ประเภท** | **ฐานที่เกิดเสียง** | **หรัสวะ**  **(สระเสียงสั้น)** | **ทีรฆะ (สระเสียงยาว)** | |
| **สมานากษะระ**  **(สระแท้)** | 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 | |
| **สันธยักษะระ (สระประสม)** | กัณฐะตาละวยะ(คอและเพดาน) |  | E e เอ | AI ai ไอ |
| กัณโฐษฐยะ(คอและริมฝีปาก) |  | O o โอ | AU au เอา |

ตารางที่ การเปรียบเทียบพยัญชนะวรรค

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

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

ตารางที่ การเปรียบเทียบพยัญชนะอวรรค

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

*\* มีในภาษาสันสกฤตแบบพระเวท*

ตารางที่ การเปรียบเทียบสัญลักษณ์พิเศษ

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

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

*\*\*ปริวรรตเป็นอักษรไทยสำหรับ จันทรพินทุ ไม่มีการกำหนดอย่างเป็นทางการโดยทั่วไปไม่แปลง คือตัดทิ้ง*

**2.2 งานวิจัยที่เกี่ยวข้อง**

**2.2.1 การพัฒนาโปรแกรมการถอดเสียงอักษรไทยเป็นอักษรโรมัน**

**ตามวิธีการของราชบัณฑิตยสถาน**

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

**2.2.2 สัททอักษรไทยปาฬิ**

วิจินตน์ ภาณุพงศ์ ​ได้​วิเคราะห์​อักขรวิธี​ใน​การ​พิมพ์​ปาฬิ​ภาสาอักษร​สยาม ใน​พระ​ไตรปิฎก จปร. ๒๔๓๖ และ​สรุป​ว่า​เป็นการ​นำเสนอ​ใน​ลักษณะ​สัททอักษร ซึ่ง​อาจ​เรียก​ว่า “สัททอักษรส​ยาม-ปาฬิ” นอกจาก​นี้ เมื่อ​ทำตา​ราง​เสียง​พยัญชนะ​ปาฬิ​แสดง​ตำแหน่ง​ฐาน​ที่​เกิด​เสียง​พร้อม​ทั้ง​ลักษณะ​การ​ออก​เสียง โดย​อาศัย​หลัก​สัท​ทนี​ติ​ของ​ตะวัน​ออก​เทียบ​กับ​หลัก​วิ​ชา​สัท​ศาสตร์​ของ​ตะวัน​ตก ทำให้​สามารถ​เลือก​ชุด​สัททอักษรสา​กล-ปาฬิ ที่​เขียน​เสียง​อ่าน​พระ​ไตร​ปิฎก​ปาฬิ​ภาสา​อักษร​โรมัน​มา​เทียบ​กับ​ปาฬิ​ภาสา-อักษร​สยาม​ได้​ด้วย จาก​หลัก​การ​ดัง​กล่าว บทความ​นี้​จึง​ได้​พัฒนา​แนว​ความ​คิด​ของ​สัททอักษรส​ยาม-ปาฬิ มา​เป็น​สัททอักษร​ไทย​ปาฬิ เพื่อ​ใช้​สำหรับ​เขียน​เสียง​อ่าน​ปาฬิ​ภาสา​ที่​เขียน​ด้วย​อักษร​ไทย​ใน​ปัจจุบัน ทั้งนี้​เพื่อ​ประโยชน์​ใน​การ​เขียน​เสียง​อ่าน​ใน​บท​สวด​มนต์​จาก​พระ​ไตร​ปิฎก​ปาฬิ​ให้​ถูก​ต้อง​หรือ​ใกล้​เคียง​ที่สุด​ตาม​วิธี​ออก​เสียง​ปาฬิ​ภาสา​ที่​ได้​สืบทอด​กัน​มากว่า​สอง​พันปี

**2.2.3 การศึกษาวิเคราะห์คีตโควินทกาวยะ**

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

**2.2.4 การถอดคำแบบถ่ายเสียงสำหรับชื่อบุคคลภาษาไทยที่เขียนด้วยอักษรโรมัน**

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

**2.2.5 Syllable-based Thai-English Machine Transliteration**

ชัย วุฒิวิวัฒน์ชัย และอัษฎางค์ แตงไทย

1. การวิจัยพัฒนานี้ มีชุดไลบรารี่ ฟังก์ชัน บางส่วนอยู่บนพื้นฐานของงานที่  
    http://learnsanskrit.org/tools /sanscript ที่อนุญาตให้ใช้ได้ตาม สัญญาอนุญาตของครีเอทีฟคอมมอนส์แบบ แสดงที่มา-ไม่ใช้เพื่อการค้า-อนุญาตแบบเดียวกัน (CC-BY-NC-SA) ดังนั้นการวิจัยพัฒนานี้ จึงต้องอยู่สัญญาอนุญาตแบบเดียวกัน