

บทที่ 2

ทฤษฎีและงานวิจัยที่เกี่ยวข้อง

เนื้อหาในส่วนนี้นำเสนอทฤษฎีและงานวิจัยที่เกี่ยวข้องกับ การปริวรรตอักษรด้วยเครื่อง สำหรับภาษาสันสกฤตแบบอักษร โรมันเป็นอักษรไทย มีดังต่อไปนี้

2.1. ทฤษฎีที่เกี่ยวข้อง

2.1.1 การปริวรรตอักษรหรือการเขียนคำทับศัพท์ (transliteration)

การปริวรรตอักษร หรือการเขียนคำทับศัพท์ [11][12] คือการดำเนินการแปลงอักษร หรืออักษรวิธีจากระบบการเขียนหรือภาษาหนึ่งไปสู่อีกระบบหนึ่งอย่างมีหลักการ เมื่อ ปริวรรตแล้วสามารถแปลงกลับเป็นอักษรหรืออักษรวิธีเดิมได้ เพื่อให้สามารถเขียนคำใน ภาษาต่างประเทศด้วยภาษาและอักษรในภาษานั้น ๆ ได้สะดวก และใกล้เคียงอักษรวิธีการ เขียนเดิมให้ได้มากที่สุด เช่น การปริวรรตอักษรโรมันภาษาอังกฤษ มาเป็นอักษรไทยเพื่อใช้ ในภาษาไทย หรือการปริวรรตอักษรภาษาไทย ไปเป็นอักษรโรมันเพื่อใช้ในภาษาอังกฤษ เป็น ปกติแล้วการปริวรรตอักษรคือการจับคู่จากระบบการเขียนหนึ่งไปยังอีกระบบหนึ่งแบบ คำต่อคำหรืออักษรต่ออักษร การปริวรรตอักษรได้พยายามที่จะสร้างความสัมพันธ์หนึ่งต่อ หนึ่งและทำให้เกิดความถูกต้องแม่นยำ เพื่อให้ผู้อ่านที่รับรู้สามารถสะกดคำต้นฉบับจากคำ ปริวรรตอักษรได้ ดังนั้นจึงมีการกำหนดหลักการปริวรรตอักษรที่ซับซ้อนในการจัดการกับ ตัวอักษรบางตัวในภาษาต้นฉบับที่ไม่สัมพันธ์กับอักษรในภาษาเป้าหมาย ความหมายอย่าง แคบของการปริวรรตอักษรคือ การปริวรรตอักษรแบบถอดอักษร (transliteration) นั้นเป็นการ คงตัวอักษรและเครื่องหมายวรรคตอนทุกอย่างเอาไว้ ทั้งนี้การถอดอักษรไม่สนใจความ แตกต่างของเสียงในภาษา เนื่องจากมีข้อจำกัดทางเทคนิค หรือการถอดอักษรโบราณเพื่อให้ ยังคงรักษารูปแบบการเขียนเดิมเอาไว้มากที่สุด การปริวรรตอักษรเป็นการถอดอักษร ต่างจาก การถอดเสียง (transcription) ซึ่งเป็นการจับคู่เสียงอ่านของภาษาหนึ่ง ๆ ไปยังรูปแบบการ เขียนของอีกภาษาที่ใกล้เคียงที่สุด ถึงแม้ว่าระบบการถอดอักษรส่วนใหญ่จะยังคงจับคู่อักษร ต้นฉบับกับอักษรในภาษาเป้าหมายที่ออกเสียงคล้ายกันในช่วงๆ ถ้าหากความสัมพันธ์ระหว่าง ตัวอักษรกับเสียงเหมือนกันทั้งสองภาษา 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 และใช้ทางศาสนา บันทึกบทสวด ปรัชญา หลักการทางศาสนา ทั้งในพหุเอกสารทั้งใน ศาสนาพราหมณ์-ฮินดู พุทธ เชน และ ซิกข์ ซึ่งในส่วน ของ ศาสนาพราหมณ์ – ฮินดูถือว่า ภาษาสันสกฤตเป็นภาษาศักดิ์สิทธิ์

คำว่า สันสกฤต แปลว่า “กลั่นกรองแล้ว” ซึ่งเป็นภาษาของชนชั้นพราหมณ์ ตรงข้ามกับภาษาพูดของชาวบ้านทั่วไปที่เรียกว่า “ปราคฤต” ภาษาสันสกฤตมีพัฒนาการในหลายยุคสมัย โดยมีหลักฐานเก่าแก่ที่สุด ภาษา ภาษาพระเวท (Vedic Sanskrit) ที่ปรากฏในคัมภีร์ฤคเวท เมื่อราว 1,200 ปี ก่อน ค.ศ. อันเป็นบทสวดสรรเสริญพระเจ้าในลัทธิพราหมณ์ในยุคต้น ภาษาพระเวทดั้งเดิมยังมิได้มีการวางกฎเกณฑ์ให้เป็นระเบียบรัดกุมและสละสลวย และมีหลักทางไวยากรณ์อย่างกว้าง ๆ ราว 57 ปี ก่อน พ.ศ. พราหมณ์ชื่อ “ปาณินิ” ชาวแคว้นคันธาระ ท่านเห็นว่าภาษาสันสกฤตแบบพระเวทนั้นมีภาษาถิ่นปนเข้ามา หากไม่เขียนไวยากรณ์ที่เป็นระเบียบแบบแผนไว้จะคละกับภาษาถิ่น ปาณินิได้ศึกษาและจัดเรียบเรียงตำราไวยากรณ์ขึ้น 8 บท ชื่อว่า “อษฎาชาสี” ภาษาที่ปรับปรุงใหม่นี้เรียกว่า “ตันดิสันสกฤต” หรือ สันสกฤตแบบแผน (Classical Sanskrit) วรรณคดีสันสกฤตแบบแผนที่สำคัญ และเป็นที่รู้จักมากคือ มหาภารตะ และ รามายณะ ภาษาสันสกฤตอีกสาขาหนึ่ง เรียกว่า ภาษาสันสกฤตผสม หรือ ภาษาสันสกฤตผสมในพุทธศาสนา (Buddhist Hybrid Sanskrit or Mixed Sanskrit) เป็นภาษาสันสกฤตยุคหลังถัดจากภาษาสันสกฤตแบบแผน พบในคัมภีร์ทางพระพุทธศาสนาทั้งในนิกาย สรวาสติวาทและพระพุทธศาสนาฝ่ายมหายาน วรรณกรรมภาษาสันสกฤตพบการใช้ที่หลากหลาย มีพัฒนาการมาอย่างต่อเนื่องสามารถ แบ่งประเภทตามเนื้อหาได้เป็น 4 ประเภทใหญ่ๆคือ

1. อาคม(āgama) มีเนื้อหาเกี่ยวกับศาสนาและปรัชญา
2. อิติहाส(itihāsa) มีเนื้อหาเกี่ยวกับประวัติศาสตร์หรือวีรชนและประเพณีที่สืบทอดกันมา
3. ศาस्ตร์(sāstra) มีเนื้อหาเกี่ยวกับศิลปวิทยาการและงานวิชาการ
4. กาวยะ(kāvya) กวีนิพนธ์หรือบทประพันธ์ที่อยู่ในรูปของศิลปะ

ภาษาสันสกฤตไม่มีอักษรสำหรับเขียนชนิดใดชนิดหนึ่งโดยเฉพาะ เป็นภาษาที่ไม่มีระบบการเขียนเป็นของตนเอง แต่จะเขียนด้วยอักษรหลายชนิด อักษรเก่าแก่ที่ใช้เขียนภาษาสันสกฤตมีหลาย

Thai

ธรรมชญะ จ กฤตชญะ จ ศคฺย วาฤโย ธฤฒ วรตตะ

International Alphabet of Sanskrit Transliteration (IAST)

dharmajñāḥ ca kṛtajñāḥ ca satya vākyaḥ dhṛḍha vrataḥ

National Library at Kolkata(kolkata)

dharmajñāḥ ca kṛtajñāḥ ca satya vākyaḥ dhṛḍha vrataḥ

Indian Languages Transliteration(ITRANS)

dharmaj~naH cha kRRitaj~naH cha satya vAkyaḥ dhRRiDha vrataH

Harvard-Kyoto(HK)

dharmajJaH ca kRtajJaH ca satya vAkyaḥ dhRDha vrataH

Sanskrit Library Phonetic Basic (slp1)

DarmajYaH ca kftajYaH ca satya vAkyaḥ DfQa vrataH

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 อื่นๆ เช่น ภาษาบาลี และ ภาษาอปรกรรมสะ เป็นต้น

แต่เดิมในการการปริวรรตอักษรภาษาสันสกฤตเป็นอักษรโรมันยังไม่ได้มีมาตรฐานกลางที่ใช้ร่วมกันแต่ใช้วิธีการปริวรรตตามข้อเสนอของผู้เชี่ยวชาญหลายๆท่าน ดังที่แสดงในรูปที่ 2.4

Devanāgarī	अ	आ	इ	ई	उ	ऊ	ऋ	ॠ	ऌ	ॡ	ऋ	ॠ
N. Halhed (1776)	ă	ā	ě	ē	ô	ō	rě	rē	lě	lē	ī	ō
M. Monier-Williams (1864)	a	ā	i	ī	u	ū	ṛi	ṛī	ḷi	e	ai	o
M. Müller (1866)	a	ā	i	ī	u	ū	ṛi	ṛī	ḷi	e	ai	o
W. D. Whitney (1879)	a	ā	i	ī	u	ū	ṛ	ṛī	ḷ	e	ai	o
Congress of Orientalists (1894)	a	ā	i	ī	u	ū	ṛ	ṛī	ḷ	e	ai	o

Devanāgarī	N. Halhed	M. Monier-Williams	M. Müller	W. D. Whitney	Orientalists
क, ख	kě, khě	ka, kha	ka, kha	ka, kha	ka, kha
ग, घ	gě, ghě	ga, gha	ga, gha	ga, gha	ga, gha
च, छ	chě, chhě	ča, čha	ch, chha	ca, cha	ca, cha
ज, झ	jě, jhě	ja, jha	ja, jha	ja, jha	ja, jha
ट, ठ	tě, thě	ta, tha	ta, tha	ta, tha	ta, tha
ड, ढ	dě, dhě	da, dha	da, dha	da, dha	da, dha
त, थ	tě, thě	ta, tha	ta, tha	ta, tha	ta, tha
द, ध	dě, dhě	da, dha	da, dha	da, dha	da, dha
प, फ	pě, phě	pa, pha	pa, pha	pa, pha	pa, pha
ब, भ	bě, bhě	ba, bha	ba, bha	ba, bha	ba, bha
ङ, ञ, ण	ṅě, gnyě, ně	n·a, ṇa, ṇa	ṇa, ṇa, ṇa	ṇa, ṇa, ṇa	ṇa, ṇa, ṇa
न, म	ně, mě	na, ma	na, ma	na, ma	na, ma
य, र, ल, व	yě, rě, lě, wě	ya, ra, la, va	ya, ra, la, va	ya, ra, la, va	ya, ra, la, va
श, ष, स, ह	shě, shě, sě, hě	śa, sha, sa, ha	śa, sha, sa, ha	śa, śa, sa, ha	śa, śa, sa, ha
ः, ॰	h, ung	h, ṁ	h, ṁ	h, ṁ	h, ṁ/ṁ

รูปที่ 2.4 การปริวรรตอักษรภาษาสันสกฤตเป็นอักษรโรมันตามข้อเสนอของบุคคล

ต่างๆก่อนมีแบบการปริวรรตอักษรภาษาสันสกฤตสากล[14]

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

จนกระทั่งในปี ค.ศ. 1894 มีการประชุมทางวิชาการเกี่ยวกับด้านตะวันออกศึกษาหรือเอเชียศึกษาในปัจจุบัน การประชุมผู้เชี่ยวชาญด้านตะวันออกศึกษานานาชาติ ครั้งที่ 10 ที่นครเจนีวา ประเทศสวิตแลนด์ (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

GENEVA, September 10th, 1894.

TRANSLITERATION OF THE SANSKRIT AND
PALI ALPHABETS.

अ	a	ऊ	ū
आ	ā	ऋ	r̄
इ	i	ॠ	r̄
ई	ī	ऌ	l̄
उ	u	ॡ	l̄
ए	e	प	p
ऐ	ai	फ	ph
ओ	o	ब	b
औ	au	भ	bh
क	k	म	m
ख	kh	य	y
ग	g	र	r
घ	gh	ल	l
ङ	ṅ	व	v
च	c	श	ś
छ	ch	ष	ṣ
ज	j	स	s
झ	jh	ह	h
ञ	ñ	ळ (in Pali ḷ)	ḷ
ट	ṭ	— (Anusvara, Niggahita) m̄	
ठ	ṭh	ॡ (Anunāsika)	m̄
ड	ḍ	ः (Visarga)	ḥ
ढ	ḍh	× (Jihvāmūliya) . . .	ḥ
ण	ṇ	× (Upadhmanīya) . . .	ḥ
त	t	ऽ (Atagraha)	ḥ
थ	th	Udatta	ˆ
द	d	Svarita	˜
ध	dh	Anudatta	˘
न	n		

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 ราชบัณฑิตยสถานยังไม่ได้วางหลักเกณฑ์การปริวรรตอักษร แต่ในบรรดาผู้ศึกษาจะยึดธรรมเนียมการปริวรรตอักษรที่สืบทอดต่อกันมา อย่างไรก็ตามหลักการยังไม่เป็นกฎเกณฑ์ที่แน่นอนตายตัวบางตำราบางอาจารย์อาจจะแตกต่างกันไป หรือมีการปรับเปลี่ยนกันเล็กน้อย ส่วนใหญ่จะปริวรรตอยู่ใน 2 รูปแบบ คือ

1. การปริวรรตอักษรแบบคงรูป หรือ แบบแผน คือการปริวรรตอักษรแบบตรงตามรูปศัพท์เดิมโดยใช้เครื่องหมายพินทุด้วย เพื่อให้ถอดกลับเป็นอักษรโรมันหรืออักษรเทวนาครีได้ถูกต้องและได้ลักษณะอักษรวิธีได้ใกล้เคียงกับคำเดิม การปริวรรตอักษรรูปแบบนี้เหมาะสำหรับใช้ในหนังสือที่ต้องการเขียนอย่างเป็นทางการหรือเอกสารทางวิชาการ

2. การปริวรรตอักษรแบบปรับรูป หรือ แบบทั่วไป คือหมายถึงการปริวรรตอักษรโดยปรับให้เข้ากับอักษรวิธีไทย เช่น โดยไม่ใช้พินทุ และเพิ่มวิสรรชนีย์หรือใส่เครื่องหมายทัณฑฆาตเพื่อให้อ่านได้ง่ายขึ้นและรูปคำกลมกลืนกับภาษาไทย จึงนิยมใช้การทับศัพท์อย่างง่ายเพื่อความสะดวกในการอ่านสำหรับบุคคลทั่วไป พบเห็นได้ตามหนังสือบทสวดมนต์ทั่วไป ลักษณะดังกล่าวทำให้การปริวรรตอักษรแบบปรับรูป นั้นใกล้เคียงกับการถอดเสียง (transcription) แต่การปริวรรตรูปแบบนี้ไม่ใช่การถอดเสียง เพราะพยัญชนะจากการปริวรรตบางตัวออกเสียงไม่ตรงมาตรา เช่น อักษร ค จากการปริวรรตออกเสียงคล้าย ก , อักษร ท ฐ จากการปริวรรตออกเสียงคล้าย ด , อักษร พ ภ จากการปริวรรตออกเสียงคล้าย บ

โรมัน(IAST)

sarvapaśyaśāraṇaṃ kuśalaśyopasaṃpadā |
svatittaparyodhanaṃ etaṃ buddhānuśāsaṇaṃ ||

ไทย-คงรูป(แบบแผน)

สรวปาปสุยากรณั กุศลสุโยปสัมปทาฯ
สวติตตปรโยทปนั เอตัม พุทธานุศาสนมฯ

ไทย-ปรับรูป(แบบทั่วไป)

สรระวาปัสยากระระณัณ กุศะลัสโยปะสัมปะทาฯ
สวะติตตะปรโยทะปะนัณ เอตัม พุทธานุศาสนัณฯ

รูปที่ 2.5 การปริวรรตอักษรภาษาสันสกฤตแบบโรมัน IAST เป็นไทยรูปแบบต่างๆ

การปริวรรตอักษรภาษาสันสกฤตแบบโรมัน IAST เป็นไทยนั้นผู้ปริวรรตต้องมีความรู้ความเข้าใจเกี่ยวกับรูปแบบอักษรในแต่ละส่วนมีตำแหน่งและหน้าที่อย่างไรจึงจะปริวรรตอักษรได้อย่างถูกต้อง สำหรับการเปรียบเทียบรูปอักษรโรมัน IAST เป็นไทย มีส่วน ประกอบดังต่อไปนี้ 4 ส่วนดังนี้

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

ตารางที่ 2.1 การเปรียบเทียบรูปสระ

ประเภท	ฐานที่เกิดเสียง (ฐานกรณ์)	หรัสวะ (สระเสียงสั้น)	ทिरมะ (สระเสียงยาว)
สระแท้ สมานาक्षระ	กัณฐะ (คอ)	A a อะ	Ā ā อา
	ตาละวะ (เพดาน)	I i อิ	Ī ī อี
	มูรันชะ (ปุ่มเหงือก)	U u อุ	Ū ū อู
	ทันดะ (ฟัน)	Ṛ ṛ ฤ	Ṝ ṝ ฤา
	โอรุชะ (ริมฝีปาก)	Ḷ ḷ ฤ	Ḹ ḹ ฤา
สระประสม (สันษยักษระ)	กัณฐะตาละวะ (คอและเพดาน)		E e เอ AI ai ไอ
	กัณ โอรุชะ (คอและริมฝีปาก)		O o โอ AU au เอา

ตารางที่ 2.2 การเปรียบเทียบรูปพยัญชนะวรรค

วรรณ	ฐานที่เกิดเสียง (ฐานกรณ์)	สปรสสะ(เสียงกัก)				อนุนาสิกะ (เสียงขึ้นจมูก)
		อโฆษะ(เสียงไม่ก้อง)			โฆษะ(เสียงก้อง)	
		ลมเบา*	ลมหนัก**	ลมเบา*	ลมหนัก**	
ก วรรณ	กัณฐะ (คอ)	K k ก	KH kh ข	G g ค	gh ฅ	Ñ ñ ง
จ วรรณ	तालव्य (เพดาน)	C c จ	CH ch ฉ	J j ช	JH jh ฌ	Ñ ñ ญ
ฎ วรรณ	मूर्च्छन् (ปุ่มเหงือก)	T t ฎ	TH th ฐ	D d ฑ	DH dh ฒ	N n ณ
ด วรรณ	तन्तय (ฟัน)	T t ด	TH th ถ	D d ท	DH dh ธ	N n น
ป วรรณ	ओष्ठ्य (ริมฝีปาก)	P p ป	PH ph ผ	B b พ	BH bh ภ	M m ม

* ลมเบา คือ อัลปพ्राणः ** ลมหนัก คือ महाप्राणः

ตารางที่ 2.3 การเปรียบเทียบรูปพยัญชนะอวรรณ

ฐานที่เกิดเสียง (ฐานกรณ์)	अन्तस्ते(เสียงเปิด)	ओष्मते(เสียงเสียดแทรก)	
	โฆษะ(ก้อง)	อโฆษะ(ไม่ก้อง)	โฆษะ(ก้อง)
	อัลปพ्राणः(ลมเบา)	มหาप्राणः(ลมหนัก)	
กัณฐะ(คอ)			H h ห
तालव्य(เพดาน)	Y y ย	Ś ś ศ	
मूर्च्छन्(ปุ่มเหงือก)	R r ร	Ṣ ṣ ษ	
तन्तय(ฟัน)	L l ล	S s ส	
ओष्ठ्य(ริมฝีปาก)	V v ว		
พยัญชนะ อวรรณอื่นๆ	L l ฬ *		

* มีในภาษาสันสกฤตแบบพระเวท และภาษาบาลี

ตารางที่ 2.4 การเปรียบเทียบรูปเครื่องหมายและสัญลักษณ์พิเศษ

ชื่อสัญลักษณ์	คำอธิบาย	สัญลักษณ์	ปริวรรตเป็นอักษรไทย
อนุสวาระ	เครื่องหมายกำกับสระใช้กำกับเสียงขึ้นจมูก	Ṁ ṁ	◌ (นิคหิต)
จันทรพินทุ/ อนุนาสิกะ	เครื่องหมายกำกับ บนเสียงอนุสวาระ ที่เน้นเสียงพิเศษคือ เน้นเสียงขึ้นจมูกกว่าปกติ	Ṃ ṃ , Ṇ ṇ*	◌̣, ◌̣̣ (อ๊, อั) มีผู้ใช้ไม้หันอากาศและมีนิคหิตข้างบน และ ใช้จันทรพินทุสำหรับอักษรโรมัน(moon dot)
วิสรรคะ	เครื่องหมายกำกับสระ ใช้เพื่อแทนเสียงลมหายใจมาก ที่สะท้อนเสียงซ้ำของสระข้างหน้า	Ḥ ḥ	ะ (วิสรรชนีย์) , ห์
อวครหะ	เครื่องหมายกำกับ ว่า สระจะมีการย่อหรือตัดออกในการสนธิในภาษาสันสกฤต	' apostrophe	' (อัญประกาศเดียว)
ทัณฑะ	เครื่องหมายจบประโยคหรือจบบาทที่สองของโคลก	 danda	๑ (อังกั่นเดี่ยว)
ทวิทัณฑะ	จบประโยคเพื่อจะขึ้นย่อหน้าใหม่หรือจบโคลก	 double danda	๑๑ (อังกั่นคู่)

* ประกาศสำนักนายกรัฐมนตรี เรื่อง หลักเกณฑ์การทับศัพท์ภาษาจีนและภาษาฮินดี ลงวันที่

22 ธันวาคม 2549 ใช้จันทรพินทุเป็น Ṁ ṁ [17]

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

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

क	क्क	क्क़	क्क्ल	क्क्व	क्क्ख	क्क्ख	क्क्
ka	kka	kkra	kkla	kkva	kkṣa	kkha	kca
कच	कण	क्त	क्त्य	क्त्र	क्त्य	क्त्	क्त्य
kcha	kṇa	kta	ktya	ktra	ktrya	ktva	ktvya
कथ	कथ्न	कथ्य	क्क	क्क्य	क्क	क्क	क्क
ktha	kthna	kthya	kna	knya	kpa	kpra	kpla
कफ	कफ्न	कफ्य	क्क	क्क्य	क्क	क्क	क्क
kpha	kma	kmya	kya	kra	krya	kla	klya

รูปที่ 2.6 ตัวอย่างลักษณะพยัญชนะสังโยค อักษรเทวนาครีเทียบกับอักษรโรมัน

การปริวรรตอักษรภาษาสันสกฤตแบบโรมัน IAST เป็นไทยทั้งสองรูปแบบ ให้ถือตามตารางเปรียบเทียบสระพยัญชนะเครื่องหมายและสัญลักษณ์พิเศษและตัวเลข ดังที่กล่าวมาเป็นหลัก ต่อไปนี้จะอธิบายถึงวิธีการปริวรรตอักษรที่ผู้ศึกษาได้พบทั้งสองรูปแบบหลัก อย่างไรก็ตามหลักการยังไม่เป็นกฎเกณฑ์ที่เป็นทางการ มีรูปแบบรายละเอียดแตกต่างกันบางประการ หรืออาจจะต้องอาศัยหลักการปริวรรตอักษรจากกลุ่มภาษาเดียวกันมีการปริวรรตที่เป็นแบบแผนและมาตรฐานอยู่แล้ว เช่น ภาษาบาลี ภาษาฮินดี และจะสามารถจำแนกวิธีการการปริวรรตอักษรภาษาสันสกฤตแบบโรมัน เป็นอักษรไทยในรูปแบบต่างๆ ได้ดังนี้

1. วิธีการปริวรรตเป็นอักษรไทยแบบคงรูป หรือ แบบแผน

ลักษณะการปริวรรตเป็นอักษรไทยแบบคงรูป หรือ แบบแผน การปริวรรตอักษรแบบตรงตามรูปศัพท์เดิมโดยใช้เครื่องหมายพินทุด้วย เพื่อให้ถอดกลับเป็นอักษรโรมันหรืออักษรเทวนาครีได้ถูกต้องและได้ลักษณะอักษรวิธีได้ใกล้เคียงกับคำเดิม นั้นเมื่อเปรียบเทียบกับการปริวรรตแบบปรับรูป นั้น จะพบเอกสารที่ใช้แบบคงรูปมากกว่าเนื่องจากถูกใช้ในวงการวิชาการ และมีหลักการปริวรรตที่เป็นแบบแผนและมีหลักการแต่อย่างไรก็ดี หลักการที่ว่าก็ไปตามแต่ละตำราอาจารย์ ยังไม่เป็นมาตรฐานเดียวกันเสียทีเดียว หลักการยังไม่เป็นกฎเกณฑ์ที่แน่นอนตายตัวบางตำราบางอาจารย์อาจจะแตกต่างไปเล็กน้อย ดังสรุปได้ต่อไปนี้

1. การถอดรูปพยัญชนะ

1.1. การถอดรูปพยัญชนะอักษรโรมันปกติ ให้ถอดตามตารางเปรียบเทียบพยัญชนะ

2. การถอดรูปสระ

2.1. ถอดรูปสระ ā ī ū ṛ ṛī (รูปสระ อา อิ อี อุ ฤ ฤ ฤ ฤ) หากตามพยัญชนะต้นให้ถอดตามปกติ

2.2. ถอดรูปสระ a (รูป สระอะ) ไม่ต้องถอดรูปหรือไม่เติม วิสรรชนีย์ (ะ) ตัวอย่างตามตารางที่ 2.6

ตารางที่ 2.6 ตัวอย่างการถอดรูปสระ a

bhagavadgītā	daśabala	śatakrata
ภควทกิตา	ทศพล	ศตกฤต

- 2.3. ถอดรูปสระ e ai o au (รูปสระ เอ ไอ โอ เอา) ที่ประสมอยู่พยัญชนะต้นอักษรโรมัน ต้องถอดรูปสระก่อนจึงค่อยถอดพยัญชนะต้น ยกเว้น รูปสระ au (รูปสระเอา) นั้นให้ใส่ไม้หน้า(่)ก่อน จึงค่อยถอดพยัญชนะต้นและปิดท้ายด้วย ลากข้าง (ำ) ตัวอย่างตาม ตารางที่ 2.7

ตารางที่ 2.7 ตัวอย่างการถอดรูปสระ e ai o au ที่ประสมอยู่พยัญชนะต้นอักษรโรมัน

teṣām	maitrī	namo	sahitau
เตษำ	ไมตรี	นโม	สहिเตา

- 2.4. หากสระอักษรโรมันขึ้นต้นคำ ให้ถอดโดยประสม อักษร อ กับสระนั้นๆเข้าไป และ a (รูปสระ อะ) ให้ใช้ อักษร อ ตัวเดียว เว้นไว้แต่สระอักษรโรมัน r ī l ī (รูปสระ ฤ ฤ ฤ ฤ) ไม่ต้องประสม อักษร อ ตัวอย่างตาม ตารางที่ 2.8

ตารางที่ 2.8 ตัวอย่างการถอดรูปสระอักษรโรมันขึ้นต้นคำ

สระเสียงสั้น		สระเสียงสั้น	
agnim	อคุณิ	ārya	อารุย
indrā	อินทุรา	īśāno	อิสานโน
usrā	อุสุรา	ūtaya	อุตย
สระประสม			
eha	เอห	aibhiragne	ไอภิริกฺเน
ojasā	โอสสา	auśijah	เอาศิชะ

3. การใช้เครื่องหมายพินทุ

- 3.1. โดยใส่ไว้ได้พยัญชนะไทยที่ถอดมาจากพยัญชนะอักษรโรมันที่ไม่มีสระอักษรโรมันกำกับหรือตามมา ตัวอย่างตาม ตารางที่ 2.9

ตารางที่ 2.9 ตัวอย่างการถอดรูปพยัญชนะอักษรโรมันที่ไม่มีสระอักษรโรมันกำกับ

sarva	dharmacakra	sattvā
สรฺว	ธรรมจักร	สตฺตฺวา

- 3.2. ในกรณีถอดรูปพยัญชนะอักษรโรมันที่ประสมด้วยสระ $r \bar{r} \bar{l} \bar{l}$ (ฤ ฤา ฤ ฤา) มีสองแนวคิด คือ ใส่พินทุไต้พยัญชนะที่ประสมด้วยสระเหล่านี้ เช่น $mr̄ga$: มฤค และไม่ใส่พินทุไต้พยัญชนะที่ประสมด้วยสระเหล่านี้ ตัวอย่างตาม ตารางที่ 2.10

ตารางที่ 2.10 ตัวอย่างการถอดรูปพยัญชนะอักษรโรมันที่ประสมด้วยสระ $r \bar{r} \bar{l} \bar{l}$

	$mr̄ga$	$kr̄ṣṇa$
แบบใส่พินทุ	มฤค	กฤษณ
แบบไม่ใส่พินทุ	มฤค	กฤษณ

4. การถอดรูปเครื่องหมายและสัญลักษณ์พิเศษ

- 4.1. รูปเครื่องหมายอนุสวาระ อักษรโรมันใช้ $m̄$ ถอดเป็นนิคหิตไว้ข้างบน(°) ลักษณะการถอดให้ถอดลำดับหลังสระอักษรโรมัน ยกเว้น รูป \bar{a} (รูปสระ อา) ให้ถอดอนุสวาระก่อน ตัวอย่างตาม ตารางที่ 2.11

ตารางที่ 2.11 ตัวอย่างการถอดรูปเครื่องหมายอนุสวาระ

$vihāram̄$	$gr̄hapatim̄$	$mahāsām̄ghikānām̄$
วิหาร	คฤหบดี	มหาสำฆิกานำ

- 4.2. รูปเครื่องหมายจันทรพินทุหรืออนุนาสิก อักษรโรมันใช้ $m̃$ มี 3 แนวคิด คือ ตัดทิ้งถอดเป็นไม้หันอากาศและมีนิคหิตข้างบน(̣) และ ใช้จันทรพินทุสำหรับอักษรโรมัน(̣ : Moon dot) ลักษณะการถอดให้ถอดลำดับหลังสระอักษรโรมัน ยกเว้น รูป \bar{a} (รูปสระ อา) ให้ถอดจันทรพินทุก่อน ตัวอย่างตาม ตารางที่ 2.12

ตารางที่ 2.12 ตัวอย่างการถอดรูปรูปเครื่องหมายจันทรพินทุหรืออนุนาสิก

	$dasyūtm̃radharām̃$	$tālam̃lokān$
แบบตัดทิ้ง	ทศยูธรา	ตาลโลกาน
แบบไม้หันอากาศมีนิคหิตข้างบน	ทศยูธรรำ	ตาลโลกาน
แบบ Moon dot	ทศยูธรรำ	ตาลโลกาน

- 4.3. รูปเครื่องหมายวิสรรคะ อักษรโรมันใช้ **h** มี 2 แนวคิด คือ ถอดเป็น วิสรรชนี (ะ) และ แบบอักษร ห ใ้สัณฑฆาต (ห) แบบนี้เป็นพระราชวินิจฉัยของพระบาทสมเด็จพระมงกุฎเกล้าเจ้าอยู่หัวที่ทรงใช้แทนวิสรรชนี (ะ) [18] ลักษณะการถอด ให้ถอดลำดับหลังสระอักษรโรมัน ตัวอย่างตาม ตารางที่ 2.13

ตารางที่ 2.13 ตัวอย่างการถอดรูปเครื่องหมายวิสรรคะ

	mahāsattvāḥ	agniḥ
แบบวิสรรชนี (ะ)	มหาสตุวะ	อคุนิะ
แบบ ห ใ้สัณฑฆาต	มหาสตุวาท	อคุนิห์

- 4.4. รูปเครื่องหมายอวครหะ อักษรโรมันใช้ ' (apostrophe) ถอดเป็น ' (ัญประกาศ เดียว) เมื่อถอดแล้วไว้ตำแหน่งเดิม บ้างก็ไม่ถอดใช้รูปเดิม บ้างก็ตัดออกแล้วเปลี่ยนเป็นเว้นวรรค
- 4.5. รูปเครื่องหมาย | (ทัณฑะ) || (ทวิทัณฑะ) ถอดเป็น ฯ (อังคั่นเดียว) ๑ (อังคั่นคู่) ตามลำดับ เมื่อถอดแล้วไว้ตำแหน่งเดิม บ้างก็ไม่ถอด ใช้รูปเดิม

5. การถอดรูปพยัญชนะสังโยค หรือ พยัญชนะสังยุคต์ ที่ประสมกับรูปสระ e ai o au (รูปสระ เอ โอ ไอ เอา) เนื่องจากการถอดพยัญชนะสังโยคจากอักษรเทวนาครีหรืออักษรที่ใช้ระบบสังโยคหรือการซ้อนของตัวพยัญชนะโดยตรง นั้นจะเห็นลักษณะการซ้อนของพยัญชนะได้ชัดเจนจะสามารถถอดรูปพยัญชนะสังโยคได้ง่าย แต่หากถอดจากอักษรโรมัน จะไม่สามารถเห็นลักษณะการซ้อนของตัวพยัญชนะ ต้องอาศัยความเชี่ยวชาญในภาษาสันสกฤต ดังนั้นการถอดรูปสระหน้าพยัญชนะสังโยค จึงมีวิธีปฏิบัติหลายแนวทางในการปริวรรตจึงแบ่งได้ลักษณะดังนี้
- 5.1. กรณีพยัญชนะสังโยคซ้อนกันมากกว่า 2 ตัว ส่วนมากเหมือนกัน ไม่คำนึงถึงพยัญชนะสังโยค ถอดรูปสระและพยัญชนะปกติทั้งหมด ตัวอย่างพยัญชนะสังโยค ttv (ตตุว) เช่น bodhisattvo ถอดเป็น โพธิสตุตุโว

กรณีพยัญชนะสังโยคซ้อนกัน 2 ตัว

- 5.2. หากทราบว่าชุดพยัญชนะอักษรโรมัน เป็นพยัญชนะสังโยคแล้ว ยกตัวอย่าง เช่น kṣ(กษ) , jñ(จญ) หากผสมกับสระอักษรโรมันดังกล่าวเมื่อถอดเป็นสระอักษรไทยแล้ว จะเขียนสระหน้าไว้พยัญชนะตัวแรกของพยัญชนะสังโยค และใส่พินทุที่พยัญชนะตัวแรก และเว้นการใส่พินทุที่ตัวสุดท้าย ตัวอย่างตาม ตารางที่ 2.14

ตารางที่ 2.14 ตัวอย่างการถอดรูปชุดพยัญชนะอักษรโรมันที่เป็นพยัญชนะสังโยค

kṣetre	saṃjñe	tumhe	prṣṭhe	draupadī
เกษเตร	สัญญ	ตุเมห	ปฤษฐ	เทราปที

- 5.3. ถอดเฉพาะชุดพยัญชนะอักษร โรมันที่เป็นพยัญชนะสัง โยคซ้อนหรือควบกับพยัญชนะโรมันเฉพาะ 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

pun̐yena	kṣetre	kleśa	aśvaiḥ	ādhyo	rātrau
ปุณฺยณ	เกษเตร	เกลศ	อไศวะ	อาโद्य	ราเตร

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

ตารางที่ 2.16 ตัวอย่างการถอดรูปพยัญชนะสังโยคที่ซ้อนเฉพาะ r

kṣetre	draupadī	rātrau	citro	traividya	idrevato
กษेत्र	เทราปที	ราเตรา	จิโตร	ไทรวิทยา	อิเทรวโต

5.5. ไม่คำนึงถึงพยัญชนะสังโยค ถอดรูปสระและพยัญชนะแบบปกติทั้งหมด ตัวอย่างตาม ตารางที่ 2.17

ตารางที่ 2.17 ตัวอย่างการถอดรูปพยัญชนะหรือสระที่ไม่คำนึงถึงพยัญชนะสังโยค

kṣetre	pr̥ṣṭhe	draupadī	traividya	tumhe	saṃjñe
กษेत्र	ปฤษฐ	เทราปที	ตฺไรวิทยา	ตุมเห	สัญญ

ตารางที่ 2.18 สรุปตัวอย่างการ ตัวอย่างการถอดรูปพยัญชนะสังโยคแบบต่างๆ

คำศัพท์อักษรโรมัน	พยัญชนะสังโยค	แบบ 5.2	แบบ 5.3	แบบ 5.4	แบบ 5.5
kṣetre	kṣ (กษ), tr (ตฺร)	กษेत्र	กษेत्र	กษेत्र	กษेत्र
saṃjñe	jñ (จฺญ)	สัญญ	สัญญ	สัญญ	สัญญ
tumhe	mh (มฺห)	ตุมเห	ตุมเห	ตุมเห	ตุมเห
pr̥ṣṭhe	ṣṭh (ษฐ)	ปฤษฐ	ปฤษฐ	ปฤษฐ	ปฤษฐ
ādhyo	ḍhy (ฒย)	อาโฒย	อาโฒย	อาฒโย	อาฒโย
pun̐yena	ny (ณฺย)	ปฺณฺยเณ	ปฺณฺยเณ	ปฺณฺยเณ	ปฺณฺยเณ
kleśa	kl (กฺล)	เกฺลศ	เกฺลศ	เกฺลศ	เกฺลศ
aśvaiḥ	śv (ศฺว)	อศฺวเ	อศฺวเ	อศฺวเ	อศฺวเ
draupadī	dr (ดฺร)	เทราปที	เทราปที	เทราปที	เทราปที
rātrau	tr (ตฺร)	ราเตรา	ราเตรา	ราเตรา	ราเตรา
citro	tr (ตฺร)	จิโตร	จิโตร	จิโตร	จิโตร
idrevato	tr (ตฺร)	อิเทรวโต	อิเทรวโต	อิเทรวโต	อิเทรวโต
traividya	tr (ตฺร)	ไทรวิทยา	ไทรวิทยา	ไทรวิทยา	ตฺไรวิทยา

2. วิธีการปริวรรตเป็นอักษรไทยแบบปรับปรุง หรือ แบบทั่วไป

วิธีการปริวรรตเป็นอักษรไทยแบบปรับปรุง หรือ แบบทั่วไป เป็นการปริวรรตอักษรโดยปรับให้เข้ากับอักษรวิไทย เช่น โดยไม่ใช้พินทุ และเพิ่มวิสรรชนีย์หรือใส่เครื่องหมายทัณฑฆาต เพื่อให้อ่านได้ง่ายขึ้นและรูปคำกลมกลืนกับภาษาไทย จึงนิยมใช้การทับศัพท์อย่างง่าย บางรูปแบบที่พบอาจมีรูปแบบการถอดเสียง (transcription) ผสมอยู่ด้วย วิธีการปริวรรตแบบนี้ยังไม่มีหลักการปริวรรตที่ชัดเจน แล้วแต่ผู้ที่จะทำการปริวรรตตามแนวทางของตน หรือถ้าหากใช้ในลักษณะทางการอาจจะต้องอาศัยหลักการปริวรรตอักษรจากกลุ่มภาษาเดียวกันที่มีแบบแผนการปริวรรตที่เป็นแบบแผนและมาตรฐานอยู่แล้ว เช่น หลักการปริวรรตอักษรไทยแบบปรับปรุงจากภาษาบาลี ยกตัวอย่างเช่น บทสวดมนต์สำหรับบุคคลทั่วไป หรือใช้หลักการเขียนคำทับศัพท์ภาษาฮินดี ตามหลักเกณฑ์การทับศัพท์ของราชบัณฑิตยสถาน[17] รูปแบบวิธีการปริวรรตเป็นอักษรไทยแบบปรับปรุงจะกล่าวแบบกว้างๆ ดังสรุปได้ต่อไปนี้

1. การถอดรูปพยัญชนะ

1.1. การถอดรูปพยัญชนะอักษรโรมันปกติ ให้ถอดตามตารางเปรียบเทียบพยัญชนะ

2. การถอดรูปสระ

2.1. ถอดรูปสระ ให้ใช้ลักษณะเดียวกันกับแบบปริวรรตเป็นอักษรไทยแบบคงรูป

2.2. ถอดรูปสระ a (รูป สระอะ) ถอดรูปโดยเดิม วิสรรชนีย์ (ะ) ตัวอย่างตาม ตารางที่ 2.6

ตารางที่ 2.19 ตัวอย่างการถอดรูปสระ a

karṇa	daśabala	śatakṛta
กะระณะ	ทะสะพะละ	สะตะกะระตะ

2.3. ถอดรูปสระ e ai o au (รูปสระ เอ โอ โอ เอา) ที่ประสมอยู่พยัญชนะต้นอักษรโรมัน ให้ใช้ลักษณะเดียวกันกับแบบปริวรรตเป็นอักษรไทยแบบคงรูป

2.4. หากสระอักษรโรมันขึ้นต้นคำ ให้ใช้ลักษณะเดียวกันกับแบบปริวรรตเป็นอักษรไทยแบบคงรูป และหากเป็นสระ a (รูปสระ อะ) ให้ใช้ อักษร อ เดิมวิสรรชนีย์ (อะ) หากตามด้วยพยัญชนะอักษรโรมันที่ไม่มีสระอักษรโรมันกำกับหรือตามมาให้เป็นตัวสะกด ให้ใช้ อักษร อ เดิมไม้หันอากาศ (อ)

ตารางที่ 2.20 ตัวอย่างการถอดรูปสระ a ขึ้นต้นคำ

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

3. ถอดรูปสระ a กับพยัญชนะตัวสะกด (การใช้ วิสรรชนี, ไม้หันอากาศ, ร หัน)

- 3.1. หากพยัญชนะอักษรโรมันประสมกับสระ a (รูปสระ อะ) ตามด้วยพยัญชนะอักษรโรมันที่ไม่มีสระอักษรโรมันกำกับหรือตามมาให้เป็นตัวสะกด ให้ถอดเป็น ไม้หันอากาศ

ตารางที่ 2.21 ตัวอย่างการถอดรูปพยัญชนะที่ประสมสระ a ตามด้วยพยัญชนะ

sarva	cakra	sattvā
สรฺว	จักระ	สัตฺตวา

- 3.2. กรณิ ร-เรพะ (รฺ) คือพยัญชนะอักษรโรมันประสมกับสระ a (รูปสระ อะ) ตามด้วยพยัญชนะอักษรโรมันตัว r ที่ไม่มีสระอักษรโรมันกำกับหรือตามมา (ar) พบ 3 วิธี ให้ใช้ คือ ยังถอดเป็น ไม้หันอากาศ กับวิธีที่โบราณจารย์ท่านนำมาภาษาสันสกฤตมาใช้ในภาษาไทย ท่านถอดเป็น ร- หัน (รร) วิธีนี้ใช้ในหลักการเขียนคำทับศัพท์ภาษาสันสกฤตของราชบัณฑิตยสถาน[17] และใช้แบบเป็น ร ตัวเดียว

ตารางที่ 2.22 ตัวอย่างการถอดรูปพยัญชนะที่ประสมสระ a ตามด้วย r

	sarva	dharma	karma	garbha
แบบ ร- หัน	สรฺวะ	ธรรมะ	กรรมะ	ครรรกะ
แบบ ไม้หันอากาศ	สัรวะ	ธัรมะ	กัรมะ	คัรรกะ
แบบเป็น ร ตัวเดียว	สรวะ	ธรมะ	กรรมะ	ครกะ
ตัวอย่างใช้ในภาษาไทย	สรรรพ	ธรรม	กรรม	ครรรภ์

- 3.3. กรณิพยัญชนะอักษรโรมันประสมกับสระ a (รูปสระ อะ) ตามด้วยพยัญชนะอักษรโรมันตัว v ที่ไม่มีสระอักษรโรมันกำกับหรือตามมา (av) เนื่องจากถอดเป็นไทยแล้วจะไปพ้องรูปสระอว พบ 2 วิธี ให้ใช้ คือ ถอดเป็น วิสรรชนี (ะ) กับ แบบตัดทิ้ง วิธีนี้ใช้ในหลักการเขียนคำทับศัพท์ภาษาสันสกฤตของราชบัณฑิตยสถาน[17]

ตารางที่ 2.23 ตัวอย่างการถอดรูปพยัญชนะที่ประสมสระ a ตามด้วย v

	bhavyā	tālavya	yamavya	gavyū
แบบ วิจารณ์ชัย	ภะวยา	ตาละวยะ	ยะมะวยะ	กะวยู
แบบ ตัดทิ้ง	ภวยา	ตาลวยะ	ยะมวยะ	กวยู

- 3.4. กรณีพยัญชนะอักษรโรมันประสมกับสระ a (รูปสระ อะ) ตามด้วยพยัญชนะอักษรโรมันตัว h ที่ไม่มีสระอักษรโรมันกำกับหรือตามมา (ah) พบ 2 วิธี คือ ยังถอดเป็นไม้หันอากาศ กับ แบบตัดทิ้ง วิธีนี้ใช้ในหลักการเขียนคำทับศัพท์ภาษาสันสกฤตของราชบัณฑิตยสถาน[17] พบการใช้ในภาษาบาลีทั้งสองแบบ

ตารางที่ 2.24 ตัวอย่างการถอดรูปพยัญชนะที่ประสมสระ a ตามด้วย h

	brahma	mahyān
แบบ ไม้หันอากาศ	พรัหมะ	มัหยาน
แบบ ตัดทิ้ง	พรหมะ	มหยาน

4. การถอดรูปเครื่องหมายและสัญลักษณ์พิเศษ

4.1. รูปเครื่องหมายอนุสวาระ(ṃ)และ เครื่องหมายจันทรพินทุหรืออนุนาสิกะ(ṅ)

อนุสวาระเป็นเครื่องหมายกำกับสระ ใช้กำกับเสียงขึ้นจมูก จันทรพินทุ/อนุนาสิกะ เครื่องหมายกำกับ บนเสียงอนุสวาระ ที่เน้นเสียงพิเศษคือ เน้นเสียงขึ้นจมูกกว่าปกติ

การถอดรูปนั้นมักจะถอดเหมือนกันทั้งสองเครื่องหมาย พบมีการใช้อักษร ง และ อักษร ม ในการถอดรูปเครื่องหมายดังกล่าว โดยเฉพาะการใช้อักษร ง หลักการเขียนคำทับศัพท์ภาษาสันสกฤตของราชบัณฑิตยสถาน ที่ให้ถอดรูปเครื่องหมายทั้งสองเป็นอักษร ง หรือหลักในภาษาบาลีการถอดรูปเครื่องหมายอนุสวาระ(ภาษาบาลีไม่มีเครื่องหมายจันทรพินทุ) กรณีใช้ปกตินั้นก็เป็นถอด อักษร ง

แต่ในภาษาสันสกฤตเครื่องหมายทั้งสองนั้นออกเสียงในกรณีปกติคล้าย ม แต่เสียงจะขึ้นนาสิกเล็กน้อย ดังนั้นจึงใช้อักษร ม

การถอดรูปแบบทั้งหมดที่กล่าวมายังมีใช้อยู่โดยทั่วไป อย่างไรก็ตามภาษาบาลีและสันสกฤตยังมีกฎไวยากรณ์ในการออกเสียงเครื่องหมายดังกล่าวมีลักษณะที่ใกล้เคียงกัน เรียกว่า การสนธิ โดยในภาษาบาลี เรียกเฉพาะเจาะจงว่า นิคหิตสนธิต่อไปจะกล่าวถึงหลักการการสนธิถอดรูปเครื่องหมายอนุสวาระ(म्)และเครื่องหมายจันทรพินทุโดยจะกล่าวถึงเฉพาะเครื่องหมายอนุสวาระ ส่วนเครื่องหมายจันทรพินทุให้ใช้หลักการเดียวกัน ในภาษาสันสกฤตมีหลักการดังนี้

- 4.1.1. กรณีใช้เครื่องหมายอนุสวาระ ไม่มีพยัญชนะโรมันอื่นตามติดมาหรือไม่มีการสนธิคำ เครื่องหมายอนุสวาระให้ถอดเป็น อักษร ม อยู่กับสระใดให้เป็นตัวสะกดสระนั้น

ตารางที่ 2.25 ตัวอย่างการถอดรูปอนุสวาระที่ไม่มีพยัญชนะโรมันอื่นตามติดมา

siddham	śaraṇam	teṣāṃ
สิทธัม	สระระณัม	เตชาม

- 4.1.2. กรณีใช้เครื่องหมายอนุสวาระและมีพยัญชนะวรรคโรมันตามติดมา คำนั้นมี การเข้าสนธิ 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	สังกะ
paṃca	c	จ	จ วรรค	ṅ	ญ	ปัญจะ
paṃḍita	ḍ	ฑ	ฎ วรรค	ṇ	ณ	ปณฺฑิตะ
saṃtoṣa	t	ต	ต วรรค	n	น	สันโตษะ
saṃbhava	bh	ภ	ป วรรค	m	ม	สัมภะวะ

4.1.3. กรณีใช้เครื่องหมายอนุสวาระและมีพยัญชนะอวรรคโรมันที่เป็นอูษมะ ได้แก่

ś ṣ ṣ h ตามติดมา ยังเป็นที่ถกเถียงกันอยู่แนวทางใหญ่ๆมี 2 แนวทาง ส่วนแนวทางอื่นๆนั้นจะใช้ 2 แนวทางนี้ปนกัน แนวทางทั้งสองมีรายละเอียดดังนี้

4.1.3.1.แนวทางแรก หากมีพยัญชนะอวรรคโรมันที่เป็นแบบอูษมะ คำนั้นมีการ

เข้าสนธิ เครื่องหมายอนุสวาระให้ถอดเป็นพยัญชนะนาสิก(อนุนาสิกะ)

ในฐานที่เกิดเสียงของตัวที่ตามมา อยู่กับสระใดให้เป็นตัวสะกดสระนั้น

ตารางที่ 2.27 ตัวอย่างการถอดรูปอนุสวาระแบบสนธิมีพยัญชนะตามแบบอูษมะ

คำศัพท์	พยัญชนะตาม		ฐานกรณ์	พยัญชนะนาสิก		ถอดเป็น
	โรมัน	ไทย		โรมัน	ไทย	
siṃha	h	ห	กัณฐะ	ภ	ง	สิงหะ
saṃśaya	ś	ศ	तालव्य	ภ	ญ	สัญชะยะ
saṃskā	ṣ	ษ	มูรฉันยะ	ภ	ณ	สันษกา
Samskṛta	s	ส	ทันตยะ	n	น	สันสกฤตะ

4.1.3.2.แนวทางที่สอง หากมีพยัญชนะอวรรคโรมันที่เป็นแบบอูษมะ ไม่มีการ

สนธิคำ เครื่องหมายอนุสวาระให้ถอดเป็นอักษร ม อยู่กับสระใดให้เป็น

ตัวสะกดสระนั้น

ตารางที่ 2.28 ตัวอย่างการถอดรูปอนุสวาระแบบไม่สนธิมีพยัญชนะตามแบบอูษมะ

คำศัพท์	พยัญชนะตาม		ถอดเป็น
	โรมัน	ไทย	
siṃha	h	ห	สิมหะ
saṃśaya	ś	ศ	สัมชะยะ
saṃskā	ṣ	ษ	สัมษกา
Samskṛta	s	ส	สัมสกฤตะ

4.1.4. กรณีใช้เครื่องหมายอนุสวาระและมีพยัญชนะอวรรคโรมันที่เป็นแบบอันต์สละ ได้แก่ *y r l v* ตามติดมา ไม่มีการสนธิคำ เครื่องหมายอนุสวาระให้ถอดเป็น อักษร ม อยู่กับสระใดให้เป็นตัวสะกดสระนั้น

ตารางที่ 2.29 ตัวอย่างการถอดรูปอนุสวาระและมีพยัญชนะตามแบบอันต์สละ

คำศัพท์	พยัญชนะตาม		ถอดเป็น
	โรมัน	ไทย	
saṃyoga	y	ย	สั้มโยคะ
saṃrakṣa	r	ร	สั้มรักษะ
samlāpa	l	ล	สั้มลาปะ
saṃvatsara	v	ว	สั้มวัตสระระ

4.2. รูปเครื่องหมายวิสรรคะ ยังถกเถียงกันอยู่มาก ใช้หลากหลายกับตัวอักษร ห , ฮ เช่น ใช้พินทุ หุ หู ใช้ทัณฑฆาต ห์ ฮ์ ใช้วิสรรชนีย์ หะ ฮะ แต่แนวคิดหลังไม่ถูกต้องเท่าไรนัก เพราะเครื่องหมายวิสรรคะ ออกเสียงลมหายใจมาก ที่สะท้อนเสียงซ้ำของสระข้างหน้า แต่สรุปวิธีการที่ยอมรับกัน ได้มี 2 แนวคิด คือ แบบใช้อักษร ห ใสทัณฑฆาต (หฺ) วิธีนี้เป็นพระราชวินิจฉัยของพระบาทสมเด็จพระมงกุฎเกล้าเจ้าอยู่หัว[19] และใช้ในหลักการเขียนคำทับศัพท์ภาษาสันสกฤตของราชบัณฑิตยสถาน [17] อีกแบบคือใช้อักษร ฮ หรือ ห แล้วถอดเสียง (transcription) เสียงซ้ำที่สะท้อนของสระข้างหน้า ตารางที่ 2.13

ตารางที่ 2.30 ตัวอย่างการถอดรูปเครื่องหมายวิสรรคะ

	แบบ ห ใสทัณฑฆาต	แบบถอดเสียง(ฮ)	แบบถอดเสียง(ห)
duḥka	ทุหฺชะ	ทุสุชะ	ทุหุชะ
rāmaiḥ	ราไมหฺ	ราไมฮิ	ราไมหิ
agniḥ	อคนิหฺ	อคนิฮิ	อคนิหิ
hetuḥ	เหโตหฺ	เหโตโฮ	เหโตโท
namaḥ	นะมะหฺ	นะมะฮะ	นะมะหะ

- 4.3. รูปเครื่องหมายอวครหะ โดยมากตัดทิ้ง
- 4.4. รูปเครื่องหมาย ı (ทัณฑะ) ıı (ทวิทัณฑะ) ให้ใช้ลักษณะเดียวกันกับแบบปวิรรต เป็นอักษรไทยแบบคงรูป
5. การถอดรูปพยัญชนะสังโยค หรือ พยัญชนะสังยุคต์ ที่ประสมกับรูปสระ e ai o au (รูปสระ เอ ไอ โอ เอา) ในแบบปรับรูปนั้นก็ยังไม่มืแบบแผนในการกำหนด แล้วแต่ผู้ปวิรรตเห็นว่าสมควร รูปแบบที่พบบ่อยเช่น หากมีการในภาษาไทยอยู่แล้วก็ใช้ลักษณะนั้น

ตารางที่ 2.31 ตัวอย่างการถอดรูปพยัญชนะสังโยคแบบคำที่ใช้ในภาษาไทย

śreṣṭha	śloka	kṣetra
เศรษฐะ	โคลก	เกษตระ

หรือพยัญชนะสังโยคตัวนั้นมีลักษณะซ้อนคล้ายคำควบกล้ำในภาษาไทย(ซ้อนกับ อักษร ร ล ว) ให้ใช้แบบคำควบกล้ำ

ตารางที่ 2.32 ตัวอย่างการถอดรูปพยัญชนะสังโยคแบบคำควบกล้ำไทย

traividya	draupadī	kleśa
ไตรวิทยา	เทราปะที	เกลสะ

2.2. ทฤษฎีการปริวรรตอักษรด้วยเครื่อง

2.2.1. การปริวรรตอักษรด้วยเครื่อง

แนวทางการปริวรรตอักษร(Machine Transliteration Approaches) สามารถแบ่งตามลักษณะการทำงานของระบบออกเป็น 2 กลุ่ม คือ

1. การปริวรรตอักษรด้วยเครื่องแบบใช้ฐานกฎ (Rule-based Approach)

เป็นปริวรรตอักษรด้วยเครื่องโดยใช้ความรู้ทางด้านกฎไวยากรณ์มาใช้ในการกำหนดกฎเกณฑ์ของระบบ ซึ่งจะต้องมีการแยกคุณลักษณะและข้อมูลทางกฎไวยากรณ์ของอักษรต้นฉบับจากนั้นจึงทำการวิเคราะห์ตามกฎไวยากรณ์ของอักษรเป้าหมาย ข้อจำกัดของการแปลภาษาด้วยเครื่องกลุ่มนี้คือ ผู้พัฒนาต้องมีความรู้ทางกฎไวยากรณ์ของทั้งสองตัวอักษรเป็นอย่างมาก

2. การปริวรรตอักษรด้วยเครื่องแบบใช้สถิติ (Statistical Approach)

เป็นการปริวรรตอักษรด้วยเครื่องโดยใช้วิธีการทางสถิติมาช่วยในการปริวรรต ซึ่งการปริวรรตกลุ่มนี้จำเป็นต้องมีฐานข้อมูลคู่ภาษาที่มีการจับคู่ประโยคในอักษรต้นฉบับและตัวอักษรเป้าหมาย เพื่อเป็นฐานความรู้ให้ระบบทำการเรียนรู้ และใช้ค่าทางสถิติ เอ็นแกรม (N-Gram Model) ซึ่งเป็นการคำนวณค่าของการที่คำเกิดขึ้นร่วมกัน ถ้าชุดคำชุดใดมีค่าเอ็นแกรมหรือค่าความน่าจะเป็น (Probability) สูง แสดงว่าชุดคำนั้นมีโอกาสเกิดขึ้นร่วมกันบ่อยครั้ง การคำนวณหาค่าเอ็นแกรมของชุดคำที่มีอยู่ในฐานข้อมูลคู่อักษรทำให้ได้ค่าความน่าจะเป็นของชุดคำต่างๆ เพื่อนำไปเปรียบเทียบและใช้ในการแปลข้อความได้ โดยสามารถเลือกใช้ค่าเอ็นแกรมได้ตั้งแต่ 2 คำ (bigrams), 3 คำ (trigrams) เป็นต้น ยิ่งใช้ค่าเอ็นแกรมมาก ยิ่งทำให้ภาษาต้นฉบับมีการได้หลากหลายและมีความละเอียดมากขึ้นข้อดีของวิธีการนี้คือ ผู้พัฒนาไม่จำเป็นต้องมีความรู้ทางด้านกฎไวยากรณ์ของทั้งสองตัวอักษรจึงทำให้ไม่เกิดปัญหาเรื่องไวยากรณ์ต่างๆ แต่ปริวรรตอักษรด้วย

วิธีการนี้จำเป็นต้องมีฐานข้อมูลคู่อักษรที่มีจำนวนข้อมูลมหาศาลเพื่อหาค่าทางสถิติที่จะนำไปใช้เปรียบเทียบกับประโยคที่จะนำมาปริวรรตได้อย่างครอบคลุม

2.2.2. ประเภทของโมเดลปริวรรตอักษรด้วยเครื่อง

ประเภทของโมเดลปริวรรตอักษรด้วยเครื่อง (Machine Transliteration Models) สามารถแบ่งตามลักษณะได้ 4 ประเภท คือ

1. โมเดลแบบใช้ตัวอักษร (Grapheme Based Model)

คือโมเดลการปริวรรตโดยจับคู่ตัวอักษรต้นฉบับกับตัวอักษรเป้าหมาย โดยไม่สนใจเกี่ยวกับการออกเสียงของคำที่ได้จากการปริวรรต มักใช้ในตัวอักษรในตระกูลเดียวกัน มีสระพยัญชนะใกล้เคียงกันหรือการปริวรรตที่ต้องการรักษารูปศัพท์เดิมไว้โดยไม่สนใจเสียงของตัวอักษรเป้าหมาย

2. โมเดลแบบใช้เสียงตัวอักษร (Phoneme Based Models)

คือโมเดลการปริวรรตโดยจับคู่เสียงของตัวอักษรต้นฉบับกับเสียงของตัวอักษรเป้าหมาย มักใช้ในกรณีที่ตัวอักษรต่างตระกูลกัน หรือมีสระพยัญชนะไม่เท่ากันอยู่มาก เป็นการปริวรรตที่ต้องการรักษาเสียงเดิมไว้ให้ใกล้เคียงที่สุดโดยไม่สนใจรูปศัพท์เดิมของตัวอักษรตัวอักษรต้นฉบับ

3. โมเดลแบบผสม (Hybrid Model)

คือโมเดลการปริวรรตที่ใช้โมเดลแบบใช้ตัวอักษรและโมเดลแบบใช้เสียงตัวอักษรรวมกันมักใช้ในกรณีที่ตัวอักษรต่างตระกูลกันแต่ต้องการรักษารูปศัพท์เดิมไว้ ถ้าส่วนใดมีสระพยัญชนะไม่ครบคู่ก็ใช้สระพยัญชนะที่เสียงใกล้เคียง หรือผสมให้เกิดเสียงคล้ายกันแทน

4. โมเดลแบบสมนัย (Correspondence Based Model)

คือโมเดลการปริวรรตที่คล้ายกับ โมเดลแบบผสม แต่ใช้ตัวอักษรต้นฉบับกับเสียงตัวอักษรต้นฉบับ ในการจับคู่เสียงสระพยัญชนะในอักษรเป้าหมาย ในกรณีที่ตัวอักษรเป้าหมายมีลักษณะเป็นคำหรือพยางค์ หรืออยู่ในลักษณะอักษรหนึ่งตัวประกอบ

ไปด้วยสระพยัญชนะ เช่นการทับศัพท์หรือการปริวรรตอักษร อักษรโรมัน ภาษาอังกฤษ เป็น อักษรจีน ญี่ปุ่น เกาหลี เป็นต้น

2.3. งานวิจัยที่เกี่ยวข้อง

ศราวุธ ปลัดทอง [20] ได้เสนอการพัฒนาโปรแกรมการถอดเสียงอักษรไทยเป็นอักษรโรมันตามวิธีการของราชบัณฑิตยสถาน โดยนำคำ วลี หรือประโยค มาแยกคำตามพจนานุกรมเพื่อแบ่งพยางค์ โดยใช้วิธีการจัดกลุ่มตัวอักษร และนำกลุ่มอักษรไปแยกคำโดยวิธีเปรียบเทียบคำกับพจนานุกรมแบบฟอร์เวิร์ดและแบ็กเวิร์ด จะได้ผลลัพธ์เป็นหน่วยเสียงระดับพยางค์ แล้วจึงนำแต่ละพยางค์ไปถอดเสียงเป็นอักษรโรมันตามวิธีการของราชบัณฑิตยสถาน และเพิ่มสัญลักษณ์แทนระดับเสียงเพื่อใช้แทนการออกเสียงวรรณยุกต์เพื่อให้การออกเสียงมีความถูกต้องมากขึ้น ผลการทดลองพบว่ามีความถูกต้อง 99.95 %จากคำในพจนานุกรมจำนวน 21,399 คำ

ชุลีกร กิตติกุล [21] ได้เสนอการถอดคำแบบถ่ายเสียงสำหรับชื่อบุคคลภาษาไทยที่เขียนด้วยอักษรโรมัน การถอดคำแบบถ่ายเสียงสำหรับแต่ละคำสามารถสร้างได้จากกฎ หรือใช้แบบจำลอง ทางสถิติ หรือค้นจากพจนานุกรม อย่างไรก็ตามการขาดมาตรฐานและความหลากหลายของการแปลงชื่อบุคคลไทยให้เป็นชื่อ ที่เขียนด้วยอักษรโรมันเป็นงานที่ท้าทาย และแม้ว่าวิธีที่ใช้พจนานุกรมเหมือนจะให้ผลที่ค่อนข้างถูกต้องที่สุด แต่ส่วนของการแปลงตัวอักษรเป็นเสียงก็ยังคงมีความจำเป็นสำหรับคำที่ไม่พบในพจนานุกรม งานวิจัยนี้เสนอวิธีการถอดคำแบบถ่ายเสียงสำหรับชื่อบุคคลภาษาไทยที่เขียนด้วยอักษรโรมันให้เป็นเสียงภาษาไทย โดยคำนึงถึงความนิยมในการใช้งาน ชื่อบุคคลภาษาไทยที่เขียนด้วยอักษรโรมันจะถูกแบ่งให้เป็นสายลำดับของแกรมโดยใช้พจนานุกรมแกรมสะสมซึ่งถูกสร้างจากชื่อมากกว่า 130,000 ชื่อ ผลการศึกษาพบว่าวิธีนี้ให้ความถูกต้องของ

วินัย มาลีลัยและพฤษดี ศิริแสงตระกูล [22] ได้เสนอการแปลงภาษาบาลีอักษรไทยเป็นสัทอักษร โดยใช้พจนานุกรม (dictionary) ร่วมกับกฎบาลีไวยากรณ์ (Pali grammar) และกฎภาษาศาสตร์ (linguistic) และได้ใช้วิธีการตัดคำแบบเลือกคำที่ยาวที่สุด(longest matching algorithm) เพื่อใช้ในการค้นหาความหมายของคำศัพท์ ในงานวิจัยนี้ผู้วิจัยได้สร้างพจนานุกรมภาษาบาลี-ไทย ประกอบด้วยคำศัพท์บาลี 14,323 คำ โดยอ้างอิงคำศัพท์บาลีจากหนังสือ

พจนานุกรมภาษาบาลี-ไทยของพระมหาไพโรจน์ พจนานุกรมที่สร้างขึ้นประกอบด้วยคำศัพท์ภาษาบาลีอักษรไทย ความหมายประเภทของคำศัพท์ คำในรูปภาษาไทย คำในรูปอักษรโรมันและคำในรูปสัทอักษร โดยจัดเก็บไว้ในระบบฐานข้อมูลและในงานวิจัยนี้ได้นำเสนอกฎในการแปลงภาษาบาลีเป็นสัทอักษรเพื่อใช้ในการแปลงคำที่ไม่ปรากฏในพจนานุกรมเพื่อทดสอบความถูกต้องของระบบ ผู้วิจัยได้ทำการทดสอบระบบที่พัฒนาโดยการแปลงเอกสารที่เขียนด้วยภาษาบาลีอักษรไทย ซึ่งประกอบด้วย บทสวดมนต์ 99 บท พุทธศาสนสุภาษิต 181 สุภาษิต และวินย. มหาวิงุโค 1 จากพระวินยปิฎก เล่มที่ 1 ของพระไตรปิฎกภาษาบาลีฉบับสยามรัฐ ผลการทดสอบความถูกต้องของการแปลงคิดเป็นร้อยละ 98.75

อัจฉราภรณ์ ธาตุวิสัย [23] ได้เสนองานการศึกษาวิเคราะห์คิตโควินทกาวะ โดยมีวัตถุประสงค์ คือ 1. เพื่อปริวรรตอักษรและแปลคิตโควินทกาวะเป็นภาษาไทย 2. เพื่อศึกษาวิเคราะห์เนื้อหาของคิตโควินทกาวะ และ 3. เพื่อศึกษาวิเคราะห์วรรณศิลป์ของคิตโควินทกาวะ ต้นฉบับที่ใช้ในการศึกษาค้นคว้าครั้งนี้คือต้นฉบับภาษาสันสกฤต ตัวอักษรเทวนาครีจากหนังสือคิตโควินทกาวะของชยเทพ ตีพิมพ์โดย สัตสันมะ เสวะ สะมิถิ ไม่ระบุปี ที่พิมพ์ การวิจัยแบ่งเป็น 2 ส่วน คือ การวิเคราะห์เนื้อหาใน 3 หัวข้อ ได้แก่ การร้อยเรียงเรื่อง องค์ประกอบของเรื่องและแนวคิดของเรื่องและการวิเคราะห์วรรณศิลป์ โดยใช้ทฤษฎีอรรถศาสตร์ 2 ทฤษฎี ได้แก่ทฤษฎีอรรถการ และทฤษฎีรส โดยมีขั้นตอนการศึกษา คือ ค้นคว้าและรวบรวมข้อมูลต่างๆ ที่เกี่ยวข้องปริวรรตคิตโควินทกาวะจากอักษรเทวนาครีเป็นอักษรไทย และแปลจากภาษาสันสกฤตเป็นภาษาไทย จากนั้นจึงศึกษาวิเคราะห์ในด้านเนื้อหาและด้านวรรณศิลป์ แล้วอภิปรายและสรุปผลการศึกษา