



Romanization, Transliteration, and Transcription for the Globalization of the Thai Language¹

Nitaya Kanchanawan*

Associate Fellow of the
Academy of Arts, The Royal
Institute, Thailand

Abstract

This paper differentiates between the similar terms of transcription, transliteration, Romanization and Anglicization. The six systems which are employed to represent Thai words through the Roman script are examined. These systems are King Rama VI system, the first Royal Institute systems (both General and Precise Systems), the second Royal Institute system, the third Royal Institute system, Anglicization system and the ISO system. Whereas each system serves a different purpose, the ISO system can be used globally as a means to exchange information both by humans and machines.

Key words : transcription, transliteration, Romanization, Anglicization, Rama VI system, Royal Institute system, ISO system

1. Introduction

Two or more different languages may use similar or different sets of symbols for writing purposes. The same symbol set may be modified to suit a particular language in question. The crucial point is that each language contains its own pronunciation. Thus, by default, the same linguistic symbol set may not always represent the same sounds for two different

languages. This can be seen in the case of the Roman script used in transcribing English, French, German, etc. or in the case of the Thai and Cambodian scripts which came from the same origin. Thus, using different sets of written symbols for the same language may cause some confusion in pronunciation or other ambiguities.

According to the International

Standard Organization (ISO), however, the need for harmonized standards in the world of information is global. The wide spread use of the Internet has strengthened the need for interoperable standards and compatible information systems. Thus, characters from different individual writing systems must be converted into a single universal alphabet to allow for alphabetical intercalation in

* Associate Professor, Faculty of Humanities, Ramkhamhaeng University

¹ Paper presented in Round Table "Romanization, Transliteration, and Transcription for Globalization." 1st World Congress on the Power of Language: Theory, Practice, and Development. In Honor of Her Royal Highness Princess Maha Chakri Sirindhorn's 50th Birthday Anniversary in the Year of Languages. 22-25 May 2006. Queen Sirikit Center, Bangkok, Thailand



bibliographies, catalogues, indexes, etc. It is essential to create an environment for the univocal transmission of written messages between two countries using completely different writing systems.

2. Conversion Methods for Writing Systems

There are two basic conversion methods for writing systems and they are transliteration and transcription.

2.1 Transliteration is the process which consists of representing the characters of an alphabetical or syllabic system of writing by using the characters of a conversion alphabet. The use of a conversion alphabet is the easiest way to ensure the complete and unambiguous reversibility of the conversion alphabet into the converted system.

In exceptional cases, e.g. when the number of characters used in the conversion system is smaller than the number of characters of the original writing system, it is often necessary to use digraphs or diacritical marks to close the gap. In this case, one must avoid as far

as possible arbitrary choice and the use of purely conventional marks, and try to maintain a certain phonetic logic in order to give the system a wide acceptance.

However, it must be accepted that the graphism employed may not always produce correct pronunciation due to the individual phonetic habits of the language (or languages) which usually use the conversion alphabet. On the other hand, this graphism must be such that the reader who has a knowledge of the original language may mentally restore unequivocally the original graphism and thus pronounce it correctly.

2.2 Transcription is the process whereby the sounds of a given language are represented by a system of signs contained in a conversion language.

A transcription system is of necessity based on the orthographical conventions of the conversion language. Transcription is not strictly reversible.

Various transcription systems may be employed for the conversion of all writing systems. It is the only method that can be used

for systems that are not entirely alphabetical or syllabic, and for all ideophonographical systems of writing such as Chinese.

2.3 Romanization is the process whereby a non-Latin writing systems is converted to the Latin alphabet. In doing so either transliteration or transcription or a combination of the two systems may be used depending on the nature of the converted system and the desired objectives.

2.4 Regarding the use of the Latin alphabet for the representation of the Thai language, there are four different systems used to transcribe: King Rama VI system (in which the Pali and Sanskrit written forms are retained), transcription system, and the Romanization and Anglicization systems.

2.4.1 King Rama VI System (1913)

With the introduction of the family name system in 1913 by King Rama VI, a conversion system from the Thai script to the Latin alphabet was developed whereby the written forms of Pali and Sanskrit were retained.

Thai Consonant	Pali and Sanskrit	Thai Pronunciation
ก	K	K
ข	Kh	Kh
ค	G	Q
ฃ	Gh	Gh
ง	N (or ng)	ng
จ	Ch	Ch





Thai Consonant	Pali and Sanskrit	Thai Pronunciation
ฃ	Chh	Chh
ช	J	X
ฌ	-	S
ฉ	Jh	-
ญ	ñ	N (or ng)
ฎ	-	D
ฏ	T	T
ฐ	Th	Th
ถ	D	D
ฒ	Dh	Dh
ณ	N	N
ด	-	D
ต	T	T
ถ	Th	Th
ท	D	Th
ธ	Dh	Th
น	N	N
บ	-	B
ป	P	P
ผ	Ph	Ph
ฝ	-	F
พ	B	Ph
ฟ	-	F
ภ	Bh	Bh
ม	M	M
ย	Y	Y
ร	R	R
ล	L	L
ว	V or W	V or W
ศ	S	-
ษ	Sh	-
ส	S	S
ห	H	H
ฬ	L	-



Thai Consonant	Pali and Sanskrit	Thai Pronunciation
ฮ	-	H
ฤ ฦ	Ri, Rî	Ri, Rî
ฦ ฦ	Li, Lî	Li, Lî

(Amondarunarak, 2511: 36-37)

Thai Vowel	Roman Script	Thai Vowel	Roman Script
ะ	a	อิ	iu
า	â	อุ	uy
ิ	i	เอ	eo
ึ	î	แ	aeo
ู	u	เ	oe
ุ	û	เ	oey
เ	e	อ	oy
โ	ô	ไ	ôy
ึ	ü	อ	ua
ึ	ûe	อ	uay
แ	ae	เ	ia
ไ	âi	เ	îa
อ	ai or ay	เ	iia
เ	au or oa	เ	iiay
อ	âo	เ	iâu

(Amondarunarak, 2511: 36-38)

Words written in this System may not be pronounced correctly. For example the new airport สุวรรณภูมิ “Suvarnabhumi” is pronounced differently from the Thai sound. The meaning, however, is retained.

2.4.2 The First Royal Institute Systems (1939)

In 1939 The Royal Institute introduced two systems for transcribing Thai characters into

Roman: The General System and the Precise System. The terms “transliteration” and “transcription” are used interchangeably since the Thai official Romanization method included elements of both these concepts.

The Precise System indicates tones and unpronounced letters thus showing the derivation of the words. In the Precise System there is a Roman equivalent for every

Thai consonant and this system employs diacritical marks, many of them being tonal marks. In the General System the sound of words is indicated by Roman letters with limited diacritical markings.

The diacritical marks, however, present obstructions for practical use, and thus, the two systems have not been popular with general audiences and language learners. Surprisingly, the





concept of the Precise System is almost identical with that of the transliteration system.

2.4.3 The Second Royal Institute System (1968)

The Royal Thai Government has adopted the General System for use in transcribing geographical names within Thailand.

The General System was chosen for Romanizing geographical names on maps and in other places because it is more practical cartographically and typographically. In order to avoid the confusion between diacritical marks of geographic names with other map symbols, the system was modified

in 1968. All diacritical markings are omitted in accordance with the recommendation of the United Nations Committee of Experts on Geographical Names for the world standardization of geographic names.

VOWELS

Transliteration	Thai Vowels ²	Illustrations
a	-ะ -ั -า -	(สะพาน SAPHAN) (ลันตา LANTA) (บาง BANG) (สมุทร SAMUT)
am	-ำ	(ลำ LAM)
i	-ิ -ี	(สิงห์ SING) (บุรี BURI)
u	-ุ -ู -อ	(สตึก SATUK) (พีช PHUT) (บรบัว BORABU)
e	-ะ -ั -ุ -ุ -	(บุรี BURI), (ภู PHU) (ปะนาเระ PANARE) (เพ็ญ PHEN) (เขน KHEN)
ae	แ-ะ แ-	(แซะ SAE), (สะแก SAKE)
o	โ-ะ โ- -าะ -อ	(พะโต๊ะ PHATO) (ลพ LOP) (สามโก้ SAMKO) (เกาะ KO), (บ่อ BO)
oe	เ-อะ เ-อ เ-ิ	(เซอะ SOE) (อำเภอ AMPHOE), (เนิน NOEN)

² The position of the consonant bearing the vowel is indicated by the hyphen; the position of the associated vowel—following, preceding, above, below—is also shown.



Transliteration	Thai Vowels	Illustrations
ia	เียะ เีย	(เียะ PHIA) (เทียน THIAN)
ua	เือะ เือ	(เือะ KUA) (เมือง MUANG) (ผัวะ PHUA) (บัว BUA) (ควน KHUAN)
ai	ไ- ไ- -ัย ไ-ย -าย	(ใหญ่ YAI) (ไผ่ PHAI), (ชัย CHAI) (ไทย THAI), (พาย PAI)
ao	เา -าว	(เจ้า CHAO) (ข้าว KHAO)
ui	-ุย	(กุย KUI)
oi	โ-ย -อย	(โดย DOI), (ดอย DOI)
iu	-ิว	(จิว NGIU)
eo	เ็ว เ-ว	(เร็ว REO), (เลว LEO)
oei	เ-ย	(เลย LOEI)
uai	เือย -วย	(เดือย DUAI) (ห้วย HUAI)
aeo	แ-ว	(แมว MAEO)
ieo	เียว	(เขี้ยว KHIEO)

CONSONANTS

Initial	Final	Thai consonants
K-	-K	ก
KH-	-K	ข ค ฃ
NG-	-NG	ง
CH-	-T	จ
CH-	-T	ฉ ช ฅ
Y-	-N	ญ
D-	-T	ด ฎ ๓ ³
T-	-T	ต ฏ
TH-	-T	ถ ฐ ๓ ๓ ๓ ๓

³ The character ๓ may be pronounced TH or D when it is in initial position.





Initial	Final	Thai consonants
N-	-N	น ฌ
B-	-P	บ
P-	-P	ป
PH-	-P	ผ พ ภ
F-	-P	ฝ ฟ
M-	-M	ม
Y-	-	ย
R-	-N	ร
L-	-N	ล พ
W-	-	ว
S-	-T	ซ ฌ ฌ ฌ ฌ ฌ
H-	-	ห ฮ
-	-	อ

The General System of transcription provides only an approximate pronunciation for non-Thai speakers. The exact pronunciation can be acquired only from the Thai characters themselves, or with the use of a dictionary. Thus, the omission of diacritic marks may cause ambiguous or mis-pronunciation, e.g. the word “south” (ใต้ - “tai”) and “to die” (ตาย - “tai”) cannot be distinguished.

2.4.4 The Current Royal Institute Romanization System (1999)

The Royal Institute introduced the currently used transliteration system in 1999. This was an adaptation of the 1968 system modified in order to eliminate some ambiguity. Due to the influence of English pronunciation, “จ” and “ฌ ฌ ฌ” cannot be distinguished. Phonetically “จ” should be transcribed as “c” and

“ฌ ฌ ฌ” should be transcribed as “ch”. Most of the Thais, however, tend to pronounce “c” as /k/like in the familiar English words, e.g., cat, code, cut. The pronunciation of “c” in the words like “cell” and “city” are surprisingly ignored. Thus, the consonants are left untouched.

Regarding the vowels, some ambiguous issues are finally resolved such as introduced in the following table:

VOWEL	OLD SYSTEM	NEW SYSTEM
อี อี	u	ue
อุ อู	u	u
เอือะ เอือ	ua	uea
อัวะ อิว	ua	ua
เอือย	uai	ueai
อวย	uai	uai

Another improvement is that all the diphthongs with final “ว” are replaced by “o” such as in the following table:



VOWEL	OLD SYSTEM	NEW SYSTEM
เอา อาว	ao	ao
เอี้ยว เอว	eo	eo
แอี้ยว แอว	aeo	aeo
อิ้ว	iu	io
เอี้ยว	ieu	iao

Notice that the written form of เอี้ยว is changed from “ieu” to “iao” because เอี้ยว = เอี้ยว + ว. Since เอี้ยว = “ia” so เอี้ยว + ว should be “iao” .

2.4.5 Anglicization

Anglicization is an attempt to replace Thai words by English words with closely similar pronunciation regardless of the meaning. For example:

THAI WORD	MEANING	ENGLISH WORD WITH SIMILAR PRONUNCIATION
วันเพ็ญ [wan-phen]	(a girl name)	one pen
สมร [sa-mǎn]	(a girl name)	small
สามารถ [sǎa-mâat]	capable	smart
สิ้นกรรม [sîn-kam]	end of karma	sin gum
ปั้นขลิบ [pân-khlìp]	(name of Thai snack)	pan clip
ฟ้า [fǎa]	sky	far
กัน [kan]	together	gun

If such words cannot be found, new words imitating the English form will be used. For example:

THAI WORD	MEANING	NEW WORD IMITATING ENGLISH WORD
น้ำ [náam]	water	num (from “number”)
เจ้า [cǎo]	you (literary term)	jow (from “how”)
เขา [khǎo]	he	kow (from “how”)
กลับ [klàp]	return	glub (from “club”)
ร้อย [rǎoi]	hundred	roy (from “joy”)
ลอย [lǎoi]	float	loy (from “joy”)



Currently this is the most popular system since it is used in Karaoke singing. The average person is interested in imitating the closest sound, without regard to either rule or meaning. This form of Anglicization may be a unique case to Thailand only.

2.4.6 ISO 11940 Information and Documentation—Transliteration of Thai (1998)

The update of the Royal Institute's Romanization system was carried out simultaneously with the creation of the ISO 11940 Information and Documentation—Transliteration of Thai.

The creation of this International Standards established a system for the transliteration of Thai characters into Roman characters. One set of rules is provided for a completely reversible stringent conversion whereby absolute principles of transliteration are applied with no modification.

This Romanized transliteration system is a means of converting the Thai writing system into a Roman alphabet writing system. Since there are fewer Roman characters than Thai characters, diacritical marks, punctuation marks, and a combination of two Roman characters, or a digraph, are needed to represent one Thai character. The aim of this system is to provide a means of international communication of written messages in a form which permits the automatic transmission or reconstitution of

either written script by humans or machines. This system of conversion is intended to provide complete and unambiguous reversibility. It is recognized that the transliterated form resulting from this system does not always provide the correct pronunciation of the original Thai text. However, the system serves as a means of finding the Thai graphisms automatically and thus to allow those with a knowledge of Thai to pronounce the Thai text correctly.

In this system 87 Thai characters, including numerical symbols and traditional symbols can be transliterated into the Roman alphabet and re-transliterated back to Thai characters. For example เตะ “to kick” is transliterated as “eta” where “eta” gives the wrong pronunciation but it can be re-transliterated back to เตะ, either by humans or machines.

These 87 Thai characters may be easily accessed and employed through the Unicode system for ease of global communication in the future.

3. Conclusion

An attempt to represent Thai words using Roman script has been endeavored since 1913. Almost a hundred years have passed in which six different systems were employed for different purposes, some with traditional practice, some with

linguistic rules, some with mechanical rules and some with no rules at all. Currently, only three systems play an important role in Thai society: King Rama VI system, The Royal Institute Romanization and Anglicization. It is left to the language educators and learners to choose the one most appropriate to their purpose. However, the system that may be used more extensively in the future on a global scale to exchange information is the ISO 11940 Information and Documentation – Transliteration of Thai which was developed from 1998 in tandem with the universally popular Unicode computer encoding system.

Bibliography

- Amondarunarak, Chamuen (Chaem Sundaravej). 2511 (1968). **The Essential Royal Works of King Rama VI, The Birth of Family Name, book 1.**
- ISO 11940 Information and Documentation – Transliteration of Thai**
ISO/TC 46 N 1854
- The Royal Institute, 1939. **Notification of the Royal Institute Concerning the Transcription of Thai Characters into Roman.**
- The Royal Institute, 1982. **Romanization Guide for Thai Script.**
- The Royal Institute, 2543 (2000). **The Proclamation of the Office of the Prime Minister, the Proclamation of the Royal Institute, and Romanization Rules.**



บทคัดย่อ การถอดอักษรไทยเป็นอักษรโรมัน การถอดอักษร และการถ่ายเสียงเพื่อโลกาภิวัตน์ของภาษาไทย
นิตยา กาญจนวรรณ

ภาควิชาอักษรศาสตร์ สำนักศิลปกรรม ราชบัณฑิตยสถาน

บทความนี้ชี้ให้เห็นความแตกต่างระหว่างการถ่ายเสียง การถอดอักษร การถอดอักษรไทยเป็นอักษรโรมัน และการลากเข้าสู่ภาษาอังกฤษ เปรียบเทียบระบบการใช้อักษรโรมันเขียนคำไทย ๖ ระบบ อันได้แก่ระบบของพระบาทสมเด็จพระมงกุฎเกล้าเจ้าอยู่หัว ระบบแรกของราชบัณฑิตยสถาน (ระบบทั่วไปและระบบพิสดาร) ระบบที่ ๒ ของราชบัณฑิตยสถาน ระบบปัจจุบันของราชบัณฑิตยสถาน ระบบการลากเข้าสู่ภาษาอังกฤษ และระบบ ISO ทุกระบบมีประโยชน์และข้อจำกัดในแบบต่างๆ กัน ส่วนระบบ ISO เป็นระบบที่ทั้งมนุษย์และเครื่องสามารถใช้แลกเปลี่ยนข้อมูลได้ทั่วโลก

คำสำคัญ : การถ่ายเสียง, การถอดอักษร, การถอดอักษรไทยเป็นอักษรโรมัน, การลากเข้าสู่ภาษาอังกฤษ, ระบบของพระบาทสมเด็จพระมงกุฎเกล้าเจ้าอยู่หัว, ระบบของราชบัณฑิตยสถาน, ระบบของ ISO

หลักเกณฑ์การถอดอักษรไทยเป็นอักษรโรมันแบบถ่ายเสียง

ข้อ ๙ การถอดคำย่อ

๙.๑ คำย่อที่มาจากคำเต็มที่อยู่จักกันดีและเป็นคำที่ไม่ยาวนัก ให้อ่านเต็มตามหลักการอ่าน และให้ถอดเป็นอักษรโรมันเต็มตามคำอ่าน เช่น

จ.	ซึ่งย่อมาจากคำ	จังหวัด	=	changwat
อ.	” ”	อำเภอ	=	amphoe
ชม.	” ”	ชัวโมง	=	chuamong
น.	” ”	นาฬิกา	=	nalika
พ.ศ.	” ”	พุทธศักราช	=	phutthasakkarat
รร.	” ”	โรงเรียน	=	rongrian

๙.๒ คำย่อที่มาจากคำประสมหลายคำและค่อนข้างยาว หรือยาวมากจะถอดตามคำอ่านของตัวย่อ หรือถอดเต็มก็ได้ เช่น

อ.ตร.	ย่อจาก	อธิบดีกรมตำรวจ	=	otoro หรือ athipbodi krom tamruat
ผอ.	”	ผู้อำนวยการ	=	pho-o หรือ phu-amnuaikan
สปช.	”	สำนักงานคณะกรรมการการประถมศึกษาแห่งชาติ	=	sopocho หรือ Samnakngan Khanakammakan Kanprathomsueksa Haeng Chat

ประกาศในราชกิจจานุเบกษา เล่ม ๑๑๖ ตอนที่ ๓๗ง ๑๑ พฤษภาคม ๒๕๕๒