

UNITED NATIONS

**Group of Experts on
Geographical Names**
Fifteenth Session
Geneva, 11-19 November 1991

WORKING PAPER

No.41

Agenda item 14

REPORT OF THE WORKING GROUP ON ROMANIZATION SYSTEMS

TRANSLITERATION OF THAI

TRANSLITERATION OF THAI

I. SCOPE AND FIELD OF APPLICATION

This Romanized transliteration system is a means of converting the Thai writing system into the Roman alphabet. 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 for international communication of written messages in a form which permits the automatic transmission or reconstitution of these by men or machines. This system of conversion is intended to provide complete and unambiguous reversibility. It is possible that the result obtained from this system will not provide for correct pronunciation of the original Thai text, but it will serve as a means of finding automatically the Thai graphism and thus to allow anyone with a knowledge of Thai to pronounce the Thai text correctly. An attempt to transpose a preposed vowel after an initial consonant may comfort some who are used to the Roman writing system where a vowel always follows a consonant to form a syllable unless it is a syllable without an initial. It is best to leave the preposed vowel in the Thai transliteration system where it is because a transliteration system is not a transcription system.

Capital letters are reserved for writing an initial in proper nouns.

It is unwise to make use of capital letters in the transliteration system for certain consonants; it has been found that common nouns and proper nouns cannot be distinguished from one another.

This system is based on phonetic logic in the selection of transliterated representations as much as possible in order to lessen the burden on those who wish to learn the system.

3.2.1.2.3 Vowels written below the line
(lower level characters):

• • •

3.2.1.3 Tone Marks and Cancellation Mark written
in the uppermost level¹ (uppermost level
characters):

• • • • •

3.2.2 Thai Digits:

๐ ๑ ๒ ๓ ๔ ๕ ๖ ๗ ๘ ๙

3.2.3 Special Markers:

๗ ๙ ๐ ๑ ๒ ๓

3.3 "Special Symbols" refers to the four symbols used to
differentiate one Roman character that represents the
same sound converted from different Thai characters:

- 3.3.1 upperscore " (ISO 5426 : 4/5)
- 3.3.2 underscore " (ISO 5426 : 5/8)
- 3.3.3 underpoint " (ISO 5426 : 5/6)
- 3.3.4 horn " (ISO 5426 : 4/14)

4. ARRANGEMENT OF CHARACTER SEQUENCE

4.1 The conversion is made character by character from left
to right.

4.2 A consonant (characters # 1 - # 48 in 6.) can have either
an upper level or a lower level character and/or an
uppermost level character; the sequence of the characters
must be as follows:

4.2.1 A consonant with either an upper level character or
an uppermost level character:

- consonant
- upper level character or uppermost level character

4.2.2 A consonant with both an upper level character and an uppermost level character:

- consonant
- uppermost level character
- upper level character

4.2.3 A consonant with a lower level character:

- consonant
- lower level character

4.2.4 A consonant with an uppermost level character and a lower level character:

- consonant
- uppermost level character
- lower level character

It must be noted that both an upper level character and a lower level character cannot concurrently appear with the same consonant.

5. TYPING SEQUENCE

5.1 If a Roman character contains one special symbol, type the character before the symbol.

5.2 If a Roman character contains two special symbols, type the character first, then the upperscore, and the underpoint or the horn respectively.

6. ROMANIZATION TABLE

NO.	THAI CHARACTER	ROMANIZED CHARACTER	CHARACTER CODE	COMMENTS
1	ก	ก		Ko Kai
2	ข	kh		Kho Khai
3	ฃ	kh		Kho Khuat
4	ค	kh		Kho Khwai
5	ฆ	kh		Kho Khon
6	ฅ	kh		Kho Rakhang
7	ง	ng		Ngo Ngu
8	จ	c		Co Can
9	ฉ	ch		Cho Ching
10	ช	ch		Cho Chang
11	ซ	s		So So
12	ฌ	ch		Cho Choe
13	ญ	y		Yo Ying
14	ฎ	d		Do Chada
15	ฏ	t		To Patak
16	ฐ	th		Tho Than
17	ฑ	th		Tho Montho
18	ฒ	th		Tho Phu Thao
19	ณ	p		No Nen
20	ด	d		Do Dek
21	ต	t		To Tao
22	ถ	th		Tho Thung
23	ธ	th		Tho Thahan

NO	THAI CHARACTER	ROMANIZED CHARACTER	CHARACTER CODE	COMMENTS
24	ธ	th		Tho Thong
25	น	n		No Nu
26	บ	b		Bo Bai Mai
27	ป	p		Po Pla
28	พ	ph		Pho Phung
29	ฟ	f		Fo Fa
30	ผ	ph		Pho Phan
31	ฝ	f		Fo Fan
32	ผ	ph		Pho Samphao
33	ม	m		Mo Ma
34	ย	y		Yo Yak
35	ร	r		Ro Rua
36	ฤ	v		Ro Ru
37	ฌ	z		Ro Ru (long)
38	ล	l		Lo Ling
39	ฬ	l	ISO 5426: 718	Lo Lu
40	ฌ	q		Lo Lu (long)
41	ว	w		Wo Waen
42	ศ	s		So Sala
43	ษ	s		So Rusi
44	ห	h		So Sua
45	ฬ	h		Ho Hip

NO	THAI CHARACTER	ROMANIZED CHARACTER	CHARACTER CODE	COMMENTS
46	น	h		Lo Chula
47	ด	x		O Ang
48	บ	p		Ho Nok Itak
49	ป	e		Sara A
50	พ	q		MaI Han Akat
51	พ	h		Sara A (long)
52	พ	h		Sara Am
53	พ	h		Sara I
54	พ	h		Sara I (long)
55	พ	q		Sara U (unrounded)
56	พ	q		Sara U (long)
57	พ	u		Sara U (rounded)
58	พ	h		Sara U (long rounded)
59	พ	e		Sara E
60	พ	ae	ISO 5426: 7/1	Sara Ae
61	พ	o		Sara O
62	พ	e	ISO 5426: 7/5	Sara Ai MaI Ruan
63	พ	h		Sara Ai MaI MaI MaI
64	พ	o	ISO 5426: 4/1	MaI Ek
65	พ	e	ISO 5426: 4/3	MaI Tho

NO	THAI CHARACTER	ROMANIZED CHARACTER	CHARACTER CODE	COMMENTS
66	๑	c	ISO 5426: 4/2	Nai Tet
67	๒	v	ISO 5426: 4/15	Nai Chattava
68	๓	c	ISO 5426: 5/13	"Thanthakhat" silences a character or a group of characters when it appears above that character or the last character of the group such as ๓๑๓ ๓๑๓ ๓๑๓ ๓๑๓.
69	๔	e	ISO 5426: 4/6	"Nai Tai Khui" shortens a long vowel syllable.
70	๕	"	ISO 5426: 3/8	"Nai Yamok" indicates a repetition of a word, a group of words or a phrase, or even a sentence.
71	๖	f	ISO 5426: 3/6	"Paiyan (oi)" indicates that the word has been shortened. It can combine with ๑ as ๑๖๑ which is equivalent to etc., i.e., et cetera.
72	๗	e	ISO 5426: 5/15	"Yamakkan" is written over the first member of a consonant cluster in old texts to indicate that it must be pronounced as a consonant cluster.

NO	THAT CHARACTER	ROMANTIZED CHARACTER	CHARACTER CODE	COMMENTS
73	⊙	⊙	ISO 5426: 2/7	<p>“Pa Kau” or “Fong Han” is used at the beginning of a stanza in poetry and also at the beginning of a paragraph in old texts.</p> <p>“Fong Han Fan Nu” occurs in older texts and can be considered a variant of ⊙.</p> <p>“Angkhan Dieo” is used at the end of a sentence, and at the end of a stanza in old texts. It is also used as a middle symbol in writing dates in the lunar calendar, for example 4 7 6 meaning Wednesday, the 15th day (Full moon) of the sixth lunar month. (Note: It looks very similar to “, “Paiyan Hoju” (in § 71.)</p> <p>“Vangkhan Khuv” is used to terminate a paragraph in old texts. It can also combine with “, “Kho Hoju” in § 76 as “ to terminate a story.</p>
74	◌	◌	ISO 5426: 2/6	
75	◌	◌	ISO 5426: 3/12	

NO	THAI CHARACTER	ROMANIZED CHARACTER	CHARACTER CODE	COMMENTS
76	◌	◌	ISO 5426: 3/11	"Kho Mut" is written at the end of a story to indicate termination in old texts.
77	◌	◌	ISO 5426: 4/10	"Nikkhahit" is written above the rightmost shoulder of a single character to represent an 'ang' of 'am' sound, such as ◌ [◌] pronouncing (sāṅkhan), and above the rightmost shoulder of a character and a vowel to represent an 'm' sound, such as ◌ [◌] pronouncing (chumnum).
78	◌	◌	ISO 5426: 5/4	"Phinthu" is written under a single consonant to signify a final or under the first member of double consonants to signify a consonant cluster in Pali texts.
79	◌	0		Sun
80	◌	1		Nung
81	◌	2		Song
82	◌	3		Sam
83	◌	4		Si
84	◌	5		Ha
85	◌	6		Hok

NO	THAI CHARACTER	ROMANIZED CHARACTER	CHARACTER CODE	COMMENTS
86	๗	7		Chet
87	๘	8		Pat
88	๙	9		Kao

ANNEX

CHARACTER AND SPECIAL SYMBOL SPECIFICATION

A1. Thai consonants with the same sound are represented by the same Roman character or a digraph of Roman characters. A special symbol or special symbols will be used in differentiating the Thai character equivalents, for example:

ก = ph

ค = ph

A2. An "h" as a second member of a digraph signifies aspiration, for example:

กห = ch

คห = ph

A3. A special symbol is placed above, below, or at the right of a character of the first member of a digraph, for examples:

ก̄ = kh

ก̇ = th

ก̂ = t'h

A4. High Class Consonants² are identified by an upperscore marker in order that automatic pronunciation rules can be easily derived from them, for example:

ก̄ = kh

ก̇ = t̄

Remark : an example of an automatic pronunciation rule :

a consonant with an upperscore occurring before a long vowel with or without a sonorant will yield a rising tone.

A5. The less frequently used consonant of two different consonants with the same sound in the same consonant class (i.e. both Middle Class Consonants or both Low Class Consonants) is identified by underpoint, for example:

š = ṣ̌
ŋ = ŋ̣

A6. A High Class Consonant is differentiated from a Low Class Consonant having the same sound by an upper score, for example:

ṣ̌ = ṣ̌̇
ŋ̣ = ŋ̣̇

A7. If more than two different consonants have the same sound, a special symbol or symbols (as indicated in 1.3) must be employed. An underpoint, a horn, or an underscore will be used in descending order respectively, according to their frequency of occurrence in the Thai language, for example:

ṣ̌̇ = tḥ
ṣ̣̌̇ = tḥ̣
ṣ̣̣̌̇ = tḥ̣̣
ṣ̣̣̣̌̇ = tḥ̣̣̣

A8. Short and long vowels are specified by the same Romanized character. An upper score mark identifies a long vowel, for example:

ə̣̇ = e
ə̣̣̇ = ē

A9. The four Roman Characters q, v, x, and z are used in representing consonants with different phonetic value as follows:

- q with a phonetic value of [ɓ] for v.
- v with a phonetic value of [ɓɓ] for z.
- x with a phonetic value of [ɓɓ] for q.
- z with a phonetic value of [ʔ] for x.

FOOTNOTES:

¹If a tone mark or a cancellation mark is directly over a consonant without an upper level character, it is adjusted to an upper level in typesetting and displaying on the computer screen for neatness of appearance.

²There are three consonant classes in the Thai language: High, Middle, and Low.

The High Class Consonants are ๑ ๒ ๓ ๔ ๕ ๖ ๗ ๘ ๙ ๑๐. They are classified as High Class Consonants because a live syllable (a syllable ending with a long vowel or a sonorant) which has a High Class Consonant as its initial is pronounced with a rising tone and can take ^๑ and ^๒ tone marks, yielding low and falling tones respectively; a dead syllable (a syllable ending in a short vowel or a stop consonant) which has a High Class Consonant as its initial is pronounced with a low tone and can take only ^๑, yielding a falling tone.

The Middle Class Consonants are ๑๑ ๑๒ ๑๓ ๑๔ ๑๕ ๑๖ ๑๗ ๑๘. They are classified as Middle Class Consonants because a live syllable which has a Middle Class Consonant as its initial is pronounced with a mid tone and can take ^๑, ^๒, ^๓, and ^๔ tone marks, yielding low, falling, high, and rising tones respectively; a dead syllable which has a Middle Class Consonant as its initial can take ^๑, ^๒, and ^๓ tone marks, yielding falling, high, and rising tones respectively.

