

Changing the Northern Khmer orthography

by Dorothy M. Thomas

© 2018 SIL International® All rights reserved.

Fair-Use Policy: Publications in SIL's Notes on Literacy series are intended for personal and educational use. You may make copies of these publications for instructional purposes free of charge (within fair-use guidelines) and without further permission. Republication or commercial use of SIL's Notes on Literacy is expressly prohibited without the written consent of the copyright holder.

Originally published as:

Thomas, Dorothy M. 1989. "Changing the Northern Khmer orthography." *Notes on Literacy* 57:47–59. See also: [Thomas, D. 1989a](#) (in *Bibliography (Literacy)*)

[Topics: Asia: Thailand: Khmer, Northern, orthography, scripts: non-Roman]

1. Introduction

Unless you have tangled with the Devanagari scripts of South and Southeast Asia, it is hard to imagine how complicated it is to write still another language based on one of the daughter scripts, such as Cambodian and Thai. But a reasonable orthography using Thai letters had been worked out for the Northern Khmer (N.K.) language by William Smalley and John Ellison and it had been used, with some adaptations, for about 15 years. However, there are so many problems and varieties of possible solutions, plus interference from the Cambodian script for Standard Khmer, that some objections had been raised recently by influential people. The hope was to get a consensus from the people themselves and to stir up interest in writing this dialect of Khmer, with a minimum of conflict with the Thai writing system, yet efficient for N.K.

On April 23–24, 1987, the Institute of Language and Culture for Rural Development at Mahidol University, Bangkok, on the invitation of David and Dorothy Thomas, SIL members and instructors at the Institute, sponsored a conference in Surin Province, the heart of the N.K. area in Thailand.

2. Background

Some of the problems of writing N.K. are accidents of history, some are because of language differences, and some are psycholinguistic. And of course these are intertwined.

Historically, this area was part of the great Khmer empire which flourished from A.D. 800 to 1400. Now all the land north of the Dangrek mountains belongs to Thailand. So the Northern Khmer people have been cut off politically from the Southern Khmer of Cambodia for a long time; the mountains form a natural barrier to frequent interaction with the Cambodian Khmer. The approximately one million speakers of N.K. are concentrated in Surin province, just north of the Cambodian border, but there are also many in the provinces east and west of Surin.

When the Thai arrived in what is now modern Thailand, the Mon and the Khmer controlled most of the land. The Thai borrowed much from their languages and cultures. Eventually, the Mon were almost completely assimilated into Thai culture, but there was a protracted period of conflict between the Khmer and the Thai. When Thailand became dominant in the area, there was much influence from Thai on Khmer, so now there are many borrowings and re borrowings in both directions, including Sanskrit and Pali words from the Buddhist literature. Furthermore, although the Mon, Khmer, and Thai scripts are all derived from the Devanagari script of India, each one is different. The ancient system of writing borrowed words seems to have been as in English: write the borrowed word as it is written in the donor language, not as it is pronounced in the receptor language. So now words borrowed at different times or from an intermediate donor are spelled differently, especially in Thai, although the meanings may be identical or closely related. So the question is, how should N.K. spell these words?

Mon and Khmer had a register distinction between breathy vowels and nonbreathy vowels, but it was written as two different sets of initial consonants. Thai may also have had the same distinction, but now it is reflected in the tonal system, further complicating the writing system of Thai. N.K. has lost the register distinction completely but has developed an incredible number of pure vowels as a result.

In more recent history, around the turn of this century, the Thai government developed a policy of having only one standard written language for the whole country. So, N.K. was not written.

Orthographically, the language with the strongest influence on N.K. is Standard Thai. The area also has a lot of Lao and Kuy villages, and some people speak all three languages (Lao, Kuy, and N.K.), plus Standard Thai (closely related to Lao). However, the Lao orthography is not well known here so it has little influence on the writing of N.K. Kuy is very nonprestigious, so it also has had little influence on Khmer. Although N.K. and Cambodian are closely related, only men who read Cambodian Buddhist texts while in the monkhood can read the Cambodian script. So the influence of the Cambodian script on the majority of the northern Khmer is minimal.

With such a large number of speakers, the N.K. language shows no signs of dying out. The people see themselves as Thai who speak Khmer. In Surin Province, they are in the majority, and although in most of the high schools the students are fined for speaking N.K., grade school teachers have a real problem with the young Khmer pupils from the villages because they really do not know Thai when they begin to study in school. Although many of the government officials and school teachers are Thai or Lao, there are also many well-educated Khmer in high positions. These include the provincial governor of Surin, the mayor of Surin City, and many local officials. So the Khmer are not embarrassed to use their language in many public situations, but for official communication they use Thai. Church services are usually in Thai because either the leaders are Thai and do not speak any N.K., or else they are Khmer but were trained

only in Thai and are embarrassed by their lack of vocabulary in N.K. However, the Khmer respond enthusiastically to songs or teaching in their own language.

The languages of the region have developed many extra letters and irregular spellings because of their histories. A Thai child takes about as long to learn to read as his counterpart does in the English language. This situation creates an unbearable load for the beginning Khmer pupils. For a Khmer child to learn to read in his own language first would be a distinct advantage, but the likelihood that he would have a long period with Khmer classes in school is rather slim. So, to have a fairly one-to-one relationship of sound and symbol in N.K. orthography would be a big help for the children.

3. Preparation for the conference

The acting head of the Institute for Language and Culture, Dr. Suwilai Premsrirat, has taught a field methods course where the students worked on N.K. She knows something about the language already, as well as having specialized in another Mon-Khmer language, so she was the natural one to head the conference. In the fall of 1986, she began making contact with various organizations and government agencies working in the N.K. area, notifying them of the proposed conference. She got official recognition of the conference by the government so that civil servants could have their expenses paid and the conference would have some prestige. She invited the governor of Surin to give the opening address and asked his wife, the principal of a prestigious high school in Surin City, to host the conference.

Stephen Teel, SIL-CMA, drafted a write-up of the orthography problems with explanations and various possible solutions. Dr. Suwilai translated this write-up into Thai and had it copied as a handout for the conference. Students from Mahidol prepared charts of the N.K. phonemes with possible Thai letters. Also available were a number of N.K. texts in the old orthography.

4. Conference deliberations

About 10 Thai, 10 Khmer, and three Westerners came for the meeting. The Thai speakers were all from Mahidol, either professors or students. All the Khmer speakers were from Surin province, although people from the other two provinces had responded that they would come. One Khmer was also a student at Mahidol, writing his thesis on a comparison of N.K. and Standard Khmer (Cambodian). One Khmer had a Ph.D. degree; some had never graduated from high school. There was one Khmer woman.

Discussions were held in Thai or N.K. Foreigners said as little as possible, but the Thai linguists contributed well to the discussion. One Khmer who has been using the Smalley-Ellison script for years was greatly pleased with their help. Final decisions were made by the Khmer speakers.

A tight schedule had been prepared, but it got off to a slow start because the Khmer had a long discussion about the necessity of a separate writing system for N.K. (Some were familiar with the Cambodian writing system and even had a vested interest in it.) However, what settled the question was the testimony of one Khmer schoolteacher. After that, there was much discussion about particular points, but a general agreement as to purpose.

One of the Thai professors remarked that she had never seen a seminar before where the participants were so committed. At the end of the second morning they stayed on until 1:30, and then even returned after lunch to finish, although the meeting was officially to end at 12:30. The hotly debated items will never be forgotten by the Khmer delegates, and this fact will stand them in good stead as they try to convince their friends concerning the orthography.

At the end, one Khmer stood up and said how revolutionary it was, that he had tried to write N.K. himself, but unsatisfactorily, and that now he felt he could write it.

Testing is the next step. With our “consistent changes” computer program, changing the orthography in old materials is a simple matter. So we prepared a booklet of folk tales on mimeograph stencils and the local office for Nonformal Education is printing it for their cadres to use in village literacy classes. The cadres will fill out a form indicating places where the villagers had problems. Hopefully, one of the Institute students or graduates will also conduct a survey on the problem points.

Dr. Suwilai sent out a report of the decisions to all the participants and asked if they wanted a follow-up meeting in six months or a year. They voted for a year, so the results of the testing should be in by that time.

5. Problems and decisions

Eventually, everything which we considered nonnegotiable was adopted, and almost everything negotiable was changed.

5.1. Word level

The most surprising decisions of the conference were 1) to write word spaces and 2) where to position certain vowels. Both these decisions are very non-Thai; however, the Khmer realized the ambiguities they would eliminate by making the decisions.

1. Thai does not write word spaces, and thus it is difficult to identify word breaks. Unless the context is clear, it is impossible to know if there is an open syllable, followed by an initial consonant cluster, or a closed syllable followed by a single (or double) consonant. Since N.K. has more possible consonant clusters and more complicated ones than Thai, the problem is multiplied if word breaks are not inserted. The computer also needs word breaks.
2. Some vowels in Devanagari-based scripts are written before the consonant that actually precedes them phonetically and phonemically. In some cases in Thai, and more often in N.K., this situation is complicated for two reasons. If there is a consonant cluster, Thai writes the vowel to the left of the whole cluster. So it is not clear if there is one word with a consonant cluster or two words, one with the fronted vowel, and one with the unwritten vowel. Since the Khmer decided to write word breaks, this problem will not exist, but there is still a problem with simple three-phoneme syllables. Are they CVC or CCV? There is one minimal pair in N.K. which really is ambiguous if vowels are written before clusters. But even for other words, it will make word attack simpler to know at a glance if

there is a consonant cluster or not. So it was decided to write the vowels before the consonant closest to the syllable nucleus.

5.2. Morphophonemically

Morphophonemically there are also a few problems; some similar to Thai, others not.

1. As in Thai, when there are two nouns in a noun phrase, if the first one has a long vowel, it is shortened. Thai write these vowels consistently as long. This problem was not actually discussed, but undoubtedly the Khmer will want “words spelled consistently.”
2. N.K. has several functors which are unstressed open syllables, and the vowels are not clear. The morphemes are really clitics, but when written as unstressed syllables of an adjoining word, they form unusual word patterns. When written as full words, they get too much stress. In Thai, similar functors can end with a final glottal when stressed, and are written with a glottal, although they frequently end in light aspiration or an open syllable. But this solution is not suitable for N.K., where /-ʔ/ and /-h/ are both in strong contrast with an open syllable.
The conference participants decided to connect the clitics to the main word. However, if the functor precedes a word with a preposed vowel, the vowel of the clitic can be dropped.
3. There is one functor /k əʔ/ which functions almost exactly like an analogous functor in Thai, but sounds a little different, especially when stressed. The Thai functor is written with no vowel but with the symbol for shortening a vowel. There is one problem for N.K., in that there is also a very common pronoun which has a somewhat similar shape. For now, we will try writing the N.K. functor as in Thai.
4. Unlike Thai, sentence intonation is very important in N.K., usually affecting the last word of a clause. Thai does not use punctuation marks, and has only a break in the type to show sentence breaks. The committee was definite that they want more breaks written than in Thai.
There is a special intonation pattern for yes-no questions. It is either a lengthened final syllable or an added vowel. The added vowel does not form a true syllable, as there is no initial consonant and it is unstressed. Instead of writing the question intonation-syllable, the conference participants decided to use the question mark.

5.3. Proper names

Proper names are very frequently borrowed from Thai, Pali, or Sanskrit. For personal names, people will probably want to write them as in Thai. But for geographical names, the conference participants decided to write them first in N.K. and then in Thai in parentheses. This procedure will be especially helpful when the N.K. name is actually different from the Thai name.

5.4. Loan words

Loan words will be written as pronounced in N.K. for now, but will be followed by the Thai spelling in parentheses. Probably, this procedure will come up later for further discussion.

5.5. Northern Khmer sounds and symbols

Consonants					Vowels		
ph	th	ch	kh	h*	i,ii	iii,uuu	u,uu
p	t	c*	k	ʔ	i,u	o,oo	v,vv
b	d				e,ee	a,aa	o,oo
m	n	ɲ	ŋ		ɛ,ɛɛ	ʌ,ʌʌ	ɔ,ɔɔ
		r*				a,aa	ɔ,bb
		l*			ɛ,iiɛ	ʌ(ɛ)uuɛ	ʌ(ɛ)uuɛ
(fin loans)	s						
w	y						

Figure 1 Phonemes of Northern Khmer (Sounds not occurring in Thai are circled.)

The phonemic inventory of consonants in N.K. is as shown in Figure 1. Consonants marked with an asterisk do not occur finally in Thai, but do in N.K. /s/ is only initial in both languages. The short vowel /ɔ/ is predictably short in Thai, so it is written as a long vowel in Thai, but the length is contrastive in N.K. The vowel /uuu/ occurs in only one word in N.K.

5.5.1. Consonants

One nonproblem is what to do with the two sets of consonants in Thai which represent tonal classes. Everyone recognizes that N.K. has no phonemic tones, so the set for mid tone is used. Another nonproblem is the final voiceless stops: in Thai they are written as voiced stops. But, in neither language is there voicing contrast in the final position so it does not matter. The real problems for the consonants are the phonemes that are only word-initial in Thai, and the one phoneme /p/ that does not exist in Thai.

Although the consonant /p/ does not occur in Thai at all, Thai does have a letter that represents /y/ initially and /n/ finally, and in some dialects of Thai it is actually pronounced as a palatal nasal. So, it is a natural symbol for the N.K. palatal nasal. However, there is a problem because, after front vowels in N.K., there is no contrast between /p/ and /ŋ/. Except for Khmer who are very phonetically sophisticated, psychologically they are convinced that they are saying a velar. The tongue tip is clearly behind the lower teeth, causing the tongue to hump in the middle; and the phonetics professor assured the delegates that they were actually saying a palatal, but they remained unconvinced. (Historically, most of these palatals after front vowels were velars.) So for now “every man will write what is right in his own eyes” for the final palatal nasal.

The final gutturals cause a problem because, in Thai, the final glottal stop sometimes is realized as [h], and in N.K. the final glottal stop is sometimes realized as [k]. Thai has no phonemic final /h/; N.K. has no

(1989). *Notes on Literacy*, 57.

phonemic final /k/. However, the k-ʔ contrast may be coming in from Thai, and there is a strong preference for writing /k/ after the short central and back high vowels. Thai has no glottal stops after long vowels, but N.K. does. So, in the Smalley-Ellison orthography, the Thai symbol for /k/ is used after long vowels. Probably, the Khmer will continue using the /k/ symbol after the long vowels and the short high ones, although the Khmer who are used to writing in the existing orthography do balk at the arbitrary spelling after the short high vowels.

The final /h/ came in for much discussion because it is so non-Thai. A few Thai words are written as if they had a silent /h/, but a different symbol is used from the one used initially for words with level tones. To complicate matters, Cambodian uses an /s/ symbol for final /h/. But, eventually, the symbol used for initial /h/ in N.K. was adopted.

Surprisingly, there was little discussion of the final /r, l, c/. Some minimal pairs were presented and everyone seemed satisfied. The Khmer know that the Thai pronounce these sounds wrongly when they try to speak N.K.

Consonant clusters are not written consistently in Thai. The conference participants decided to write them consistently in N.K., regardless of whether there is a slight open transition between the consonants or not.

Syllabic nasals are not very common in N.K., but they do occur before consonants in a dozen or so words. The “obvious” solution of just writing a simple nasal is not acceptable because it would be read with an intervening (unwritten) vowel. The Smalley-Ellison solution of writing as if there were a presyllable /ʔam/ was adopted, although it does take some learning.

	labial	alv.	palatal	velar	glottal
asp. stop	𑄢	𑄣	𑄤	𑄥	𑄦/-𑄧,-𑄨
unasp. stop	𑄩	𑄪	𑄫	𑄬	
vd. stop	𑄭	𑄮			
nasal	𑄯	𑄰	𑄱	𑄲	
trill		𑄳			
lateral		𑄴			
fricative	(𑄵)	𑄶	𑄷		𑄸
semivowel	𑄹		𑄺		

Figure 2 1987 Northern Khmer orthography—consonants

5.5.2. Vowels

The one vowel that was radically changed was the / ʌ/. Formerly, it was written as the central counterpart of / ɛ/ but it looks strange in loan words, so the conference participants chose to write it as a lowered / ə/.

The special symbols for short /a/ (before /m/, /w/, /y/ and /ʔ/) will be used as in Thai.

The central off-glided vowel is very short in Khmer, but long in Thai. Time will tell whether people will consistently write the symbol for vowel shortening or not.

For the lowered vowels, the present system was accepted: a dot below the line, except when the vowel itself is already written below the line. Then a circle above the vowel is used. The only thing this changed from the present orthography is that we had been using the raised circle instead of the lowered dot for the only lowered vowel which is written on the line /ɨ, ʉ/.

Thai usually writes the vowel shortener above the initial consonant, but this procedure is difficult if there is already a mark there. For the computer, this is a disaster. However, there is some precedence from Thai to shift the shortener to the right (for newspaper headlines), and since we began doing this several years ago, there has been little reaction one way or the other from the Khmer readers. Probably for handwriting, people will put the shortener on top of other diacritics, but they accepted the shifted position for printed materials.

The symbols chosen for the vowels are shown in Figure 3. As an aid to visualizing the positions of the vowels, the examples show an initial glottal and either a final *d* or a final glottal. (Note that the symbol for initial glottal is the same as for parts of the mid-central vowels and the low back vowels.) Final glottal is unwritten after some short vowels. For open syllables (indicated here with #), some long central vowels add the symbol for initial glottal/low back vowel after the vowel.

	short		long	short		long	short		long	
	ʔVd	ʔVʔ	Vd	ʔVd	ʔVʔ	ʔVd	ʔV#	ʔVd	ʔVʔ	Vd
high	ឺត/ឺ		ឺត	ឺត/ឺតា		ឺត/ឺត#		ឺត/ឺ		ឺត
mid-high	ឺត/ឺ		ឺត	ឺត/ឺតា		ឺត/ឺត#		ឺត/ឺ		ឺត
mid	ឺត/ឺត		ឺត	ឺត/ឺតា		ឺត/ឺត#		ឺត/ឺ		ឺត
mid-low	ឺត/ឺត		ឺត	ឺត/ឺតា		ឺត/ឺត#		ឺត/ឺ		ឺត
low		(ឺត, ឺត, ឺត, ឺត)	/ឺត*			ឺត/ឺត#		ឺត/ឺ		ឺត
.
glided	ឺត		ឺត	ឺត		(ឺត)	ឺត/ឺត#/ឺត			ឺត

-d = any consonant (except ʔ following a short vowel)

*The variant symbols for short /a/ are:

/ʔad, ʔaw, ʔam, ʔay and ʔaʔ/

See [Smalley1976](#)

(1989). *Notes on Literacy*, 57.

Page 8.

6. Update on the Northern Khmer orthography

The second Northern Khmer Orthography Conference was held on May 26–27, 1988, in Surin. It, too, was sponsored by Mahidol University, with Dr. Suwilai leading the meetings.

This time there was a good representation from the local offices of Nonformal Education from all three of the main provinces where the Northern Khmer people live. (Last year, due to a conflicting schedule, only two came, and for only one day.)

There were several samples of written N. K. for the delegates to compare

- a folk tales booklet and a diglot health care booklet, in the agreed-on orthography
- a list of plant names in a slightly divergent orthography, and
- some proverbs which had been rewritten hastily (with many slips) by a man who had only just heard about the agreed-on orthography.

Although the folk tales booklet was prepared early, it had not been distributed and tested in time to get valid reactions to the orthography. But, reports did come in about what style of books the villagers want (about eight lines to the page, diglot, lots of pictures, and a better print job).

The main thrust of this year's meeting was getting written Northern Khmer into use. Those present decided to meet again in October to begin producing materials, and discussed the possibility of having a pilot project to see if their children would get a better start in school with initial literacy in Northern Khmer.

The more balanced representation of dialects affected the one decision on spelling. The all-Surin group, the year before, had sidestepped the decision on the neutralized final palatal-velar after front vowels. This time, the participants voted overwhelmingly in favor of writing it as a velar, as it is sometimes pronounced in the other provinces, and frequently *perceived* to be pronounced in Surin. Subsequently, the man with the longest experience with using the palatal symbol made the switch to both reading and writing it as a velar with no noticeable problems at all.

Reference

- See Smalley, William. 1976. "The problem of vowels: Northern Khmer." In *Phonemes and orthography: Language planning in ten minority languages of Thailand*, edited by William A. Smalley. *Pacific Linguistics Series C* 43:43–83.
- Citations