



Deseret Language and Linguistic Society Symposium

Volume 19 | Issue 1

Article 23

4-2-1993

Problems in the Compilation of an English–White Hmong Dictionary

Brian McKibben

Follow this and additional works at: <https://scholarsarchive.byu.edu/dlls>

BYU ScholarsArchive Citation

McKibben, Brian (1993) "Problems in the Compilation of an English–White Hmong Dictionary," *Deseret Language and Linguistic Society Symposium*: Vol. 19 : Iss. 1 , Article 23.

Available at: <https://scholarsarchive.byu.edu/dlls/vol19/iss1/23>

This Article is brought to you for free and open access by the Journals at BYU ScholarsArchive. It has been accepted for inclusion in Deseret Language and Linguistic Society Symposium by an authorized editor of BYU ScholarsArchive. For more information, please contact scholarsarchive@byu.edu, ellen_amatangelo@byu.edu.

Problems in the Compilation of an English—White Hmong Dictionary

Brian McKibben

The Hmong live in the mountains of southern China, northern Vietnam, northern Laos, Thailand, and Burma. During the Vietnam war, Hmong guerrilla soldiers, backed by the CIA, fought to keep the North Vietnamese out of Laos. Hmong people also helped rescue downed American pilots in Laos. When the United States pulled out of Vietnam and the surrounding areas, the Hmong were left to themselves. The communist Pathet Lao regime in Laos persecuted the Hmong relentlessly. Many Lao and Hmong refugees crossed the border into Thailand where they lived in refugee camps. The United States agreed to accept many of these refugees (Garrett 633–642).

As the Hmong refugees began to relocate in America, problems occurred as a pre-literate society met with twentieth-century American technology. Government, social, school, and religious workers found it difficult to communicate with the waves of refugees (Garrett 633–642). Many of these workers wanted to learn to speak Hmong, but had no access to a bilingual reference dictionary.

Beginning in 1990, I began teaching an intensive course to Hmong missionaries in the Missionary Training Center located in Provo, Utah. After over a decade of Hmong refugees living in America, only a few English to Hmong wordlists had been published. I began to work on a more complete pocket dictionary for the missionaries to use. As I worked, I realized that the students could benefit greatly from a more extensive dictionary—one which included explanations and example sentences. After almost three years of research, I finished the **English—White Hmong Dictionary**, currently the only reference from English to White Hmong of its kind.

During the research for this dictionary, I encountered a number of problems, two of which were more significant than the others. The first problem was how to decide on proper spellings of

Hmong words. The second was how to decide which terms to add to the dictionary. I will discuss in this paper how I dealt with these hurdles. In order to discuss these two problems further, it is first necessary to say something about the Hmong language.

The Hmong spoken in Laos and Thailand is made up of two major dialects: green (or blue) and white (Heimbach vii). These names reflect the colors of Hmong women's traditional skirts (Hawj 10). I chose White Hmong as the dialect for my dictionary because it is the most prominent variety.¹

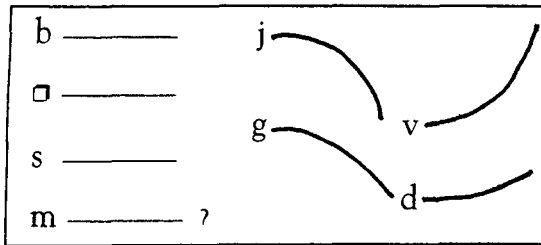
William Smalley, Linwood Barney, and Yves Bertrais developed the romanized Hmong orthography in Laos between 1951 and 1953. They needed a way to write Hmong so they could translate the Bible and other literature. Other systems of writing Hmong exist, but this romanized orthography has become the most widely used system (Smalley 13).

Hmong is a tonal language with eight tones represented in written text by a syllabic final consonant (see Figure 1): b (high level), _ (mid level), s (lower mid level), m (low level with a terminal glottal stop), j (high falling), g (falling breathy), v (mid rising), and d (low rising). There are 55 consonants and 14 vowels (Heimbach xvii–xxiii).

As stated above, I began teaching a Hmong class in December of 1990. No English Hmong reference had been created, yet. I began to compile a wordlist using merge files, making English to Hmong and Hmong to English a one-step procedure. This wordlist evolved into a small pocket dictionary which was printed and distributed to the students learning Hmong. This pocket dictionary consisted of two columns: an English column and a Hmong column. If a student wanted to know a quick translation for an English word, they only needed to look it up.

Brian McKibben is a junior from West Virginia majoring in linguistics at BYU. He spent two years in Sacramento working with Hmong people and three years after that compiling an English—White Hmong Dictionary.

Figure 1: The White Hmong tonal system



As the students began using the pocket dictionaries, we teachers noticed the students becoming confused over some of the words. They did not know how to use some words in a sentence or used the wrong Hmong word for an English homonym. So I began to think that examples and explanations would help the students even more.

So I began planning. The purpose of the dictionary was to help English speakers produce Hmong. It would include explanations to distinguish subtle differences in meanings between words. Sample sentences would show how to use the words and phrases in a sentence, and all of these sentences would be composed by native Hmong speakers.

I remembered a filmstrip I saw once in elementary school which explained how dictionaries were written. It showed a lonely lexicographer leaning over a table, reading a book under the light of a single candle. When he came to a new word, he wrote it down on a 3" by 5" card and carefully stored it away. Although I saw this in second grade, the idea had remained intact. I decided to make a corpus of Hmong material on disk to retrieve and cite example sentences more easily. The sources I drew from included books, audio and video tapes, magazines, newsletters, newspapers, interviews, and letters. In total, nearly 500,000 words were either typed or scanned into computer files for the corpus.

This introduced the first major problem of the project. Hmong spelling has not yet been completely standardized. I decided to standardize the spellings based on Ernest Heimbach's spellings in his *White Hmong—English Dictionary* to create KWIC concordances.

The second case of non-standardization involved the compounding of syllables to form polysyllabic words. For example, "sij hawm" (time) means the same thing as "sijhawm." The only difference is the appearance of the word in print. Again, I decided to follow the precedent set by another reputable Hmong work—in this case, the Hmong Bible. By going through the entire Hmong Bible, I found 180 compound words which were consistently spelled as a compound word. These agreed with Heimbach's dictionary. However, as I corresponded with Yves Bertrais, he expressed his concern over the compounding of these words for two main reasons: first, it was an arbitrary system, and second, that some words in Hmong, such as

"phooj ywg" (friend) were borrowed from Chinese where they are written as two separate characters: 朋友 (péngyu).

I researched these two concerns and found that the reasoning for the compounding was the result of the two syllables being said closer together in running speech than other words (Heimbach xvii), and when Chinese is transferred to a romanized script, such as pinyin, some of the two-character words are written as a single word-unit called a *cī zǐ* (Smith). In addition to these two reasons, compounding the words makes spell-checking and searching easier and more accurate.

Figure 2: Microstructure of an Entry

clean v. ntxuav, ntxhua, cheb
 "Ntxuav" refers to cleaning dishes and parts of the body:
 ex. Some went to *clean* themselves. Ib txhia los mus ntxuav tes taw.^{P,34}
 "Ntxhua" refers to cleaning clothing:
 ex. Some went to *clean* clothing. Ib txhia los mus ntxhua khaubncaws.^{P,34}
 "Cheb" refers to cleaning or straightening a place:
 ex. The courtyard had to be *cleaned*. Yuav tsum cheb tej vaj loog.^N

Having dealt with the first major problem, I went on to decide on the macro and microstructures of the dictionary. The macrostructure has to do with the content of the dictionary, or the lexemes that compose the body of the dictionary. The microstructure of the dictionary deals with what information is presented for each lexeme (Béjoint 70–71). For example, the macrostructure of this dictionary would include the basic words and phrases as might be needed by an English speaker learning to speak Hmong. The microstructure, on the other hand, would include the lexeme itself, the part of speech, the corresponding Hmong word or words, any explanations, English and Hmong example sentences, and citations stating where those sentences were obtained (see Figure 2).

I developed a flow chart for adding words to the dictionary (see Figure 3). First, I read the Hmong material. When I encountered a new Hmong term or phrase, it was first accurately defined by interviewing Hmong people. If only a few Hmong people understood the term, a further test ensued. If the Hmong language had another term available for the same word in English, it was not included in the dictionary. If, however, Hmong had no other term available for the same English word, it was included with an explanation. After that, any secondary meanings of the word were determined by comparing other instances of the word in the corpus.

At this point, the second major problem occurred. Many words and phrases which Hmong people use are actually Lao and Thai borrowings (Hawj 4). In fact, some Hmong people use Lao

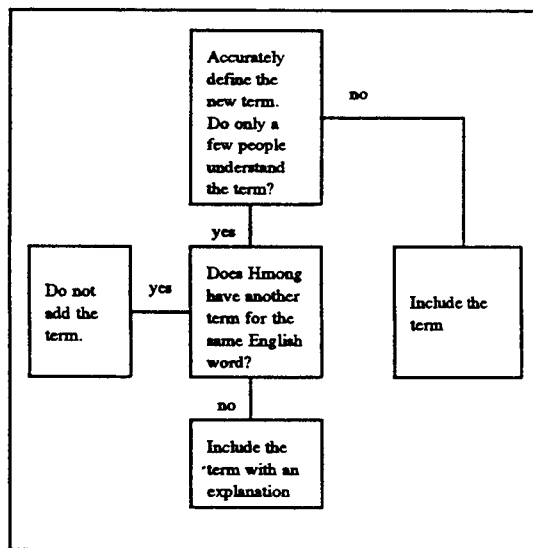


Figure 3: Flow chart for adding words to the dictionary

and Thai borrowings so much, that we sometimes find ourselves slipping into Lao when speaking with informants. However, there is a tendency, almost an unwritten “rule,” to resist using these borrowings in written Hmong texts. This may be because of the difficulty of writing Lao and Thai words using the Hmong orthography. I had to decide whether or not these terms were important enough to include in the dictionary.

In addition to this, here in America there are many things for which no Hmong words exist (i.e., “insurance”). Rather than invent a new Hmong word, many Hmong people use a “Hmongified” version of the English word, often twisting the English word so as to be unintelligible by native English speakers who speak Hmong.

In any case, I had to develop a set of criteria for adding Lao/Thai and English borrowings to the dictionary. I decided that if Hmong people understood and used the Lao/Thai borrowing, and if the Lao/Thai term were used as much as or more than the Hmong term, or the Lao/Thai term had undergone some kind of semantic or phonological change, or if no other Hmong term existed for the same English word, then I would add the Lao/Thai term (see Figure 4).

In deciding whether or not an English borrowing should be added, I concluded that since the original aim of the dictionary was for English speakers to produce Hmong, and since most English speakers would encounter Hmong people in predominantly English-speaking areas (i.e., North America and Australia), the English borrowings could be included based on the same criteria as used for Lao/Thai terms. Furthermore, if the dictionary were used in Laos, the Hmongified English words would probably not come up in conversation (i.e.

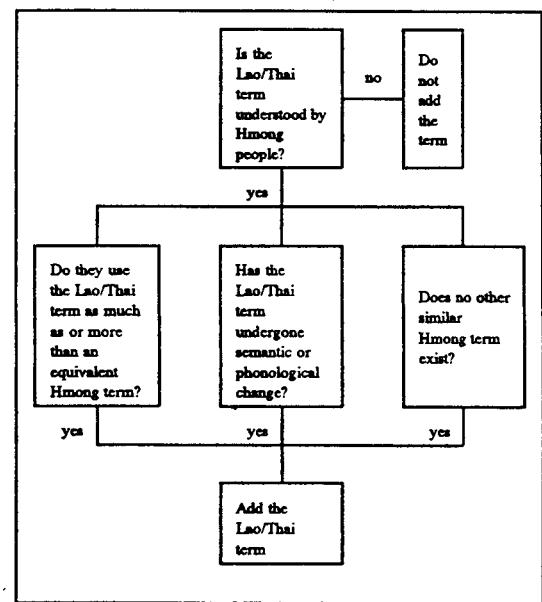


Figure 4: Flow chart for adding Lao/Thai borrowings

“insurance”, “snow”, “ice cream”). In most cases I tried to make sure another equivalent term was available in Lao/Thai or Hmong.

After a period of time using these processes, I would print a prototype and distribute it to the students we taught. They gave me their suggestions and comments and I made corrections or enhancements based on their feedback. This went on for two and a half years, until finally, I printed the finalized version of the English—White Hmong Dictionary in November of 1992.

Solving the two biggest problems of the dictionary—how to standardize the spellings and how to decide whether or not to include borrowings from Lao/Thai and English—involved research, collaboration, and thought. Research helped the dictionary to be accurate. Collaboration helped to insure that the dictionary could help learners of Hmong in the best way possible. It also helped me to know what other ideas authorities in the Hmong field have. With the help of decision-making flow charts and judgement, I was able to evaluate and solve these problems. I hope that this dictionary will fill a void which has existed for too long and help English speakers who want to communicate with Hmong people do so more easily.

Works Cited

- Béjoint, H. “On Field-Work in Lexicography: Field Dictionaries.” *Lexicography: Principles and Practice*. Ed. R. R. K. Hartmann. London: Academic Press, Inc., 1983. 53–76.
- Garrett, W. “Thailand: Refuge from Terror.” *National Geographic* (1980): 633–642.

- Hawj, Soobleej. *Tswv Yim Moob*. Milwaukee: Soobleej Hawj, 1990.
- Heimbach, Ernest E. *White Hmong—English Dictionary*. Ithaca, NY: Cornell University Press, 1969.
- Smalley, William. "Hmong Writing." *Haiw Hmoob* Dec. 1987: 13+.
- Smith, Kim. Personal interview. 11 June 1993.

End Note

¹From now on, unless specifically referring to another dialect, Hmong will refer to White Hmong.