Chapter I

Introduction

Background and Rationale

The Tai family of languages has been divided into three branches: the Northern, Central and Southwestern (SW) branches (Li 1960). Thai belongs to the SW branch along with Lao, Tai Yuan and among others, the Shan languages. All of these languages have a genetic relationship; some languages being more closely related than others.

Chamberlain (1972, 1975) claims that the Tai languages of Li's SW branch can be divided into two groups. These are a P and a PH group. The distinction between the two groups is based on an individual language's development of the Proto-Tai (PT) voiced stop consonant series. In languages which fall into the P Group (PG): Shan, Tai Lü, Black Tai, White Tai, Tai Khamti, Tai Mao, Tai Yuan, etc., the proto voiced stops have developed into homorganic voiceless unaspirated stops. In languages of the PH group: Thai, Lao, Phu Thai, Phuan, etc., the proto voiced stops have developed into the homorganic voiceless aspirated stops. Gedney (1991) is among those who disagree with this subdivision of Southwestern Tai (SWT) languages believing that the P verses PH distinction is a fairly recent development. Generally it has been assumed that the languages of the SW branch are more or less homogeneous.

The historical comparative method for determining the genetic relationship between languages compares cognate sets of words. If two words have similar semantic and phonetic content they are considered as possible cognates. Cognates being words in daughter languages which have evolved from a single word in a common parent language. Most of the proposed classifications of SW languages proceeded Li's (1977) reconstructions and cognate lists for Proto-SWT sounds. A comparison of sound correspondences utilizing cognates from Li's Proto-SWT sounds for the various languages of the SW branch may lead to further insight into the languages' genetic relationships.

Purpose

The purposes of this work is to determine criteria for the sub-classification of PG† languages and to suggest a sub-classification of these languages based on those criteria.

Hypothesis

Further subdivision of the PG languages of the SW Branch of the Tai family of languages is possible through application of the comparative method to cognate sets drawn from the languages' lexicons.

---

† Unless otherwise indicated PG will refer to the PG languages which fall within SWT.

Scope

Approximately 1000 PT cognate sound correspondence sets compiled from the following sources: Tai Mao (author's field notes), Tai Khamti (Harris 1976, Weidert 1977), Tai Lü (Yu Tsui Nung 1979, Williams 1986), Tai Nüa (Yu Tsui Nung 1979, Gedney 1976), Tai Yai (Poo-Israkij 1985), Tai Yuan (Rungruengsri 1991, Bunphan 1980), Tai Khün (Petsuk 1978), Black Tai (Fippinger 1989, Gedney 1964) White Tai (Donaldson 1961,1963, Gedney 1964), and Red Tai (Gedney 1964).

Methodology and Procedures

The methodology for this study will include four steps: data collection, compiling of sound correspondences, analysis of sound correspondences and interpretation of sound correspondences. Language data will be gathered from the cited published sources and through elicitation from Tai Mao speakers residing in Bangkok.

1. Data Collection

From the published sources and my own field notes comparisons are made of words with similar meaning to find sets of cognate words. This process involves considering the similarity of the initial consonants, final consonants, tones and vowels. Tone comparisons are based on the proto tone categories rather than phonetic shape. This will be discussed in the section: Presentation of Data. Words which have similar or related meanings in the modern languages and which show regular sound

correspondences are considered Tai cognates to be used in comparison of PG languages.

2. Presentation of Data

Part of the process of comparing cognate words from various lexicons requires a regularization of the conventions used by the various authors in the presentation of their data. Each author has presented his own data in such a way as to best display the phonetic and phonological aspects of the language which he desires to highlight. Some researchers favor the phonetic realization of a sound in their choice of phonemic symbols while others favor comparative or other criteria. In order to avoid confusion when comparing cognates from the various sources some minor changes in the use of phoneme symbols have been made.

Poo-Israkij in her representation of Tai Yai vowels is an example of a situation where an author has favored a more phonetic transcription. Poo-Israkij has written phonetically long vowels in Tai Yai with the IPA symbol for length: /i:/, /e:/, etc. This is in spite of the fact that none of the vowels in Tai Yai except /a/ and /a:/ have a phonemic length contrast. Poo-Israkij's reasoning is that vowels which are not contrastive for length are closer in phonetic length to /a:/ than to /a/, therefore she writes all vowels which are not contrastive for length with the length symbol. Certainly Poo-Israkij's choice of symbol does not affect her analysis of Tai Yai but for our purposes in comparing sound changes among different Tai PG languages it would be unreasonably burdensome to require the reader to remember the idiosyncratic motivation for each author's set of phonemic symbols. Poo-Israkij's choice of the representation of vowel length and the resultant regularization is

language specific but there are other examples of difference in the choice of phonemic symbols which apply more widely such as the representation of syllable final semi-vowels.

The phonologies of all of PG languages contain syllable final semi-vowels /y/ and /w/. Some authors, such as Li (1977) and Yu Tsui Nung (1979), represent these as /-i/, and /-u/ whereas others such as Gedney (1964) and Harris (1976), represent these as /-y/ and /-w/. In this thesis we will consistently represent these two syllable final semi-vowels as consonants with the symbols /-y/ and /-w/. As will be discussed in Chapter II this decision to consider /-i/ and /-u/ as final consonants rather than vowels reduces the number of language diphthongs and eliminates triphthongs producing a more economical phoneme system.

The phonologies of several of the PG languages also have a syllable final semivowel /-y̶/ which is represented as /-y̶/ or /-ɯ/ by different researchers. Although unlike the syllable final /-y/ and /-w/, which occur with many vowels including phonemically long vowels and diphthongs, the distribution of /-y̶/ is extremely limited. /-y̶/ only occurs following the vowel /a/. Whether to segment [ay̶] which is phonetically a vowel diphthong as a single vowel phoneme, or as a vowel plus a final consonant is not clear. The diphthong could be considered a short /a/ plus an additional final consonant /-y̶/, but the /-y̶/ would then have a distribution unlike any other consonant in that it can only occur syllable final and only after a short /a/. On the other hand, if [ay̶] is considered to be a vowel, /aɯ/, then its distribution is unlike any other vowel in that it cannot be followed by a consonant. To consider the phoneme /aɯ/ to be a single vowel phoneme (as we have done with /ia/, /ɯa/ and /ua/) seems to be the better solution as it requires only one distribution restriction rather

than the two distribution restrictions required for the final consonant solution. In this thesis the vowel diphthong will be represented by the symbol /aɯ/ and is included in the discussion of the languages' vowel inventories.

The palatal semi-vowel /y/ is represented by different researchers with the symbol /y/ or /j/. In this thesis the symbol /y/ will be used in both the syllable initial and syllable final positions.

The palatal unaspirated affricate, /c/, is represented by the symbols /c/, /ts/, and /tɕ/. In this thesis the symbol /c/ will be used, although this choice of symbol is not meant to imply that there are no differences in the tongue position used by various dialects in the pronunciation of this phoneme.

It should be noted at this point that Poo-Israkij (1985) uses both the symbols /t͡s/ (with the ligature) and the symbol /c/. Poo-Israkij's /ts/ corresponds with /s/ in all other writings on PG languages. Unfortunately Poo-Israkij gives neither a phonetic description of her two "affricate" phonemes /t͡s/ and /c/ nor an explanation for her choice of symbols. Therefore, in this thesis, since her /t͡s/ corresponds to /s/ in other Tai dialects and in order to avoid confusion with /ts/, and for convenience in typing (the author cannot type the symbol /t͡s/ with the ligature), Poo-Israkij's /t͡s/ will be replaced with /s/. The choice of symbol here is not meant to imply that there is not a difference in the pronunciation of Poo-Israkij's /t͡s/ and the /s/ phoneme in other PG languages.

Different researchers have different methods of representing tone. In order to present the data in this thesis which focuses on comparative matters with a single system of tone representation, the phonemic tone systems of the various authors have been exchanged for a

system based on the PT tone boxes created following the methodology of determining tones outlined by Gedney (1972), a methodology based on tonal development in Tai languages. (Tonal development in SWT is discussed in more detail in Chapter II.)

Using Gedney's methodology a tone correspondence figure is created for each language. This figure contains the 20 different tone boxes for PT cognate words and is shown below as Table 1. In Table 1 each of the PT Tones; A, B, C, DL, and DS are subdivided into 4 tone boxes based on PT initial consonant classes.

Table 1 PT Tone Boxes

	PT Tones				
PT Initial Consonants	A	B	C	DL	DS
1. Voiceless friction sounds *s *hm, *ph, etc.	A1	B1	C1	DL1	DS1
2. Voiceless unapirated stops *p, *t, etc.	A2	B2	C2	DL2	DS2
3. Glottal sounds *ʔ, *ʔb, etc.	A3	B3	C3	DL3	DS3
4. Voiced sounds *b, *m, *l, *z, etc.	A4	B4	C4	DL4	DS4

For practical purposes as a discovery procedure Gedney (1972) provides a list of common Tai cognates for each of these 20 PT tone boxes. By eliciting the items for each proto tone box from native speakers the tonal inventory of a language may be determined as well as the

relationship of the modern tones to the classes of PT initial consonants. Once the relationship between the modern tones and the initial consonant classes has been determined it is usually a relatively straightforward procedure to reconstruct to which proto tone box a word in the modern language lexicon belongs. Words may then be labeled according to their historical position in the PT tone system rather than synchronically with phonemic tones. For instance, in a PG language, LIVE syllables with an initial aspirated stop must belong to either proto tone box A1, B1 or C1. If a syllable's tone matches with the tone of other words from Gedney's list for the B1 tone box, this word is specifically a B1 word. The syllable's tone may then accordingly be labeled as B1. There are of course exceptions and irregularities, especially between the Northern branch and the other Tai languages, but these are beyond the scope of this thesis.

Synchronically speaking, in Tai Mao if a word is a LIVE syllable and begins with an /m/ it could belong to any one of nine tone boxes namely A1, A3, A4, B1, B3, B4, C1, C3 and C4. These nine tone boxes represent all of the six Tai Mao tones. So in this case whichever tone the word has, it is labeled correspondingly with the appropriate tone box. That almost works except that there are 9 possible boxes and only 6 tones. For Tai Mao an ambiguity exists because the A3 and B4 tone boxes, the B1 and B3 tone boxes and the C1 and C3 tone boxes have the same tones. Words from each of the three pairs of the PT tone boxes have the same tone in the modern language. If we were considering Tai Mao in isolation, a word which had an initial /m/ and a mid level tone could be labeled as belonging to either tone box A3 or B4, a word which had an initial /m/ and a low level tone could be labeled as belonging to either tone box B1 or B3, and so forth. In most cases though, these types of ambiguities can be resolved by

comparisons with other Tai languages which have not produced parallel coalescences in their tone development.

Throughout this thesis we will label tones based on these 20 tone boxes which we refer to respectively as A1, A2, A3, A4, B1 ... DS4. The phonemic representations for each of these proto tone box labels are provided in the descriptions of individual languages in Chapter III.

3. Comparison of Sound Correspondences

A comparative table of cognate sets of words in PG languages has been generated based upon sound correspondences and semantic similarity. Words having the same or similar meanings and related phonological structure are considered as cognates. When one element may seem too dissimilar for the words to be cognates, other sources such as Li (1977) have been consulted. Such is the case with the word for 'flower' which in Tai Mao is /mɔk DL3/ while in Tai Yuan it is /dɔ:k DL3/. Normally one would not expect an initial /m/ to correspond with an initial /d/ although when considering the entire word, other similar pairs of words in the languages' lexicon and the correspondences with other PG languages, it becomes obvious that /m/ in Tai Mao and /d/ in Tai Yuan correspond in a certain set of PT words. Li reconstructs the PT initial in this case as *ʔbl/r. A reconstruction further supported by the SWT language Ahom (Barua 1964) which preserves the cluster in /blak/: 'flower', and in the NT language Saek (Gedney 1993) which has similarly preserved the cluster in /blɔ:k/: 'flower'.

4. Analyses of Sound Correspondences

Analyses of sound correspondences are used to postulate genetic relationships among PG languages. Languages which share sound changes show the potential for forming a seperate genetic subdivision of PG languages, especially when several unique sound correspondences coincide.

Significance of Research

The product of this research will be a genetic classification of PG languages, and a summary of the sound changes which support this subclassification of SWT languages. Furthermore some insights into the history, inter-relationships and migrations of the Tai peoples should be possible based on the genetic relationships of the languages. As Gedney has put it: "Linguists are blessed above students of other aspects of earlier culture in being able to utilize a wonderful general principle that sound change is always regular" (Gedney 1991: 193). Another way of putting this is that language change does not lie, if two groups of people are proposed as having a common history then their languages will contain the evidences of that common history.