Lao, Thai, and their Next of Kin: The Peninsular Branch of Southwestern Tai

Pittayawat Pittayaporn[1]

Abstract—This article reevaluates the subgrouping of Southwestern Tai languages using early phonological innovations reflected in Old Thai and Old Lao inscriptions. By analyzing four shared consonantal changes, simplification of *kʰr-, coalescence of *ɓl-, aspiration of *ʰr- and occlusivization of *x-, the research results propose a "Peninsular Southwestern Tai" subgroup encompassing Thai, Lao, Southern Thai, and most of the so-called "PH" varieties. This proposal offers a more empirical analysis of the genealogical relationships among these languages and suggests a southward spread of Peninsular Southwestern Tai from northern Laos to southern Thailand.

Keywords: Old Thai and Old Lao Inscriptions; Peninsular Southwestern Tai Subgroup; Phonological Innovations; Southwestern Tai Languages

Introduction

In the absence of sufficient textual evidence to trace the spread of ethnolinguistic groups, language genealogy offers valuable insights. Historical linguistics, when classifiying languages belonging to the same family, uses the concept of a "subgroup" to describe a set of languages that share a common parent language. These languages are thought to have branched off from a common intermediate language (Campbell 1998: 166–167; Harrison 2003).

Southwestern Tai (SWT) languages, a subset of the broader Tai languages of the Kra–Dai language family, form one such subgroup. This subgroup includes Thai (also known as Siamese) and Lao, the national and modern languages of the Kingdom of Thailand and Lao PDR, and several related languages spoken in the region (Li 1960; Strecker & Pittayawat 2018). Although historical records of SWT languages appear only after the 14th century, linguistic evidence suggests an earlier presence in the Indochinese Peninsula. In a previous study, I proposed that these languages began spreading from southern China into mainland Southeast Asia no later than the 10th century (Pittayawat 2014). However, the exact relationship among these languages and the routes of their expansion remain unclear.

Several scholars, including Marvin Brown (1965), James Chamberlain (1975), John Hartmann (1980), and Nanna Jonsson (1991: 148–151) have suggested grouping Thai, Lao, Southern Thai, and lesser-known languages of Thailand and Laos such as Phu Thai and Phuan together within the SWT subgroup. They base this classification primarily

[1] Chulalongkorn University's Center of Excellence in Southeast Asian Linguistics, Bangkok.
Email: pittayawat.p@chula.ac.th.

https://doi.org/10.69486/113.1.2025.2

on a shared sound change: Proto-Southwestern Tai (PSWT) voiced stops became voiceless aspirated stops, i.e., *b- *d-, *ɟ- and *g- changing to /pʰ-/, /tʰ-/, /cʰ-/, and /kʰ-/, respectively. However, recent research by Shinnakrit (2020) shows that this sound change was completed only in the 15th century, supporting the earlier view espoused by Burnay & Cœdès (1927–28), and Gedney (1991: 208), among others. William Gedney, in particular, argues that such a recent change is not reliable for determining deeper historical relationships between languages. Consequently, the validity of Chamberlain's (1975) "PH" group, i.e., SWT languages in which the voiced stops became voiceless aspirated, is now in question.

Another set of criteria for classifying SWT languages involves analyzing their tonal systems. Many scholars have followed Chamberlain's (1975) approach, using tonal criteria to group Lao, Thai, Southern Thai, and other related varieties such as Tak Bai, Khorat, Nyo, Phuan, and Phu Thai into a major subgroup within SWT. However, I have warned against relying heavily on tonal developments for language classification (Pittayawat 2556). This caution is well founded, as disagreements persist about the relationships between Lao and Thai. Brown (1965; 1966) argues that Lao and Thai are closely related to each other, while Chamberlain (1975; 1991) claims that they are more distantly related as Lao is closer to Southern Thai.

This debate highlights the complexity of tracing language relationships and the importance of empirical evidence in historical linguistics. These issues shed light on the history of Tai-speaking populations and the spread of these languages in Southeast Asia. Chamberlain's (1975) classification of Lao and Southern Thai as one subgroup and Thai and SWT varieties of northeast Laos and northwest Vietnam as another raises geographical concerns. The separation of Lao and Southern Thai by Thai in the middle and the separation of Thai and its supposed relatives in northeast Laos by Lao requires significant historical speculation to make sense. Although not impossible, the proposed scenario is problematic and lacks clear evidence (see below).

Therefore, this article aims to reassess the genealogical relationship among the so-called PH varieties, focusing on Lao and Thai, to determine whether they genuinely form a valid subgroup. My findings suggest that Lao, Thai, and many other PH languages indeed belong to the same subgroup, which can be labeled "Peninsular SWT". This classification is based on shared early phonological innovations reflected in Lao and Thai inscriptions from the 14th to early 15th centuries. The proposed subgroup structure suggests that the precursor of Lao, Thai, and its closest relatives spread southward from northern Laos to southern Thailand.

Assessment of Previous SWT Subgrouping Proposals

The generally accepted method for subgrouping languages in historical linguistics involves analyzing shared innovations. These are sound changes shared by a group of daughter languages, distinguishing them from other related languages (Campbell 1998: 70; Harrison 2003).

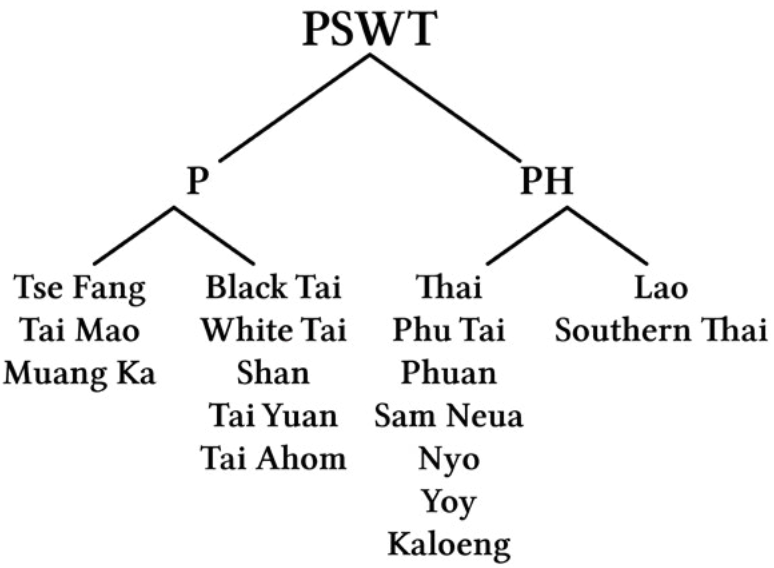


Figure 1: Chamberlain's subgrouping of SWT languages
© Chamberlain 1975: 50 (Adapted)

For instance, Proto-Tai (PT), the hypothetical ancestor of all Tai languages, had a *-ɤn sound that later became *-on in PSWT. The PT word *sɤn^{C} 'heel' evolved into *son^{C} (ส้น, son) in PSWT, a change still evident in all SWT languages. In Shan and Black Tai, words with PT *-ɤn became /-on/, whereas in White Tai and Lue, they further developed into /-un/ (Pittayawat 2556). While *-ɤn > *-on is a shared innovation among SWT languages, subsequent changes of /-on/ to /-un/ represent either parallel innovations or cases of diffusion, making them unreliable for subgrouping purposes.

For SWT languages, most studies primarily use patterns of splits and mergers of PSWT *A, *B, *C and *D tones as the classification criteria. This approach is exemplified by the widely cited subgroup structure proposed by Chamberlain (1975), as shown in **Figure 1**. Other notable studies following this method include those by Hartmann (1980), and Pranee & Theraphan (2000), among many others.

A family tree is an effective graphical representation of the genealogical relationships among languages. Each branch of the tree represents a subgroup, generally understood as a cluster of languages descended from an immediate common parent language, which, in turn, is derived from an earlier proto-language. Chamberlain's (1975: 62–63) proposed family tree for SWT presents a chronological ordering of innovations that is inconsistent with known phonological history. A critical issue in his model is the assumption that the devoicing of PSWT voiced stops represents the oldest innovation within the family. This assumption leads to a problematic classification system based on the reflexes of these stops.

In Chamberlain's model, SWT varieties are divided into two main subgroups, based on the outcome of the devoicing process. The "PH" group includes varieties that show voiceless aspirated reflexes of PSWT voiced stops, for example, where PSWT *b- evolved into /pʰ-/. The "P" group comprises varieties where PSWT voiced stops developed into unaspirated counterparts, such as PSWT *b- becoming /p-/. This subgroup structure incorrectly assumes that the P/PH distinction represents the primary and earliest split in the SWT family when, in fact, other phonological innovations likely preceded this change.

Multiple lines of evidence indicate that devoicing occurred relatively late. The Thai and Lao writing systems likely developed when the voicing distinction was still intact (Li 1977: 22–23; Ferlus 1999). Indic loanwords with voiced initials were borrowed before devoicing, resulting in /pʰ/, /tʰ/, /cʰ/, and /kʰ/ in Thai for Indic *b*, *d*, *j*, and *g* (Gedney 1965). Thai name transcriptions in the *Veritable Record of the Ming Dynasty* (明實錄, *Ming Shilu*; 1368–1644) suggest the devoicing was completed in the mid-14th century (Shinnakrit 2020). Additionally, some Tay and Zhuang dialects, close relatives of SWT languages, maintain phonemic voiced stops (Theraphan 1997; Pittayawat & Kirby 2017).

Furthermore, Chamberlain (1975) employs patterns of tonal development as criteria for subgrouping. However, it is crucial to recognize that these patterns result from a series of mergers and splits, rather than single events. For example, the B≠D pattern used by Chamberlain to group Lao and Southern Thai results from separate sets of three distinct tonal changes in each language. While Lao arrived at the pattern through the merger of C23 and C4, merger of C1 and DL123 and merger of C234 and DL4, Southern Thai varieties underwent merger of C1 and DL1, optional merger of C23 and DL23, and optional merger of DL23 and DL4. Similarly, the B=D pattern is not a merger but a retention from the proto-language (Gedney 1989; Pittayawat 2009a: 239–254). Therefore, previous SWT subgrouping proposals are problematic, as they equate apparent patterns of mergers and splits with shared innovations.

Furthermore, although tonal developments can inform genealogical classifications (Dockum 2019: 74–111), they are more suitable for determining shallower-level subgroups. To illustrate the inadequacy of subgrouping based on patterns of tonal splits and mergers, consider Chamberlain's model in **Figure 1**, which groups Tai Mao with the Shan dialects of Tse Fang and Muang Ka within one subgroup, while placing Southern Shan in another. This arrangement presents a significant challenge, as research conducted by Edmondson & Solnit (1997) establishes these dialects as closely related and coherent subgroups. Moreover, Chamberlain's model leads to incongruities, as seen in the case of the Lue dialects of Chiang Rung and Mueang Yong. These dialects would be placed in different branches due to the distinct splits of PSWT *A tones, despite their close relationship (Pittayawat 2556).

Specific to the relationships among Lao, Thai, and other PH languages are two key tonal patterns: the connection between *B and *D patterns and the split

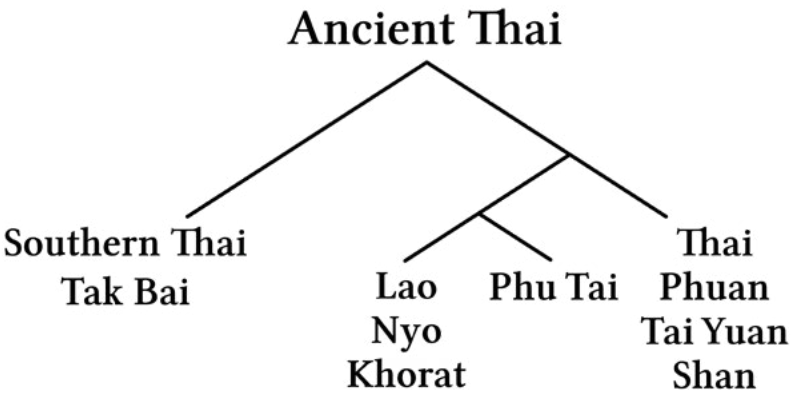


Figure 2: Brown's subgrouping of SWT languages © Brown 1965: 66–69 (Adapted)

of *B, *C, and *D tones. In this model, Lao and Southern Thai are subgrouped to the exclusion of Thai, Phu Thai, and Phuan. This classification is based on Lao and Southern Thai exhibiting *B≠*D and *BCD1-23-4 patterns, while the others share *B=*D and *BCD123-4 patterns. However, this model implies problematic geographical and historical scenarios. This implies that Thai separates Lao and Southern Thai, while Lao sits between Thai and its presumed close relatives, Phuan and Phu Thai. This arrangement presents a significant challenge for our understanding of the historical spread and development of these languages.

To address these geographical incongruities, Chamberlain (1972: 239; 1991: 471–474) proposes a speculative historical scenario. He suggests that Southern Thai, which he claims to be a descendant of the Sukhothai language, should be considered a remnant of the westerly expansion from northern Laos. Thai, claimed to have descended from the language of Ayutthaya, is proposed to have arrived in present-day central Thailand from northeast Laos around Xiang Khoang (ຊຽງຂວາງ). Finally, Chamberlain posits that the Lao arrived in northeast Thailand from northern Laos a few centuries later, effectively breaking the Thai–Phuan–Phu Thai continuum. However, given the problems associated with using such tonal criteria for subgrouping, the bifurcation of the PH subgroup must be abandoned.

Brown's (1965) influential subgrouping hypothesis, which predates and likely influenced Chamberlain's work, offers an alternative subgroup structure for SWT. As illustrated in **Figure 2**, Brown's model proposes that Southern Thai and Tak Bai, viewed as descendants of Sukhothai, branched off earliest from other languages. The remaining PH languages, along with P languages such as Shan and Tai Yuan, form a separate branch. Within this branch, Brown suggests a different arrangement from that of Chamberlain. Lao and Phu Thai constitute one subgroup, while Thai, Phuan, Shan, and Tai Yuan form another.

Although Brown considers consonantal and vocalic changes in his classification, he appears to implicitly prioritize tonal development. However, this approach has significant weaknesses.

The most problematic aspect is his lack of justification for the selection and chronological ordering of innovations. For example, he (1965: 114) suggests that the change *ɣ- > g- was an early innovation that separated Southern Thai from other languages. Yet, inscriptional evidence shows that *ɣ persisted in Thai after the 13th-century development of the Thai script (Li 1977: 214–219; Ferlus 1988).

Brown's model also faces chronological challenges. For the stage he calls "950 Yunnan", ancestral to all varieties in his study except Southern Thai and Tak Bai, he hypothesizes two consonantal and six tonal shared innovations. The tonal changes are particularly noteworthy, including both phonetic alterations and phonemic splits and mergers. For instance, he proposes that this hypothetical language underwent the *ABCD > *ABCD1-23-4 split and changed the B4 tone to a falling tone. This model's reliance on tonal changes for deep-level subgrouping and its chronological assumptions cast doubts on its validity.

It is important to note that Brown (1966), followed by Chamberlain (1972; 1991), posits that the language of Sukhothai and the language of Ayutthaya were distinct, rather than representing a single linguistic continuum. Brown's (1966; 1985) interpretation of tone marks in Sukhothai inscriptions is particularly noteworthy. He views these symbols as reflecting the sound system of an earlier stage of Southern Thai, which he believes had three lexical tones. Furthermore, Brown speculates that the modern Thai orthographic practice of classifying consonant symbols into "high", "mid", and "low" categories was devised to correspond with a presumed three-way consonant distinction in the Sukhothai period. He uses this presumed correspondence as evidence for a particularly close relationship between the Sukhothai-period Thai and Southern Thai. However, this hypothesis is problematic despite having gained some popularity among researchers of Tai languages (e.g., Chanida 1991; Dueanpen 2541). As noted by Sireemas & Pittayawat (2019: 32–33), no substantial evidence has been put forward to support the proposed close relationship between Sukhothai Thai and Southern Thai. This lack of concrete evidence casts doubt on the scenario that the Sukhothai and Ayutthaya inscriptions represent two relatively different languages. This critical reassessment of Brown and Chamberlain's proposals invites a more nuanced understanding of the historical development of Thai and its related languages, emphasizing the need for rigorous evidence-based approaches to linguistic historiography.

In summary, tonal splits and mergers, while valuable for studying Tai dialects, are unreliable for subgrouping SWT varieties because of their recent nature and the tendency to cross subgroup boundaries. To accurately determine the genealogical relationships among SWT languages, it is essential to identify early segmental changes that can be verified empirically. This approach will provide a more robust foundation for understanding the historical development and the relationships of SWT languages, avoiding the pitfall of over-reliance on tonal patterns.

Table 1: List of inscriptions analyzed
(1.1-1.11: Old Thai inscriptions; 2.1-2.4: Old Lao inscriptions)

	Inscription	FAD no.	Date	Language
1.1	Wat Si Chum	สท. 2	14th–15th c.	Thai
1.2	Nakhon Chum	กพ. 1	1357	Thai
1.3	Wat Hin Tang	สท. 37	15th c.	Thai
1.4	Lan Ngoen Phasa Thai Wat Mahathat	อย. 9	15th c.	Thai
1.5	Praditsathan Phra That	กท. 59	15th c.	Thai
1.6	Phaen Dibuk Wat Mahathat	อย. 2	1374	Thai
1.7	Wat Chang Lom	สท. 18	1384	Thai
1.8	Wat Song Khop I	ชน. 13	1408	Thai
1.9	Chao Then Si Thep Khirimanon I	ชน. 5	1413	Thai
1.10	Wat Sorasak	สท. 25	1417	Thai
1.11	Wat Taphan	สฎ. 5	15th c.	(Southern) Thai
2.1	Phra That Rang Ban Rae	สน. 4	1350	Lao
2.2	Wat Luang (Nong Khai)	นค. 9	1414	Lao
2.3	Wat Daen Mueang (Wat Patchantaburi) I	นค. 1	1530	Lao
2.4	Wat Daen Mueang (Wat Patchantaburi) II	นค. 2	1535	Lao

Methodology

The primary data source for this epigraphic study is the *Inscriptions in Thailand Database* (Princess Maha Chakri Sirindhorn Anthropology Centre). In order to consider the Thai language of the early period, the analysis included both Sukhothai and Ayutthaya inscriptions, in both Khom and Thai scripts, dated before 1450. This cutoff date is significant because the devoicing of PSWT obstruents in Thai was completed in the 15th century (Shinnakrit 2020). The results demonstrate that the two corpora of Sukhothai and Ayutthaya inscriptions reflect the same set of sound changes, suggesting a relatively unified dialect continuum. To facilitate referencing, all inscriptions used in this article were assigned specific codes, as listed in **Table 1.**

It is important to note that while the Wat Taphan Inscription (1.11) may provide possible evidence for 15th-century Southern Thai, its linguistic affiliation is not definitive. The language of this inscription may be more closely related to the prestigious language of the Central Plain than the southern vernacular. This ambiguity highlights the challenges in definitively classifying historical language varieties and emphasizes the need for a cautious interpretation of inscriptional evidence.

In order to consider the Thai language of the early period, this article examines early inscriptions written in the Lao script (also known as Thai Noi) found in northeast Thailand. In addition to two pre-1450 inscriptions, the

corpus includes two items dating from the 1530s. This later cutoff date is necessary because the oldest discovered inscription written in the Lao script dates only from the end of the 15th century (Lorrillard 2018). The limitation to northeast Thai inscriptions is due to limited access to data from Laos itself.

This article adopts the reconstruction of the PSWT sound system presented by myself (Pittayawat 2009b). While this proposal largely aligns with previous reconstructions by scholars such as Sarawit (1973), Li (1977), and Jonsson (1991), it introduces several important differences in both consonant and vowel systems. The focus is specifically on consonantal changes, as the interpretation of vowel graphs presents significant challenges, a point emphasized in the work of Sireemas & Pittayawat (2019). Most notably for our purposes, I proposed a separate series of uvular consonants, including the voiceless stop *q- and the voiceless fricative *χ-. In Li's reconstruction, etyma that I reconstructed with *q- and *χ- are listed under *k^h- and *x-, respectively.

The approach of this article, which identifies early innovations in the Thai and Lao languages by examining how ancient inscriptions represent PSWT consonants (Li 1977; Pittayawat 2009b), is similar to the influential works of Diller (1988) and Duangduen & Pranee (1976). A phonemic merger was inferred when a PSWT sound was represented by a graph originally assigned to a different sound, suggesting that the two sounds were no longer distinct. The presence of shared innovations provides compelling evidence for considering these varieties to be particularly close. If Lao and Thai share a number of innovations, they can be considered as belonging to the same immediate subgroup. An illustrative example of this approach can be seen in the treatment of words with PSWT initial *χ- in modern Thai. Currently, PSWT etyma such as *χaw^C 'to enter' (เข้า, *khao*) and *χɯːn^C 'to ascend' (ขึ้น, *khuen*) are consistently spelt with the graph transliterated as *kh*, which originally represented only *k^h-. This pattern suggests that PSWT *χ- has become /k^h-/, merging with PSWT *k^h-.

The phonemic values of the graphs used in this analysis are based on Michel Ferlus's (1988) study, which is summarized in the **Appendix**. While Ferlus (1988) does not give exact phonemic values for Old Thai and Old Lao graphs, he provides their counterparts in what he calls *thaï commun* or "common Thai", which can be considered his understanding of the PSWT sound system (Ferlus 1988). This article assumes that the original phonemic values of the graphs were identical to their PSWT correspondents. The transliteration of consonant and vowel graphs follows the systems of Ferlus (1988; 1999) and Uraisi (1984; 2553)[2] because they accurately represent both the graphemic contrasts and the correspondences to the Indic and Khmer scripts. Note that while some vowel transliterations may appear complex, they can be disregarded as they are not pertinent to the present analysis.

[2] The systems are largely identical, differing primarily in their notation of modified letters that were invented to represent Thai sounds absent in Old Khmer. Ferlus employs an apostrophe (ʻ), while Uraisi uses an underline; for example, Ferlus's *ʻkh-* corresponds to Uraisi's kh-. Both systems are essentially derived from Cœdès's original transliteration scheme.

Table 2: Data sources for modern PH and P varieties

PH/P	Variety	Source
PH	Thai (Bangkok)	Hudak (2008)
	Lao (Nong Khai)	Hudak (2008)
	Southern Thai (Songkhla)	Apichaya & Pittayawat (fieldnotes)
	Tak Bai (Su-ngai Padi)	Puttachart & Thananan (2541)
	Khorat Thai (Noen Sung)	Kanitha & Pittayawat (fieldnotes)
	Nyo (Sakon Nakhon)	Thepbangon (1984)
	Yoy (Akat Amnuai)	Thepbangon (1984)
	Kaloeng (Dong Mafai)	Areeluck (1985)
	Phuan (Hat Siao)	Chalong (1986)
	Phu Thai (Waritchaphum)	Pittayawat (fieldnotes)
	Phu Thai (That Phanom)	Wilaiwan (2520)
P	Tai Yuan (Chiang Mai)	Hudak (2008)
	Shan (Möng Nai)	Hudak (2008)
	White Thai (Mường Tè)	Hudak (2008)
	Black Thai (Sơn La)	Hudak (2008)

Furthermore, these early shared innovations serve as valuable markers for identifying modern SWT varieties that belong to the same subgroup. By examining the modern reflexes of PSWT sounds, inferences can be made regarding whether specific sound changes have occurred. For instance, Southern Thai exhibits /r-/ for PSWT *r-, as in /rɯə^{5}/ < PSWT *rɯə^{A} 'boat' (เรือ, *ruea*), maintaining a distinction from /h-/, which is the regular reflex of PSWT *h-. This indicates that Southern Thai did not undergo the change *r- > h- that occurred in Lao.

By comparing the sound patterns of these modern varieties with the established early innovations, it becomes possible to reconstruct their historical relationships in the absence of direct ancient evidence. The main sources of data on the modern PH and P varieties used in this article are listed in **Table 2.** These sources form the basis for my comparative analysis of contemporary language forms, complementing historical data gleaned from ancient inscriptions.[3]

The language varieties were selected based on two primary criteria: the availability of extensive vocabulary data containing specific Tai words needed for analysis and the presence of precise phonemic transcriptions. Despite substantial research on some varieties, such as Tai Yo (e.g., Ferlus 2008) and Khorat (e.g., Wanna 1979), many were excluded due to a lack of comprehensive wordlists. For Southern Thai, while numerous studies exist on varieties spoken in Thailand and Malaysia (e.g., Diller 1979; Theeraphan 2521; Umar 2003), the

[3] The semantic glosses provided in single quotation marks (' ') represent approximate meanings of the reconstructed PSWT forms rather than the precise meanings of their reflexes in individual languages or their contextual usage in inscriptions. These standardized glosses facilitate cross-linguistic comparison of modern forms. For precise semantic interpretations in specific languages or inscriptional contexts, please consult the original sources.

absence of crucial lexical items necessitated original data collection. Notable gaps in documentation also persist for several PH languages in central and northeast Laos and northwest Vietnam, including Thai Muong Vat, Sam Neua, and Tai Et among others, despite pioneering fieldwork reports (e.g., Chamberlain 1975; 1983; Sasithorn 2553; Theraphan 2003; Warunsiri 2023).

Each language in this article encompasses multiple varieties spoken in different geographical locations. Although these varieties exhibit lexical and phonological differences, they share identical developments with respect to the early innovations attested in inscriptions (see below). Only varieties that demonstrate distinct reflexes of relevant PSWT sounds are included. For instance, while Southern Thai varieties differ in tonal patterns and reflexes of PSWT vowel and consonants such as *ml- and *ŋ- (e.g., Theraphan 2521; Umar 2003), they remain uniform regarding early innovations. The Nyo-Yoy-Kaloeng cluster (e.g., Kanjana 2524; Mudjalin 2559; Phinnarat 2000) presents notable variations in tonal mergers and splits, alongside minor divergences with regards to consonants such as the development of PSWT *f-. Similarly, Phuan varieties (e.g., Phinnarat 2010) show diversity in tonal patterns and the development of *χ-, while maintaining consistent reflexes for *x-, which merged with *k^h- in early inscriptional evidence. Particularly noteworthy are Phu Thai varieties (e.g., Phinnarat 2000), which are largely uniform except for Phu Thai Kapong spoken in Waritchaphum district, Sakon Nakhon province. This variety uniquely shows /h-/ for PSWT *x- rather than the typical /k^h-/ in other Phu Thai varieties (Pittayaporn 2009b). Two varieties with /h-/ and /k^h-/ reflexes are included to represent this diversity.

Map 1 shows the geographical distribution of PH dialects examined in this article. The Phuan, Phu Thai, Nyo, Yoy, and Kaloeng dialects are notably scattered across northeast and central Thailand, having arrived there through forced migrations in the 19th century. This study relies on data from varieties spoken in Thailand, as comprehensive wordlists containing the relevant lexical items are not available for their counterparts in Laos.

Early Innovations Shared by Lao and Thai

Four shared consonantal changes in early Lao and Thai emerge from analysis of the inscriptional data. These shared innovations offer strong evidence for grouping these two major PH varieties as a subgroup within SWT. It is noteworthy that evidence for all the identified innovations is present in the Old Thai inscriptions from both the Sukhothai and Ayutthaya periods, consistent with the assumption of a unified linguistic continuum.

The first notable innovation in both Old Lao and Old Thai is the simplification of PSWT *k^hr- to *k^h-. **Table 3** shows that all words with *k^hr- onsets in PSWT consistently appear with *kh*- (corresponding to *k^h-) in both Thai and Lao inscriptions.[4] This change had

[4] Tone marks are omitted in our transliterations because they are not relevant to the analysis and

Map 1: Location of SWT varieties discussed in this article
© Pittayawat Pittayaporn & Krittin Klaimee

Table 3: Examples of words attesting to the change PSWT *k^{h}r- > *k^{h}-

Language	Inscription	'to beg' ขอ, *kho*	'slave' ข้า, *kha*	'other people' เขา, *khao*	'to drive' ขับ, *khap*
Thai	1.1	***kh**a*	—	***kh**au*	***kh**app*
	1.2	***kh**a*	***kh**ā*	***kh**au*	—
	1.3	***kh**a*	—	***kh**au*	—
	1.4	***kh**a*	***kh**ā*	—	—
	1.6	***kh**a, **kh**ạ*	***kh**ā, **kh**ā*	—	—
	1.7	—	***kh**ā, **kh**ā*	—	—
	1.8	***kh**a, **kh**a*	***kh**ā*	—	—
	1.10	***kh**ạ, **kh**ạ*	***kh**ā*	—	—
	1.11	***kh**a*	—	—	—
Lao	2.3	—	***kh**ā*	—	—

occurred by the time these texts were inscribed, probably in the 14th to early 15th century. Importantly, this simplification occurred earlier than the changes to other clusters. For example, *kl- and *gr- were still regularly preserved in the Old Thai inscriptions, as seen in *klai* 'far' (< PSWT *klaj^{A}; ไกล) and *grva* 'household' (< PSWT *gruə^{A}; ครัว, *khrua*) in Wat Si Chum Inscription (1.1) and Wat Song Khop Inscription I (1.8), respectively. Similarly, *pl- were variably maintained graphically in Old Lao, as seen in *pluk* 'to root' (< PSWT *plu:k^{D}; ปลูก) in Wat Daen Mueang Inscription I (2.3). This early simplification of *k^{h}r-, contrasted with the retention of other clusters, marks a significant shared innovation in Lao and Thai phonological development.

The second innovation observed is the coalescence of PSWT *ɓl- into *ɗ-. As evidenced in **Table 4**, words that originated from *ɓl- are consistently spelled with the single graph given by Ferlus as t-, which represents /ɗ-/, in both Old Thai and Old Lao inscriptions. This indicates that the change had occurred by the 14th century, when the oldest texts were inscribed. Despite the limited occurrence of words with *ɓl-, this innovation is unmistakable as such words never appear with the graphs transliterated as *p-* (representing /ɓ-/) or *l-* (representing /l-/) in any Old Thai or Old Lao inscriptions.

Another shared innovation is the aspiration of PSWT *^{h}r-. **Table 5** shows that all instances of the voiceless rhotic *^{h}r- are represented as *h-* in both Old Thai and Old Lao inscriptions. This contrasts with other voiceless sonorants, such as *^{h}l- and *^{h}n-, which are attested with *h-* combined with graphs representing their corresponding voiced sonorants. For instance, in the Thai Wat Si Chum Inscription (1.1), we find ***hn**ieạ* 'upstream' (< PSWT *^{h}nuə^{A}; เหนือ, *nuea*) and ***hl**ieạṅ* 'yellow' (< PSWT *^{h}luəŋ^{A}; เหลือง, *lueang*). Similarly, in the Lao Wat Daen Mueang Inscription I (2.3), we see ***hm**aị* 'new' (< PSWT *^{h}mauᵂ; ใหม่,

Ferlus (1988) does not provide guidance on how to transliterate them.

Table 4: Examples of words attesting to the change PSWT *ɓl- > *ɗ-

Language	Inscription	'moon' เดือน, *duean*	'flower' ดอก, *dok*
Thai	1.1	***t**ean, **t**ian*	***t**ạk*
	1.2	***t**i̯ean, **t**ian*	—
	1.3	***t**ian*	***t**ạk*
	1.7	***t**i̯ean, **t**ian*	—
	1.9	***t**ian*	—
	1.10	***t**ian*	***t**ạk*
	1.11	***t**ein, **t**i̯ean*	***t**ạk*
Lao	2.1	***t**ein*	—
	2.2	***t**ian*	—
	2.3	***t**ein*	—

Table 5: Examples of words attesting to the change PSWT *^{h}r - > *h-

Language	Inscription	'six' หก, *hok*	'head' หัว, *hua*	'stone' หิน, *hin*	'five' ห้า, *ha*	'stream' ห้วย, *huai*
Thai	1.2	***h**ak*	—	—	—	—
	1.3	—	—	***h**in*	—	—
	1.4	***h**ak*	—	—	—	—
	1.5	—	***h**vă*	—	—	—
	1.7	***h**ak*	—	***h**in*	—	—
	1.9	***h**ak*	—	—	—	—
	1.10	***h**ak*	***h**vvă, **h**vav*	—	—	—
	1.11	***h**ak*	—	—	—	—
Lao	2.2	—	—	—	***h**ā*	***h**vay*
	2.3	—	—	—	***h**ā*	***h**vay*

Table 6: Examples of words attesting to the change PSWT *x- into *k^{h}-

Language	Inscription	'white' ขาว, *khao*	'green' เขียว, *khiao*
Thai	1.1	*khāv*	*khyav*
	1.5	*khāv*	—
	1.7	—	*khyav*
	1.8	*khāv*	—
	1.11	*khāv*	—

mai), and ***hl****vaṅ* 'big' (< PSWT *^{h}luəŋ^{A}, หลวง, *luang*).

These examples demonstrate the retention of conservative representations for *^{h}l-, *^{h}n-, and *^{h}m-, contrasting with the simplification of *^{h}r- to /h-/ in both languages. This pattern of differential treatment between *^{h}r- and other voiceless sonorants provides strong evidence for a shared innovation in the early stages of Thai and Lao. This innovation is of particular significance since it is not found in the early stages of Tai Yuan (Bauer 2015; Shinnakrit 2021; 2023) and Shan/Tai Nüa (Shinnakrit & Buragohain 2024).

The final innovation found in Thai inscriptions is the occlusivization of PSWT *x- into *k^{h}-. In Thai, words that originally had *x- in the proto-language were primarily written with *kh-* as onsets, as shown in **Table 6**. While it has been established that the graph *'kh-* was created by modifying *kh-* to represent a voiceless velar fricative /x-/ (Li 1977: 207–214; Diller 1988), words spelled with this graph in Thai inscriptions all go back to either the voiceless uvular stop *q- or the voiceless uvular fricative *χ- in PSWT (Pittayawat 2009b: 123–124). This strongly suggests that the merger of *x- with *k^{h}- had already occurred before the merger between *q- and *χ- into /x-/ filled the gaps left behind. The change *x- > *k^{h}- must have taken place very early in the Thai language, as evidenced by the Old Thai inscriptions that reflect them. The words 'white' (ขาว, *khao*) and 'green' (เขียว, *khiao*) share the reconstructed initial consonant *x- in PSWT.

This innovation is also particularly significant since it is not found in the earlier stages of Tai Yuan (Shinnakrit 2021; 2023) and Shan/Tai Nüa (Shinnakrit & Buragohain 2024). Unfortunately, it is not possible to ascertain whether these changes also occurred in Lao, as the graph *'kh-* does not exist in the available Old Lao inscriptions, perhaps because of the inscriptions' significantly younger age. It is also important to distinguish it from the second round of occlusivization that occurred in the 14th–15th century. This subsequent change eventually led to the disuse of the *'kh-* graph, as discussed by scholars such as Diller (1988).

Peninsular SWT

This section demonstrates how early shared innovations reflected in Old Thai and Old Lao inscriptions provide compelling evidence for grouping not only Thai, Lao, and Southern Thai together, but also most PH varieties. These innovations are also found in lesser-known varieties that have no pre-modern written records such as Southern Thai, Tak Bai, Khorat, Nyo, Yoy, Kaloeng, and Phuan. However, Phu Thai appears to be more distantly related, as it possibly lacks one of the shared innovations. The proposed name for the subgroup is "Peninsular SWT" or PSWT, reflecting its geographical distribution across much of the Indochinese peninsula.

The most readily discernible shared innovation among PH varieties is the coalescence of PSWT *ɓl-. This phonological change is evident across all PH varieties examined in this article, including Southern Thai, Tak Bai, Khorat, Phuan, Phu Thai, Nyo, Yoy,

Table 7: Reflexes of PSWT *ɓl- in PH varieties

Language	'moon' เดือน, *duean*	'flower' ดอก, *dok*
Thai	duən^{1}	dɔ:k^{2}
Lao	duən^{2}	dɔ:k^{2}
Southern Thai	duən^{3}	dɔ:k^{4}
Tak Bai	duən^{1}	dɔ:k^{3}
Khorat Thai	diən^{5}	dɔ:k^{2}
Nyo	duən^{1}	dɔ:k^{3}
Yoy	duən^{1}	dɔ:k^{3}
Kaloeng	duən^{5}	dɔ:k^{2}
Phuan	duən^{1}	dɔ:ʔ^{2}
Phu Thai (Waritchaphum)	dɤ:n^{3}	dɔʔ^{5}
Phu Thai (That Phanom)	dɤ:n^{3}	dɔʔ^{5}

and Kaloeng. As illustrated in **Table 7**, words that originally had initial *ɓl- in PSWT now consistently exhibit /d-/ as onsets in these varieties. This reflex suggests that *ɓl- merged with *ɗ-, mirroring the change observed in Lao and Thai inscriptions.

Notably, this innovation extends beyond PH varieties. Some P varieties, such as Yuan, Khuen, and Lue also display reflexes identical to *ɗ-. For instance, Yuan /**d**uan^{1}/ exemplifies this shared development. This broader distribution suggests that while the coalescence of PSWT *ɓl- provides strong evidence for grouping PH varieties together, the resulting subgroup may be more extensive than initially thought. It potentially encompasses not only Peninsular SWT but also some P varieties. This finding underscores the complex interrelationships within SWT and highlights the need for a nuanced approach to subgrouping on the basis of early shared innovations.

The simplification of PSWT *k^{h}r- is less straightforward. While all PH varieties display a simple /k^{h}-/ as a reflex of the velar cluster as illustrated in **Table 8**, Lao and Phu Thai, Phuan, Nyo, Yoy and Kaloeng have simplified all initial clusters. It is unclear whether *k^{h}r- was simplified earlier, with the other clusters simplified later, or whether they were all simplified simultaneously. If the former is true, the simplification of *k^{h}r- would support grouping Phu Thai, Phuan, Nyo, Yoy and Kaloeng with Lao, Thai, Khorat, and Tak Bai.

Similarly unclear is the interpretation of data regarding the aspiration of PSWT *^{h}r-. As illustrated in **Table 9**, Thai, Southern Thai, Khorat, and Tak Bai display /h-/ as the reflex of the voiceless rhotic while retaining *r-, whereas Lao and other nearby languages exhibit /h-/ for both PSWT *^{h}r- and *r-. Although the Phuan variety spoken in Suphanburi lacks direct reflexes for the particular words in the following table,

Table 8: Reflexes of PSWT *kʰr- in PH varieties

Language	'to beg' ขอ, *kho*	'slave' ขา, *kha*	'other people' เขา, *khao*	'to drive' ขับ, *khap*	'egg' ไข่, *khai*
Thai	**k**ʰɔː⁵	**k**ʰaː³	**k**ʰaw⁵	**k**ʰap²	**k**ʰaj²
Lao	**k**ʰɔː¹	**k**ʰaː⁴	—	**k**ʰap¹	**k**ʰaj³
Southern Thai	**k**ʰɔː¹	**k**ʰaː²	**k**ʰaw⁷	—	**k**ʰaj¹
Tak Bai	**k**ʰɔː¹	—	**k**ʰaw³	—	**k**ʰaːj³
Khorat Thai	**k**ʰɔː⁵	—	—	—	**k**ʰaj²
Nyo	**k**ʰɔː¹	—	**k**ʰaw¹	**k**ʰap⁴	**k**ʰaj³
Yoy	**k**ʰɔː¹	—	**k**ʰaw¹	**k**ʰap⁴	**k**ʰaj³
Kaloeng	**k**ʰɔː⁵	—	—	—	**k**ʰaj²
Phuan	**k**ʰɔː⁵	**k**ʰaː⁶	—	**k**ʰap⁵	**k**ʰaj²
Phu Thai (Waritchaphum)	**k**ʰɔː¹	—	—	**k**ʰap⁴	**k**ʰaj³
Phu Thai (That Phanom)	**k**ʰɔː¹	**k**ʰaː⁵	—	—	**k**ʰaj³

Table 9: Reflexes of PSWT *ʰr- in PH varieties

Language	'six' หก, *hok*	'head' หัว, *hua*	'ear' หู, *hu*	'stone' หิน, *hin*
Thai	**h**ok²	**h**uə⁵	**h**uː⁵	**h**in⁵
Lao	**h**ok¹	**h**uə¹	**h**uː¹	**h**in¹
Southern Thai	**h**ɔk¹	**h**uə¹	**h**uː¹	**h**in¹
Tak Bai	**h**ok¹	**h**uə¹	**h**uː¹	**h**iːn¹
Khorat Thai	**h**ok⁴	—	—	**h**in⁵
Nyo	**h**ok⁴	**h**uə¹	**h**uː¹	**h**in¹
Yoy	**h**ok⁴	**h**uə¹	**h**uː¹	**h**in¹
Kaloeng	**h**ok⁵	**h**uə⁵	**h**uː⁵	**h**iːn⁵
Phuan	**h**ok⁵	**h**uə⁵	**h**uː⁵	**h**in⁵
Phu Thai (Waritchaphum)	**h**ok⁵	**h**oː¹	**h**uː¹	**h**in¹
Phu Thai (That Phanom)	**h**ok⁵	**h**oː¹	**h**uː¹	**h**in¹

other words that had *ʰr- in PSWT consistently show /h-/ as expected. Examples include /hiw⁶/ 'to carry' (หิ้ว, *hiu*), /huːʔ²/ 'loom' (หูก, *huk*), and /hɔːʔ²/ (หอก, *hok*) 'spear'. Hence, it remains unclear whether *ʰr- in the latter set of dialects underwent a direct change to /h-/ early on or if it first changed to /r-/ before becoming /h-/ at a later stage.

While inscriptional data suggests the latter for Lao, epigraphical evidence is not available for the other languages. Therefore, similar to the uncertainty

Table 10: Reflexes of PSWT *x- in PH varieties

Language	'white' ขาว, *khao*	'green' เขียว, *khiao*	'galangal' ข่า, *kha*	'hook' ขอ, *kho*
Thai	**k**^{h}a:w^{5}	**k**^{h}iəw^{5}	**k**^{h}a:^{2}	**k**^{h}ɔ:^{5}
Lao	**k**^{h}a:w^{1}	**k**^{h}iəw^{1}	**k**^{h}a:^{3}	**k**^{h}ɔ:^{1}
Southern Thai	**k**^{h}a:w^{1}	**k**^{h}iəw^{1}	-	**k**^{h}ɔ:^{1}
Tak Bai	**k**^{h}a:w^{1}	**k**^{h}iəw^{1}	**k**^{h}a:^{3}	—
Khorat Thai	**k**^{h}a:w^{5}	**k**^{h}iəw^{5}	—	—
Nyo	**k**^{h}a:w^{1}	**k**^{h}iəw^{1}	**k**^{h}a:^{3}	—
Yoy	**k**^{h}a:w^{1}	**k**^{h}iəw^{1}	**k**^{h}a:^{3}	—
Kaloeng	**k**^{h}a:w^{5}	**k**^{h}iəw^{5}	**k**^{h}a:^{2}	—
Phuan	**k**^{h}a:w^{5}	**k**^{h}iəw^{5}	**k**^{h}a:^{2}	**k**^{h}ɔ:^{5}
Phu Thai (Waritchaphum)	**k**^{h}a:w^{1}	**k**^{h}e:w^{1}	**k**^{h}a:^{3}	**k**^{h}ɔ:^{1}
Phu Thai (That Phanom)	—	**h**e:w^{1}	**h**a:^{3}	—

surrounding the simplification of *k^{h}r- clusters, the aspiration of *^{h}r- is ambiguous, making it unclear whether Phu Thai, Phuan, Nyo, Yoy, and Kaloeng should also be classified in this same subgroup.

The three phonological innovations discussed so far provide strong evidence for grouping Lao, Thai, Southern Thai, Khorat, and Tak Bai together in Peninsular SWT. However, determining the status of Phu Thai, Phuan, Nyo, Yoy, and Kaloeng requires examining the fourth innovation, the early occlusivization of *x-. As shown in **Table 10**, the reflex of PSWT *x- in Nyo, Yoy, Kaloeng, and Phuan is /k^{h}-/.

This change presents ambiguity in the history of these varieties. Since PSWT *q- as reconstructed by me (Pittayawat 2009b: 123–124) presumably also became /x-/ and eventually /k^{h}-/, it is not clear whether PSWT *x- became /k^{h}-/ early on, before *q- changed to /x-/, or if the two dorsal sounds first merged to /x-/, which became /k^{h}-/ only later. This ambiguity obscures their precise historical relationships with Lao and Thai. Therefore, Nyo, Yoy, Kaloeng, and Phuan may only be provisionally placed within Peninsular SWT, as there is no conclusive evidence against their sharing of these innovations.

The status of Phu Thai dialects within the Peninsular SWT subgroup presents an interesting challenge. While the Phu Thai Kapong dialect of Waritchaphum does not exhibit the early occlusivization change from *x- to /k^{h}-/, the dialect spoken in That Phanom shows this change. This variation might suggest a potential split within Phu Thai varieties. However, grouping some Phu Thai dialects with Lao and Thai while excluding others is unlikely, given the overall homogeneity of Phu Thai dialects considered in this article. The most plausible hypothesis is that Phu Thai, as a whole, did not undergo the early occlusivization. Instead, it appears that PSWT *x- only recently became /k^{h}-/ in some dialects and /h-/ in others. Given this hypothesis, the Phu Thai dialects are tentatively

placed outside the Peninsular SWT subgroup. The occurrence of /kʰ-/ as a reflex of PSWT *x- requires careful evaluation, as some instances may represent lexical borrowings from Lao or other languages Phu Thai is in contact with rather than inherited forms.[5]

Discussion

In contrast to earlier studies on the subgroup structure of SWT, which were often restricted in their analysis by considering only tone-related developments, the current research concentrates on innovations related to consonants and vowels. The article proposes that Thai, Lao, Southern Thai, Tak Bai, Khorat, Nyo, Yoy, Kaloeng, and Phuan form a subgroup within SWT called "Peninsular SWT". Notably, I argue that Phu Thai is outside this subgroup.

From a broader SWT perspective, an important question arises regarding the relationship between the Peninsular SWT and other SWT languages. This discussion begins with Phu Thai, which, as discussed earlier, may not have undergone the early occlusivization of PSWT *x- to *kʰ, potentially excluding it from the Peninsular SWT subgroup. However, Phu Thai varieties share the coalescence of PSWT *ɓl- into *ɗ-, indicating a close relationship with Peninsular SWT. As so-called PH varieties, they are provisionally considered the closest relatives of Peninsular SWT, despite not fully participating in all its defining innovations.

The discussion then extends to P languages, particularly to Tai Yuan and Shan. These languages did not undergo either the aspiration of PSWT *ʰr- or the early occlusivization of PSWT *x- to *kʰ-. Like Nyo, Yoy, Kaloeng, and Phuan, it is not possible to determine from modern reflexes whether Shan and Tai Yuan share these innovations. However, the reflexes of PSWT *ʰr- and *x- in these two languages in the 16th-century were /r-/ and /x-/ respectively, remaining distinct from the reflexes of PSWT *h- and *kʰ- (Bauer 2015; Shinnakrit 2021; 2023; Shinnakrit & Buragohain 2024). This evidence indicates that Tai Yuan and Shan did not share two of the four innovations that define Peninsular SWT.

Interestingly, Tai Yuan and Shan differ in their treatment of PSWT *ɓl-. Tai Yuan underwent the coalescence of PSWT *ɓl- into *ɗ-, aligning it more closely with Peninsular SWT. In contrast, most modern Shan/Tai Nüa varieties exhibit /m-/, the regular reflex of PSWT *ɓ-, in most words with PSWT *ɓl-, as evidenced by examples such as /mɔk²/ 'flower' < *ɓlɔːkᴰ (ดอก, *dok*) and /mɔŋ¹/ < *ɓlɔːŋᴬ (ดอง, *dong*). This distinction suggests that Tai Yuan has a closer genetic relationship to Peninsular SWT than does Shan. Intriguingly, Harris (1975) documents a "Tai Nüa" dialect spoken in Van Pong Tong in Xishuang Banna that exhibits a distinctive pattern: unlike other Shan/Tai Nüa dialects, it shows /l-/ as the regular reflex of both PSWT *ɗ- and *ɓl-. The status of this variety remains uncertain as it may represent either a divergent Shan dialect or a non-Shan variety that has undergone contact-induced changes from surrounding Shan communities.

As for the SWT dialects spoken in Vietnam, Black Tai and White Tai

[5] See Phinnarat (2000) for discussion of contact among Lao, Phu Thai, and Nyo varieties.

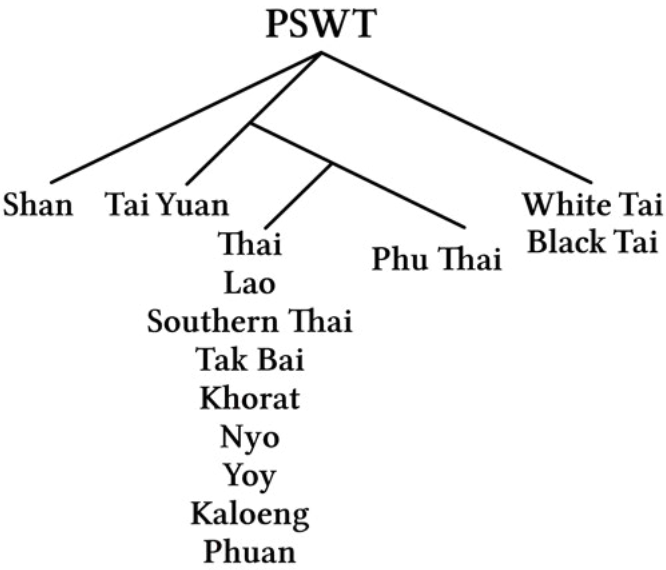


Figure 3: Tentative SWT family tree © Pittayawat Pittayaporn

do not belong to the Peninsular SWT subgroup, as they lack the key innovations characteristic of the group. These include the early simplification of *kʰr-, and the early occlusivization of *x-. Instead, Black Tai, White Tai, and their close relatives in Vietnam exhibit distinct phonological developments that set them apart from Peninsular SWT.

One notable feature is Black Tai and White Tai treatment of PSWT *kʰr-. In these varieties, *kʰr- coalesced into /cʰ-/, which, in some cases, further evolved into /s-/. This development is evidenced by etyma such as PSWT *kʰrapᴰ 'to drive' (ขับ, *khap*), realized as /cʰap²/ in White Tai and /sap²/ in Black Tai, and PSWT *kʰrɔːᴬ 'to beg' (ขอ, *kho*), reflected as /cʰɔ¹/ and /sɔ¹/, respectively. This pattern contrasts sharply with the simplification observed in Peninsular SWT languages. Moreover, White Tai and Black Tai both show /b-/ as the reflex of PSWT *ɓl-, a development distinct from the Peninsular SWT varieties. Examples include White Tai /bɤn¹/ and Black Tai /bɯən¹/ from PSWT *ɓlɯanᴬ 'moon' (เดือน, *duean*), and both varieties showing /bɔʔ²/ 'flower' from PSWT *ɓlɔkᴰ (ดอก, *dok*). An additional distinguishing feature is observed in White Tai, which clearly maintains distinct reflexes for PSWT *x- and *kʰ-. This is demonstrated by words such as /xaːw¹/ 'white' from PSWT *xaːwᴬ (ขาว, *khao*) and /kʰaw³/ 'rice' from PSWT *kʰawᶜ. (ข้าว, *khao*). This retention of distinction contrasts with the early occlusivization observed in Peninsular SWT languages.

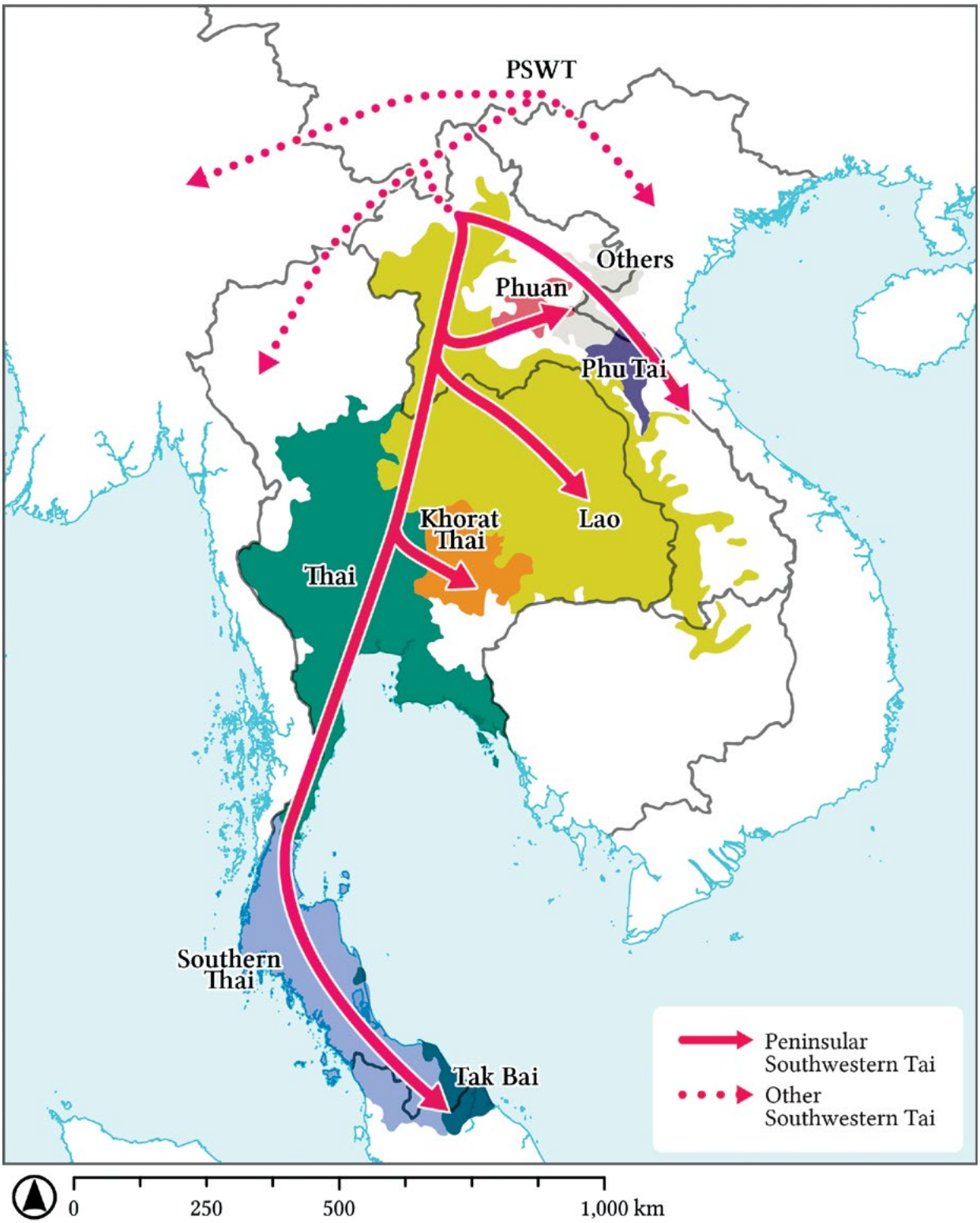
The proposed tree in **Figure 3** diverges notably from that of Chamberlain (1975). While Chamberlain's PH languages are indeed subgrouped together, forming the Peninsular SWT subgroup in our model, his P languages do not constitute a unified branch. The proposed subgroup structure also diverges from Brown's model in its treatment of Southern Thai and Tak Bai. Contrary to Brown's classification, which positions these varieties as the

most divergent in a separate branch, the current model incorporates them within a broader group of PH languages. The proposed subgroup model is predicated on the assumption that Old Thai varieties of Sukhothai and Ayutthaya are closely related. This assumption is corroborated by the analysis of early phonological innovations reflected in the inscriptions from both kingdoms, lending credence to its validity. The inscriptional evidence suggests a linguistic continuum rather than two distinct languages.

This perspective challenges earlier models that posited a more significant linguistic divide between Sukhothai and Ayutthaya (Brown 1965; 1966; Chamberlain 1972; 1991). This suggests a more gradual and unified development of Thai and related languages across these historical kingdoms. Nevertheless, additional research is necessary to fully elucidate the diachronic relationships among these varieties and refine our understanding of the historical development of SWT. It is important to note that the precise nature of the relationship between these historical varieties and their connection to modern PH languages requires further investigation. The current hypothesis posits that the Thai inhabitants of the Sukhothai and Ayutthaya kingdoms spoke variants of the same language, which is ancestral not only to modern Thai but also to the Southern Thai, Tak Bai, and Khorat varieties. This position stands in contrast to Dueanpen's (2541) and Chanida's (1991) acceptance of Brown's and Chamberlain's hypotheses (see above).

Furthermore, the Peninsular SWT hypothesis presents a historical scenario that, in contrast to previous subgrouping proposals, aligns closely with the existing historical findings given below. This hypothesis, which places Thai and Lao within the same SWT subgroup, suggests a pattern of Tai-speaking population movement from Laos to central Thailand, which then served as an intermediary dispersing center for further expansion into southern Thailand.

Assuming that the language of Sukhothai is ancestral to modern Thai, this model proposes that the pre-13th century precursor of both Lao and Thai likely originated in the Luang Prabang area, spreading southward to Sukhothai via the present-day areas of Xayabuli (ໄຊຍະບຸລີ) in Laos and Nan in Thailand, along the Mekong, and then the Nan rivers. This scenario, illustrated in **Map 2**, finds support in historical and geographical factors. The chronicles of Nan and inscriptions from Sukhothai portray the rulers of Sukhothai, Nan, and Luang Prabang as belonging to the same lineage (Prasert 2541: 401–404; Masuhara 2546: 25–27; Sarasawadi 2561: 74–75). Additionally, ancient routes connected Sukhothai to the Mekong River via the Nan River, linking to either Hongsa (ຫົ້ງສາ) or Pak Lai (ປາກລາຍ) in Laos before continuing upriver to Luang Prabang (Teerawatt 2558). These routes may have served as migration pathways for Peninsular SWT speakers as they dispersed from the middle Mekong region. The convergence of linguistic evidence with historical and geographical data presents a compelling case for the proposed SWT subgrouping model.



Map 2: Proposed dispersal scenario for Peninsular SWT
© Pittayawat Pittayaporn & Krittin Klaimee

****

In the absence of sufficient textual evidence to trace ethnolinguistic group dispersal, linguistic genealogy offers valuable insights. For SWT languages, genealogical relationships can inform hypotheses about their spread across the Indochinese Peninsula. The proposed subgrouping model challenges previously held views on the historical development of these languages. These views have posed significant obstacles to advancing our understanding of how Tai-speaking populations arrived and fundamentally transformed the ethnolinguistic landscape of Southeast Asia. The divergent conclusions of this article stem primarily from its methodological emphasis on empirical evidence and systematic historical linguistic analysis. This study thus provides a new framework for understanding language spread and population movement in the region.

Acknowledgments

I would like to express my gratitude to Maria Timpika Chomcha, Chaimongkol Jarudamrongsak, Santhawat Thanyawong, and Theera Roungtheera for their assistance with manuscript preparation, and to Kanitha Putthasatien, Apichaya Kaewuthai, and Rungwimol Pintasaard for their collaboration in collecting field data. I am grateful to Nattanan Sawatdiwong and Krittin Klaimee for creating the beautiful diagrams and maps and to Thanachot Keatnapat for sharing his knowledge of Thai inscriptions. I thank the audience at the 28th Annual Meeting of the Southeast Asian Linguistic Society and two anonymous reviewers for their valuable comments and suggestions. Any shortcomings remain entirely my responsibility.

References

• Areeluck Tisapong. 1985. *A Phonological Description of the Kaloeng Language at Ban Dong Ma Fai, Sakon-Nakhorn*. MA Thesis. Nakhon Pathom: Mahidol University.

• Bauer, Christian. 2015. Vestiges of Proto-Tai Clusters in 15th Century Inscriptions from Northern Thailand. *Zeitschrift der Deutschen Morgenländischen Gesellschaft* 165(1): 211–214.

• Brown, J. Marvin. 1965. *From Ancient Thai to Modern Dialects*. Bangkok: Social Science Association Press of Thailand.

———. 1966. The Language of Sukhothai: Where Did It Come From? And Where Did It Go? *Social Science Review* 3: 40–42.

———. 1985. Historical Explanations for the Peculiarities of the Thai Writing System. In *From Ancient Thai to Modern Dialects and Other Writings on Historical Thai Linguistics*, ed. by J. Marvin Brown, 5–17. Bangkok: White Lotus.

• Burnay, Jean & Cœdès, George. 1927–28. Note sur les tons et les initiales du vieux-siamois à l'époque de Sukhodaya. *Journal of the Siam Society* 21(2): 103–117.

• Campbell, Lyle. 1998. *Historical Linguistics: An Introduction*. Cambridge, MA: MIT Press.

• Chalong Eam-Eium. 1986. *A Phonological Description of Phuan at Tambon Hat Siaw, Si Satchanalai District, Sukhothai Province*.

MA Thesis. Nakhon Pathom: Mahidol University.
• Chamberlain, James R. 1972. The Origin of the Southwestern Tai. *Bulletin des amis du royaume lao* 7–8: 233–244.
———. 1975. A New Look at the History and Classification of the Tai Languages. In *Studies in Tai Linguistics in Honor of William J. Gedney*, ed. by Jimmy G. Harris & James R. Chamberlain, 49–66. Bangkok: Central Institute of English Language.
———. 1983. The Tai Dialects of Khammouan Province: Their Diversity and Origins. *The Science of language* 4: 62–95.
———. 1991. The Efficacy of the P/PH Distinction for Tai Languages. In *The Ram Khamhaeng Controversy: Collected Papers*, ed. by James R. Chamberlain, 453–486. Bangkok: The Siam Society.
• Chanida Rod-in. 1991. *A Phonological Reconstruction of the Sukhothai Initial Consonants in 3 Inscriptions of the 14th Century A.D., with a Comparison to Inscription 1.* MA Thesis. Nakhon Pathom: Mahidol University.
• Diller, Anthony. 1979. Tones, Segments, and Thai Regional Society. In *Studies in Tai and Mon-Khmer Phonetics and Phonology in Honour of Eugénie J.A. Henderson*, ed. by Theraphan L-Thongkum et al., 60–93. Bangkok: Chulalongkorn University Press.
———. 1988. Consonant Mergers and Inscription One. *Journal of the Siam society* 76: 46–63.
• Dockum, Rikker. 2019. *The Tonal Comparative Method: Tai Tone in Historical Perspective*. PhD Dissertation. New Haven: Yale University.
• Duangduen Suwattee & Pranee Kullavanijaya. 1976. A Study of <ʻy>, <y>, <hñ>, and <ñ>. In *Tai Linguistics in Honor of Fang-Kuei Li,*, ed. by Thomas W. Gething et al. Bangkok: Chulalongkorn University Press.
• Dueanpen Rakpaet (เดือนเพ็ญ รักษ์แพทย์). 2541 BE (1998 CE). การศึกษาระบบเสียงภาษาสุโขทัยและเปรียบเทียบกับภาษานครศรี-ธรรมราช [*A Phonological Study of the Sukhothai Dialect with Comparison to the Nakhon Srithammarat Dialect*]. MA Thesis. Nakhon Pathom: Mahidol University.
• Edmondson, Jerold A. & Solnit, David B. 1997. Comparative Shan. In *Comparative Kadai: The Tai Branch*, ed. by Jerold A. Edmondson & David B. Solnit, 337–359. Dallas: University of Texas Press.
• Ferlus, Michel. 1988. Langues et écritures en Asie du Sud-Est : les écritures d'origine indienne et leur adaptation aux langues de l'Asie du Sud-Est continentale : les consonnes. Paper Presented at the *21st International Conference on Sino-Tibetan Languages and Linguistics*, Lund, Sweden, 7–9 October.
———. 1999. Sur l'ancienneté des écritures thaï d'origine indo-khmère. Paper Presented at the Symposium *Georges Cœdès aujourd'hui*, Bangkok, 9–10 September.
———. 2008. The Tai Dialects of Nghệ An, Vietnam (Tay Daeng, Tay Yo, and Tai Muong). In *The Tai-Kadai Languages*, ed. by Anthony Diller et al., 298–316. London: Routledge.
• Gedney, William J. 1965. *Indic Loanwords in Spoken Thai*. PhD Dissertation. New Haven: Yale University.
———. 1989. Speculations on Early Tai Tones. In *Selected Papers on Comparative Tai Studies*, ed. by Robert J. Bickner et al., 207–228. Ann Arbor: University of Michigan Press.
———. 1991. Comments on Linguistic Arguments Relating to Inscription One. In *The Ram Khamhaeng Controversy*, ed. by

James R. Chamberlain, 193–226. Bangkok: The Siam Society.
• Harris, Jimmy G. 1975. A Comparative Word List of Three Tai Nüa Dialects. In *Studies in Tai Linguistics in Honor of William J. Gedney*, ed. by Jimmy G. Harris & James R. Chamberlain, 202–230. Bangkok: Central Institute of English Language.
• Harrison, Shelly P. 2003. On the Limits of the Comparative Method. In *The Handbook of Historical Linguistics*, ed. by Brian D. Joseph & Richard D. Janda, 213–242. Oxford: Blackwell.
• Hartmann, John F. 1980. A Model for Alignment of Dialects in Southwestern Tai. *Journal of the Siam Society* 68(1): 72–86.
• Hudak, Thomas J. 2008. *William J. Gedney's Comparative Tai Source Book*. Honolulu: University of Hawai'i Press.
• Jonsson, Nanna L. 1991. *Proto-Southwestern Tai*. PhD Dissertation. Albany: State University of New York.
• Kanjana Koowatthanasiri (กาญจนา คูวัฒนะศิริ). 2524 BE (1981 CE). วรรณยุกต์ในภาษาญ้อ [*The Tones of Nyo*]. MA Thesis. Bangkok: Chulalongkorn University.
• Li, Fang-Kuei. 1960. A Tentative Classification of Tai Dialects. In *Culture in History: Essays in Honor of Paul Radin*, ed. by Stanley Diamond, 951–959. New York: Columbia University Press.
———. 1977. *Handbook of Comparative Tai*. Honolulu: University of Hawai'i Press.
• Lorrillard, Michel. 2018. Research on the Inscriptions in Laos: Current Situation and Perspectives. In *Writing for Eternity: A Survey of Epigraphy in Southeast Asia*, ed. by Daniel Perret, 87–107. Paris: EFEO.
• Masuhara, Yoshiyuki (โยชิยูกิ มาซูฮารา). 2546 BE (2003 CE). ประวัติศาสตร์เศรษฐกิจของราชอาณาจักรลาวล้านช้าง สมัยคริสตศตวรรษที่ 14–17 จาก "รัฐการค้าภายในภาคพื้นทวีป" ไปสู่ "รัฐกึ่งเมืองท่า" [*On the Economic History of the Kingdom of Laos from the 14th to 17th Centuries: From "Hinterland State" to "Semi-Port State"*]. Bangkok: Matichon Press.
• Mudjalin Luksanawong (มุจลินทร์ ลักษณะวงษ์). 2559 BE (2016 BE). การจัดกลุ่มย่อยภาษากะเลิงในประเทศไทยและสาธารณรัฐประชาธิปไตยประชาชนลาว [*A Sub-grouping of Kaloeng in Thailand and Lao PDR*]. PhD Dissertation. Mahasarakham University.
• Phinnarat Akharawatthanakun. 2000. Tonal Variations and Changes in a Language Mixture Area: A Case Study of Northeastern Thailand (Isan). *MANUSYA: Journal of Humanities* 5(2): 30–51.
———. 2010. Phonological Variation in Phuan. *MANUSYA: Journal of Humanities* 13(2): 50–87.
• Pittayawat Pittayaporn (พิทยาวัฒน์ พิทยาภรณ์). 2009a. *The Phonology of Proto-Tai*. PhD Dissertation. Ithaca: Cornell University.
———. 2009b. Proto-Southwestern Tai Revised: A New Reconstruction. *Journal of the Southeast Asian Linguistics Society* 2: 119–143.
———. 2556 BE (2013 CE). พัฒนาการวรรณยุกต์กับการแบ่งกลุ่มย่อยภาษาไทตะวันตกเฉียงใต้ [Tonal Developments and Southwestern Tai subgrouping]. วารสารอักษรศาสตร์ [*Journal of Letters*] 42(2): 305–339.
———. 2014. Layers of Chinese Loanwords in Proto-Southwestern Tai. *MANUSYA: Journal of Humanities*, Special Issue 20: 47–68.
• Pittayawat Pittayaporn & Kirby, James. 2017. Laryngeal Contrasts in the Tai Dialect of Cao Bằng. *Journal of the International Association of Phonetics* 47(1): 65–85.
• Pranee Kullavanijaya & Theraphan L-Thongkum. 2000. Linguistic Criteria for Determining Tai Ethnic Groups: Case Studies on Central and Southwestern Tais. In *Proceedings: The International*

Conference on Tai Studies, July 29–31, 1998, ed. by Somsonge Burusphat, 273–298. Nakhon Pathom: Mahidol University.

• Prasert Na Nagara (ประเสริฐ ณ นคร). 2541 BE (1998 CE). สัมพันธภาพระหว่างราชวงศ์พระร่วงและราชวงศ์น่าน [The Relationship between the Phra Ruang and Nan Dynasties]. In สารนิพนธ์ ประเสริฐ ณ นคร [*Prasert Na Nagara's Writings*], ed. by Niyada Laosunthon, 401–404. Bangkok: Kasetsart University Press.

• Puttachart Potibal (พุทธชาติ โปธิบาล) & Thananan Trongdee (ธนานันท์ ตรงดี). 2541 BE (1998 CE). สถานะของภาษาตากใบในภาษาไทถิ่น (โครงการระยะที่ 2) [*The Position of Tak Bai Tai in Tai Dialects (Phase 2)*]. Pattani: Prince of Songkhla University.

• Sarasawadi Ongsakul (สรัสวดี อ๋องสกุล). 2561 BE (2018 CE). ประวัติศาสตร์ล้านนา [*History of Lanna*]. Bangkok: Amarin Printing [12th Ed.]

• Sarawit, Mary E. 1973. *The Proto-Tai Vowel System*. PhD Dissertation. Ann Arbor: University of Michigan.

• Sasithorn Onlao (ศศิธร อ่อนเหลา). 2553 BE (2010 CE). การจัดกลุ่มภาษาไทในลุ่มน้ำพาว เมืองคำเกิด แขวงบอลิคำไซ สาธารณรัฐประชาธิปไตยประชาชนลาวโดยใช้เกณฑ์การเปลี่ยนแปลงร่วมกันของเสียงวรรณยุกต์ [Grouping of Tai Languages in Nam Paw Basin, Khamkeut District, Bolikhamxai Province, Lao People's Democratic Republic by Using Criteria of Tonal Sound Changes]. *Journal of Mekong Societies* 6(3): 47–72.

• Shinnakrit Tangsiriwattanakul. 2020. *Ming Shilu* as Evidence of Devoicing of Voiced Obstruents in Siamese. *Journal of the Southeast Asian Linguistics Society* 13(2): 67–95.

———. 2021. Lanna Tai of the 16th Century: A Preliminary Study of the *Sino-Lanna Manual of Translation*. *Journal of the Southeast Asian Linguistics Society* 14(2): 1–19.

———. 2023. From Proto-Southwestern Tai to Modern Lanna Tai: Implications from the 16th-Century Phonology. *MANUSYA: Journal of Humanities* 26: 1–33.

• Shinnakrit Tangsiriwattanakul & M. Buragohain. 2024. The Phonological History of Tai Nüa: Implications from the *Sino-Baiyi Manual of Translation*. *Journal of the Southeast Asian Linguistics Society* 17(1): 19–45.

• Sireemas Maspong & Pittayawat Pittayaporn. 2019. Length Contrast of High Vowels in the Thai Language of Sukhothai Period: What Do the Inscription Says? *Cahiers de linguistique Asie orientale* 48(1): 30–60

• Strecker, David & Pittayawat Pittayaporn. 2018. Tai Languages. In *The World's Major Languages*, ed. by Bernard Comrie, 672–678. London: Routledge.

• Teerawatt Sankom (ธีระวัฒน์ แสนคำ). 2558 BE (2015 CE). เส้นทางความสัมพันธ์ระหว่างรัฐสุโขทัยกับกลุ่มบ้านเมืองในลุ่มน้ำโขง [Routes of Relationships between Sukhothai and Towns in the Mekong Basin]. In สุโขทัยกับอาเซียน: มองปัจจุบันผ่านอดีตจากมิติประวัติศาสตร์ ศิลปะ โบราณคดี *[Sukhothai and ASEAN: Viewing the Present through the Past in History, Art, and Archaeology]*, ed. by Pipad Krajaejun, 53–73. Bangkok: Thammasat University Press.

• Thepbangon Boonsner. 1984. *The Phonological Description of Yooy*. MA Thesis. Nakhon Pathom: Mahidol University.

• Theraphan L-Thongkum (ธีระพันธ์ ล.ทองคำ). 2521 BE (1978 CE). เสียงและระบบเสียงในภาษาไทยถิ่นใต้ จังหวัดสุราษฎร์ธานี 16 อำเภอ [*Sounds and Sound Systems of the Sub-dialects of the Suratthani Dialect of Southern Thai*]. Bangkok: Chulalongkorn University Press.

———. 1997. Implications of the Retention of Proto-Voiced Plosives and Fricatives in the Dai Tho Language

of Yunnan Province for a Theory of Tonal Development and Tai Language Classification. In *Comparative Kadai: The Tai Branch*, ed. by Jerold A. Edmondson & David B. Solnit, 191–220. Dallas: University of Texas Press.

———. 2003. The Tai Muong Vat Do Not Speak the Black Tai Language. *MANUSYA: Journal of Humanities* 6: 74–86.

• Umar, Umaiyah H. 2003. *A Classification of Thai-Kedah Dialects Using Phonological Characteristics*. Bangkok: Nutcha Printing.

• Uraisi Varasarin. 1984. *Les Éléments khmers dans la formation de la langue siamoise*. Paris: Société d'études linguistiques et anthropologiques.

———. (อุไรศรี วรศะริน). 2553 BE (2010 CE). ร่องรอยภาษาเขมรในภาษาไทย [*Khmer Elements in the Formation of the Thai Language*]. Nakhon Pathom: Silpakorn University Press.

• Wanna Phudhitanakul. 1979. *A Tonal Comparison of Dialects in Nakhon Ratchasima Province, Thailand.* M. Phil. Thesis. London: SOAS.

• Warunsiri Pornpottanamas. 2023. Classifying 'Lao' Ethnic Languages Spoken in Central Thailand Using Tonal Criteria: A Case Study of 'Chachoengsao Lao'. *MANUSYA: Journal of Humanities* 26: 1–24.

• Wilaiwan Khanittanan (วิไลวรรณ ขนิษฐานันท์). 2520 BE (1977 CE). ภาษาผู้ไท [*The Phu Thai Language*]. Bangkok: Thammasat University Press.

Website

• Princess Maha Chakri Sirindhorn Anthropology Centre. *Inscriptions in Thailand Database*: https://db.sac.or.th/inscriptions/.

Appendix: Phonemic Values of Old Thai and Old Lao Graphs

Phonemic values of Old Thai and Old Lao graphs are based on correspondences between the graphs and the PSWT phonemes established by Ferlus (1988: 22–25). The sibilants *ś* and *ṣ* lack PSWT correspondences and are thus treated in this article as phonologically equivalent to *s*. Similarly, the graphemes *gh*, *ḍh*, *dh*, *bh*, and *ḷ* are considered phonologically identical to *g*, *ḍ*, *d*, *b*, and *l*, respectively. These phonological equivalences do not affect the present analysis. Vowel graph transliterations are omitted as they are not pertinent to the present analysis.

Transliteration	Old Thai Thai Script	Old Thai Khom Script	Old Lao Thai Noi Script	Old Lao Dhamma Script	Phonemic Value
k	ก	ក	ก	๓	/k-/
kh	ข	ខ	ຂ	๒	/kh-/
‘kh	ฃ		຃	๖	/χ-/
g	ค	ត	ค	๑	/g-/
‘g	ฅ		ฅ	๕	/ɣ-/
gh	ฆ	ឍ		๗	/g-/
ṅ	ง	ង	ง	๘	/ŋ-/
c	จ	ច	ຈ	๐	/c-/
ch	ฉ	ឆ		๙	/ch-/
j	ช	ជ	ຊ	๘	/ɟ-/
‘j	ฌ	ឝ	ຝ	๕	/z-/
ñ	ญ	ញ	ญ	๖	/ɲ-/
ṭ	ฏ	ដ		๑	/ɗ-/
ṭh	ฐ	ឋ	ฐ	๓	/th-/
ḍ		ឌ		๒	/ɗ-/
ḍh		ឍ		๗	/d-/
ṇ	ณ	ណ	ณ	๙	/n-/
t	ด	ត	ດ	๒	/ɗ-/
‘t	ต		ຕ	๐	/t-/
th	ถ	ថ	ຖ	๑	/th-/
d	ท	ទ	ທ	๓	/d-/
dh	ธ	ធ	ธ	๘	/d-/
n	น	ន	ນ	๔	/n-/
jh		ឈ		๓	/ɟ-/
p	บ	ប	ບ	๖	/ɓ-/
‘p	ป		ປ	๖	/p-/
ph	ผ	ផ	ຜ	๗	/ph-/
‘ph	ฝ		ຝ	๙	/f-/

Transliteration	Old Thai Thai Script	Old Thai Khom Script	Old Lao Thai Noi Script	Old Lao Dhamma Script	Phonemic Value
b	พ	ព	ພ	ᨻ	/b-/
ʻb	ฟ		ຟ	ᨼ	/v-/
bh	ภ	ភ	ຠ	ᨽ	/b-/
m	ม	ម	ມ	ᨾ	/m-/
y	ย	យ	ຍ	ᨿ	/j-/
r	ร	រ	ຣ	ᩁ	/r-/
l	ล	ល	ລ	ᩃ	/l-/
v	ว	វ	ວ	ᩅ	/w-/
ś	ศ	ឝ	ຨ	ᩆ	/s-/
ṣ	ษ	ឞ	ຩ	ᩇ	/s-/
s	ส	ស	ສ	ᩈ	/s-/
h	ห	ហ	ຫ	ᩉ	/h-/
ḷ		ឡ	ຬ	ᩊ	/l-/
ʔ	อ	អ	ອ	ᩋ	/ʔ-/