Identifying the Tai Substratum in Cantonese[1]
Robert S. Bauer
Institute of Language and Culture for Rural Development
Mahidol University at Salaya, Thailand

1.0 Introduction

A number of common Cantonese words seem to lack any etymological relationship with semantically-equivalent words in Mandarin. How can these lexical differences be accounted for? Given that the same phenomenon occurs in other Chinese or Han dialects, the question holds much interest for anyone studying the historical development of the Han dialects. It thus seems quite natural that the question should also be an issue in Sino-Tibetan historical comparative linguistics. Indeed, twenty years ago Hashimoto (1976) believed it was too early to talk about a Sino-Tibetan language family on a par with Indo-European until Sinologists had adequately answered the fundamental question, what accounts for the apparent diversity of basic vocabulary within the Sinitic half of the family? He (p. 2) drew attention to the fact that the comparison of basic words in the Han dialect families of Mandarin (northern group), Wu, Gan, Xiang (central group), and Min, Yue, and Kejia (southern group) indicates a fair portion of it derives from diverse sources. Table 1 below lists forms from the seven major Han dialect families for four lexemes for which we observe several patterns of relationship: for 'to cover' there appears to be a clear division between the northern and central groups with their open syllables, on the one hand, and the southern group with syllables closed by the bilabial nasal, on the other; in the case of 'frog' all dialect families except Kejia and Wu may share one root morpheme which had a velar initial and a bilabial stop final consonant; for the third person singular pronoun Mandarin and Xiang share the same form, but forms in the other dialect families clearly derive from different etyma; the morpheme for 'this' is shared by Xiang and Gan and possibly Wu but takes quite diverse phonetic shapes in the other dialects.

Table 1. Basic vocabulary in 7 Han dialect families (from HFCH).

English Gloss	Mand.- Beijing	Wu- Suzhou	Xiang- Changsha	Gan- Nanchang	Yue- Guangzhou	Kejia- Meixian	Min- Xiamen
cover, to	kai⁵	kE⁵	kai⁵	kai⁵	khɐm³	khɛm²	khɐm⁵
frog	xa¹ma⁰	tsʰin¹o¹	ka⁷mau⁰	ha²ma⁰	kɐp⁷ᵃna:³	tsʰiaŋ¹kuai³ε⁰	tsʰan²kap⁷a
he, she	tʰa¹	li¹	tʰa¹	tɕʰiɛ²	khɵy⁴	ki²	i¹
this	tʂei⁵	kE¹	ko⁷	kɔ³	ni:¹	e³	tsit⁷

This bit of comparative lexical material is enough to demonstrate that the so-called Han "dialects" are not as homogeneous as the term *dialect* might suggest. However, what is more important and relevant to the purpose of this paper is the question, what has created this etymological diversity among the various Han dialect families? As will become clear in the following discussion, some scholars believe that historical contact among Han dialects and non-Han languages must account for some of it. In a wide-ranging paper on this very topic Wang (1995:18) has noted that "the deeper we probe into these questions of contact and differentiation, the more we are likely to realize the importance of the non-Han languages in the formation of the Han dialects, and that China is indeed a complex Sprachbund . . ." Determining the contribution of the non-Han languages to the historical development of the Han dialects is a subject of great significance; yet its magnitude and complexity may help explain why we still do not have an etymological dictionary of Han dialectal vocabulary (as opposed to the Chinese characters which are not necessarily equivalent to words). Working with the reading pronunciations of the Chinese characters in the different dialects rather than vocabulary items in speech, Sinologists have generally focused their attention on the changes that have taken place in the pronunciations of the standard Chinese characters; as a result, they have tended to shy away from tackling the more difficult questions associated with the emergence and evolution of the Han dialects.

Over the last few years my own interest in this subject has been mainly limited to the southern group of Han dialect families of Yue, Min, and Kejia. In my approach I have focused on a particular kind of basic vocabulary, viz., characterless morphemes -- words which are not etymologically associated with standard Chinese characters. I have begun with Yue, the dialect family with which I am most familiar, and then have systematically expanded the scope of my research by tracing lexical connections to the neighboring Min and Kejia dialects, and then to the non-Han languages of Southeast Asia. This paper attempts to identify the origins of 29 Cantonese words which appear to lack any etymological relationship with standard Chinese characters.

2.1 Language Families of the Southeast Asia Sprachbund

From their early beginnings down to the present day the language groups of Southeast Asia, i.e., Chinese (or Sinitic), Tibeto-Burman, Austronesian, Tai, Mon-Khmer, and Miao-Yao, have been in contact with each other and have been influencing each other's development. This long history of mutual influence has resulted in areal convergence which has made working with these languages a particularly complicated matter. Gedney (1976:66), a specialist in the description of Tai dialects, has stated the difficulty of sorting out relationships among Southeast Asia's languages as follows:

"Linguists who study Southeast Asia, one of the most complex areas in the world, are constantly struck by instances of apparent convergence, where languages or language groups which are genetically unrelated (or if related, the relationship is so far back in time as to [be] irrelevant), e.g. Thai and Mon-Khmer, or Thai and Tibeto-Burman, or Thai and Vietnamese, or any Southeast Asian language and Chinese, show similarities in grammatical structure or the organization of the semantics of the lexicon. These convergences are clearly the result of contact; how much former bilingualism must be assumed in order to explan these similarities is not yet clear."

A topic of particular interest to me is how the various language groups of Southeast Asia have influenced the development of the southern Han dialect families of Yue, Min, and Kejia. At the same time, I am also deeply concerned with their historical origin and how they have been influencing the development of each other. I have not been alone in pursuing this line of research. With respect to the development of Yue, Li J-z. (1990:28) has delimited the problem as follows:

"In a nutshell, the Yue dialect area encompasses Min, Zhuang, Yao, and Kejia languages. In that case, under the circumstances in which numerous languages are enclosed within the same area, how exactly has the Yue language emerged and developed?"

2.2 The Sino-Tai Relationship

Benedict (1975:32, 35, 135), a pioneer in the historical comparative linguistics of Southeast Asian languages, has proposed that Tai, Austronesian, and Miao-Yao belong to a larger family which he has named Austro-Tai, while Tibeto-Burman, Karen, and Chinese constitute the Sino-Tibetan family; Mon-Khmer, a sub-branch of Austroasiatic, is a separate family from these other two. Linguists in China have generally recognized that Sino-Tibetan includes Chinese, Tai, Tibeto-Burman, Karen, and Miao-Yao. However, they have not defined the Sino-Tai genetic relationship according to the traditional sense of this term. Chinese and Tai have been described as not originally sharing a genetic relationship but as having developed such a relationship as a result of their convergence through areal contact. Based on a detailed comparison of basic vocabulary from Chinese, Tibeto-Burman, and Tai, the eminent Chinese Tibeto-Burmanist, Dai Qingxia (1991:63), concluded that Chinese and Tai share a "'genetic' link" which is of a different type than the one which joins Chinese with Tibeto-Burman. Dai's comparison of basic vocabulary clearly showed that Chinese and Tibeto-Burman share cognates with each other but not with Tai. On the other hand, Chinese and Tai share some non-basic vocabulary with each other but not with Tibeto-Burman. He concluded that Sino-Tibetan languages are related through two kinds of genetic relationship:

"one is formed by direct descent from the proto-language (like that of Chinese and T[ibeto]-B[urman], and the other is forged under language contact (like that of Tai and Chinese). This is then one of the major points on which the formation of Sino-Tibetan is different from other families, such as the Indo-European. . . while Chinese and T-B are outwardly dissimilar but have the same origin, Tai and Chinese are similar in appearance but have different origins. . . We can . . . treat the relationship between Chinese and Tai as a genetic one, since the latter has experienced a qualitative change under language contact and has organically assimilated into the family of the influencing language" (Dai 1991:63).

The Tai and Chinese historical linguist, Li Fang Kuei (1976:231-237), compared a number of items of basic vocabulary from modern Thai, modern Chinese, and his reconstructed Old Chinese which he claimed showed regular phonetic correspondences. He made clear his belief that these shared words must have been inherited from the same source language and not borrowed. Prapin (1976:12-25), a student of Li, presented an even larger set of Sino-Tai lexical comparisons for which she claimed the phonetic correspondences established that Old Chinese and Proto-Tai were related not only through their descent from a common ancestor but also contact borrowing. She stated that this borrowing occurred in both directions but did not specify for which words Tai or Chinese was the donor.

In the eyes of most Western linguists the convergence between Chinese and Tai in monosyllabicity, tones, vocabulary, grammatical categories, and word order has resulted from their areal contact. Egerod (1976:52), the prominent Danish sinologist who worked for decades on both Tai and Han dialects, described the Sino-Tai relationship as follows (names of language families have been adjusted to correspond with those used in this paper):

"The similarities between [Tai] and Chinese are non-genetic and just one symptom of the tremendous consequences of the rise, intermingling, and fall of three major language families in present-day China: [Sino-Tibetan], Mon-Khmer, and [Austronesian]. And it is toward [Austronesian] and not toward [Sino-Tibetan] that we have to look for the genetic relationship with [Tai]."

2.3 Sino-Tai Lexical Borrowing

Regardless of whether Tai and Chinese have separate origins, there is no question that their close contact has resulted in each language group influencing the development of the other. The traditional, sinocentric view (much revised by Benedict 1976:60) holds that the Han civilization historically dominated East Asia for more than two millennia, and that the Chinese language exerted great influence on its neighbors, e.g., by exporting vocabulary, tones, and monosyllabicism to Tai and Miao-Yao (Benedict 1976:62) and Vietnamese.

Nevertheless, Chinese has not always occupied the position of the donor language. There is evidence that lexical borrowing moved in both directions -- not only from Chinese to Tai but also from Tai to Chinese. Benedict (1976:87-99) has assembled a lengthy list of basic Chinese words which he has claimed were loans to Archaic Chinese from his reconstructed Proto-Austro-Tai. Included in his list were terms for edible and useful plants (bean, rice, garlic, ginger, hemp, mustard, peach), both domestic and wild animals (cattle, horse, elephant, tiger, rabbit, fowl, bee), metals (copper/bronze, iron, gold), tools (crossbow, hat, quiver, needle, knife, writing stylus, boat, kiln, stairs, tripod), higher numbers (thousand, ten thousand), natural phenomena (lightning, thunder, wind), and various verbs some of which were associated with trade (change, dye, escort, hull grain, open, pass, sell, weave). Chinese characters are associated with all these purported loanwords from Austro-Tai into Archaic Chinese.

3.1 Criteria for Identifying Tai Loanwords in Cantonese

My principal concern in this paper is in identifying Tai loanwords which have been borrowed into Cantonese. In order to identify what is and is not a Tai loanword it is necessary to develop some reliable criteria on which to establish such identification. At the outset one should recognize the difficulty of distinguishing between loanwords and autochthonous vocabulary in Southeast Asian languages. Matisoff (1983:62-63), a trailblazer in Sino-Tibetan studies, has applied an apt metaphor to the linguistic process which makes the linguist's task even more arduous "in the hot-house homogenizing atmosphere of Southeast Asia":

"Just as it takes a very few years for a patch of cleared land to revert to tropical rainforest, so do borrowed words or grammatical features quickly lose any aura of foreignness as they become 'overgrown' by the articulatory or grammatical habits of the native language."

That Tai and Cantonese show a strong tendency toward monosyllabism means that the maximum phonetic material we have to compare are the initial and final consonants of the syllable. Monosyllabism thereby introduces its own special problems which must be confronted. In order to raise the reliability quotient of comparisons I believe they must be based on fairly rigorous phonetic and *semantic criteria. Allowing too much leeway in the similarity of sounds and* meanings diminishes the scholarly appeal of the comparison (as well as the fun of *the enterprise). Monosyllabic comparisons also raise the possibility that what* appears to be a very attractive match is really just an accidental similarity. One way to avoid spurious 'lookalikes' is to compare phonosemantically similar forms from more than two Yue dialects and more than two Tai languages.

In my study of the historical comparative semantics of southern Han dialects I have been mainly working with characterless words or *chorphans* ([kʰór-fənz],

Bauer 1992b), i.e., morphemes which lack etymological relationships with standard Chinese characters and thus may not have written forms (hence orphaned morphs). By excluding morphemes which are etymologically associated with standard Chinese characters, I have to some extent narrowed the scope of my study to words whose origins most probably lie outside Chinese. However, I do recognize that it is quite possible that some chorphans may have had at some earlier time perfectly good etymological associations with Chinese characters, but for one reason or another the Chinese characters have been lost or forgotten (a similar phenomenon has occurred in the history of English; inspection of Old English texts shows that a large portion of vocabulary has fallen from use and been replaced). The attempt to make associations between chorphans and old Chinese characters found in the ancient dictionaries is a field of study called 本 字 考 and is exemplified by Bai (1980) and Sin (1994). However, even if such a link can be established between a Cantonese chorphan and a Chinese character from one of the early riming dictionaries, e.g., the *Guangyun* 廣 韵 (1008 A.D.) or the *Jiyun* 集 韵 (1037 A.D.), or from China's first etymological dictionary, the *Shuowen Jiezi* 説 文 解 字 (100 A.D.), must we conclude automatically that the chorphan has a Chinese origin? Of course not. As we know from our observations of modern Cantonese, the phonetic forms of words are borrowed from English and then the phonetically (and sometimes semantically) appropriate Chinese characters are selected for transcribing the syllables which represent the loanwords. I will mention here that of the 31 lexemes presented in this paper only two -- $nɐm^3$ 'to think, ponder deeply' and $k^hɐm^3$ 'to cover' -- also appear in Bai's set of several dozens of Cantonese words and expressions for which she sought to establish etymologies based on the occurrence of phonosemantically similar Chinese characters found in the ancient rime dictionaries.

As another criterion for establishing that a word in a Yue dialect has been borrowed from Tai, I propose that a comparison needs to be relatively broad-based to include not only Yue's neighboring and related Han dialect families of Min and Kejia, but also the non-Sinitic language groups of south China and Southeast Asia, namely, Miao-Yao, Tai, Austronesian, Tibeto-Burman, and Mon-Khmer, which have been or are still in contact with Yue. The broad-based comparison is intended to help exclude words from non-Tai sources. On the one hand, if a word is found in one or more of the Yue dialects and in many Tai languages, but not in the Min and Kejia dialects and not in other Southeast Asian languages (except possibly Vietnamese which has been much infuenced by Chinese), then there seems to be a strong likelihood that the word has been borrowed from Tai. However, one result of the historical contact among Cantonese, Tai, and Vietnamese has been that they share some words in common, so that determining the direction of borrowing can be a difficult task. On the other hand, if the word is widely found in a number of dialects of two or all three Chinese dialect families, then it is probably an old Chinese word with a relatively long history. Further, if the word occurs in Yue and Kejia or Yue and Min and

also in several or even all Southeast Asian language groups, then it is unlikely to be a Tai loanword in Yue, but the history of the word must go back to a time when the speakers of the ancestral forms of these languages were in close contact.

3.2 Origin of Yue

According to Norman (1988:210-213), the three southern Han dialect families of Kejia, Min, and Yue have descended from the same historical source language, namely, Old Southern Chinese, which was itself derived from varieties of Archaic Chinese brought to south China approximately 2,000 years ago. Before the arrival of the ethnic Han immigrants to south China during the Han dynasties (206 B.C. - A.D. 220), peoples who spoke non-Sinitic languages occupied the area. Early Chinese historians referred to some of these people as the Yue (Yue-Hashimoto 1991:294-298). The term has become another name for Guangdong and Guangxi Provinces and the region's major Han dialect family. Based on what we know about the mutual influence of languages in contact, we can reasonably assume that the historical contact among speakers of Han dialects and the non-Han languages of the Yue people must have resulted in mutual lexical borrowing and even the emergence of pidgins and creoles, but in time the numerically dominant Han assimilated most of the non-Han populations. It is possible that Cantonese originated through the mixing of an early form of Old Southern Chinese with an early form of Tai. Li J-z. (1990:32) has already proposed a a rather similar hypothesis and has set out a procedure for investigating it:

"As modern Cantonese is a kind of language which has formed and developed on the basis of the mutual blending of the "ancient Hua-Xia language" (archaic Chinese) and the "ancient Man-Yi language" (ancient Yao language and ancient Tai language). So, it can be determined that even if Cantonese is now the result of developing according to its own internal laws of development and long ago changed beyond recognition in relation to the differences of these languages mentioned above, nevertheless, because at its earliest stage of development it had absorbed special features of the above languages (which included such features of phonetics, vocabulary, and grammar) and gradually developed on the basis of these, therefore, all that is needed is for us to carry out detailed research on Yue and then we will discover without difficulty certain structural arrangements in the vocabulary and grammatical structure of Yue -- whether many or a few in which vestiges of the features of the above-mentioned languages have still been preserved. To use the same kind of terminology, in the "substratum" of Yue we can still find the residue of the above-mentioned languages".

3.3 Tai Substratum in Yue

One possible explanation for the occurrence of chorphans (and other distinctive linguistic characteristics) in the Yue dialects is that they constitute non-Han substrata. It has been recognized for some years that Cantonese displays

some interesting and unusual differences in grammar, phonology, and lexicon in comparison to standard Mandarin and other Chinese dialects, and it is just these differences which turn out to be similarities Cantonese shares with Tai languages. The origin of these differences and similarities has attracted the attention of scholars who have attempted to identify neighboring minority language groups, such as Zhuang-Dong or Miao-Yao, as their source. Li J-f. (1992:330) has stated that certain phonetic aspects associated with the tone systems of the Yue dialects has been influended by Tai. For example, the distinction between long and short vowels found in the Yue dialects sets them apart from all other Han dialect groups which do not make use of differences in vowel length. The Tai languages also distinguish between long and short vowels but do so on an even larger scale than is found in Yue. The length difference in Yue is prominently associated with the Ru tone category in Yue in which syllables end in final stop consonants -p, -t, -k; in many Yue dialects morphosyllables with short vowels generally occur in the Upper Yin Ru tone category (high level tone contour), and corresponding morphosyllables with long vowels in the Lower Yin Ru tone category (mid level tone contour) and morphosyllables with both long and short vowels in the Yang Ru category (low level tone contour), e.g., $m\!et^{7a}$ ˥55 乜 'what', $ma\!:t^{7b}$ ˧33 抹 'to wipe', $m\!et^{8}$ ˩22 物 'thing', $pa\!:k^{8}$ ˩22 白 'white'. Zhuang shows a similar feature with short vowels occurring in morphosyllables with the high level stopped tone (with a tone contour which is almost identical to the corresponding Cantonese tone) and long and short vowels occurring in morphosyllables with the mid level and high rising stopped tones (also practically the same as the corresponding Cantonese tones), e.g., mat^{7a} ˥55 'flea', $ma\!:t^{7b}$ ˧35 'inclined', mat^{8} 'sealed tight', $ma\!:t^{8}$ ˧33 'to wipe' (a Cantonese loan?).

Another feature shared by both Yue dialects and Tai languages is the occurrence of the sonorant initials m-, n-, ŋ-, l- with both the Yin and Yang (upper and lower register) tones which generally correspond to historically voiceless and voiced initial consonants, respectively. Given that in other Han dialects these voiced initials generally occur with the Yang tones, Li J-f. (1992:331) has claimed that some of these words with the sonorant initials and the Yin tones in the Yue dialects have entered Yue as the result of the early contact relationship between the Yue and Tai language groups.

With the long history of Tai languages as well as other non-Sinitic languages in South China, one should not be surprised if the development of Cantonese has been influenced in various ways by its neighboring languages. Before the arrival of the Han Chinese in south China sometime in the latter part of the 1st millenium B.C., this area was inhabited by aboriginal tribes, some of which were referred to as Bai Yue 百 越 or "Hundred Yue". The language(s) spoken by the Bai Yue is (are) believed to be the source language of the modern Tai languages (Li J-f. 1992:330). In the Yue dialects there are a number of words which are not recorded in the ancient written records, and some of these words do not appear to be related to any semantically-equivalent words in other Han dialects. According

to Li J-f. (1992:334), some words in the Yue dialects which correspond phonetically and semantically to words in the Tai languages can be explained as follows:

". . . these are language elements which the Han language had absorbed from the Bai Yue language during the long period in ancient times when the Han and Yue languages blended together. At the same time these [words] are also the language substratum which has remained after the hundreds of thousands of Bai Yue and their descendants abandoned their own language and adopted the Han language".

3.4 Cantonese-Tai Lexical Comparison

Over the last few decades several publications have appeared in which linguistic forms from Cantonese (and other Yue dialects) have been compared with phonosemantically similar material from Tai and/or Miao-Yao languages, e.g., Bauer (1987:100-107); Egerod 1967:117 (his translation of Yuan 1983); Li J-f. (1990:72-76; 1992:335-337); Li J-z. (1990:39, 40); Ou-yang (1989:609-611); You (1992:174; 1995:259-260); Yuan (1983:179); and Yue Hashimoto (1976:2; 1991:305-307). On the basis of their comparisons scholars have concluded that Cantonese acquired its forms through its historical contact with languages belonging to these groups. While the results of some of this comparative work look quite promising, I believe one should approach these conclusions with some caution and carefully evaluate on a case-by-case basis the individual items which Cantonese is claimed to have borrowed from the Tai or Miao-Yao languages through their historical contact. In my view the main problem that weakens these claims is the narrow base upon which this comparative work has been based.

Some scholars, e.g., Li J-z. (1990:39-40), have compared Cantonese with only one or two other languages belonging to the Tai or Yao language groups. Such narrow comparisons can tell us nothing about how widespread the distribution of an item is within Tai or Yao. The extent to which an item occurs in a group of related languages can help us determine whether or not the item belongs to the autochthonous vocabulary or is a loanword borrowed from some other language. This matter is closely related to another problem, namely, the lack of any explicitly stated criteria with which to decide that an item is indeed a loan into Cantonese rather than a loan from Cantonese or some other Han dialect into the Tai or Yao languages. As is well-known by anyone who works in this area, the Tai and Yao languages have been strongly influenced by the Han dialects with which they have been in contact (for example, in Bangkok a large portion of the ethnic Chinese community is from Chaozhou, and Egerod (1959) has listed a large number of Chaozhou words which have been borrowed into standard Bangkok Thai; the borrowing of Cantonese words by Yao speakers has been partly influenced through the Yao's use of religious texts written in Chinese characters which are typically read with Cantonese pronunciation; the Yao also

employ Cantonese speakers as well as speakers of other Han dialects to teach their children to read the Chinese characters). Very often items which one finds in these languages can also be found in Cantonese and *vice versa*. So, how does one determine the direction of borrowing of a lexical item? We return to this question in a following section.

3.5 Chinese Characters and Cantonese Words

On the one hand, if a word in Cantonese is a so-called colloquial word and does not have a standard Chinese character etymologically associated with it as its written form, then this naturally raises the questions, where did this colloquial word come from and why is it not associated with a standard Chinese character? As mentioned in the Introduction, a significant portion of the basic vocabulary of Cantonese is not cognate with the semantically equivalent words in standard Chinese (or in other Han dialects), and many of these colloquial words are not written with standard Chinese characters. There are several reasons for this state of affairs, but the main reason is that Cantonese-speakers have been taught to read and write the standard Chinese languge and to believe that it is the one and only form for Chinese writing which is based on Mandarin, the national language of China. Since Cantonese speakers have not been taught to write Cantonese, they believe that it is not a written language and that there is no need for a standardized written form of Cantonese, which after all is a regional dialect. However, Cantonese speakers do write Cantonese words with both standard and non-standard Chinese characters in formal and informal texts, such as personal letters, newspaper cartoons, advertising posters, graffiti, etc. Although similar conventions for representing Cantonese words have evolved informally, yet the written form of Cantonese remains nonstandardized.

This brings us to the rather interesting phenomenon of writing colloquial Cantonese morphemes with standard Chinese characters which have been borrowed to represent the colloquial morphemes because of the homophony (or near homophony) between the reading pronunciation of the standard Chinese characters and the colloquial morphemes. Although a Cantonese morpheme is written with a Chinese character, this does not necessarily mean that the Cantonese morpheme is a Chinese word with a Chinese etymology. For example, Cantonese has a well-developed propensity for borrowing words from English and transcribing them with standard Chinese characters; witness tek^{7a}-si^{6*} 的 士 'taxi' which Hong Kong Cantonese has borrowed from English and officially writes with two standard Chinese characters (on street signs, the rear fenders of taxicabs). Now, if one did not know English, one might be rather puzzled about the origin of this word, since the sum of the meanings of these two standard Chinese characters (的 'possessive marker' + 士 'bachelor (in ancient China); a social stratum in ancient China between senior officials (大夫) and the common people (庶民), etc.', Wu 1992:142, 623) have nothing to do with the meaning of 'taxi'. The fact that words borrowed into Cantonese from other languages can be

written with standard Chinese characters can tell us nothing about the origin of the word.

Now, another problem in Cantonese etymology arises with the colloquial Cantonese morpheme which is not cognate with the semantically equivalent word in Mandarin but is witten with a Cantonese character, that is one devised by Cantonese-speakers to represent the Cantonese word. Let us consider the colloquial morpheme $k^h \partial y^1$ 'he, she, it' which is written with the Cantonese character 佢 and which corresponds to Mandarin $t^h a^1$ 他 (with variant homophonous characters, e.g., 她 to refer to female persons and 它 to refer to non-human nouns). Given the wide divergence in the modern pronunciations of these two words and the use of different but semantically equivalent Chinese characters in ancient Chinese texts, it is highly doubtful that anyone would claim that the two have derived from the same historical source. According to Egerod (1974:803), the history of the Cantonese word is indeed very ancient and extends. back over a period of 3,000 years or more; its source has been traced to the nominative form of the third person pronoun in Early Archaic Chinese (the language of approximately 1000 B.C.); this word has been reconstructed as $*g h y \partial g$ or $*k y \partial g$. As for its written form, the association of the character 渠 $k^h \partial y^1$ with the third person pronoun lexeme has been attested since the period of the Three Kingdoms (三國 ca. 220 A.D.) (p.c., Sin Chow-yiu). Interestingly, the reflex of this same etymon also occurs in standard Bangkok Thai as $k^h \ddot{a} w$ 'he, she'. If one believes that Thai is historically related to Chinese, then $k^h \ddot{a} w$ is a distantly inherited form; if not, then $k^h \ddot{a} w$ is a loanword from a Han dialect -- but probably not Cantonese. (If one accepts the Chinese etymology of Cantonese $k^h \partial y^1$, then the question of origin rebounds onto Mandarin $t^h a^1$ -- is it Chinese? -- but this is not a concern here).

3.6 Direction of Borrowing

Another question we must deal with in regard to linguistic forms common to both Cantonese and Tai languages is, how do we determine the direction of borrowing? This is a rather important question but may be next to intractable given the present state of our knowledge of Southeast Asian linguistic history. A reliable answer to this question would certainly be most welcome and help raise the level of our discussion above speculation. As mentioned above, Chinese scholars comparing Cantonese and Tai languages have not addressed this question. Yet they must be operating with some criteria for determining the direction of borrowing, even if these have not been stated explicitly. One criterion seems to be that (1) if the word is not cognate with the semantically equivalent form in standard Chinese and therefore does not have an etymology tied up with a standard Chinese character, and (2) if a phonosemantically similar word appears in at least one Tai language, then *ipso facto*, Cantonese must have borrowed the word from Tai. However, it should be obvious that matching up a Cantonese word with a phonosemantically similar form found in only one or two other Tai

languages cannot settle the question of the direction of borrowing with any reliability because the basis of two-language comparison is far too narrow.

A case in point is found in a very interesting article written by Li J-z. (1990) and provocatively entitled "Is Cantonese an independent language in the Han language group? [粵語是漢語族中的獨立語言嗎?]". In his article Li (p. 39) has claimed that many colloquial Cantonese items cannot be traced back to ancient Chinese nor to any neighboring Chinese dialects, but that their sources can be found in the contemporary Zhuang language. In support of his claim Li compares about a dozen and a half colloquial morphemes in Cantonese with phonosemantically similar forms from only the Yongning dialect of Zhuang. Ouyang (1989:609-610) and Li J-f. (1990; 1992:335-337) have done better by including five or more Tai languages in their comparisons. In my view this is still not good enough. Before we can identify the provenance of a possible Tai loanword in Cantonese our comparison must include as many Southeast Asian languages as possible in order to determine the distribution of the lexical root.

4.0 Comparison of 31 Lexical Roots across Southeast Asian Languages

The remainder of this paper compares 31 lexical roots (#4 and #20 which are phonetic variants do not occur in Yue but do occur in Kejia) across the major language groups of Southeast Asia. Table 2 gives at a glance the linguistic distribution of these 31 lexical roots. The first column lists the lexical root and its English gloss; transcription in upper case letters is intended to represent the general phonetic shape of the root at both historical and contemporary stages. In succeeding columns there follow the lexical forms representing the three southern Han dialect families of Yue, Min, and Kejia, and the major Southeast Asian language groups of Tai, Yao, Austronesian, Tibeto-Burman, and Mon-Khmer. The *Appendix*, Comparative Language Data, lists all of the language material on which this paper has been based.

As for the semantics underlying the cross-linguistic lexical comparisons, the meanings of the compared words from various languages appear to be either quite close or even identical. However, I acknowledge that the glosses given here omit fine semantic differences, e.g., whether the action was done with the sole or the heel of the foot, whether or not the action was done with force, intention, continuously, or just once, etc. Such details are relevant in context, but I do not think their omission significantly affects the comparisons presented below. Since many of my sources are in Chinese, I have cited in the Appendix original Chinese glosses; my English translations are based on Wu 1992.

Among some lexical roots we find systematic phonetic variation of three types: the palatal nasal initial consonant may vary with the palatal glide; the aspirated velar stop initial consonant may vary with its nonaspirated counterpart; and the bilabial nasal final consonant may vary with its homorganic stop. Homorganic alternation of nasal and stop final consonants is a common characterisic of some Southeast Asian languages.

Table 2. Distribution of lexical roots across language groups of Southeast Asia.

Root and Gloss	Yue	Min	Kejia	Tai	Yao	Austronesian	Tibeto-Burman	Mon-Khmer
1. HVŋ *young chicken which has not laid eggs*	$kɐj^1\ hɔ:ŋ^2$ (鷄)口 廣州			$kai^5\ haŋ^6$ Zhuang-Wm	$ka^1\ yaŋ^6$ Bunu			
2. J/Ñ/ɲVŋ *to step on, tread (1)*	$jaŋ^5$ 口 廣州		$naŋ^5$ 梅縣	$jâŋ$ Thai	$jaŋ^2$ Mian		$naŋ$ Atsi	
3. ɲVM *to step on, tread (2)*	$njum^5$ 信宜		$njɛn^3$ 梅縣	$ŋam$ Lao				
4. ɲVP *to step on, tread (3)*	$njap^8$ 梅縣		$njap^8$ 梅縣	$jiàp$ Thai				
5. TSJVP *to wink, blink (1)*	$tsaap^{7b}$ 眨 廣州		$tsiap^8$ 河源	$dzap^7$ Buyi	$dziep^7$ Mian	$kəɟap$ Malay	$tsʰab\text{-}tsʰab$ Tibetan-Wrtn	$cəp^7$ Vietnamese
6. JVP *to wink, blink (2)*	$jəp^{7a}$ 弇 廣州		iap^7 陸川	jap^7 Dong			(mik) jap Luoba	kip-jap Khmu-Thailand
7. KHV *excrement*	$kʰɛ:^1$ 口 廣州			$kʰɨ:$ Thai	kai^3 Biao Min		$kʰje$ Burmese-Wrtn	
8. KHVM *to cover (1)*	$kʰam^3$ 冚 廣州	$kʰam^5$ 廈門	$kʰɐm^5$ 梅縣	$kʰɔm^6$ Li-Quandui	hom^6 Mian		$kʰam$ Limbu	
9. KHVP *to cover (2)*	$kʰɐp^{7a}$ 扱 廣州	$kʰap^7$ 廈門	$kʰɛp^7$ 梅縣	$kʰop^7$ Li-Tongshi			$kʰop^l$ Tibetan-Sgar	
10. KVM *to cover (3)*	kem^3 冚 台山	$kiem^2$ 遂溪	kem^3 梅縣	$kəm^{5a}$ Dong	kom^3 Mian		kum^3 Jinghpo	gam Jeh

Root and Gloss	Yue	Min	Kejia	Tai	Yao	Austronesian	Tibeto-Burman	Mon-Khmer
11. KVP *to cover (4)*	$kɐp^{7a}$ 陽江	ka^{7} 廈門		kap^{8} Maonan		kop-kóp Ilokano	kap^{1} Tibetan-Lhasa	$kɔ̀p$ Mon
12. K(H)L/RVM *to cover (5)*	$kʰiam^{3}$ 恩平	$kaŋ^{5}\ laŋ^{6}$ 松溪	$kʰiam^{2}$ 從化	$kʰlum$ Thai			$kʰrum$ Burmese-Wrtn	klap Jeh
13. K(H)VM *to press down*	$kɐm^{6}$ □ 廣州		$kʰim^{5}$ □ 梅縣	kam^{6} Buyi				
14. KVP *frog*	$kɐp^{7b}\ naːi^{3}$ 蛤蟆 廣州	$uoŋ^{6}\ kap^{7a}$ 中山		kòp Thai	kop^{7} Mian			
15. K(H)VP *to bite*	$kʰɐp^{8}$ □ 廣州			$kʰòp$ Thai				
16. K/LVK *pen, coop*	lok^{8} □ 廣州			$kʰɔ́k$ Thai			kok^{55} Achang	
17. LVK *to* *scald, boil, burn*	lok^{8} 渋 廣州	lua^{7} 潮州	luk^{8} 梅縣	luâk Thai	$łu^{7}$ Mian		$lɔk^{55}$ Achang	$luak^{8}$ Vietnamese
18. LVM *to* *collapse, topple,* *fall down (bldg.)*	lam^{5} 冧 廣州			lóm Thai				rəlum Khmer
19. LVM *to step across (1)*	lam^{5} □ 廣州	lam^{5} 中山	lam^{5} 香港	$ɲam^{2}$ Li-Xifang	lam^{5} Mian			
20. KHVM *to* *step across (2)*			$kʰiam^{5}$ 梅縣	$kʰâm$ Thai	$tɕʰam^{5}$ Mian			kham Khmu-Mamnai
21. LVT *to slip* *off, come off*	$lɐt^{7a}$ 甩 廣州	lut^{7} 廈門	lut^{7} 梅縣	lùt Thai	$łut^{7}$ Mian	lolos Indonesian	lut^{21} Zaiwa	lot^{8} Vietnamese

Root and Gloss	Yue	Min	Kejia	Tai	Yao	Austronesian	Tibeto-Burman	Mon-Khmer
22. MVT *to peel off, pinch, nip, pluck*	mit^{7a} 搣 廣州		mit^7 河源	mit^7 Be-Lingao	pit^7 Mian	mengepit Indonesian		but^7 Vietnamese
23. NV *this, here*	ni:^{1}/nej^{1}呢 廣州		li^3 梅縣	ni: Thai	nai^3 Mian	ini Indonesian	ni:^{4} Tibetan-Sgerrtse	nih Khmer
24. NVM *to think, ponder*	nɐm^3 諗 廣州		njam^5 梅縣	nam^5 Zhuang-Wm	nam^3 Mian			
25. NVM *soft, tender*	nɐm^2 腍 廣州		njam^2 惠州	nûm, nîm Thai			nám Atsi	ka^{231} ɲɔm^{35} Plang-Mane
26. ŋVŋ *stupid, muddled*	ŋɔŋ^6 戇 廣州	ŋoŋ^6 莆田	ŋɔŋ^5 永定	ŋoŋ Thai	ŋɔŋ^6 Bunu			ŋɔŋ^1 Jing
27. TVM *to hang down; low*	tɐm^5 □ 廣州		tem^1 梅縣	tâm Thai				tăm Khmu-Nanqian
28. TVM *to pound, beat (1)*	tɐm^3 抌 廣州	tam^3 中山		tam Thai				dam^1 Vietnamese
29. TVP *to pound, beat (2)*	tɐp^8 揸 廣州		tep^8 西河	tap^8 Buyi	ta:p^8 Mian	tatab Javanese	tap^2 Tibetan-Lhasa	dap^8 Vietnamese
30. TVM *to stamp (foot) (1)*	tɐm^6 □ 廣州	tɐm^3 中山	tɔm^3 梅縣	tam^6 Buyi	dam^6 Mian			zamp^3 Vietnamese
31. T(H)VP *to stamp (foot) (2)*	tap^8 □ 英德	ta:p$^{7b/8}$ 踏 中山	tʰap^8 梅縣	ta:p^{8a} Dong		tebteb Balinese	tap^1 Tibetan-Lhasa	da:p^8 Vietnamese

Li Fang Kuei's Proto-Tai initials, tone categories, and glosses have been cited where it is possible to compare this material with Cantonese. Although Li did not explicitly reconstruct Proto-Tai rimes and combine them with the initials into lexical roots, I have taken this further step on my own. As the following discussion of the individual roots will show, phonetic correspondences among Cantonese, Tai languages, and Proto-Tai tend to be relatively close or the same; e.g., the sonorant initials generally agree; Cantonese velar stop initials usually correspond to Proto-Tai velar fricatives; Cantonese alveolar nasal and stop initials correspond to Proto-Tai alveolar initials; and Cantonese and Proto-Tai final consonants are identical.

The tone systems of both Proto-Tai and Cantonese have four tone categories. As stated earlier, some historical linguists believe that Tai (along with Miao-Yao, Vietnamese, Tibeto-Burman) acquired the four-tone system under the influence of Chinese. Table 3 below sets out the correspondence between the four tone categories of Cantonese and Li's Proto-Tai. The Cantonese tone categories are named Level, Rising, Going, and Entering (the English translation of the traditional Chinese names) and are numbered 1 through 8. The odd numbered tones correspond to an upper register associated with historically voiceless initial consonants, and the even numbers correspond to a lower register associated with historically voiced initial consonants. However, as mentioned in a previous section, the Cantonese sonorants (m-, n-, η-, l-) occur in both registers (it is significant that words with this set of initials occurring in the upper register generally belong to the colloquial layer of Cantonese). Li labeled the four Proto-Tai tones as A, B, C, and D. However, for a reason only known to himself Li made Tai tone C correspond to the Chinese Rising tone (second tone category) and Tai tone B to the Chinese Going tone (third tone category), so that the two systems do not coincide without adjustment (p.c., Chris Court). In Li's system 1 indicates that the Proto-Tai initial was voiceless and 2 that it was voiced.

Table 3. Correspondence of Cantonese and Proto-Tai Tone Categories.

	Level 平	A	Rising 上	C	Going 去	B	Entering 入	D
Historically Voiceless Initials	1	A1	3	C1	5	B1	7a 7b	D1S D1L
Historically Voiced Initials	2	A2	4	C2	6	B2	8	D2S D2L

A few words regarding my sources: the bulk of the Yue material comes from the surveys of the Yue dialects by Zhan and Cheung (1988, 1994), but I have also

relied heavily on the Cantonese-Putonghua dictiónary of Rao et al 1981, Meyer and Wempe's 1947 Cantonese-English dictionary, Zeng's 1986 Cantonese-English glossary, Lau's 1977 Cantonese-English dictionary, and Luo's 1987 study of Yue-Xinyi; for the Kejia dialects I have drawn upon the survey of the Gan-Kejia dialects by Li and Zhang, Maciver's 1926 Kejia-English dictionary, and Hashimoto's 1972 study. My main sources for the Tai material have been the 1985 *Collected Vocabulary of the Zhuang-Dong Language Group* (ZDYC), Wang et al 1984, Haas 1985, Li 1977, Benedict 1975, the 1959 Buyi dialects survey (BYDB), Ouyang and Zheng's 1983 survey of Li dialects, the study of Be dialects by Zhang et al 1985, Hashimoto's 1980 study of Be, the 1984 *Zhuang-Han Dictionary* (ZHCH); my student Ittisak very helpfully clarified for me the differences in the meanings of similar Thai words. For Miao-Yao languages I have made use of the 1987 survey of Miao-Yao dialects (MYFC), Mao's 1992 Chinese-Yao dictionary, and Downer's 1973 study of Yao-Mian; for Austronesian languages I have used Echols and Shadily's 1989 Indonesian-English dictionary, the Chinese-Malay dictionary of Tan et al 1986, and various other sources; for Tibeto-Burman languages Qu and Tan 1983, Huang 1992, Yabu's 1982 Atsi vocabulary, and the 1992 Chinese-Zaiwa dictionary. For Mon-Khmer languages I have consulted the very comprehensive survey of these languages in China by Yan and Zhou (1995), Nguyen's 1966 Vietnamese-English dictionary, Mrs. Le Thip Phuong Mai, a postgraduate student from Hanoi and now studying at Mahidol University, an early draft of Amphon's 1995 M.A. thesis on the Ban Nawattai variety of Khmer, Jacob 1968 and 1985, my colleague Naraset who teaches Khmer, and my former student Prakop who is a native-speaker of Khmer.

1. **HVŋ** *young chicken, pullet*

As indicated by the material in the Appendix, this root only occurs in Yue dialects, e.g., Guangzhou *hɔːŋ³*. Functioning as a modifier, the morpheme follows the noun *kɐjˡ* 鷄 'chicken'; although this is an unusual sequence for standard Chinese, Yue precisely mimics the Tai pattern. The table below lists semantically equivalent forms from Min and Kejia which must have had a phonetically quite different source (reconstructed here as **luan*) and are unrelated to the Yue root:

Table 4. Kejia and Min forms for 'young (chicken) which has not yet laid eggs'.

Kejia-Huizhou	(ke[1]) lan[6] (Zhan, Cheung 1988:95)
Kejia-Dongguan	(kai[1]) lɔn[5] (Zhan, Cheung 1988:95)
Kejia-Zhongshan	(kai[1]) nɔn[5] (Zhan, Cheung 1988:95)
Min-Zhongshan	(kai[1]) nan[1] kiɛn[3] (Zhan, Cheung 1988:95)
Min-Quanzhou	(kue[1]) luã[5] (Lin 1993:215)
Min-Xiamen	(kue[1]) nuã[5] (Zhou 1993:108)]

Li (1977:209, #32; 282, #5) has reconstructed Proto-Tai voiceless velar fricative initial *x-, B1 tone, and diphthong *-ɯa- which yield *xɯaŋ 'young (chicken)'. As indicated in the Appendix, the root occurs in a number of Tai languages (e.g., Wuming haːŋ⁶, Dehong xəŋ⁵). The initials in Proto-Tai and Cantonese are close, but not the tones (Cantonese 3 corresponds to Tai C1). However, on the basis of its widespread distribution in Tai, its limited distribution in Yue, and the unrelated root found in Min and Kejia, I believe that the Yue root has been borrowed from Tai.

2. J/NVŋ (1), 3. ɲVM (2), 4. ɲVP (3) *to step on, tread*

These three roots appear to be phonetic variants with alternation of the initial and final consonants. The first two are found in Yue (Guangzhou jaːŋ⁵, Xinyi njam⁵) and the third in Kejia (Meixian njap⁸). On the basis of the Yue dialect material available to me, I have only found the first root in standard Cantonese. Possibly showing a phonetic and semantic correspondence are Kejia forms (Meixian naŋ⁵) with the alveolar nasal initial and rimes similar to the Cantonese form. In several Tai languages we find phonetically similar forms with corresponding palatal glide and palatal nasal initials and rimes similar to the Chinese dialect forms, e.g., Thai jâːŋ 'to step', Lao ɲaŋ 'to walk', 'to step on'. Li (1977:178, #5) has reconstructed Proto-Tai *j- and B2 for the first root to give *jaŋ 'to take a step, step'. While the Cantonese and Tai initials and rimes agree, the tones do not (Cantonese 5 = Tai B1). Several Tibeto-Burman languages also show semantically similar forms with nasal initial (e.g., Zaiwa naŋ²¹). I believe the Cantonese form is a Tai loan. However, the closely similar forms in Kejia and Tibeto-Burman suggest that a deeper relationship may be involved here.

Li (1977:173, #1;181, #9) has also reconstructed initials and tones for two phonosemantically related roots: Proto-Tai *ɲam B2 tone 'to tread upon' and *ʔjam B1 tone 'to step on'. With these roots we compare Yue-Xinyi njɛm⁵ 'to step on, trample'. We also note the similarity of forms with homorganic final stop; cf. Kejia njap⁸ 'to step, tread', Thai jiàp and Lao ɲiap¹ 'to take a step, step on'.

5. TSJVP (1), 6. JVP (2) *to wink, blink*

In Bauer 1992a ("Winkin', blinkin', and nod") I proposed that these two roots are derived from an early Southeast Asian root KVSLVP which is associated with the meanings 'to move up and down, to flash on and off' and lexemes denoting the blinking of the eye, the flashing of lightning, and the waving of the hand. In that paper I examined a set of phonetically-related but hypothetical lexical roots from which may have been derived a number of phonosemantically similar forms for to blink (eye), to flash (lightning), and to wave (hand) which are found in many contemporary Southeast Asian languages. I coined the term *Panasea* (Pan-Southeast Asianism) to refer to the "same" word which has widespread geographical and linguistic distribution. The present paper is concerned with only the first meaning 'to blink, wink'.

Yang (1968:22) has reconstructed Old Chinese *ktsiap 'to wink, blink' to which I believe the Yue and Kejia forms with affricate and fricative initials are related. The Yue and Kejia forms with palatal approximant initials belong to the JVP root which is a phonetic variant and ultimately derived from TSJVP. Li (1977:181, #12) has reconstructed Proto-Tai *ʔj- and D1S tone to give *ʔjap 'to move quickly up and down (like jaws in chewing)' which I believe is the source of phonetically similar forms meaning 'to blink' in Tai languages. Li J-f. (1990:73) has also observed that Cantonese jɐp⁷ᵃ closely resembles forms in Zhuang, Dai, Dong, and Shui (and Buyi dialects as well). However, phonetically similar forms occur not only in Tai but also in Yao, Austronesian, Tibeto-Burman, and Mon-Khmer. TSJVP and JVP must be very old Southeast Asian roots, and it is unlikely that the Yue forms descend from Tai.

7. KHV excrement

Although my Yue dialects material lists this root in only two dialects, it is probably common throughout Yue. It also seems to be quite widespread in Tai and Tibeto-Burman. On the basis of the close phonosemantic similarity of Yue-Guangzhou kʰɛ⁻¹, Thai kʰîː, Zaiwa kʰji²¹, Burmese-Written kʰje³, and Leqi kʰjeⁱ⁵⁵, I believe that the history of this word involves all three language groups and that the Cantonese word is not a Tai loan. Li (1977:209, #28) has reconstructed a pair of voiceless and voiced velar fricative initials in variation, *x- ~ *γ-, C1 tone, and diphthong *-ei for Proto-Tai *xei ~ *γei. The Cantonese and Tai initials share the same place of articulation but not the manner; the tones do not agree (Cantonese 1 = Tai A1) but the rimes are similar. Benedict (1972:39, #125) has reconstructed Proto-Tibeto-Burman *klij (and *kləj). The occurrence of the velar stop initial in Yue and numerous Tai and Tibeto-Burman languages suggests to me that the Tai source also had a velar stop intial; in addition, the change of a fricative initial into a stop seems less natural than the reverse.

8. KHVM (1), 9. KHVP (2), 10. KVM (3), 11. KVP (4), 12. K(H)L/RM/P (5)
to cover

Luo (1996:856-858) has produced a detailed analysis based on a number of Tai languages to show that all of these roots belong to a Tai word family. However, in a previous paper (1994b) I demonstrated that these five roots (and a few others) constitute a set of closely related semantic and phonetic variants that extends across all the southern Han dialects as well as all the major Southeast Asian language families. A good part of the material from the earlier paper appears in the Appendix at the end of this paper.

In my 1987 study I carefully steered away from giving a Tai etymology to Cantonese kʰɐm³ and related forms in the Min, Kejia, and Wu dialects; I accepted (but expanded) the etymological relationship claimed by Li and Chen between the Min forms with aspirated velar initials and the Chinese character 軗 for which the following entry appeared in the 11th century rime dictionary, Jiyun: 軗, 陷韵, 物

相 值 合, 口 陷 切. On the basis of these phonetic specifications, the Ancient Chinese reconstruction for this character is *k^hæm which may correspond closely with the modern dialect forms; however, the gloss given in the entry does not precisely mean 'to cover' but for 'two things to come together and be closed up'. Bai (1980:217) suggested there was a connection between Cantonese k^hɐm³ and another entry from the *Jiyun*, namely, 籲, 感 韵, 古 禫 切, 蓋 也; this indicates that the character 籲 can be reconstructed as *kəm and that its gloss is 'to cover'. Bai noted the discrepancy between the historical initial which was unaspirated and the aspirated initial of the modern Cantonese form.

You (1992:174-175), noting that the character does not occur in the *Guangyun* or in other historical documents or in the colloquial language of north China but is limited to the southern dialects, rejected this ancient Chinese character as the source of the Cantonese, Min, and Wenzhou forms. Comparing Wu-Wenzhou kaŋ³ to Proto-Tai *gum (no source cited for this reconstruction), he implied that it may have had a Tai origin. As evidence that Cantonese k^hɐm³ is a Tai loan, Ouyang (1989:609) compared it with similar forms in several Tai languages. For the first root Li (1977:215, #19) has reconstructed Proto-Tai *ɣ- and C2 tone for *ɣuɰ, the Cantonese and Tai initials agree in the place of articulation but not in the manner; the tone categories coincide but not the voicing of the historical initials; the vowels differ but the final consonants are identical.

The fruits of my own subsequent research have led me to believe that whether or not there is any link between either of the two ancient Chinese characters and the modern Chinese dialect forms is largely irrelevant to their etymology. On the basis of the numerous phonosemantically similar forms found in a variety of Southeast Asian languages, I have set up two pairs of roots with alternation of homorganic nasal and final stop consonants and aspirated and unaspirated voiceless velar initial consonants. Bauer 1994 ("*Cover, *Conceal their secrets reveal") proposed that these must be very old Southeast Asian roots or Panaseas and not Tai loanwords in the Han dialects, as only a very long history can explain their widespread linguistic and geographical distribution.

From his comparison of Cantonese k^hɐp⁷ᵃ 罩 住 to be covered' with Zhuang-Yongning k^hɐp⁵⁵ (with the same meaning), Li (1990:39) was led to believe that the Cantonese word must represent a Tai substratum since he could not relate the Cantonese word to any etymon in Ancient or Archaic Chinese, or to phonosemantically similar forms in any other Han dialects or non-Sinitic languages. However, the limitation of his very narrow base of comparison is readily apparent. Inspection of the list of lexical items in the Appendix for the Kejia and Min dialects and Tai, Austronesian, Tibeto-Burman, and Mon-Khmer languages unequivocally demonstrates the inaccuracy of his statement. Cantonese k^hɐp⁷ᵃ may be related in some way to the Chinese graph 蓋 which has been reconstructed for the Archaic Chinese period by Karlgren (1957:170-171, GSR #642q) as *g^hap with the meanings 'to thatch, cover' and *kâb with the meanings 'to cover, conceal, a cover (of a car)'. Bodman (1980:49) has reconstructed the

same graph for the Pre-Chinese (or Proto-Chinese) period as *kap/kâp 'to cover, thatch' and for Old Chinese (= Karlgren's Archaic Chinese) as *kaps 'to cover, conceal; a cover, lid'. Bodman has also compared his reconstructed Chinese forms with Written-Tibetan figebs 'to cover' and khebs 'a cover' to suggest a historical relationship. Chinese forms with aspirated velar initial are found not only in Yue but also in some Kejia and Min dialects. Tai languages with similar forms include Zhuang-Yongning and several varieties of Li (e.g., Tongshi k^hop^7). On the Tibeto-Burman side, phonetically similar forms with the meanings 'a cover' and/or 'to cover' occur in some Tibetan dialects (e.g., Lhasa k^hep^1) and Limbu (k^hap). However, a search through the reference materials on the Miao-Yao languages which are available to me has turned up no comparable items.

KVM is represented with Yue-Taishan kem^3, Kejia-Meixian kem^3, Min-Sixian $kiem^2$. A number of Tai languages have corresponding forms, e.g., Mulao $kəm^6$, Buyi kom^6. Li did not reconstruct a comparable Proto-Tai root, but he (1977:250, #14; 272, #4) has Proto-Tai *h-, B1 tone, and *-o- which give *hom 'to cover up' based on forms with glottal fricative initial and bilabial nasal final in several Tai languages. Tibeto-Burman is represented by Jinghpo $mă^1$ kum^3 and Mon-Khmer by Jeh gam.

Finally, on the basis of forms with consonant cluster initials in Tai (Thai $k^hlu:m$, $k^hrɔ:m$, $k^hrɔ́:p$) and Tibeto-Burman (Burmese-Written k^hrum, Kachin grup), I have set up the root K(H)L/RVM/P to which may be related Yue-Enping k^hiam^3 and Kejia-Conghua k^hiam^2 with the palatal onglides corresponding to the second consonant of the cluster, and Min bisyllabic forms with alternating k- and l- initial consonants, e.g., Jianou $kaiŋ^5$ $laiŋ^6$. Li (1977:231, #1) has reconstructed Proto-Tai cluster initial *gr- and B2 tone for *grɔm 'to put over, cover up'. The Yue, Kejia, and Tai initials agree in the place and manner of articulation but not in the voicing or in the tone category, but the Min form may correspond with the Tai initial, final, and tone category.

13. K(H)VM to press down

For this word Yue-Guangzhou kem^6 has an unaspirated velar stop initial, Kejia- Meixian k^him^5 an aspirated velar stop initial, and the Tai languages both unaspirated and aspirated velar stop and fricative initials, e.g., Wuming kam^6, Thai $k^hôm$, Be-Lingao $xɔm^4$. Li (1977:208, #16) has reconstructed Proto-Tai voiceless velar fricative initial *x- and B1 tone for *xom 'to press down from top, oppress'. For the same reason stated above, I believe the etymon is more likely to have had a velar stop initial, since the change of a stop into a fricative seems more natural than the reverse. The Tai and Han tone categories coincide, although Cantonese tone 6 indicates a historically voiced initial. I believe that both the Kejia and Yue reflexes represent different phonetic developments from the same etymon which may have been borrowed from Tai.

14. KVP *frog*

According to Karlgren (1957:179), 蛤 is reconstructed as *$k\partial p$* for Archaic Chinese and as *$k\hat{a}p$* for Ancient Chinese and meant 'oyster, mussel'. The Chinese character 蛤 is used to represent the morpheme 'frog' in Mandarin and Yue, but it has been borrowed for its sound and not its meaning. This root KVP has widespread distribution in both Yue and Tai, cf. Guangzhou *kɐp⁷ᵃ na³* (the second syllable may correspond to the Tai morpheme for 'field') and Wuming *kop⁷*. Li (1977:188, #42; 272, #11) has reconstructed Proto-Tai *k-, D1S tone, and *$-o$- which yield *kop. The initial, final, and tone agree in Yue and Tai, and I believe the Yue forms go back to a Tai source. Li J-f. (1990:73) has also compared the Cantonese word with forms from five Tai languages to support this view. However, the origin of the root may possibly lie in sound symbolism whereby it was an attempt to represent the sounds that frogs make.

15. K(H)VP *to bite*

On the basis of my Yue sources, I have found this root represented only in standard Cantonese (*kʰɐp⁸*), but it occurs quite widely among the Tai languages (cf. Thai *kʰòp*). Li (1977:209, #31; 269, #13) has reconstructed Proto-Tai *x-, D1S, and *$-e$- for *xep 'to bite'. The Cantonese and Tai initials agree in their place of articulation but not in manner; the tone categories coincide but the low tone in Cantonese indicates a historically voiced initial. To support his claim that the Cantonese word has a Tai source, Li J-f. (1992:336) compared it with similar forms in several Tai languages. While I agree, I also think the phonetic shape of the word in both languages may be the product of coincidental sound symbolism.

16. K/LVK *pen, coop for domestic animals*

Alternation between the lateral approximant and velar stop initials of this root is intended to imply a relationship between Yue and Tai; cf. Yue-Guangzhou *lok⁸*, Be-Lingao *luk⁸*, Thai *kʰɔ́:k* 'enclosure, pen, sty (for animals)'. Li (1977:215, #16) has reconstructed Proto-Tai initial *y- and D2L tone for *$yɔ:k$ 'enclosure for animals'. Although the Cantonese and Tai initials do not agree in either their place or manner of articulation, the tone categories do agree. Apparently influenced by the Cantonese initial, Li J-f. (1990:75; 1992:337) has reconstructed Proto-Tai *rok and has indicated that the Cantonese word has a Tai source. However, given that forms with velar stop initials are also found in Tibeto-Burman languages (cf. Lahu *kʰɔ́ʔ*, Achang *kok⁵⁵*), this language group must also figure into the relationship. In this regard, Matisoff (1972:31, #16) has reconstructed Proto-Lolo-Burmese *$krok$ which has led me to conclude that the Yue, Tai, and Tibeto-Burman forms are all ultimately related.

17. LVK *to scald, boil, burn*

In an earlier publication (Bauer 1987:104) I had proposed that Yue-Guangzhou *lok⁸*, Min-Chaozhou *luaī⁸*, and Kejia-Meixian *luk⁸* all derived from

a Tai source. However, additional research has prompted me to change my mind.
Table 2 and the Appendix indicate phonosemantically similar forms occur in Tai,
Yao, Tibeto-Burman, and Mon-Khmer; such wide distribution leads me to
believe this is a very old Southeast Asian root. Li (1977) did not reconstruct this
root for Proto-Tai.

18. LVM *to collapse, topple, fall down (building)*

I have only found this root represented in standard Cantonese with *lɛm⁵* 'to
collapse, topple down, go bankrupt' in my search through the Yue dialect
materials available to me. As indicated in the Appendix, it is well-represented in
the Tai languages (cf. Thai *lóm*, Maonan *lam⁵*) and occurs in two Mon-Khmer
languages (Khmer *rəlum*, Jeh *rɔ:lǎm*). Li (1977:134, #19; 269, #11) has
reconstructed Proto-Tai initial **l-*, C2 tone, and **-e-* for **lem* 'to fall, topple'.
Cantonese and Tai agree in the initial and final but not in the tone (Cantonese 5 =
Tai B1). Both Ouyang (1989:609) and Li J-f. (1990:72-73) have compared the
Cantonese word with forms in several Tai languages to indicate that it is a Tai
loan and I agree with them.

Forms in the table below indicate there is a second and probably related Tai
root with identical phonetic shape, but different tone, and the meaning 'to sink,
cave in' which is rather similar to the first root (the change in tone may
correspond to the change in meaning. Li (1977:138, #25) has reconstructed Proto-
Tai **ɬ-* and B1 tone for **ɬom* 'mud; sink in mud' to which the following forms
may be related. Old Javanese *kelem* 'to sink in' (Blust 1986:6) would seem to
provide evidence in support of the Tai-Austronesian relationship.

Table 5. Tai forms for 'to sink, cave in' possibly derived from **ɬom*.

Buyi-Dujunfuxi	luam⁵ '陷 to sink, cave in' (BYDB 1959:273)	
Zhuang-Wuming	lom¹'陷下去 to sink, cave in' (ZDYC 1985:252)	
Dai-Dehong	lom⁵ '陷下去 to sink, cave in' (ZDYC 1985:252)	
Thai	lôm 'to sink, capsize, overturn' (Haas 1985:475)	
Buyi-Chengguan	lɔm⁵ '陷 (下去) to sink, cave in' (ZDYC 1985:252)	
Shui	ləm⁶ '陷下去 to sink, cave in' (ZDYC 1985:252)	
Maonan	ləm¹ '陷下去 to sink, cave in' (ZDYC 1985:252)	
Dong	ləm⁶ '陷下去 to sink, cave in' (ZDYC 1985:252)	
Li-Xifang	ɬom¹ '陷落 to sink down' (Ouyang, Zheng 1983:500)	
Li-Baisha	ɬom¹ '陷落 to sink down' (Ouyang, Zheng 1983:500)	

In several Tai languages [but not in the Southwest branch (Li 1977:101, note
#32)] we find still a third root with the same meanings as the other two but with
the voiceless dental stop initial *t-*; with this root we can compare Li's
reconstructed Proto-Tai initial **t-* and B1 tone for **tom* 'to collapse, fall' (Li
1977:99, #32). Interestingly, all three roots are represented in Zhuang-Wuming.

Table 6. Tai forms 'to collapse, topple down' possibly derived from **tom.*

Zhuang-Longzhou	tum⁵ '倒塌 to collapse, topple down' (Wang et al 1984:854)

Zhuang-Longzhou tum^5 '倒塌 to collapse, topple down' (Wang et al 1984:854)
Zhuang-Wuming tom^5 '倒塌 to collapse, topple down' (Wang et al 1984:854)
Zhuang-Wuming tom^5 '塌 (下去) to collapse, fall down, cave in'
　　　　　　　　(ZDYC 1985:243)
Buyi-Chengguan tɔm^5 '塌 (下去) to collapse, fall down, cave in'
　　　　　　　　(ZDYC 1985:243)
Po-ai tɔm B1 'to collapse, fall' (Li 1977:99, #32)
Be-Lingao dum^3 '塌 (下去) to collapse, fall down, cave in'
　　　　　　　　(ZDYC 1985:243)

Although two different Proto-Tai initials have been reconstructed, all three sets of forms with the lateral and dental intials may have descended from the same etymon which had the consonant cluster **tl-* and existed in a language that preceded Proto-Tai and was in contact with Old Southern Chinese. Luo (1996:854-855) has shown that these roots constitute a Tai word family.

Finally, we note the existence in a few Tai languages of phonetically distinct words which mean 'to collapse, fall down' and are unrelated to the above roots:

Table 7. Phonetically diverse forms for 'to collapse, fall down' in Tai.

Dai-Xishuangbanna luʔ^7, kun^6 '倒塌 to collapse, fall down' (Wang et al 1984:854)
Dai-Dehong lan^5 '倒塌 to collapse, fall down' (Wang et al 1984:854)
Dong paŋ^1 '倒塌 to collapse, fall down' (Wang et al 1984:854)
Shui paŋ^1 '倒塌 to collapse, fall down' (Wang et al 1984:854)
Li-Tongshi thou^5 '倒塌 collapse, fall down' (Wang et al 1984:854)

19. LVM (1), 20. KHVM *to step across* (2)

The first root is well-represented in Yue, e.g., Guangzhou *la:m⁵*, but in only a few Kejia dialects, e.g., Xianggang *lam⁵/⁶*; on the Tai and Yao sides, we find Li-Xifang *ɲa:m²* and Yao-Mian *la:m⁵*. The second root with the aspirated velar stop initial occurs in seven Kejia dialects, e.g., Meixian *kʰiam⁵*, and it is also widely found in Tai where it appears with the aspirated velar stop, velar fricative, and glottal fricative initials, cf. Thai *kʰâ:m*, Tai Lo *xa:m³*, Buyi *ɣa:m³*, Li-Zhongsha *ha:m²*. It also occurs as a probable Tai loan in a few Mon-Khmer languages, e.g., Khmu-Manmai *kham*. Li (1977:208, #2) has reconstructed Proto-Tai **x-* and C1 tone for **xa:m* 'to cross, step across'; however, for the same reasons stated above I think the initial was a velar stop rather than a fricative. The Kejia and Tai initials agree in the place of articulation but not in the manner, and the tone categories do not agree. However, it is my view that the Yue, Kejia, and Tai reflexes represent different developments of the same etymon which had the

cluster initial *$k^h l$- and occurred in a language that preceded Proto-Tai and was in contact with Old Southern Chinese: Yue lost k^h- which left l-; Kejia retained k^h- and palatalized -l- to the i-onglide; when Tai inherited the root, -l- had been lost, so that it either kept k^h- or fricativized it. Yao-Mian has an alternant form, $t\varphi^h a\!:\!m^5$ (Mao 1992:134-135), which reflects a further development of both affrication and palatalization caused by the i-onglide which was lost.

A third root, LVP 'to step across', with the homorganic final stop occurs in Yue (e.g., Huaxian lap^{7b}) and Kejia (e.g., Huizhou lap^7) but not in any other language group.

21. LVT *to slip off, come off*

Tone 7 associated with Yue-Guangzhou $l\emph{ɐt}^{7a}$, Min-Xiamen lut^7, and Kejia-Meixian lut^7 'to slip off' indicates that the initial of the etymon was a voiceless lateral; this agrees with Li's Proto-Tai *l- and D1L tone which give *lut 'to slip off, come off' (Li 1977:138, #32). The very close phonosemantic similarity of modern Han and Tai forms led me to propose in an early publication (Bauer 1987:105) that the forms in the southern Chinese dialect families derived from a Tai source; however, in a 1989 paper ("Slip, slide, 'n' away") I changed my mind after the discovery that this root is another *Panasea* (cf. TSVP, JVP 'to wink, blink' and K(H)VM, K(H)VP 'to cover' above). LVT not only has a very archaic history with Proto-Chinese *$hlot$ and Old Chinese *lot 脱 'to peel off, take off clothes, relieve, careless' (Bodman 1980:103-104), but it also has a very wide distribution across all the major language groups of Southeast Asia (cf. Table 2 and the Appendix). Benedict (1975:384) has reconstructed Proto-Austro-Tai *(q/)luts/luts 'to slip off, slip away'; and Blust (1988:124, 148) has reconstructed Proto-Austronesian *lus 'to slip off' and *rus 'to slip or slide off'.

22. MVT *to peel off, pluck, nip, pinch*

The surveys of the Yue dialects did not include this lexeme in their set of basic verbs, so we only have the standard Cantonese form. However, the Gan-Kejia dialects survey did include it and it is well represented in both dialect families. We also find the root widespread in Tai, Li, and Be languages, and this led Benedict (1975:355) (but not Li) to reconstruct Proto-Tai *$\mathcal{P}bit$ and Proto-Li *$mbi\!:\!t$. The root also occurs in Yao and possibly in two Austronesian languages and in Mon-Khmer. The widespread distribution of this root leads me to believe it must be an old word in Southeast Asia and not a Tai loan. However, Ouyang (1989:609) has claimed that Tai is the source of Cantonese $mi\!:\!t^{7a}$ and compared it with forms in six Tai languages.

23. NV *this*

This root is found in most of the Yue dialects around Guangzhou and north of the provincial captial but not in the west, and it occurs in only a few Kejia dialects. The root is widely distributed in the Tai languages, and Benedict

(1975:408) has reconstructed Proto-Tai *ni. Li (1977:111, #14; 286, #1) reconstructed the initial and tone as *n- and C2, respectively, but concluded the diphthong was uncertain. The Cantonese word and Benedict's Tai root agree in the phonetics but not in the tones (Yue 1 = A1). Because of this root's skewed distribution in the Han dialects, I (1987:106) have previously claimed that the Yue and Kejia forms were Tai loanwords. Believing Tai has been its source, Ouyang (1989:609) has also compared Cantonese *ni:^l* with forms in several Tai languages. However, my subsequent research has found that this root occurs in all the major language groups of Southeast Asia. Does this mean that Cantonese *ni:^l* cannot be a Tai loanword? Let us take a close look at the word for 'this' in the Yue dialects and other language groups before answering this question.

The Yue dialects show very interesting developments in the demonstrative pronouns and these are worth examining in detail. The Appendix indicates the central and northern Yue dialects all have phonetically similar forms of ni, nei, li, lei, lɐi with tone 1 which usually occurs with voiceless initials; however, this is not an anomaly because, as mentioned earlier, colloquial Yue words with sonorant initials typically occur with the upper register tones. In phonetic contrast, the corresponding forms in the Siyi group of Yue dialects and those of neighbornning areas and western Guangdong must have derived from a different source, since they have velar stop initials which are aspirated in the Si-yi dialects but unaspirated in the western dialects:

Table 8. Forms for 'this' in Yue-Siyi group and neighboring dialects (Zhan, Cheung 1988:425).

Yue-Xinhui	khui^4 kɔ^1 □ 個
Yue-Taishan	kh^{u}ɔi^4 k^{u}ɔi^1 □□
Yue-Kaiping	kh^{u}ɔi^4 k^{u}ɔi^1 □□
Yue-Enping	kh^{u}a^5 k^{u}a^1 □ 個
Yue-Doumenzhen	kh^{u}ɔ^4 k^{u}ɔ^1 □ 個
Yue-Jiangmen	khɐt^8 kai^3 □□
Yue-Heshan	khai^4 kɔ^1 □ 個

Table 9. Forms for 'this' in Yue dialects of western Guangdong (Bai 1994:42).

Yue-Huaiji (冷坑)	kɔ^5(⊣)-kɔ^5 個 個
Yue-Xinyi	kɔ^5(⊣)-tsek^7 個 隻
Yue-Gaozhou	kei^3(⊣)-kɔ^5 己 個
Yue-Lianjiang	kɔ^5(⊣)-kɔ^5 個 個
Yue-Yangchun	kɔ^3(⊣)-kɔ^5 個 個

In contrast, the western Yue dialects use three different forms for 'that': Huaiji uses the same syllable but changes the tone on kɔ⊣33 個 'this' from mid-level to low rising kɔ⊿13 個 'that'; this dialect's morphological use of tone preserves a

pattern that was probably quite common at an earlier stage of Yue. On the other hand, the Xinyi, Gaozhou and Yangchun dialects all use the same morphosyllable *na* 那 (which belongs to a different tone category in each dialect and may be a borrowing from standard Chinese); Lianjiang uses the morphosyllable *nu¹* which does not have a Chinese character associated with it and its origin is uncertain.

The above material provides us with highly useful pieces of evidence in the task of identifying the origin of Cantonese *ni:¹* and its related forms. First, we note that this root is restricted to the Yue dialects in and around the Guangzhou area and north of it. Second, the Siyi group of Yue dialects and those of western Guangdong use a phonetically different form to represent 'this'. Furthermore, forms phonetically similar to Cantonese *ni:¹* are not typically found in other dialect families of Min and Kejia [but we note the occurrence of Kejia-Huizhou *ni¹ tsak⁷* and Kejia-Conghua *tɨ³ tsak⁷* which are probably the result of contact borrowing in Yue-speaking areas (Zhan, Cheung 1988:425)]. These facts suggest that Cantonese *ni:¹* must be a later innovation. Given the relatively widespread distribution of phonetically similar forms in the Tai languages, one is tempted to conclude that Cantonese has most likely borrowed the form from Tai. However, before we jump to this conclusion, we must consider material from Austronesian, Tibeto-Burman, and Mon-Khmer.

The root **ini* has been reconstructed for Proto-Austronesian (Dahl 1977:70). Forms in some modern Tibetan dialects are clearly derived from Written Tibetan *ñdi* 'this' (Qu, Tan 1983:348, #1106) but the development can vary; in some dialects the word has developed a consonant cluster with homorganic nasal initial (*ndɨ²*, *ntɨ²*), or the nasal has replaced the original stop (*ni:*), or the the originally voiced stop initial is devoiced (as in Lhasa *ti¹*). Finally, the root also occurs in Mon-Khmer, cf. Khmer *nìh* and Vietnamese *naɨ²*. The widespread linguistic and geographical distribution of this root indicates it must have a very old history in Southeast Asia. I have concluded that Cantonese *ni:¹* is not a loan from Tai.

24. NVM *to think, ponder*

Neither the Yue nor Gan-Kejia dialects surveys asked for this lexeme, so we only have a few examples of it. Bai (1980:217) associated Cantonese *nɐm³* 'to think, ponder deeply' with the following entry in the *Jiyun*: 恁：侵韵，尼心切，思也． According to this entry's phonetic specifications, the Ancient Chinese pronunciation of 恁 is reconstructed as **niem*, and the ancient gloss given for the character was 'to think'. The root is found in several Zhuang dialects and Yao. Apparently because of its very limited occurrence, Li did not reconstruct this root. Ouyang (1989:611) compared Cantonese *nɐm³* with similar forms in two Zhuang dialects but held back making a claim that the Cantonese word was borrowed from Tai.

If we look at this lexeme in other Tai languages, we find that it is typically associated with phonetic shapes rather different from Cantonese and Zhuang.

Table 10. Forms for 'to think, consider' in Tai languages other than Zhuang (Wang et al 1984:852-853).

Li-Tongshi	buɯ⁵ '想, 思 考 to think, consider'
Mulao	tjaːŋ³, mjɛ⁶ '想, 思 考 to think, consider'
Dai-Xishuangbanna	kɯt⁸ '想, 思 考 to think, consider'
Buyi	tsim¹ '想, 思 考 to think, consider'
Shui	fa³, n̥i³ '想, 思 考 to think, consider'
Dai-Dehong	xaɯ³ tsaɯ⁶ '想, 思 考 to think, consider'
Dong	mjɛ⁶, ɕaːŋ⁵' '想, 思 考 to think, consider'
Maonan	nai⁶, sɔːŋ³ '想, 思 考 to think, consider'
Li-Baoding	ŋop⁷ '想, 思 考 to think, consider'

Given this wide phonetic variation in the forms and the limited occurrence of the NVM root in Chinese dialects and Zhuang, I do not believe that at the present time we can say anything about the direction of borrowing with any degree of reliability; it is just as likely that Zhuang and/or Yao borrowed *nam³* from Cantonese as Cantonese borrowed its form from Zhuang or Yao. This root does not seem to occur in the Tibeto-Burman languages.

25. NVM *soft*

This root is found in a few Yue and Kejia dialects, cf. Guangzhou *nɐm²* and Huizhou *ɲiam⁵* 'soft, well-cooked', and Tai, Tibeto-Burman, and Mon-Khmer languages, cf. Thai *nîm*, *nûm*, Atsi *nám*, Plang-Mane *kaʔ³¹ ɲɔm³⁵* 'soft'. Li (1977:111, #23) has reconstructed Proto-Tai initial *n- and A2 tone for *num* 'soft, tender'. With this root we find agreement between the Cantonese and Tai initials and tones. Li J-f. (1990:73) has compared the Cantonese word with forms from Zhuang and Dai to show that its source has been Tai. However, given its presence in both Tai, Tibeto-Burman, and Mon-Khmer, I believe that it must be an old Southeast Asian root.

26. ŋVŋ *stupid, muddled*

This root is widely represented in the Yue, Min, Kejia, and Wu dialect families and in a number of Tai languages (but not reconstructed in Li 1977. The root also shows up in two dialects of Yao, one dialect of Khmer spoken in Thailand, and in the Jing language which is related to Vietnamese. To establish that Cantonese *ŋɔːŋ⁶* has a Tai source, both Ouyang (1989:609) and Li J-f. (1990:74) have compared it with forms from a range of Tai languages. In my early publication on Tai loanwords in southern Han dialects (Bauer 1987:105) I had proposed that the Han forms were ultimately derived from Tai. My further research and reflection suggest to me this root may have entered Old Southern Chinese during the period of its early contact with Tai in south China.

27. TVM *to hang down, be low*

Cantonese *tɐm⁵* means 'to hang or drop down, droop, sag, let fall' (cf. the various expressions included in the Appendix). Maciver's Kejia dictionary listed *tem¹* 'to come down' (1926 [1982]:827) which may be cognate. This root occurs in most Tai languages and in a few languages belonging to subbranches of Mon-Khmer and means 'low, short', e.g., Zhuang-Wuming *tam⁵* and Khmu-Nanqian *tăm.* The sense 'to drop' is found with standard Thai *tàm* 'to be low, drop, decline' (Haas 1985:203). Li (1977:98, #7) has reconstructed the Proto-Tai initial as *t-* and B1 tone for *tam* 'low'. The Cantonese initial, rime, and tone all agree with the Proto-Tai root. The Cantonese and Tai meanings are close enough to make me believe that the Cantonese word has come from Tai. I also suspect the Khmu, Deang, Plang, and Wa forms are Tai borrowings.

28. TVM (1), **29. TVP** *to beat, pound* (2)

These are two related roots 'to beat, pound (with hand or instrument)' with final homorganic nasal/stop alternation. The first root or both roots in variation occur in a number of Yue dialects, e.g., Guangzhou *tɐm³*, *tɐp⁸* (the Yue dialects surveys elicited the forms in the expression 'to pound the back', a kind of massage). TVM and TVP occur in many Tai languages with a voiceless alveolar stop initial (recognized as a Tai word family by Luo 1996:866), but in Be, Vietnamese, and Khmer-Ban Nawattai the initial consonant is voiced. For the Proto-Tai root TVM 'to pound especially in a mortar', Li (1977:98, #9) has reconstructed the voiceless alveolar stop initial *t-* and A1 tone for *tam*. The Cantonese initial and rime agree with the Proto-Tai root but not the tone (Cantonese tone 3 = Tai C1).

For Proto-Tai TVP 'to hit, strike' both voiced and voiceless alveolar stop initials have been reconstructed: *d-* and D2S tone (Li 1977:105, #30) give *dup*, and Benedict (1975:229) has *top*. Cantonese tone 8 indicates the initial was originally voiced which agrees with the Proto-Tai root. In my early publication (Bauer 1987:100) I proposed that Cantonese *tɐp⁸* was a Tai loanword because the root was so widespread in the Tai languages but not in the other southern Han dialects. Also believing its source was Tai, Ouyang (1989:609) compared it with forms from several Tai languages. However, my further research has revealed that in addition to Tai and Mon-Khmer this root also occurs in some Kejia dialects (but apparently not in Min) and in the Yao, Austronesian, and Tibeto-Burman language groups. I believe that TVM and TVP are both old Southeast Asian roots whose nasal/stop alternation may have once corresponded to some kind of grammatical derivation or semantic difference.

30. TVM (1), **31. T(H)VP** *to stamp (foot)* (2)

In my early publication (Bauer 1987:105) I had proposed that Yue-Guangzhou *tɐm⁶* and Kejia-Meixian *tɔm³* 'to stamp (foot)' (the only two dialect forms available to me then) were derived from a Tai source. As indicated in Table 2 and

the Appendix, subsequent research has shown that the TVM root is widespread in the Yue and Kejia dialect families (but apparently not in Min: Min-Zhongshan *tɐm³* is probably a loanword from Yue) and is also found in a few Gan dialects (some forms in all three dialect families show a change in the place of articulation of the final nasal from bilabial to alveolar or velar). Among Tai languages the root is also widely attested, and a comparable form is found in Yao-Mian with the voiced dental stop initial. Both Ouyang (1989:609) and Li J-f. (1990:72) have also compared Cantonese *tɐm⁶* with phonosemantically similar forms in several Tai languages. Li (1977:104, #8) has reconstructed a voiced alveolar stop **d-* as the Proto-Tai initial and B2 tone for **dam* 'to trample on'. The Cantonese tone 6 indicates the initial was originally voiced and this agrees with the Proto-Tai root. The word occurs in Vietnamese; in the orthography the word is spelled *dậm* or *giậm* and is pronounced *zɐmp⁶*. Given the wide distribution of TVM in Chinese, Tai, and Vietnamese, I believe it must have a long history in Southeast Asia.

No morphological process links *tɐm⁶* 'to stamp (foot)' with *tɐm³* 'to beat, pound (with hand or instrument)', nor do Cantonese speakers perceive any semantic relationship between these two words. However, they are phonosemantically similar enough to raise the possibility that at an earlier historical stage of the language the change in tone and/or voicing of the initial corresponded to a change in meaning -- from pounding with the hand or instrument to pounding with the foot.

There appears to be a related, variant root T(H)VP with and without aspiration of the initial and with the homorganic bilabial stop final; however, its distribution is rather different than that for TVM and makes the nature of the relationship unclear. The standard Chinese character 踏 'to step on, tread, stamp' may or may not be the etymon of the Chinese forms. Despite the variation in the tone category, the aspirated initial in Kejia *tʰap⁷* and *tʰap⁸* represents the regular development of the historically voiceless aspirated initial; but the Cantonese reading pronunciation *taːp⁸* is anomalous: on the one hand, the unaspirated initial, low tone, and long vowel indicate a historically voiced initial and rime with low vowel; yet on the other hand, the ancient rime books specified a voiceless aspirated initial and short central vowel which yield for us an Ancient Chinese reconstruction of **tʰɔp*. The T(H)VP root occurs in Tai, but neither Li nor Benedict reconstructed it for Proto-Tai. It also turns up in Balinese tɛbtɛb, Tibetan *tɔpᴶ*, and Vietnamese *daːp⁸*. This fairly wide distribution has led me to conclude it is an old Southeast Asian root.

5.0 Conclusion

On the basis of this comparative study of a selected set of words from Cantonese and Southeast Asian languages, I believe I have identified Tai as the source of seven words out of the 29 which have been investigated. These words are *hɔːŋ³* 'young (chicken) which has not yet laid eggs', *jaːŋ⁵* 'to step on, step off, kick off', *kɐm⁶* 'to press down', *kɐp⁷ᵃ* 'frog', *kʰɐp⁸* 'to bite', *lɐm⁵* 'to collapse,

fall down', *təm⁵* 'to hang down, droop'. The other lexical items appear to have more complicated historical relationships that involve the Min and Kejia dialects as well as other Southeast Asian language groups. This group of words suggests to me that we still have a very long row to hoe before we will understand the formation and early historical relationships of these language groups in the Southeast Asia Sprachbund. I will save for a future paper speculation on what the nature of these early relationships may have been.

After I had delivered this paper at the Pan-Asiatic Linguistics Symposium, Paul Benedict who was in my audience commented on it. As I recall, he said that the lexical forms I had pulled together as the "same" word from languages he regards as genetically unrelated were really just "lookalikes". I presume he meant by this term that they were similar in sound and meaning as the result of chance. While this is certainly a possibility for a few items from a couple of languages, I think it is highly unlikely that the phenomenon of simple coincidence explains the occurrence of so many phonosemantically similar (and identical) forms in so many different languages. At any rate, he went on to say that the more one looks at lookalikes the less alike they look. And then, after looking closely at my handout (Table 2 of this paper), he added that historical linguists like himself do not like finding lookalikes. I assume this is because so many lookalikes such as mine raise troublesome questions about their origins and routes of transmissions.

The traditional approach to the historical development of the Chinese language takes for granted that the Chinese characters *equal* the Chinese language. In this historical comparative study of Han dialects and Southeast Asian languages, I have made use of colloquial vocabulary from the southern Han dialects which typically is not etymologically associated with standard Chinese characters. Over the course of my research on the historical comparative semantics of southern Chinese dialects and Southeast Asian languages, this kind of lexical material has come to play the leading role. As the old saying in historical linguistics goes, Every word has a history of its own. And we can add, But one has to be able to look for it. If we limit the purview of our research to only those words of Cantonese (or Kejia or Min or any Han dialect) which are etymologically tied to standard Chinese characters, then we ignore and place out of bounds a sizeable portion of basic vocabulary from daily life. As I hope to have shown in this study, the etymologies of some of these words turn out to be quite interesting and illuminating in regard to the development of the Sino-Tibetan, Austro-Tai, and Mon-Khmer language families.

This is *not* to say that we can or should ignore the Chinese characters in favor of ordinary words from the spoken language, just that we need a shift in focus in order to complement any approach which attempts to explicate the historical relationships of Han dialects and other Southeast Asian languages. As Yue-Hashimoto (1991:314) has pointed out, "[t]o understand the southern [Han] dialects, one must take into account both . . . an archaic Han aspect and a non-

Han aboriginal aspect." I interpret this to mean that our study must include both the Chinese characters and the chorphans.

Because the Han dialects and minority languages of China and Southeast Asia hold a critical key to our understanding of their early formation and development, there is a sense of urgency in collecting full descriptions of them before they fade away into history. In this age of the global village most people do not question the value of instant communication, the promotion of national languages and English as the international language, or the sale of Coca-Cola in every corner of the world; yet sadly, there has been a tradeoff, namely, increasing cultural and linguistic homogenization or assimilation and the loss of the world's linguistic diversity. Fortunately, resources are being devoted to recording dialects and languages of the Southeast Asian area. I gratefully acknowledge that my own research has been as fruitful and rewarding as it has because of the phenomenal growth in publications on Han dialects, minority languages of China, and other Southeast Asian languages produced just in the last decade and a half in Hong Kong, Taiwan, Japan, India, Thailand, and (especially) China. I along with other linguists welcome the outpouring of Han dialect materials, such as the surveys of Yue by Zhan and Cheung (1988, 1994) and Gan-Kejia by Li and Zhang (1992), and detailed dialect dictionaries and glossaries, such as Cai's for Chaozhou (1991) and Zeng's for Cantonese (1986). Our research work has benefited tremendously from publications on China's minority languages cited in detail in section 4.0 above.

Combined together, these resources on the Han dialects and ethnic minority languages present us with new, challenging opportunities for mapping out the formation and historical contact and development of languages in the Southeast Asia Sprachbund. In the process of culling words from my various sources, I believe I have found a few of the hitherto missing pieces for filling in holes of the fascinating linguistic jigsaw puzzle of Southeast Asia. Undoubtedly, I and other linguists working on the language families of this area would much prefer a time machine which would make possible our direct observation of their historical stages. The idea of time travel no longer seems so farfetched (as the British physicist, Stephen Hawking, would have us believe), but until the day it comes we will gratefully make do with these more mundane but precious materials on the modern languages.

1838

BIBLIOGRAPHY

Blust, Robert. 1988. *Austronesian Root Theory*. Amsterdam: John Benjamins.

Bodman, N. C. 1980. Proto-Chinese and Sino-Tibetan. In Frans van Coetsem and Linda R. Waugh, eds. *Contributions to Historical Linguistics*. Leiden: E. J. Brill. Pp. 34-199.

Burling, 1969. *Proto-Karen: a Reanalysis*. In A. L. Becker, ed. *Occasional Papers of the Wolfenden Society on Tibeto-Burman Linguistics*. Ann Arbor: The University of Michigan.

BYDB 中國科學院少數民族語言研究所. 1959. < 布 依 語 調 查 報 告>. 北京: 科學院初版社.

Cai 蔡俊明. 1991. <潮州方言詞匯>. 香港: 香港中文大學.

Chen, Li 陈章太, 李如龙. 1991. <闽语研究>. 北京:语文初版社.

Dahl, Otto C. 1977. *Proto-Austronesian*. Lund, Sweden: Curzon Press.

Dai Qingxia. 1991. On the affiliation of the Kadai group: indications from the nature of the relationship between Tibeto-Burman and Chinese. *Kadai, Discussions in Kadai and SE Asian Linguistics*. III:51-63. (Translated by Wu Hede).

Dai 戴庆厦, 主编. 1992. <汉语与少数民族语言关系概论>. 北京: 中央民族学院初版社.

Dai 戴慶廈. 1994. 漢藏語和粵語. 載單周堯, 主編. <第一屆國際粵方言研討會論文集>. 香港: 現代教育研究社. Pp. 65-73.

Diffloth, Gérard. 1984. *The Dvaravati Old Mon Language and Nyah Kur*. Bangkok: Chulalongkorn University Printing House.

Douglas, Carstairs. 1873. *Chinese-English Dictionary of the Vernacular or spoken language of Amoy*. London: Trübner & Co. [Reprinted by Ku-T'ing Book Store, Taipei, Taiwan, 1970].

Downer, G. B. 1973. Strata of Chinese loanwords in the Mien dialect of Yao. *Asia Major*. XVIII.1:1-33.

Driem, George van. 1987. *A Grammar of Limbu*. Berlin: Mouton de Gruyter.

Echols, John and Hassan Shadily. 1961. *An Indonesian-English Dictionary*. Ithaca, New York: Cornell University Press.

Egerod, Søren. 1959. Swatow loan words in Siamese. *Acta Orientalia* (Cophenhagen). 23.3-4:137-156.

Egerod, Søren. 1967. Dialectology. *Current Trends in Linguistics*. II:9-129.

Egerod, Søren. 1974. Sino-Tibetan languages. *Encyclopedia Britannica*. Pp. 796-806.

Egerod, Søren. 1976. Benedict's Austro-Thai hypothesis and the traditional view on Sino-Thai relationship. *Computational Analyses of Asian & African Languages*. 6:5-59.

Egerod, Søren and Jørgen Rischel. 1987. A Mlabri-English vocabulary. *Acta Orientalia* (Copenhagen). XLVIII:35-88.

Eitel, Ernest John. 1910. *A Chinese-English Dictionary in the Cantonese Dialect, Part II*. Hong Kong: Kelly & Walsh (revised and enlarged by Immanuel Gottlieb Genähr).

Elkins, Richard E. 1968. *Manobo-English Dictionary*. Honolulu: University of Hawaii Press.

English, Leo James. 1965. *English-Tagalog Dictionary*. Clayton, Victoria: Wilke and Co., Ltd.

Gedney, William. no date. A Comparative Tai wordlist. Mimeographed ms.

Gedney, William. 1976. On the Thai evidence for Austro-Thai. . *Computational Analyses of Asian & African Languages*. 6:65-81.

Gradin, Dwight and Thong. 1979. *Jeh Vocabulary*. Manila: Southeast Asia Publications. (Edited by Patrick Cohen).

Haas, Mary R. 1985. *Thai-English Student's Dictionary*. Stanford: Stanford University Press.

Hale, Austin. 1973. *Clause, Sentence, and Discourse Patterns in Selected Languages of Nepal, Part IV, Word Lists*. Kathmandu: Tribhuvan University Press.

Harris, Jimmy G. 1975. A Comparative word list of three Tai Nua dialects. In Jimmy G. Harris and James R. Chamberlain, eds. *Studies in Tai Linguistics in Honor of William J. Gedney*. Bangkok: Central Institute of English Language, Office of State Universities. Pp. 202-230.

Harris, Jimmy G. 1976. Notes on Khamti Shan. In Thomas W. Gething et al, eds. *Tai Linguistics in Honor of Fang-Kuei Li*. Bangkok: Chulalongkorn University Press. Pp. 113-141.

Hashimoto 橋本萬太郎. 1972.<客家語基礎語彙集>. 東京: アジア-アフリカ言語文化研究所.

Hashimoto Mantaro. 1976. The Agrarian and the pastoral diffusion of languages. *Genetic Relationship, Diffusion and Typological Similarities of East & Southeast Asian Languages. Papers for the 1st Japan-US Joint Seminar on East and Southeast Asian Linguistics*. Tokyo: The Japan Society for the Promotion of Science. Pp. 1-14.

Hashimoto Mantaro.1980. *The Be Language: a Classified Lexicon of its Limkow Dialect*. Tokyo: Institute for the Study of Languages and Cultures of Asia and Africa.

HZCD 中国社会科学院民族研究所主编. 1992. <汉载词典, *Myiwa Zaiwa Dangzhi Zhvum*>. 成都: 四川民族初版社.

He et al 何汝芬, 曾思奇, 田中山, 林登仙. 1986.<高山族语言简志(阿眉斯語)>.北京: 民族初版社.

He 贺嘉善. 1983. <仡佬语简志>.北京: 民族初版社.

HFCH 北京大学. 1964. <汉语方言词汇>.北京: 文字改革初版社.

Huang 黄布饭, 主编. 1992. <藏绵语族语言词汇>.北京: 中央民族大学初版社.

Huang, Parker Po-fei. 1970. *Cantonese Dictionary, Cantonese-English, English-Cantonese*. New Haven: Yale University Press.

Huang 黄雪贞.1983. 永定(下洋)方言詞匯 (二). <方言>. 3:220-240.

Hudak, John. 1993. *William J. Gedney's The Saek Language, Glossary, Texts, and Translations*. Ann Arbor: Center for South and Southeast Asian Studies, The University of Michigan.

Huffman, Franklin E. 1990. Burmese Mon, Thai Mon, and Nyah Kur: a synchronic comparison. *Moh-Khmer Studies*. XVI-XVII:31-84.

Jacob, Judith M. 1968. *Introduction to Cambodian*. London: Oxford University Press.

Jacob, Judith M. 1985. *A Concise Cambodian-English Dictionary*. London: Oxford University Press.

Jäschke, H. A. 1985. *A Tibetan-English Dictionary*. London: Routledge & Kegan Paul.

Jensen, John. 1977. *Yapese-English Dictionary*. Honolulu: University of Hawaii Press.

Jenner, Philip N. and Saveros Pou. 1980-81. *A Lexicon of Khmer Morphology. Mon-Khmer Studies* IX-X. Honolulu: University Press of Hawaii.

Karlgren, Bernhard. 1957. *Grammata Serica Recensa*. Stockholm: The Museum of Far Eastern Antiquities, Bulletin 29.

Lau, Sidney. 1977. *A Practical Cantonese-English Dictionary*. Hong Kong: Government Printer.

Li et al 李道勇, 聂锡珍, 邱锷锋. 1986. <布朗语简志>. 北京: 民族初版社.

Li Fang Kuei. 1965. The Tai and Kam-Sui languages. *Lingua*. 14:148-179.

Li Fang Kuei. 1976. Sino-Tai. *Genetic Relationship, Diffusion and Typological Similarities of East & Southeast Asian Languages. Papers for the 1st Japan-US Joint Seminar on East and Southeast Asian Linguistics*. Tokyo: The Japan Society for the Promotion of Science. Pp. 230-239.

Li Fang Kuei. 1977. *A Handbook of Comparative Tai*. Honolulu: The University Press of Hawaii.

Li J-f. 李锦芳. 1990. 粤语中的壮侗语族语言底层初探. 中央民族学院学报. 6:71-76.

Li J-f. 李锦芳. 1992. 粤語中的少数民族語言成分. 载戴庆厦, 主编. <汉语与少数民族语言关系概论>. 北京: 中央民族学院初版社. Pp. 327-343.

Li J-z. 李敬忠. 1990. 粤語是漢語族旱中的獨立語言嗎? <語文建設通訊>. 27:28-48.

Li, Zhang 李如龙, 张双庆. 1992. <客贛方言調查報告>. 廈門: 廈門大學初版社.

Li, Paul Ren-kui. 1978. A Comparative vocabulary of Saisiyat dialects. *Bulletin of the Institute of History and Philology, Academia Sinica*. 49.2:133-199.

Liang Min. 1990a. The Lachi language. *Kadai, Discussions in Kadai and SE Asian Linguistics*. II:35-44.

Liang Min. 1990b. On the affiliation of the Ge-Yang branch of Kadai. *Kadai, Discussions in Kadai and SE Asian Linguistics*. II:45-55.

Lin 林连. 1993. <泉州市方言志>. 北京: 社会科学文献初版社.

Lin Shaomian. 1994. Questions concerning the affiliation of the Lachi. *Kadai, Discussions in Kadai and SE Asian Linguistics*. IV:59-78.

Luce, G. H. 1985. *Phases of Pre-Pagán Burma*. Vols. I, II. Oxford: Oxford University Press.

Luo 罗康宁. 1987. <信宜方言志>. 广州: 中山大学初版社.

Luo Yongxian. 1996. Word families in Tai: a preliminary account. *Proceedings of the Fourth International Symposium on Languages and Linguistics, Pan-*

Asiatic Linguistics. Volume III:850-882. Salaya: Institute of Language and Cultural for Rural Development.

Ma 马学良, 主编. 1991. <汉藏语概论>. 北京: 北京大学初版社.

Maciver, D. 1926. *A Chinese-English Dictionary, Hakka Dialect*. Shanghai: Presbyterian Mission Press. [Reprinted 1982, Taipei: Southern Materials Center, Inc.].

Maclay, R. S. and C. C. Baldwin. 1870. *An Alphabetic Dictionary of the Chinese Language in the Foochow Dialect*. Foochow: Methodist Episcopal Mission Press.

Mao 毛宗武. 1992. <汉瑶词典>. 成都: 四川民族初版社.

Mao, Meng 毛宗武, 蒙朝吉. 1986. <畲语简志>. 北京: 民族初版社.

Mao et al 毛宗武, 蒙朝吉, 郑宗泽. 1982. <瑶族语言简志>. 北京: 民族初版社.

Marcus, Russell. 1983. *English-Lao, Lao-English Dictionary*. Tokyo: Charles E. Tuttle Company.

Matisoff, James A. 1972. *The Loloish Tonal Split Revisited*. Berkeley: Center for South and Southeast Asia Studies, University of California.

Matisoff, James A. 1983. Language diversity and language contact. In John McKinnon and Wanat Bhruksasri, eds. *Highlanders of Thailand*. Kuala Lumpur: Oxford University Press. Pp. 56-86.

Meyer, Bernard F. and Theodore F. Wempe. 1947. *The Student's Cantonese-English Dictionary*. Hong Kong: St. Louis Industrial School Printing Press.

MYFC 苗瑶语研究室编. 1987. <苗瑶语方言词汇集>. 北京: 中央民族学院初版社.

Nakajima 中道幹起. 1977. <東山島漢語方言基礎語彙集>. 東京: アジア-アフリカ言語文化研究所.

Nakajima 中道幹起. 1979. <福建漢語方言基礎語彙集>. 東京: アジア-アフリカ言語文化研究所.

Nguyen-Dinh-Hoa. 1966. *Vietnamese-English Dictionary*. Rutland, Vermont: Charles E. Tuttle Co.

Norman, Jerry. 1988. *Chinese*. Cambridge: Cambridge University Press.

Ouyang et al 欧阳觉亚, 程方, 喻翠容. 1984. <京语简志>. 北京: 民族初版社.

Ouyang 欧阳觉亚. 1989. 汉语粤方言里的古粤语成分. 载倪海曙, 主编. <语言文字学术论文集--庆祝王力先生学术活动五十周年>. 上海: 知识初版社. Pp. 605-616.

Ouyang, Zheng 欧阳觉亚, 郑贻青. 1983. <梨语调查研究>. 北京: 中国社会科学院初版社.

Peng 彭小川. 1990. 广东南海(沙头)方言音系. <方言>. 1:21-31.

Prapin Manomaivibool. 1976. Thai and Chinese -- are they genetically related? *Computational Analyses of Asian & African Languages*. 6:11-31.

Pukui, Mary Kawena and Samuel H. Elbert 1971. *Hawaiian Dictionary, Hawaiian-English, English-Hawaiian*. Honolulu: University of Hawaii Press.

Pulleyblank, E. G. 1986. Tonogenesis as an index of areal relationships in East Asia. *Linguistics of the Tibeto-Burman Area.* 9.1:65-82.

Qu, Tan 瞿靄堂, 礼克让. 1983.<阿里藏语>.北京: 中国社会科学院初版社.

Rao et al 饒秉才, 歐陽覺亞, 周無忌. 1981.<廣州話方言詞典>. 香港:商務印書館.

Sagart, Laurent. 1982. *Phonologie du dialecte Hakka de Song Him Tong.* Paris: Cahiers de Linguistique Asie Orientale.

Shi Lin and Cui Jianxin. 1989. Ai-Cham English glossary. *Kadai, Discussions in Kadai and SE Asian Linguistics.* I:25-63.

Sin 單周堯, 主編. 1994.<第一屆國際粵方言研討會論文集>.香港: 現代教育研究社.

Sin 單周堯. 1994.<説文>所見粵方言本字零拾. 載單周堯, 主編. <第一屆國際粵方言研討會論文集>. 香港:現代教育研究社. Pp. 179-187.

Sri Reshi Anandakusuma. 1986. *Kamus Bahasa Bali, Bali-Indonesia, Indonesia-Bali.* Undang: CV. Kayumas.

Suwilai Premsrirat. 1993. *Thai-Khmu-English Dictionary.* Salaya: Institute of Language and Culture for Rural Development.

Tan Ta-sen, Cheng Song Huat, and Tan Ah Leong. 1986. *Malaysia -Tionghua -Inggeris, Kamus Am Terbaru Bahasa Malaysia.* Kuala Lumpur: The Shanghai Book Company.

Thong and Dwight Gradin. 1979. *Jeh Vocabulary.* Manila, Southeast Asia Publications, Summer Institute of Linguistics.

Thongkum Theraphan. 1984. *Nyah Kur (Chao Bon) - Thai - English Dictionary.* Monic Language Studies, Volume II. Bangkok: Chulalongkorn University Printing House.

Tsuji 辻伸久. 1985.<湖南省南部中國語方言語彙集>. 東京: アジア-アフリカ言語文化研究所.

Wang et al 王均等. 1984.<壮侗语族语言简志>.北京:民族初版社.

Wang, William S-Y. 1995. Linguistic diversity and language relationships. In James Huang and Audrey Li, eds. *New Horizons in Chinese Linguistics.* Kluwer Academic Publishers.

Wu 吳景蓉. 1992. *The Pinyin Chinese-English Dictionary,* 漢英詞典. 香港 商務印書館.

Xie 谢永昌. 1994.<梅縣客家方言志>.廣州:暨南大學初版社.

Xing 邢公畹. 1985.<三江侗语>.天津:南开大学初版社.

Xu et al 徐悉艰, 肖家成, 岳相昆, 戴庆厦. 1983. <景汉词典>.昆明: 云南民族初版社.

Yabu Shiro. 1982. *A Classified Dictionary of the Atsi or Zaiwa Language (Sadon Dialect).* Tokyo: Institute for the Study Languages and Cultures of Asia and Africa.

Yan, Zhou 顏其香, 周植志. 1995.<中国孟高棉语族语言与南亚语系>.北京:中央民族大學初版社.

Yang, Paul. 1968. Modern dialects and the reconstruction of Archaic Chinese (initial consonant clusters). Mimeographed ms.

Yap, Elsa P. and Maria V. R. Bunye. 1971. *Cebuano-Visayan Dictionary.* Honolulu: University of Hawaii Press.

You 游汝杰. 1992. <汉语方言学导论>. 上海: 上海教育初版社.

You 游汝杰. 1995. 中国南方语言里的鸟虫类名词词头及相关问题. In William S-Y. Wang, ed. *The Ancestry of the Chinese Language.* Monograph Series Number 8, *Journal of Chinese Linguistics.* Pp. 253-268.

Yu 余秉昭. 1982. <同音字彙>. 香港: 光華圖書初版公司.

Yuan 袁家华.1983. <汉语方言概要>. 北京: 文字改革初版社.

Yue-Hashimoto, Oi-kan. 1972. *Studies in Yue Dialects 1: Phonology of Cantonese.* Cambridge: Cambridge University Press.

Yue-Hashimoto, Oi-kan. 1976. Southern Chinese dialects: the Tai connection. *Computational Analyses of Asian & African Languages.* 6:1-9.

Yue-Hashimoto, Anne O. 1985. *The Suixi Dialect of Leizhou, a Study of its Phonological, Lexical, and Syntactic Structure.* Hong Kong: The Ng Tortai Chinese Language Research Centre, Institute of Chinese Studies, The Chinese University of Hong Kong.

Yue-Hashimoto, Ann. 1991. The Yue dialects. In William S-Y. Wang, ed., *Languges and Dialects of China.* Monograph Series Number 3, *Journal of Chinese Linguistics.* Pp. 294-324.

ZDYC 民族语研究中心, 第五研究室, 主编. 1985. <壮侗语族语言词汇集>. 北京: 中央民族学院初版社.

Zeng Zifan. 1986. *Colloquial Cantonese and Putonghua Equivalents.* Hong Kong; Joint Publishing Co. [Translated by S. K. Lai].

Zhan, Cheung 詹伯慧, 張日昇. 1988. <珠江三角洲方言詞匯對照, *A Survey of Dialects in the Pearl River Delta, Comparative Lexicon* >. 香港: 新世紀初版社.

Zhan, Cheung 詹伯慧, 張日昇, 主編. 1994. <粵北十縣市粵方言調查報告, *A Survey of Yue Dialects in North Guangdong*>. 廣州: 暨南大學初版社.

Zhang 張振興. 1983. <臺灣閩南方言記略>. 福州: 福建人民初版社.

Zhang et al 张元生, 马加林, 文明英, 韦星朗. 1985. <海南临高话>. 南宁: 广西民族初版社.

ZHCH 广西壮族自治区少数民族语言文字工作委员会研究室, 主编.1984. <壮汉词汇>. 南宁: 广西民族初版社.

Zhou 周長楫. 1993. <廈門方言詞典>. 南京: 江蘇教育初版社.

ZMYC 中央民族学院, 苗瑶语研究室, 主编. 1987. <苗瑶语方言词汇集>. 北京: 中央民族学院初版社.