The Role of Migration and Language Contact in the Development of the Sino-Tibetan Language Family

Randy J. LaPolla

1. Introduction

A strong case can be made for a genetic linking between the Sinitic languages (the Chinese 'dialects') and the Tibeto-Burman languages. There are hundreds of clear cognates of basic vocabulary (Benedict 1972; Matras 1978; Baxter 1995; see LaPolla 1994a for a list of two hundred of the most uncontroversial) as well as some derivational morphology that can be reconstructed to the Proto-Sino-Tibetan level. Within Tibeto-Burman again we find hundreds of cognates of basic vocabulary, and there are some relatively uncontroversial groupings based on shared innovations, such as Lolo-Burmese, Bodish, Qiangic, and Karenish, but subgrouping within Tibeto-Burman (and to some extent within Sinitic) is quite problematic. Benedict (1972; see the Figure) had Tibeto-Karen as one of two branches of Sino-Tibetan (the other being Chinese), with Tibeto-Burman and Karen being the two highest branches of Tibeto-Karen. Karen was given this position because it has verb-medial word order rather than the usual verb-final order of Tibeto-Burman. However, most linguists working on Tibeto-Burman now consider Karen to be a branch within Tibeto-Burman, as they assume that Karen word order changed due to contact with Mon and Tai, and therefore is not an important factor to be used in genetic grouping. As can be seen from the...

1. What can be reconstructed is an *-t* causative and denominate prefix (Mei 1982), possibly alternation of voice and/or aspiration of initials for causatives, a *-t* suffix for translative suffixes (Benedict 1972: 98-100; Michalowski 1985; van Driem 1988), and a nominalizing *-n* suffix (see LaPolla 1994a; Jin 1998). There is no evidence of relational morphology at the Proto-Sino-Tibetan or Proto-Tibeto-Burman levels (for discussion see LaPolla 1992a, 1992b, 1994a, 1995).

2. Forrest 1973 had suggested Karen was so similar to Mon that it could just as easily be a Mon language influenced by Tibeto-Burman as a Tibeto-Burman language influenced by Mon. Luce (1978: 31) states that Karen is neither a Tibeto-Burman language nor a Mon-Khmer language, though it has been heavily influenced by both Tibeto-Burman and Mon-Khmer. He says it is 'pre-Tibeto-Burman.' Not many scholars working on Karen would agree with these assessments now.
languages to Chinese or Sino-Tibetan as a whole, that is, whether we consider the similarities among Chinese, Tai, and Hmong-Mien to be due to contact or due to genetic inheritance. Many scholars in China argue that the languages are related, but most linguists outside China feel the shared words are very old loans, and the other features, such as the similarities in the tone systems, the use of the classifier for definite marking, etc. spread areally. This makes it similar to the case of Vietnamese, which at one time was also thought to be related to Chinese, due to its many Chinese-like features and words, but is now thought to be a Mon-Khmer language heavily influenced by Chinese.

Three main factors have been involved in the formation of the present-day Sino-Tibetan language family: a shared genetic origin, divergent population movements (i.e. innovations appearing after these splits), and language contact. Population movements and language contact have in fact generally been two aspects of a single phenomenon. It is this fact that is the link between Dixon's (1997) view of rapid change due to non-linguistic causes and Heath's (1997, 1998) view of rapid change due to intense language contact, discussed by Watkins (this volume). The present chapter will look at the history of the development of this family from the point of view of population movements and language contact, to show the role language contact has had in the formation of the family as we know it today.

2. The migrations and their effects

From what we can piece together from the archeological and linguistic evidence (see for example Chang 1986, Treisman 1972, Pulleyblank 1983, Fairbank, Reischauer, and Craig 1989, Xing 1996, Ran and Zhou 1983), it seems the Sino-Tibetan-speaking people (if we associate the Neolithic Yang-shao culture with the Sino-Tibetans) originated in the central plains of what is now north China, in the valley of the Yellow River. At least 6,500 years ago, some members of the original group moved largely south and east, while others moved largely westerly at first, then moved in a southerly or south-westerly direction. Differences in identity and possibly language were evident at the time of the earliest Chinese writing, about 3,000 years ago, but there continued to be contact between the two related groups and others that surrounded them in the early period (see, for example, Wang Huiyin 1989), and frequent mixing of peoples (for example, the ancestors of some early Chinese rulers are said to have been from the western group—Ran and Zhou 1983, Ran, Li, and Zhou 1984, FitzGerald 1961). The group that stayed in the central plains, including those members of the western group that stayed in the central plains and nearby areas, as well as those who moved south-easterly, eventually became what we think of as the Chinese, while the group that moved south-westerly became what we think of today as the Tibeto-Burmans.

---

2.1. SINITIC

The movements in both directions were not single movements, but consisted of larger or smaller waves of movement, often into the same areas. Government-encouraged migration was practised as early as the Yin dynasty (roughly 1600–1027 BC), and has been practised by all Chinese governments up to the present day. There have also been massive private migrations and shifts of national or regional capitals due to natural disasters, war, and the pull of new economic opportunities (Ge, Wu, and Cao 1997).

The movement of the Chinese has almost never been to an area where there were no people. Splitting of the language by migration almost always involved language contact, either with non-Chinese languages or other Chinese dialects, and very often in government-sponsored migrations there was purposeful mixing of peoples. What we now think of as the Han Chinese have from very early on continually absorbed other peoples into the race (Wang Mingke 1992, Wiens 1967, Xu 1989). As the Chinese moved into new areas, they often absorbed the peoples there into the Han (Chinese) nationality, or, in some cases, were absorbed by the local nationalities (see, for example, Dai, Liu, and Fu 1987 and He 1989, 1998 for a case of Mongolian soldiers and settlers sent to the south-west in the Yuan dynasty (1234–1368) being absorbed into the Yi culture and developing a new language).

Table 1 summarizes the major movements, giving the time period, the place the population moved from and the place they moved to, the number of people who moved, if it is known from government records, and the original inhabitants of the area they moved to (data mainly from Lee 1978, 1982, Lee and Wong 1991, Zhou 1991, Ge, Wu, and Cao 1997).

It can be seen from Table 1 that many of the movements were chain movements. For example, the movement of over two million non-Chinese people into the central plains from the northern steppes in the second and third century caused at least three million Chinese to flee south. To give one example of how drastically these movements affected the populations, according to Lee (1978: 29), in one county (Bingzhou in Shanxi), two-thirds of the population emigrated between 289 and 312. This not only affected the population of the north, but also of the south, as one out of every six people in the south was a displaced northerner after the movement. Nanjing became the capital of the Eastern Jin (317–420) and Southern (420–581) dynasties; it attracted over 200,000 migrants, a figure greater than the original local population. The form of speech in the area then changed from a Wu dialect to a northern dialect. The speech of another Wu area, Hangzhou, became what Zhou and You (1986: 15) call a 'half-Guanhua (Mandarin)' area because of the shift of the Song dynasty capital from the north to Hangzhou in 1127 and the resulting massive influx of northerners. While the phonology is basically that of a Wu dialect, it is lexically and grammatically more similar to the northern dialects, and does not have the usual literary/colloquial reading distinction of characters that the other Wu dialects have.
The movements were often so massive that they caused major shifts in the overall demographics and language distribution of the entire country. For example, in the seventeenth century, north-east China, south-west China, and the upper Yangtze comprised only about five per cent of the population of China and ten per cent of the Mandarin speaking population, but the movement of people from the middle Yangtze and north China was so massive that by 1982 these three areas included one third of China's population and about half of the Mandarin speaking population (Lee and Wong 1991: 55). In some areas the movements have meant almost an entire displacement of the original population. For example, since 1949 there has been a massive Government-orchestrated movement of Han Chinese people into the minority areas of Inner Mongolia, Xinjiang, and Tibet. In Inner Mongolia the population is now less than twenty per cent Mongolian, and the capital, Huerhot, is less than two per cent Mongolian. This of course had a drastic effect on the use of Mongolian in the capital.

Aside from migrations of Chinese into other parts of China (or what later became part of China), there was also a great influence from non-Chinese speakers moving into areas of China, particularly north China, where for more than half of the last thousand years the Chinese were under the control of Altaic invaders. Beijing, for example (see Lin Tao 1991), was a secondary capital of the Liao dynasty (Khitans people; 907-1125) and the early Jin dynasty (Jurchens; 1115-1234), and was capital of the Jin from 1153 to 1234. Beijing was again the capital of the Yuan (Mongols; 1234-1368), Ming (Han; 1368-1644), and Qing (Manchus; 1644-1911) dynasties. Except for three hundred years during the Ming dynasty, Beijing was a political centre of non-Chinese peoples for the last thousand years. The populations changed, though, as the Jin government almost emptied the city in 1123, moving the people to the north-east. In 1368, the Ming government moved large numbers of people mainly from Shansi and Shandong into Beijing to populate the city. In 1644, the Manchu rulers moved most of the original inhabitants out of the inner city and moved the Eight Banner army and their family members into the inner city. While many of the invaders assimilated, they also had an effect on the language of the north. Mantaro Hashimoto (e.g. 1976, 1980, 1986) has talked about this as 'the Altaicization of Northern Chinese', and has argued that a continuum of features from north to south, such as the northern dialects having fewer tones, less complex classifier systems, and an inclusive/exclusive distinction in the first plural pronoun, while the southern dialects have more tones, more complex classifier systems, and other features similar to the Tai and Hmong-Mien languages (You 1982, 1985; Zhou and You 1986, Wang Jun 1991), is due to Altaic influence in the north, and Tai/Hmong-Mien influence in the south. He also suggests (Hashimoto 1976, 1992: 18) that the preservation of final -n and -g in Mandarin while all the stop endings and -m were lost might be due to the fact that these two finals are found in Manchu. Li Wen-Chao (1995) argues that the inventory of vowels and the syllable structure of Chinese changed after the Tang period due to the Altaicization of the language, that is, the adoption of the Chinese lexicon and grammar by Altaic speakers, but with Altaic phonology.

The resulting mixtures of the people from these migrations with the people originally in the areas they moved into are what give us the dialects we have today (cf. Zhou and You 1986, Wang Jun 1991). For example, the early Wu dialect had formed from a south-eastern migration into an Austroasiatic area, and the Chu dialect (a precursor to the Xiang dialect) formed from a very early southern migration into a Tai/Hmong-Mien area, and then the Gan dialect formed in the area where the Wu and Chu dialects had contact with each other in central and northern Jiangxi because of a later migration during the Han dynasty (206 BC-AD 220). Later migrations brought successive waves of immigrants into the area from the north, and then there was a split of this dialect into the Gan and Hakka dialects by migration of what became the Hakka to the east and south, and then later to the west. Contact with languages in each area where the Hakka migrated to resulted in varieties of Hakka that reflect features of those languages (see Hashimoto 1992). In Fujian (Bienenstein 1959, Norman 1991) the language was that of the Min-Yue (a subgroup of the Bai Yue) before any Chinese came into the area, and then the first Chinese settlers in the Eastern Han Dynasty (25-220) brought with them the older dialect of the Wu area, as colonization was from Zhejiang in the north. The original Wu dialect in Zhejiang changed quite a bit after that period due to the massive immigration from the north after the fall of the Western Jin Dynasty in the fourth century. Many of these latter Wu speakers again migrated south into Fujian, and so now the Fujian (Min) dialect shows evidence of influence from at least the following languages: the Min-Yue language, the Chinese language of the Han period, the post-Han stratum brought in by later immigrants, a Tang dynasty (post-eighth century) literary form of the Tang koine, and Modern Mandarin (Norman 1988, 1991). Lien (1987; discussed in W. S.-Y. Wang 1991b, ch. 4) has discussed the complicated interactions of these various strata, and has shown how these interactions led to an ongoing gradual bidirectional diffusion of features (of tones and segments) among the different strata, which has been creating forms that are not identifiable as originating from one particular source language, such as the word for 'thank' in the Chaozhou dialect, which has a segmental form, [sia], which derives from the Tang dynasty literary layer, but a tone that the form would have in the colloquial layer. There are also cases of different combinations, such as colloquial initial with literary final and tone, and literary initial with colloquial final and tone (see also Lien 1993, 1997, 2000).
Wang and Lien (1993). The initial discovery of this phenomenon led to the development of the theory of lexical diffusion (see, for example, Chen and Wang 1973), of which Lien's work is an extension. An important point to note is that while the initial strata were the result of language contact (massive borrowing of literary forms or substrate/superstrate influence), the gradual bidirectional diffusion of features has been occurring over a long period of time and is a language-internal phenomenon (though one which of course may be influenced by other factors, such as new superstrate influence).

While in Chaozhou there was a mixing of pre-existing phonemes to create new morphemic forms, there are also cases of the creation of new phones or phonemes because of contact influence, such as in the creation of voiced aspirates for phonemes in a particular tone category in the Yongxing form of the Xiang dialect spoken in Sichuan. Ho (1988; also discussed in W. S.-Y. Wang 1991b) suggests that these voiced aspirates arose because of contact between this dialect and the surrounding Mandarin dialects. In these Mandarin dialects, words that formerly had voiced initials and were in the level-tone category became voiceless aspirates, while in the Xiang dialect in general they continued to be voiced. In Yongxing, due to the competing influences of the Mandarin feature of aspiration and the Xiang feature of voicing, about 86% of the initials of morphemes in that tone class have become voiced aspirates, a new type of initial for that language.

Compare these phenomena with Dixon's (1997) discussion of the gradual diffusion of linguistic features in a linguistic area. This same sort of bidirectional diffusion among different languages of a bilingual population (rather than strata within a single language) can lead to the areal similarities associated with a linguistic area. Chen Baoya (1996) is a careful study of the bidirectional diffusion of features between Chinese and Tai in Dehong Prefecture of Yunnan Province in China. Chen has shown that in some cases there has been simplification of the sound system of a native language due to the influence of the contact language, e.g. the loss of the distinctions between /l/ and /n/ and between /ts/ and /tsʰ/ in the Chinese spoken by ethnic Chinese, as these distinctions do not exist in Tai, and the loss of certain vowel distinctions in the Tai of ethnic Tai (e.g. between /u/ and /y/), because these sounds are not distinguished in Chinese. In other cases there has been an increase in phonemes due to the influence of loanwords in the language, e.g. the development of /kʰ, tsʰ, tsʰ/ in the Tai of Luxi county. Chen argues that much of the influence is through an interlanguage he calls 'Tai-Chinese', so in a sense there is a tridirectional diffusion in this context.

In Table 1 it is stated that many of the early movements were into areas inhabited by the Bai Yue (Hundred Yue). From the linguistic evidence, it seems there were at least two subgroups of the Hundred Yue, one which spoke Austroasiatic-related languages (mostly along the coast from possibly as far north as Shandong), and another that spoke Tai and Hmong-Mien-related languages (mostly the interior of the south up to the Yangtze and as far west as present-day Sichuan province) (Pulleyblank 1985, Li Jingzhong 1994, Bellwood 1992, Tong 1998).

Norman and Mei (1976; see also Norman 1988) give words for 'die', 'dog', 'child', and others that seem to be cognate with words in Austroasiatic rather than Sino-Tibetan. Yue-Hashimoto (1967, 1991) and others (e.g. Baron 1973, You 1982, 1995, Zhou and You 1986, Huang 1990, Cao 1997, and Meng 1998) give evidence of contact influence between Cantonese and the Tai and Hmong-Mien languages, including not only lexical evidence, but structural evidence, such as word order, the specifics of the tone system, marked phonetic patterns, special uses of the classifiers, etc. In the prehistoric period, the Hundred Yue may have included speakers of the precursors of Austroasiatic, Tai, Hmong-Mien, and possibly Austronesian (see, for example, Blust 1984–5, 1994).

There has also been influence from national and provincial prestige dialects on other dialects throughout Chinese history. Centres of population concentration developed, and languages in those centres came to be quite distinct from each other, with each having prestige within its own area, and then spread out from those centres. The result is languages forming something like prototype categories rather than areas with sharp boundaries (see, for example, Iwata 1995). For example, comparing Guangzhou city Yue with Xiamen city Southern Min (each the prototype of its category), the differences are quite clear, and the languages are easily distinguishable, but in the areas of Guangdong where the two languages meet, there are many forms of each dialect that to different degrees differ from the prototype of their category while having characteristics of the other category. In some cases it is difficult to distinguish whether a certain form of speech is a Yue dialect or a Southern Min dialect, as the two have leached into each other to form something that cannot be uncontroversially put into either category. These major centres have also influenced each other in various ways: see for example Yue-Hashimoto (1993) on the spread of certain patterns of interrogative syntax and other constructions among the Yue, Min, and Beijing dialects, Chappell, this volume, on the creation of 'syntactic hybrids' in the southern dialects due to the influence of Mandarin, and Chang Kuang-yu (1994) on the spread of features of the Wu dialect.

In Modern times there has been quite a bit of influence on the dialects from the Common Language (Mandarin). There has been a strong effort to teach the Common Language, and this has been very successful in some areas, with the result often being influence on the local dialect. For example, children in Shanghai often speak Mandarin amongst themselves, as that is what they speak in school, even if they speak Shanghaiese with their parents. This has caused some changes within Shanghaiese, such as the levelling of vocabulary and phonology in terms of becoming more like Mandarin (see, for example, Qian 1991, 1997). In Taiwan, many young people of Taiwanese descent do not learn Taiwanese well (if at all),
and even when they speak it, it is often a somewhat levelled form, where, for example, a Mandarin-based compound word will be pronounced in Taiwanese rather than using the traditional Taiwanese form (e.g. instead of [sin^3 ku^5] for 'body', you often hear [sin^3 te^5], based on Mandarin shéntrí). There is also loss of distinctions in some semantic areas, such as the differentiation of verbs used for the sounds animals make.

In areas where Mandarin is a well-established second language, regional varieties are forming, such as the many varieties of Mandarin developing in the northwest of China because of influence from various Altaic, Turkic, or Tibeto-Burman languages (e.g. Dwyer 1992, Chen 1982). Another interesting example is Taiwanese Mandarin, which can be said to have created to some extent out of an interlanguage. After 1949, there was a large influx of people from the mainland because of the Communist takeover of the mainland. These people were mostly from Wu dialect areas, and spoke Mandarin only imperfectly as a second language. The Wu speakers attempted to teach the Taiwanese population Mandarin, and forced the Taiwanese to speak it even amongst themselves. The Taiwanese did not generally have access to native speakers, and so did the best they could with what they had, and often added pieces from their native language, Japanese, and English, forming an interlanguage heavily influenced by Taiwanese (see Kubler 1985, Hansell 1989 for examples). For the Taiwanese this remained a second language, but for the sons and daughters of the mainlanders, who generally did not learn their parents' dialects, and did not learn Taiwanese, this interlanguage became their first language. This group then became the first generation of native Taiwanese Mandarin speakers. There may eventually be a coalescence of the Taiwanese Mandarin and the Mandarinized Taiwanese.

2.2. TIBETO-BURMAN

Turning to Tibeto-Burman, the major migrations were west into Tibet and south-west into Burma, but there were also minor movements into northern Thailand, Laos, and Vietnam. Two large subgroupings formed by areal contact can be distinguished within Tibeto-Burman: the 'Sinosphere' and the 'Indosphere' (these terms from Matisoff, e.g. 1990). One reason for the differences between the two spheres is the objective dominance of Chinese or Indic languages over different subsets of the Tibeto-Burman languages; another is the subjective analysis of those languages falling within the scope of work by Chinese-trained or Indic-trained linguists. There are certain features that we frequently find in languages in the Indosphere that we do not find in the Sinosphere. In phonology we find, for example, the development of retroflex stop consonants. In syntax we find, for example, post-head relatives of the Indic type (relative clauses are generally pre-head and without relative pronouns in Sino-Tibetan languages). For example, in Chaudangli (ShreeKrishan, forthcoming a), of the Pithoragarh District of Uttar Pradesh, India, a relative clause is formed using one of two borrowed (Indo-Aryan) relative pronouns, /s/ (with human subjects) or /s/ (with non-human subjects).

1. (a) hidi ati shi le jo nyaro ra-s this that boy is who yesterday come-PAST 'He is the same boy who came yesterday.'
   (b) hidi ati trang le jai be or gun-ci this that horse is which mountain from fall-PAST 'It is the same horse which fell from the mountain.'

Another feature of the Indosphere, discussed by Saxena (1988a, b), is the frequent grammaticalization of a verb meaning 'say' into a quotative, causal, purpose, or conditional marker, a complementizer, or an evidential particle, due to areal influence from the Indic and Dravidian languages, whereas languages in the Sinosphere are less likely to do this. In (2) are examples of the use of the verb for 'say' as part of a causal connective in Nepali, an Indo-European language, and Newari, a Tibeto-Burman language of Nepal.

2. (a) Nepali (Saxena 1988a: 376):
   timharu midle ek jana marb ho kinabne hoy hichori hoina you(pl) among-LOC one CL fool is why+say+part this tower be+NEG 'One of you is a fool because this is not a tower.'
   (b) Newari (Saxena 1988a: 379):
   chi-pi cho-mho marb kho cho-da-se ya tho dhohora noo-khu you-pl one-CL fools are why+say+INF if this tower NEG is 'One of you is a fool because this is not a tower.'

In Sino-sphere languages we often find the development of tones. For example, among the Qiang dialects of north-western Sichuan, there is a north-west to south-eastcline in the degree to which tones are a stable and important part of the phonological system: the closer the dialect is to the Chinese areas, generally the stronger and more developed the tone system is (Sun 1981, Liu 1998, Evans 1999). Contact with Chinese can also result in monosyllabicity and an isolating structure, the most extreme example of this being Vietnamese. It also seems to be
the case that languages in the Sinosphere have simpler grammatical systems, but this brings us to the second part of this question of spheres: the subjective analysis of the linguists doing the recording of the languages. In India, linguists are trained in Sanskrit grammar, and so are familiar with paradigms and participles. They generally look for them in the Tibeto-Burman languages they describe, and often find them. They are not very familiar with tones, and do not consider them that important, and so even if the language has tones, they often will not be included in the description. On the other hand, the Chinese linguists are trained in Chinese linguistics, and so are often not familiar with paradigms and participles, but are very familiar with tones. They then generally do not describe the languages as having tight paradigms, etc., but very often find and describe Chinese-like tonal systems, even in languages (e.g. Burmese, rGyalrong) that could be argued to have register or pitch-accent systems.

As mentioned above, the Tibeto-Burman speakers followed two main lines of migration: west into Tibet and then down into Nepal, Bhutan, and northern India; and south-west down the river valleys along the eastern edge of the Tibetan plateau through what has been called the 'ethnic corridor' (Fei 1980, Sun 1983, Hoffman 1990). This split in the migration is responsible for the split between the Bodic languages and the rest of Tibeto-Burman. There is little information about the spread of Tibeto-Burman speakers into the Tibetan plateau other than that they spread from the north-east of Tibet (that is, the north-west of China; Stein 1966, Snellgrove and Richardson 1986, Ran and Zhou 1983, Hoffman 1990), but from the present wide geographic spread of the Tibetan dialects, from the closeness of the dialects, and from the fact that all dialects show some of the same uncharacteristically Tibeto-Burman features (such as non-Tibeto-Burman words for 'horse' and 'seven'), there must have been contact with non-Tibeto-Burman languages before the spread of Tibetan, and then the spread was relatively rapid, and into an area where there were no (or few) earlier inhabitants. There has also been quite a bit of contact with northern and central Asian languages since the original spread of the Tibetan dialects as well (see for example Lauffer 1916).

There is also not much we can be sure of about the early history of Burma. It is assumed that the original inhabitants were Negritos. The migration of Tibeto-Burman speakers south into Burma must have started at least by the first century AD. Fourth-century Chinese records already talk of a barbaric tribe we might identify with the modern Jinghpaw in the far north of Burma and a civilized kingdom known as Pyu which controlled central and upper Burma. The Pyu were Tibeto-Burman speakers who had come down into Burma along the Irrawaddy valley. They adopted Theravadan Buddhism and their writing system (seventh century)

---

10 See Hoffman (1990: ch. 4) on the prehistoric contact influence from the 'steppe peoples' (northern non-Sino-Tibetan) on the group that became the Tibetans.

11 This section is a synthesis of information in Luce (1937, 1976), Luce and De Maung Tin (1939), Leach (1954), Hall (1960), FitzGerald (1972), Chen Xujing (1992), and Chen Boxing (1995).
much more a part of Burma, and this began another period of Mon influence on the Burmese. Much of what we think of as Burma, such as the Irrawaddy Delta and Rangoon, was for most of its history part of a Mon kingdom.

Because of this legacy, there has been heavy influence of the Mon language on Burmese (Bradley 1980). Aside from the script and a large number of lexical loans, there has also been Mon influence on the suprasegmentals, in that Burmese 'tones' are unlike the usual Sino-Tibetan type of tones in being more like a register contrast (and in the Arakanese dialect of Burmese show vowel-height differences related to the registers), as is the case in Mon and other Mon-Khmer languages. There has been convergence in the vowel systems of Mon and Burmese, and to some extent the consonant system, where there has been a loss of contrast between alveolar fricatives and affricates versus palatal or alveopalatal fricatives and affricates, as in Mon. In Written Burmese there are also palatal finals (most finals have been lost from the spoken language), which do not usually occur in Tibeto-Burman languages, but do occur in Mon-Khmer languages. In terms of the word structure, Burmese has the typical sesquisyllabic structure of Mon-Khmer languages where the first 'half-syllable' or 'minor syllable' is unstressed and the second syllable is stressed (e.g. the Burmese pronunciation of the word Burma: [bòma]). This is a feature that characterizes a number of the languages in the area, as opposed to the Tibeto-Burman languages still in China, which generally do not show this pattern. Bradley (1980) attributes these influences to the fact that so many Burmese speakers were originally Mon speakers. Many of them are now monolingual in Burmese. In fact Burmese is spoken by many different ethnic groups, and so shows varieties in each area due to the influences of the local languages (Bradley 1996) has produced a map (with discussion) of the use of Burmese by different ethnic groups.

Another language which has had a major influence on Burmese is written Pali. Many Burmese texts are what is known as 'Nissaya Burmese', word for word translations of Pali texts which try to accommodate Burmese word order and grammar to the Pali original, and this led to influence on purely Burmese texts. 'Pali was regarded as the model of correctness in language, so that the closer to Pali one's Burmese was, the purer it seemed to be' (Okell 1965: 188). This written form eventually influenced the spoken form as well because of the influence of reading, education, and religion (Okell 1965).

The north of Burma continued to be populated by the Shan and the Tibeto-Burmans (principally Jinghpaw), and there has been much mutual cultural and linguistic influence, in some cases with subgroups of the Jinghpaw becoming Shan in language and culture and vice versa (Leach 1954: 293). In the eighteenth and nineteenth centuries these two groups extended into Assam, and the Jinghpaw brought thousands of Assamese slaves back into Burma. These formerly Indo-European-speaking slaves eventually assimilated to the Jinghpaw culture and language (Leach 1954: 294).

We see another type of contact situation in northern Burman, that is where two or more languages are in close contact, but no language is dominant, such as is the case with the Jinghpaw people (Dai 1987), which is similar to the situation that Dimmendaal (this volume) describes for the Suri group of Surmic. There are four subgroups within the Jinghpaw nationality, and each subgroup has its own language. These four groups often live together in the same villages and intermarry, and have very similar cultures, but keep the languages distinct in terms of exogamy, marrying other-language speakers, the children being considered speakers of the father's language even though they may speak one language to the father, one to the mother, and a third to the grandmother. Living in such a situation the people come to think in similar patterns and have similar cultures, and this leads to certain types of lexical and usage convergences among the languages. This is a clear case of adstratum influence. In the case of other Tibeto-Burman languages, contact has been not because they live within the same villages, but live relatively close to each other, and so become bilingual, and this can affect the languages. For example, in Lisu dialects in general, interrogatives are marked by a sentence-final particle, while in Yi dialects interrogatives are marked by reduplication of the verb. But in the Luquan dialect, the Lisu dialect closest to the Yi-speaking area, interrogatives can be, or always are, marked by reduplication of the verb (CASIML 1999: 3).

Tibeto-Burman migration into Nepal, Sikkim, and Bhutan was originally almost entirely from directly north, that is, Tibet (Poffenberger 1980), and so the earlier languages generally show a close relation to Tibetan. In Nepal (see Kansakar 1996) there are now over seventy different languages, possibly as many as a hundred (Grimes 1991). According to Kansakar (1996: 17), these include about fifty-six Tibeto-Burman languages, fourteen Indo-Aryan languages, one Austroasiatic language, one Dravidian language, and one isolate (Kusunda). Of the Tibeto-Burman languages, the Kiranti languages and what Bradley (1997) calls the Central Himalayan languages (Magar, Khamp, Chepang, Newari) came into Nepal relatively early, and the Newars (now 3.7% of the population) had a kingdom in the Kathmandu Valley from at least the eleventh century until they were conquered by the Nepali-speaking Gurkhas in the eighteenth century. A large group of Tibetans moved into Nepal during the reign of the Tibetan leader Strong-Bstan-Gampo in the seventh and eighth centuries and after, when the whole area down to the Bay of Bengal was part of the Tibetan kingdom; the Tamangs are said to be remnants of these people. Quite a few members of the Tamang–Gurung group have in the last one or two hundred years emigrated to north-east India or other areas (e.g. eastwards into Nepal) and now speak only Nepali. Among the Gurungs there is something of a cultural continuum of Buddhists in the north and Hindus in the south due to contact with Hindus in the south (Poffenberger 1980). Of those Gurungs still living in Nepal, 49.2% (221,271) no longer speak the Gurung language (Kansakar 1996: 23). The Sherpas came little by little into the eastern part of the country (Solu-Khumbu) from the Khams region of Tibet (the eastern part of Tibet) starting in the sixteenth century (Oppitz 1974, cited in Nishi
There was also a relatively large migration of Tibetans from Krong to Langtang in the 1730s. These Tibeto-Burman speakers live mostly in the northern hills of the country, while the lowlands are now inhabited by Hindu Indo-Aryan speakers, many of whom migrated there between the eleventh and thirteenth centuries. A large number of Central Tibetan (Lhasa) speakers have come to Nepal and India since the failed 1959 uprising against Chinese rule in Tibet.

Nepali, an Indo-Aryan language, is the official language of Nepal, and so is used for official purposes and in education, law, and the media. Fifty per cent of the population are said to be native speakers of Nepali (Kansakar 1996). While all indigenous languages are recognized as national languages by the 1990 constitution, aside from Nepali, only two other languages (Maithili, Indo-Aryan; Newar, Tibeto-Burman) are offered in school (as electives) beyond the primary level. Nepali is clearly the dominant language, and 'non-Nepali speakers have been at a disadvantage in education, employment and other social benefits' (Kansakar 1996: 18).

There is then great pressure to learn Nepali and this has caused an increase in bilingualism and language shift. Most of the people of the country now are bilingual in Nepali, and many languages show influence from Nepali, particularly the development of a dative human patient ('anti-ergative', LaPolla 1992b) marker [i], and in some cases convergence of grammatical categories and use, such as the convergence of the tense and ergative marking systems in Nepali and Newari (see Bendix 1974). Some of these convergences may be assisted somewhat by what Jakobson (1938) called 'linguistic affinity'; for example, the Tibetan dialects already had a locative marker [la] that could be used for dative and human patient marking. Quite a few of the languages, in fact almost all of the Kiranti (Rai) languages are endangered. In Bhutan, where there were in the past only Southern Tibetan (west) or Monpa (east) speakers, there are now a large number of Nepali speakers (though in recent years many have been expelled from Bhutan because of conflicts with the Bhutanese).

Quite a large number of Tibeto-Burman languages are found in the northwestern and north-eastern parts of India and in Bangladesh, mainly languages that came from Burma in the east, but also some from Tibet in the north. They have been greatly affected by the cultures they have come into contact with. To give a few examples, in Kashmir two varieties of Tibetan have developed: Balti and Ladakhi. Balti is spoken in the (Pakistan-controlled) Moslem Baltistan area of northern Kashmir. The speakers of Balti are now also Moslems and write their language, which is a Western Tibetan dialect, with the Arabic script. Ladakhi is in the Indian-controlled area of Kashmir, and the speakers are still more culturally

12 A competing theory, also mentioned by Nishi (1986), citing Qu (1985), is that the Sherpas migrated in the early to mid-thirteenth century. Qu (1985) also says that a part of this Sherpa population in Nepal moved back to Tibet (Shigatse in Central Tibet) about three hundred years ago, and that the speech of these migrants, due to influence from the surrounding central dialects, is now classified as a central dialect (rather than an eastern dialect), but still retains elements of the tone system of the eastern dialects.
discussion of areal features shared by the different languages in northern Thailand in Matisoff (1986), and see Bradley (1998b) on the language change and shift in progress of the Gog language; see Bradley (1986, 1998b) on the factors involved in the persistence or non-persistence of minority languages in Thailand).

3. Metatypy

I have argued elsewhere (LaPolla 1998) that language is not something separate from culture or cognition. How we represent some state of affairs represents how we conceive of that state of affairs, and how we conceive of it is related to cultural norms and experiences. When people learn some aspect of another language, if the influence of the culture associated with that language is not great, the borrowers will assimilate the borrowed form to their way of thinking. An example of this might be the distinction of animate and inanimate in relative pronouns in Chaudangsi, even though that distinction was not part of the borrowed structure. If there is heavy enough cultural contact, the contact may slowly change the way the borrowers conceptualize certain events, such that they develop what Bhattacharya (1974) has called 'new agreements in their outlook of life', thereby creating 'a common cultural core'; what Ross (1998) gives as the reason for metatypy: speakers 'increasingly come to construe the world around them in the same way' as some other group. This common cultural core or construal of the world can then lead to the spread of certain constructions or linguistic patterns. For example, in the Wutun language (Chen Naixiong '982), which is a heavily Tibetanized form of Chinese in Qinghai, rather than using two words for 'widow' and 'widower', as is standard in Chinese, the speakers of Wutun have come to agree with the Tibetans in not differentiating widows and widowers linguistically, and so use the Chinese form for 'widow' for both. The development of an inclusive/exclusive distinction in the first person plural pronoun in Northern Mandarin due to Altaic influence is another example, as making this distinction means having a clear cognitive category distinction that would lead to the use of different forms. This is true also of the example Ross (Chapter 6) gives of the development of the formal distinction betweenalienable and inalienable possession in Proto-Oceanic because of Papuan contact.

When people are used to using a particular linguistic category in a language they use regularly, they will try to use it in any language they speak. In other words, if some category or lexical item they are used to using is not in one of the languages they are using, there is a perceived gap. Many Cantonese speakers in Hong Kong, when they speak English, will frequently use then (generally said with a rising tone) at the beginning of discourse segments or speech turns. They do this because there is a particle in Cantonese, [sam][3], used in this way, and they feel the need for something with that function when they speak English. Substratum influence, such as the development of the aspect and complementizer patterns that have developed in Taiwanese Mandarin on the model of the Taiwanese dialect (Chappell, this volume) are of this nature. Heine (1994, see also 1997a, b, Heine and Kuteva, this volume) has talked about the importance of event schemas for determining the type of grammaticalization you find in a language. These event schemas are ways of conceptualizing states of affairs. An example Heine discusses is comparatives. How speakers view a comparative situation, whether as a locational schema, an action schema, or whatever, will determine what sort of structure they use to express that situation. This way of thinking can change through contact with another culture, and lead to the development of what are commonly called calques, but are better seen as examples of metatypy. Matisoff (1991) discusses several types of grammaticalization common to the languages of South-East Asia that are based on particular types of schema, such as locative verbs becoming progressives, a verb meaning 'get' becoming an auxiliary meaning 'have to/ must, able to' (see also Enfield, this volume), and a verb meaning 'give' becoming a causative or benefactive auxiliary. We can see how similar the ways of thinking and structure can become from the description of Ronggo (Chamoli District of Uttar Pradesh, India), a language that has been very heavily influenced by Hindi and Garhwali (Indo-Aryan), in Sharma (forthcoming). In discussing a particular participial form, Sharma (p. 38) first gives the English translation in (3a), but says: 'In fact this translation is not very close in its meaning. The Hindi sentence is more appropriate', and then gives the sentence in (3b).

(3) (a) di phal gyi-ta japing ya
       this fruit 1-DAT eaten is
       'This fruit was eaten by me.'

       (b) yah phal mera: khaya: hua: hai
       this fruit 1+poss eaten be +past is
       giving the sense—'I have the experience of eating this fruit in the past.'

One phenomenon in Tibeto-Burman that I think is a case of contact-induced metatypy is the parallel development of person-marking in a large number of Tibeto-Burman languages. The languages with person-marking are almost all spoken around the edge of the Tibetan plateau from north-west China down along the southern edge of the plateau, in an area of large-scale language contact, multilingualism, and mutual influence. I have given arguments elsewhere (LaPolla 1992a, 1994b) why person-marking should not be considered an archaic feature of Tibeto-Burman. Here I will just cite some examples of the person-marking forms in a number of languages to show how the same pattern of grammaticalization was followed in the different languages (similar to what happened in Australia—see Dixon 1980: 363, this volume).

The earliest example we have of person-marking in Tibeto-Burman is in Tangut, a dead language in which there are texts dating back to the eleventh century. In Tangut the optional verbal suffixes have the same phonetic form, including the tone, as the free pronouns (adapted from Kepping 1975, 1979, 1981, 1982, 1989; there is also a first and second person plural marker ni?; third person is not marked): see Table 2.
TABLE 2. Tangut person-markers and free pronouns

<table>
<thead>
<tr>
<th>Free pronouns</th>
<th>Verb suffixes</th>
</tr>
</thead>
<tbody>
<tr>
<td>1sg</td>
<td>*ga</td>
</tr>
<tr>
<td>2sg</td>
<td>*na</td>
</tr>
</tbody>
</table>

In the Kuki-Chin branch of Tibeto-Burman we find a person-marking system very similar to that in Tangut. In this system we find the Proto-Kuki-Chin pronouns *kai '1sg', *na '2sg', and *a-'3sg' grammaticalized into the person-marking prefixes *ka-, *na-, and *a- respectively (Thurgood 1985). Yet from the fact that the system is prefixal, and the fact that the pronouns that were the source of the prefixes are not the same as the Tangut forms (at least the 1sg and 3sg forms), and from the fact that the languages are not closely related, we can say that this system clearly developed independently of the Tangut system.

A middle case is the Kanauri-Almora branch, which has person-marking that is suffixal, like the Tangut system, but has a first person suffix derived from an innovative pronoun somewhat similar to that in Kuki-Chin. The forms are *-ga (< *gai) and *-na (< *naa) (there is no third person agreement suffix) (Thurgood 1985). We can still be confident of the independent origin of this system, though, because the source of the first person affix is different from that of Tangut, and though it may be similar to that of the Kuki-Chin system, it is a suffixal system.

A fourth case of clear independent development is the person-marking system of Angami Naga (Giridhar 1980), which involves prefixes clearly derived from the independent pronouns. The verbal prefixes are also isomorphic (except for the tone on the first person prefix) with the pronominal genitive noun prefixes (22 ff.): see Table 3. Again we see that not only is this a prefixing system, unlike the Tangut system, but it also derives from a set of free pronouns unique to Angami.

A fifth case is the person-marking prefixes of Mikir (Hills Karbi; Jayapaul 1987). Again we have a prefixing system, but one quite different from those discussed above: see Table 4. That this system is a recent development can be seen not only from the fact that the free pronouns and the prefixes are so similar in form, but also from the fact that the verb prefixes retain the inclusive/exclusive distinction of the free pronouns.

In the Kuki-Chin branch of Tibeto-Burman we find a person-marking system very similar to that in Tangut. In this system we find the Proto-Kuki-Chin pronouns *kai '1sg', *na '2sg', and *a-'3sg' grammaticalized into the person-marking prefixes *ka-, *na-, and *a- respectively (Thurgood 1985). Yet from the fact that the system is prefixal, and the fact that the pronouns that were the source of the prefixes are not the same as the Tangut forms (at least the 1sg and 3sg forms), and from the fact that the languages are not closely related, we can say that this system clearly developed independently of the Tangut system.

A middle case is the Kanauri-Almora branch, which has person-marking that is suffixal, like the Tangut system, but has a first person suffix derived from an innovative pronoun somewhat similar to that in Kuki-Chin. The forms are *-ga (< *gai) and *-na (< *naa) (there is no third person agreement suffix) (Thurgood 1985). We can still be confident of the independent origin of this system, though, because the source of the first person affix is different from that of Tangut, and though it may be similar to that of the Kuki-Chin system, it is a suffixal system.

A fourth case of clear independent development is the person-marking system of Angami Naga (Giridhar 1980), which involves prefixes clearly derived from the independent pronouns. The verbal prefixes are also isomorphic (except for the tone on the first person prefix) with the pronominal genitive noun prefixes (22 ff.): see Table 3. Again we see that not only is this a prefixing system, unlike the Tangut system, but it also derives from a set of free pronouns unique to Angami.

A fifth case is the person-marking prefixes of Mikir (Hills Karbi; Jayapaul 1987). Again we have a prefixing system, but one quite different from those discussed above: see Table 4. That this system is a recent development can be seen not only from the fact that the free pronouns and the prefixes are so similar in form, but also from the fact that the verb prefixes retain the inclusive/exclusive distinction of the free pronouns.

One last example is from the Delugong dialect of Sgaw Karen (Dai et al. 1991: 400); third person is unmarked: see Table 5. This system of verbal prefixes is very clearly of recent origin, being in the singular simply unstressed copies of the free pronouns, and unique to this dialect of Karen.

It is unlikely that so many languages developing person-marking in the same way is a coincidence, even given the fact that they are in most cases typologically similar. There must be some other factor, and I believe that factor is language contact, much as the Vietnamese development of tones in a way parallel to that of Chinese is at least partially due to contact with Chinese.

4. Conclusion

I have tried to show in this chapter that the history of the Sino-Tibetan-speaking peoples is one of frequent migration and contact with other languages and cultures, and each other, and that this contact has been a major influence on the development of the Sino-Tibetan language family. To understand why the languages of the family have the forms they do, and why there are difficulties in assigning a clear family-tree structure to the family, language contact must not only be taken into account, but must be considered a fundamental factor in the formation of the family.

But this then brings up a question. Those who do subgrouping (see note 3) often do not give the reasons for their groupings. In some cases there are clear isoglosses, but often subgrouping is affected by the author's subjective 'feel' of the language, shared features, or shared vocabulary, which are all often
influenced by its geographic location. Bradley (1997) is the most straightforward in this regard, as most of the names for his subgroups are geographic (e.g. 'Central Himalayan'). While some may argue that what is at issue is genetics, not location, there is value in grouping the languages geographically because contact has been so important in the development of the languages. This then brings us to a question raised in Dai (1997). Dai argued that the family tree model alone is not sufficient to account for the facts of Sino-Tibetan; we need to take into account language contact that has led to what he called 'language coalescence'. He asks, 'Is it not possible for two languages that were not originally related to become related through intense contact?' For example, could we not resolve the question of the relationship between Tai–Hmong–Mien and Chinese by saying they were not originally related but now are? If we accept geographic groupings that are most probably the result of areal contact, what does that mean for the concept of 'relatedness'?

References


Cao Guangqi, 1997. 'Zhuang-Dongyu he Hanyu Min, Yue fangyan de gongtong dian' [Commonalities among Zhuang-Dong languages and the Min and Yue dialects of Chinese], Minzu Yuyan 2.54-60.


Chen Ruxing. 1995. 'Zhuang-Dongyu he Hanyu Min, Yue fangyan de gongtong dian' [Thoughts on subgrouping in Sino-Tibetan], Yuyan Jiaoxue yu Yuyan 4.4-10.
(Republished in JCL 11–13.)
PULLEYBLANK, E. G. 1983. 'The Chinese and their neighbors in prehistoric and early historic
Tibetan languages 1,
LI SHAOMING,
-- 1988b. Qu
-- 1966. Introduction to Sino-Tibetan,
RAN GUANGRONG PULLEYBLANK, STEIN, SHARMA, S. SUN edited by 6.1-14·
Chubanshe.
Tomiin, Eugene: University of
times'. pp. 411-66 of
nationalityJ, pp. 215-34 of
Proceedings
Universitaires
[On
HONGKAI.1981.
H. 1967. Han Chinese Expansion in
Hamden, CT: Shoe String Press,
XU JIESHUN. 1989. 'Zhongguo guadui manbi minzu guanzhi shi jiao yu yanju duanxiang' [A comparative study of the histories of the northern and southern nationalities in Ancient China], Sizheng Zhanxian 1.55-61.
YEO RUI. 1982. 'Lun Taiyu Hangci zai Hanyu nanfang fangyan zhong de diceng de yicun' [On the substrata and remnants of Tai classifiers in Chinese southern dialects], Minzu Yuwen 3.
WANG HUITHIN. 1989. 'Chunqiu Zhangguo shihui de minzu yuyan gudi kuang he yu yu yu yu lyou jian yi yu' [A brief discussion of the nationality languages and their relationships in the Spring and Autumn and Warring States periods], Zhongguo Minzu Yuwen Xuebao 6:7-9, 79.
XU JIESHUN. 1989. 'Zhongguo guadui manbi minzu guanzhi shi jiao yu yanju duanxiang' [A comparative study of the histories of the northern and southern nationalities in Ancient China], Sizheng Zhanxian 1.55-61.
YEO RUI. 1982. 'Lun Taiyu Hangci zai Hanyu nanfang fangyan zhong de diceng de yicun' [On the substrata and remnants of Tai classifiers in Chinese southern dialects], Minzu Yuwen 3.
On Genetic and Areal Linguistics in Mainland South-East Asia: Parallel Polyfunctionality of ‘acquire’

N. J. Enfield

This chapter raises questions concerning genetic and areal relatedness among languages of Mainland South-east Asia (hereafter MSEA), mainly with reference to a widespread pattern of grammatical polyfunctionality involving a verb ACQUIRE. Although data are mostly from Sinitic and Tai, the issues of genetic versus areal relatedness arise across and throughout the five or more language families in the region.

I begin with introductory comments on the geographical, linguistic, and cultural situation of MSEA, including discussion of MSEA as a linguistic area. In §2, I present data from a synchronic case study of a polyfunctional verb ACQUIRE in MSEA languages, concentrating on two Tai languages (Lao and Northern...