Gair J W (1998). *Studies in South Asian linguistics: Sinhala and other South Asian languages.* Oxford: Oxford University Press.

Gair J W & Karunatillake W S (1974). *Literary Sinhala.* Ithaca, NY: Cornell University South Asia Program.

Gair J W & Karunatillake W S (1976). *Literary Sinhala inflected forms: a synopsis with a transliteration guide to Sinhala script.* Ithaca, NY: Cornell University South Asia Program.

Gair J W & Paolillo J C (1997). *Sinhala (Languages of the world/materials 34).* München: Lincom.

Gair J W, Karunatillake W S & Paolillo J C (1987). *Readings in colloquial Sinhala.* Ithaca, NY: Cornell University South Asia Program.

Geiger W (1938). *A grammar of the Sinhalese language.* Colombo: Royal Asiatic Society.

Godakumbura C E (1955). *Sinhalese literature.* Colombo: Colombo Apothecaries Ltd.

Gunasekara A M (1891). *A grammar of the Sinhalese language. Adapted for the use of English readers and prescribed for the Civil Service Examination.* Colombo: Government Press. [Reprinted Sri Lanka Sahitya Mandalaya, Colombo: 1962.]

Karunatillake W S (1992). *An introduction to spoken Sinhala.* Colombo: Gunasena.

Karunatillake W S (2001). *Historical phonology of Sinhalese: from old Indo-Aryan to the 14th century AD.* Colombo: S. Godage and Brothers.

Macdougall B G (1979). *Sinhala: basic course.* Washington D.C.: Foreign Service Institute, Department of State.

Matzel K & Jayawardena-Moser P (2001). *Singhalesisch: Eine Einführung.* Wiesbaden: Harrassowitz.

Reynolds C H B (ed.) (1970). *An anthology of Sinhalese literature up to 1815.* London: George Allen and Unwin (English translations).

Reynolds C H B (ed.) (1987). *An anthology of Sinhalese literature of the twentieth century.* Woodchurch, Kent: Paul Norbury/Unesco (English translations).

Reynolds C H B (1995). *Sinhalese: an introductory course* (2nd edn.). London: School of Oriental and African Studies. [1st edn., 1980.]

Sinhalese *See:* Sinhala.

Sino-Tibetan Languages

R J LaPolla, La Trobe University, Bundoora, Australia

The Sino-Tibetan (ST) language family includes the Sinitic languages (what for political reasons are known as Chinese 'dialects') and the 200 to 300 Tibeto-Burman (TB) languages. Geographically it stretches from Northeast India, Burma, Bangladesh, and northern Thailand in the southeast, throughout the Tibetan plateau to the north, across most of China and up to the Korean border in the northeast, and down to Taiwan and Hainan Island in the southeast. The family has come to be the way it is because of multiple migrations, often into areas where other languages were spoken (LaPolla, 2001). Proto-Sino-Tibetan (PST) would have been spoken in the Yellow River valley at least 6000 years ago. Waves of migration followed: to the southeast, forming the Sinitic languages, and to the west and southwest, forming the TB languages (the speakers of what became the Bodish languages migrated west into Tibet and then south, all the way to the Bay of Bengal, while the speakers of what became the rest of the TB languages followed the river valleys down along the eastern edge of the Tibetan plateau and across into Burma, India, and Nepal). The large spread of Mandarin Chinese to the northwest, southwest and northeast, giving it its large population and geographic spread, happened only in the last few hundred years.

In the past, and to some extent in China still today (e.g., Ma, 2003), this family was also said to include the Tai-Kadai (Zhuang-Dong) and Hmong-Mien (Miao-Yao) languages of southern China and Southeast Asia, but the resemblances found among Sinitic, Tai-Kadai, and Hmong-Mien are now understood to be a result of contact influence (these peoples originally inhabited southern China). Sino-Tibetan has the second largest number of speakers of any language family in the world, due largely to the over one billion Sinitic speakers; except for Burmese (see Bradley, 1996), most Tibeto-Burman languages have relatively few speakers.

Subgroupings within ST are still controversial, due to differences in criteria for subgrouping, a paucity of reliable data, particularly on morphosyntactic
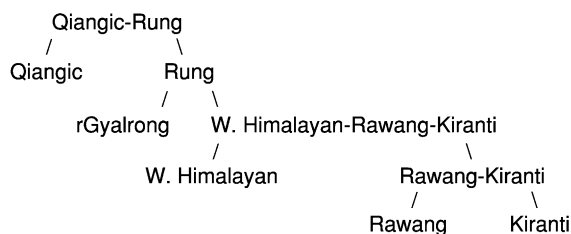
```
            Qiangic-Rung
           /            \
    Qiangic              Rung
                        /    \
              rGyalrong      W. Himalayan-Rawang-Kiranti
                            /
                  W. Himalayan          Rawang-Kiranti
                                       /              \
                                  Rawang          Kiranti
```

Figure 1 The subgrouping of Qiangic-Rung.

patterns, and the fact that the development and distribution of these languages has been greatly influenced by migration and language contact. Some of the influential proposals for subgrouping within TB are Grierson 1909, Shafer 1955, Benedict 1972, DeLancey 1987, Sun 1988, Dai, Liu & Fu 1989, Bradley 1997, Matisoff 2003, and Thurgood 2003 (see Hale 1982 for comparison of the older proposals). There is now general agreement on the existence of the following groupings (individual languages listed are only representative; see Matisoff 1996 for the many different names used for TB languages and groupings).

- Qiangic (Qiang, Pumi, Muya, Namuyi, Shixing);
- Lolo-Burmese, comprising the Burmish languages (Burmese, Lawngwaw [Maru], Ngo Chang [Achang], Zaiwa, Lachik [Lashi]) and the Loloish languages (further divided into Northern: Nosu [Yi, Yunnan or Sichuan], Nasu, Nisu; Central: Lahu, Lisu, Nusu, Jinuo; and Southern: Hani, Bisu, Phunoi, Mpi;
- Bodish (Tibetan, Dzongkha, Tamang (several varieties), Tshangla, Takpa);
- Kuki-Chin (Lushai, Asho Chin, Tiddim [Chin, Tedim], Anal, Hmar);
- Bodo-Koch (Bodo, Garo, Dimasa, Kachari, Koch, Rabha);
- Konyak (Tangsa [Naga, Tangsa], Chang [Naga, Chang], Konyak [Naga, Konyak], Nocte [Naga, Nocte], Wancho [Naga, Wancho]);
- Tani (Apatani, Mising [Miri], Adi); and
- Karenic (Pwo [Karen, Pwo], Karenni, Sgaw [Karen, S'gaw]).

There is much controversy over the affiliations of many of the languages of Northeast India and whether they all form a group together (see Burling, 1999; Matisoff, 1999), as well as the positions of the Bai language of Yunnan, China, Newari and the Kiranti languages of Nepal, Dulong-Rawang-Anong (Rawang) of Burma and China, the extinct Tangut language of northwest China, and the rGyalrong language of Sichuan, China, among many others. The latter two are most often said to be part of the Qiangic group, and the Kiranti languages are often seen as forming a higher grouping with the Bodish

languages, but LaPolla (2003a), with reference to the morphological paradigms, argued that rGyalrong, the Kiranti languages (Bantawa, Athpare [Athapariya], Dumi, Khaling, Camling), Dulong-Rawang-Anong, the Kham languages, and the Western Himalayan languages (Kinnauri, Rongpo, Chaudangsi, Darmiya; also often grouped with Bodish) should be seen as forming a single higher-level grouping. This grouping was given the name 'Rung' because of the similarity (but not identity) of this proposal to an earlier one by Thurgood (1985). The Rung languages most likely split off from an even higher-level grouping with the Qiangic languages, then rGyalrong split off from the group as migrations moved south, then Western Himalayan split off from Kiranti and Rawang, and then these two groups split ([Figure 1](); see LaPolla, 2003a, for the evidence).

Within Sinitic, it is generally agreed there are at least six major dialect groups, initially distinguished on the basis of the reflexes of the historically voiced initial consonants (Li, 1936–1937): Mandarin (northern and southwestern China), Wu (Jiangsu and Zhejiang), Xiang (Hunan), Gan (Jiangxi), Yue (Guangdong and Guangxi), and Min (Guangdong, Fujian, Hainan Island, and Taiwan). The Hakka group of dialects (Guangdong, Fujian, Jiangxi, Sichuan, and Taiwan) is seen by some as part of the Gan group and by others as a separate group. Another three groups were proposed by Li (1987): the Jin group (Shanxi and Inner Mongolia), the Hui group (Anhui and Zhejiang), and the Pinghua group (Guangxi), but these groupings are not universally accepted. Norman (1988, 2003), based on a paradigmatic set of lexical and grammatical items, further grouped the dialect groups into the Northern (Mandarin) group, the Central group (some Xiang dialects, Wu, Gan), and the Southern group (Yue, Hakka, and some Xiang dialects). He left out the Min group because he felt that the Min dialects lay "outside the mainstream of Chinese linguistic development" (2003: 81). That is, they cannot be reconciled with the reconstructed Middle Chinese system (seventh century A.D.) to which the other dialect groups can be traced.

Mandarin has the largest geographic spread and population, and can be subdivided into as many as eight subgroups (see Li, 1987; cf. Ho, 2003), based largely on the reflexes of the stopped tone category. Of these, the Southwestern (Sichuan, Yunnan, Guizhou), Central Plains, and Jianghuai (Southeastern) groups are generally recognized.

One variety of Mandarin, *P≪ut – onghu\a*, the 'Common Language' of China today, was developed in the early 20th century (and dubbed *Gul̇óy≪ǔ*, 'National Language,' at that time), taking the

phonology of the Beijing dialect but the lexicon and grammar from a more generalized Mandarin and from the vernacular literature of the time. Standardization and spread of the standard through aggressive educational programs continues today.

Min does not have a large spread and population, but because of the complex nature of its historical development (multiple migrations into the area, causing multiple strata, even within a single variety), it can be subdivided into as many as seven subgroups: Southern, Northern, Central, Eastern, Puxian, Shaojiang, and Qiongwen (Li, 1987). For an excellent book-length synchronic and historical overview of Sinitic, see Norman, 1988; for the best detailed analysis of a single dialect, see Chao (1968).

Proto-Sino-Tibetan was monosyllabic, but with a much more complicated syllable structure than most of the modern languages: $*$(PREF) (PREF) C_i (G) V (:) (C_f) (s) (Matisoff, 1991: 490; Ci = initial consonant, G = glide, : = vowel length, C_f = final consonant, s = suffixal $*$-s; parentheses mark items that do not appear in all syllables). The modern languages have moved much more toward bisyllabic or polysyllabic words, although they are often reduced again to sesquisyllabic (syllable and a half) or monosyllabic forms, and tone systems have developed in Sinitic and many of the TB languages (either through contact, through independent innovation, or a combination of the two). For example, in Sinitic the tones developed out of consonant suffixes ($*$-s, $*$-ʔ) and loss of initial voicing (Baxter, 1992: 8.2), and in Lhasa Tibetan the tones developed independently, out of loss of initial voicing and the influence of final consonants. Within this general commonality there is also diversity in phonemic inventories and syllable structures, with, for example, the Qiang language (LaPolla, 2003b) having 36 initial consonants, a complex system of consonant clusters in initial and final position, and no tones, while Lahu (Matisoff, 1973) has only 24 consonant initials, a simple (C)V syllable structure (no consonant clusters), and seven phonemic tones.

Proto-Sino-Tibetan morphology included derivational prefixes and suffixes and a voicing alternation of the initial consonant of some verbs that could affect the valency or form class of a word, but no relational morphology. Many of the modern languages have grammaticalized person-marking affixes on the verb and/or semantic role marking on nouns, but these cannot be reconstructed to the PST level (see LaPolla, 2003a, and references therein). The clause was verb focused, in that the verb was the key element, and noun phrases were optional. This is still the case in most languages. Most have not grammaticalized the kind of constraints on referent identification we associate with the concept of 'subject' and other grammatical relations. If noun phrases appeared in the clause, the verb would have been clause final. In Sinitic the clause is largely verb medial, as the verb has come to function as the divider between topical (preverbal) and nontopical (postverbal) elements (there has clearly been a progressive change away from verb-final order over time). This change has happened to a large extent in Bai and Karen as well. With morphology as with phonology we find diversity of types. Using our examples of Qiang and Lahu again, we find Qiang is agglutinative, whereas Lahu is isolating. Qiang has complex affixal systems of direction marking, person marking, and evidential marking on the verb and definite marking in noun phrases, whereas Lahu has none of these features. Both languages have developed complex sortal classifier systems – a common, but not universal, trait among ST languages. All ST languages have modifier-modified order in noun–noun structures (with genitive-head order being a subtype of this – there was no genitive marking in PST, but some languages have developed genitive marking), as well as relative-head order (Karen has a secondary head-relative order as well). Proto-Sino-Tibetan had negative-verb order, and this is still true of most ST languages.

Matisoff (2003) grouped the languages in the family into the 'Sinosphere' and the 'Indosphere' due to the linguistic and political influence of China and India, respectively, on the languages. In Indospheric languages, such as the TB languages of Northeast India and Nepal, for example, we often find the development of relative pronouns and corelative structures, and also of retroflex initial consonants. In the Sinosphere we often find the development of tone systems and more analytic structure. We also find contact influence from the Altaic languages in the north (Altaic speakers controlled large parts of northern China for long periods over the last thousand years) and the Austroasiatic, Tai-Kadai, and Hmong-Mien languages in the south. For example, there is a cline from north to south in terms of complexity of tone and also classifier systems (greater in the south, less in the north), and influence on prosody and word structure where the sesquisyllabic light-heavy structure of Austroasiatic languages is also found in many of the southern TB languages, such as Burmese and Jinghpaw (Jingpho), often leading to the reduction of the first syllable in a compound, in contrast to a trochaic stress pattern in northern TB and northern Sinitic, which often leads to the reduction of the second syllable in compounds.

See also: Bangladesh: Language Situation; Burma: Language Situation; China: Language Situation; Chinese; India: Language Situation; Nepal: Language Situation; Thailand: Language Situation; Vietnam: Language Situation.

Bibliography

Baxter W H (1992). *A handbook of Old Chinese phonology.* Berlin & New York: Mouton de Gruyter.

Benedict P K (1972). *Princeton-Cambridge studies in Chinese linguistics 2: Sino-Tibetan: a conspectus.* Cambridge: Cambridge University Press.

Bradley D (1996). 'Burmese as a lingua franca (and associated map, #87).' In Wurm S A, Mühlhäusler P & Tryon D T (eds.) *Atlas of languages used for intercultural communication in the Pacific, Asia, and the Americas, vol. II.1.* Berlin: Mouton de Gruyter. 745–747.

Bradley D (1997). 'Tibeto-Burman languages and classification.' In Bradley D (ed.) *Papers in Southeast Asian linguistics No. 14, Pacific Linguistics Series A–86: Tibeto-Burman languages of the Himalayas.* Canberra: Australian National University. 1–71.

Burling R (1999). 'On Kamarupan.' *Linguistics of the Tibeto-Burman Area 22(2),* 169–171.

Chao Y R (1968). *A grammar of spoken Chinese.* Berkeley & Los Angeles: University of California Press.

Dai Q, Liu J & Fu A (1989). 'Guanyu woguo Zangmian yuzu xishu fenlei wenti (On the problem of genetic subgrouping within the Tibeto-Burman languages of China).' *Yunnan Minzu Xueyuan Xuebao 3,* 82–92.

DeLancey S (1987). 'The Sino-Tibetan languages.' In Comrie C (ed.) *The world's major languages.* New York: Oxford University Press. 799–810.

Grierson G A (ed.) (1909). *Linguistic survey of India, III, Parts 1–3: Tibeto-Burman family.* Calcutta: Superintendent of Government Printing.

Hale A (1982). *Research on Tibeto-Burman languages. Trends in linguistics, state of the art report, 14.* Berlin & New York: Mouton.

Ho D (2003). 'The characteristics of Mandarin dialects.' In Thurgood & LaPolla (eds.). 126–130.

LaPolla R J (2001). 'The role of migration and language contact in the development of the Sino-Tibetan language family.' In Dixon R M W & Aikhenvald A Y (eds.) *Areal diffusion and genetic inheritance: case studies in language change.* Oxford: Oxford University Press. 225–254.

LaPolla R J (2003a). 'An overview of Sino-Tibetan morphosyntax.' In Thurgood & LaPolla (eds.). 22–42.

LaPolla R J, with Huang C (2003b). *A grammar of Qiang, with annotated texts and glossary.* Berlin: Mouton de Gruyter.

LaPolla R J & Lowe J B (1994). *STEDT monograph series 1A: Bibliography of the international conferences on Sino-Tibetan languages and linguistics I-XXV.* Berkeley: Center for South and Southeast Asian Languages.

Li F -K (1936–1937). 'Languages, dialects.' *The Chinese, Year Book,* 121–128. [Reprinted (1973) in *Journal of Chinese Linguistics 1(1),* 1–13.].

Li R (1987). 'Chinese dialects in China.' In Wurm S A *et al.* (eds.) *Pacific linguistics, series C, no. 102: Language atlas of China parts I and II.* Hong Kong: Longman Group (Far East). [Map and Text A–2.]

Ma X (ed.) (2003). *Han Zangyu gailun (An introduction to Sino-Tibetan languages)* (2nd edn.). Beijing: Minzu Chubanshe.

Matisoff J A (1973). *University of California publications in linguistics, 75: The grammar of Lahu.* Berkeley & Los Angeles: University of California Press.

Matisoff J A (1991). 'Sino-Tibetan linguistics: present state and future prospects.' *Annual Review of Anthropology 20,* 469–504.

Matisoff J A (1996). *STEDT monograph series 2: Languages and dialects of Tibeto-Burman.* Berkeley: Center for South and Southeast Asian Languages.

Matisoff J A (1999). 'In defense of Kamarupan.' *Linguistics of the Tibeto-Burman Area 22(2),* 173–182.

Matisoff J A (2003). *Handbook of Proto-Tibeto-Burman: system and philosophy of Sino-Tibetan reconstruction.* Berkeley, Los Angeles, & London: University of California Press.

Norman J (1988). *Chinese.* Cambridge: Cambridge University Press.

Norman J (2003). 'The Chinese dialects: phonology.' In Thurgood & LaPolla (eds.). 72–83.

Shafer R (1955). 'Classification of the Sino-Tibetan languages.' *Word 11(1),* 94–111.

Sun H (1988). 'Shilun Zhongguo jingnei Zang-Mian yu de puxi fenlei (A classification of Tibeto-Burman languages in China).' In Eguchi P K *et al.* (eds.) *Languages and history in East Asia: a festschrift for Tatsuo Nishida on the occasion of his 60th birthday,* vol. I. Kyoto: Shokado. 61–73.

Thurgood G (1985). 'Pronouns, verb agreement systems, and the subgrouping of Tibeto-Burman.' In Thurgood G, Matisoff J A & Bradley D (eds.) *Linguistics of the Sino-Tibetan area: the state of the art.* Canberra: Department of Linguistics, Australian National University. 376–400.

Thurgood G (2003). 'A subgrouping of the Sino-Tibetan languages: the interaction between language contact, change, and inheritance.' In Thurgood & LaPolla (eds.). 1–21.

Thurgood G & LaPolla R J (eds.) (2003). *The Sino-Tibetan languages.* London & New York: Routledge.

Relevant Websites

http://socrates.berkeley.edu/~ltba/ – *Linguistics of the Tibeto-Burman Area.*

http://socrates.berkeley.edu/~jcl2/ – *Journal of Chinese Linguistics.*

http://stedt.berkeley.edu/ – Sino-Tibetan Dictionary and Thesaurus Project.

http://tibeto-burman.net – Tibeto-Burman Linguistics Domain.