Introduction to Volume I: Establishing the relationships

In this volume we present several of the articles that led to the understanding of the Sino-Tibetan family that we now have (Section A), plus some of the reconstruction efforts based on that understanding (Section B).

Section A: Establishing the relationships

The first article, "Languages and dialects", is from Fang-Kuei. Li, who laid out a basic outline of the languages of China in an article for the *Chinese Year Book* published by Commercial Press in Shanghai, 1937 (pp. 59-65). This became a very influential view of the family, and so when William S-Y. Wang established the *Journal of Chinese Linguistics (JCL),* he included a slightly revised version of this paper as the first article of the first issue of *JCL* (Li 1973). In the article, Prof. Li includes Hmong-Mien and Kam-Tai within what he calls the Indo-Chinese family together with Chinese and Tibeto-Burman as major branches. Within Chinese he recognizes the Northern Mandarin group, the Eastern Mandarin group, the Southwestern Mandarin group, the Wu group, the Gan-Hakka group, the Yue group, the Xiang group, and the Min group, and also mentions isolated unnamed groups in Anhui, Guangxi, and Hunan, the latter recently investigated by Hillary Chappell (e.g. 2015). Prof. Li divides the Tibeto-Burman (TB) branch into the Tibetan group, the Bodo-Naga-Katchin group (presaging comments in Benedict 1976 and Burling 1983), the Burmese group (within which he includes the Kuki-Chin languages), and the Lolo group.

Also in the late 1930's there was a government sponsored make-work project created by Robert Shafer to collect and compare materials on Sino-Tibetan languages. Paul Benedict later led the project, and writes about the project in the second article in this volume, "Where It All Began: Memories of Robert Shafer and the 'Sino-Tibetan Linguistics Project'" (Benedict 1975). It is included because of the pioneering nature of the work that was done on that project, and the foundation it gave to the development of Sino-Tibetan historical linguistics. Shafer and Benedict had somewhat different methodologies and came to different conclusions about some of the relationships within the family, reflected in the following two articles, Benedict 1942, "Thai, Kadai, and Indonesian: a new alignment in Southeastern Asia"which argues that the Kam-Tai and Hmong-Mien languages should not be included in the Sino-Tibetan family, and Shafer 1955, "Classification of the Sino-Tibetan languages", which is something of a response to that proposal, arguing that what Benedict was saying was

actually his idea. Shafer argues that Chinese and the Tai languages ("Daic") do not form a unit, and he expresses doubt that they are related at all, yet he still includes Tai within the Sino-Tibetan family, arguing that the main divisions of the family are "Sinitic (Chinese), Daic, Bodic, Burmic, Baric, and Karenic" (Shafer 1955: 99). He gives cognate counts to argue for a more or less close relationship between certain divisions of the family, for example claiming that Bodic is closest to Sinitic, but his high cognate counts for Sinitic and Tibetan are simply due to a bibliographic bias, as they are the two languages with the largest dictionaries and so it is easier to find cognates. Fang-Kuei Li and several others still had trouble accepting Benedict's exclusion of Tai from Sino-Tibetan, pointing to unanswered questions about correspondence sets. See for example Li 1976 (in Volume 2—see the discussion in the introduction), Li 1978, as well as Luo 2008 for the history of the debate and new data that Prof. Luo brings to bear on the question.

The work Benedict did on the project led to the monograph *Sino-Tibetan: A Conspectus* (initially drafted in the 1940's but edited by James A. Matisoff and published in 1972), which in turn stimulated the whole field of Sino-Tibetan comparative studies, and led to the development of a similar project on a greater scale, the Sino-Tibetan Etymological Dictionary and Thesaurus project led by Prof. Matisoff (see Matisoff 2003 and the STEDT website: http://stedt.berkeley.edu).

The fifth article in the volume is a follow-up to the re-printing of Li 1937 by Prof. Matisoff, "Notes on Fang-Kuei Li's 'Languages and dialects of China'", published in the third issue of *JCL* (Matisoff 1973a). It is a critique of Prof. Li's outline of the Indo-Chinese family, arguing that based on Benedict's work, the Tai and Hmong-Mien languages should not be included in the family, now called Sino-Tibetan. Matisoff also argues that the outline of Tibeto-Burman given by Prof. Li is out of date, and gives the view current at the time, with Burmese and Lolo being closely related, and Kuki-Chin being more closely related to the Naga languages than Burmese or Lolo, Bodo-Garo being a separate group, and Katchin (Kachin/Jinghpaw/Chinghpaw) being "in a class by itself" (p. 473). Since that time much new data has appeared, and Matisoff himself has revisited the question of the position of Jinghpaw, arguing it is a Luish language (Matisoff 2013).

The sixth article, "Sino-Tibetan: Another look" was published by Paul K. Benedict (1976) four years after the publication of his *Conspectus*, and is due to the feedback and criticisms

he received in reviews of that book and also from discussions at the annual Sino-Tibetan conferences (see Matisoff 1973b for a review of the first five conferences, now called the International Conference on Sino-Tibetan Languages and Linguistics (ICSTLL), and LaPolla & Lowe 1994 for a bibliography of the first 25 conferences—the 50th ICSTLL is to be held in Beijing in November 2017). In the article he responds to critics and presents a large amount of data, with the article structured around the following key questions: a) Is Sino-Tibetan clearly a family?, b) If so, does it include the Tai and Hmong-Mien languages?, c) Should Chinese be set off against Tibeto-Burman (Tibeto-Karen) and then Karen from Tibeto-Burman? Benedict uses comparisons of core vocabulary (Swadesh 100 list) to assert that Sino-Tibetan does in fact form a clear family, but Tai and Hmong-Mien should not be included, and Chinese should be set off as a major branch from Tibeto-Burman, though there was contact with Tibeto-Burman languages after the split, and Karen should not be set off from the rest of Tibeto-Burman. There is also a lengthy discussion updating the *Conspectus* with new findings related to the influence of prefixes in Chinese and TB languages.[1]

The seventh article, Matisoff 1990, "On megalocomparison", is largely a criticism of Greenberg 1987, but it is included here because it is a discussion of the methodology used in determining genetic relatedness, and also compares and contrasts Greenberg's methodology with that of Paul K. Benedict, treating Benedict's methods quite critically. It also introduces the concepts of "Indosphere" and "Sinosphere" to identify the cultural spheres of the two major influences on the Tibeto-Burman languages. This paper is frequently cited in the Sino-Tibetan literature on methodology in historical and contact linguistics. The next item, "Comment on Matisoff's comparison between Greenberg and Benedict" (Benedict 1991) is a two-page reply by Benedict to Matisoff's comments about his methodology.

The ninth article, "Sino-Tibetan linguistics: Present state and future prospects", is an overview of the state of the art of Sino-Tibetan linguistics around 1990, published by Prof. Matisoff in 1991. It covers the history of the development of the field and what had been done on all aspects of the family up to that point. Aside from genetic relations, it also includes discussion of areal and typological relations, including a discussion of the position of Tai and Hmong-Mien languages. It has copious (216!) references, giving a very good

---

[1] For other views of the structure of Tibeto-Burman or Sino-Tibetan not represented here, see DeLancey 1987, Matisoff 1991, Bradley 1997, 2002.

introduction to the literature up to that point. (See also the bibliography at http://tibeto-burman.net/bib/ for more references.)

The tenth to twelfth articles are influential articles that deal with establishing subgroups within Tibeto-Burman. The first of these, "The "Sal" languages", is Robbins Burling's often-cited paper establishing the "Sal" languages as a separate grouping (1983). This grouping includes the Bodo-Garo and "(North)-Eastern Naga" (Konyak) languages, as well as Jinghpaw, and he compares forms from these languages with reconstructions in Benedict 1972 to show the uniqueness of some of the lexical items or their reflexes found in these languages. See Coupe 2012 for an up-to-date critical assessment of the different schemes for sub-grouping these languages.

Following that is Karen Ebert's (1990) paper, "On the evidence for the relationship Kiranti-Rung", which shows the remarkable similarities in the person marking affix paradigms found in Kiranti languages, rGyalrong (Gyarong), and Rawang-T'rung (the latter problematically referred to as "Nungish"). She is very conservative in her conclusions, simply stating that she found "striking similarities in the verbal paradigms of Gyarong and Eastern Kiranti on the one hand and Nungish and Khaling-Dumi on the other hand . . . an independent invention of the complex verbal paradigms of Gyarong and Eastern Kiranti is most unlikely." (p. 76). She gives reasons why they are unlikely to be retentions from Proto-Tibeto-Burman, and argues that they must be a shared innovation. This hypothesis was later used as a test case for a more rigorous approach to establishing genetic relatedness in LaPolla 2013 (first presented at ICSTLL33 in 2000; see also LaPolla 2012). The methodology is based on the fact that in Indo-European linguistics what was used to establish the family was not arbitrary word lists, but morphological paradigms (Nichols 1996). The results show strong statistical evidence that these three groups, plus the Western Himalayan languages, form one group within Tibeto-Burman due to the shared innovation of this paradigm, and the results also dovetail with work on the migrations of the Tibeto-Burman people (LaPolla 2001), where these groups are seen to have migrated from north-western China down along the river valleys skirting the eastern and southern sides of the Tibetan Plateau (see also LaPolla 2006, 2017 for

the broader picture). The idea that this particular person marking paradigm[2] might be used to establish a sub-group within Tibeto-Burman was first mentioned by Paul K. Benedict in the sixth article in this volume (Benedict 1976), and LaPolla argues it defines what he calls the "Rung" group (distinct from the grouping by that name proposed in Thurgood 1984, 1985).

The last article in this set, Sun 1993a, "The linguistic position of Tani (Mirish) in Tibeto-Burman: a lexical assessment", is a summary of the findings of the UC Berkeley PhD dissertation of Jackson T-S. Sun (Sun 1993b). Through careful fieldwork and comparative work, it establishes the Tani languages (formerly known as Mirish or Abor-Miri-Dafla) as a branch within Tibeto-Burman. The dissertation also includes reconstructed Proto-Tani forms. Prof. Sun's work on this group has been accepted by all scholars in the field, and it has laid the foundations for work done since its publication by Mark Post and others in elaborating on this family (see Post & Sun 2017 for a summary, and Post & Burling 2017 for the larger picture of the languages of Northeast India).[3]

Section B: Sino-Tibetan historical reconstruction
The first paper of this section, "The number 'a hundred' in Sino-Tibetan", is a two-page note by Jean Przyluski and Gordon H. Luce (1931) on reconstructing the Proto-Sino-Tibetan (PST) form for 'hundred'. It is included here because it is one of the earliest published attempts at such reconstruction, and it is surprisingly modern in its approach, accepting the possibility of bisyllabic roots in the proto-language, when most scholars had assumed the roots to be monosyllabic.

The next paper, "Concerning the variation of final consonants in the word families of Tibetan, Kachin and Chinese", by Stuart N. Wolfenden (1937), is a broader application of Karlgren's idea (1933—see Volume 3: Sinitic) of comparing word families as a first step in identifying

---

[2] Person marking has been independently innovated many times in Tibeto-Burman (see LaPolla 2001); the relevant paradigm used here to identify what is called the "Rung" group in LaPolla 2013 is only one of them, though is the oldest identified.

[3] After his work on the Tani languages, Prof. Sun turned to fieldwork on and comparison of the rGyalrong languages of northern Sichuan and has similarly given us a much better understanding of that whole group of languages (e.g. Sun 2000, 2017), and in the process has also been investigating the Tibetan and Qiang dialects in the same area (e.g. Sun 2014), improving our understanding of the language situation of the whole of northern Sichuan.

cognate forms. In this paper, Wolfenden compares word families in Kachin, Tibetan, and Chinese with a view to understanding variations in the final consonants of the forms, and argues, for example, that Written Tibetan -*s* represents *-ds in a number of cases, based on comparisons with stop-final forms in the other languages. The idea of using word families as a way of identifying roots and cognates has become standard practice in the field, and as evidenced in many works assuming this methodology, it is the first step in identifying possible morphological forms (suffixes and prefixes and initial alternations), in that it gives us the forms we need to explain.

The particular question addressed by Wolfenden is returned to in LaPolla 1994, "Variable finals in Proto-Sino-Tibetan", our fourth paper in this section, though aside from giving many examples of word families, it discusses methodological issues related to the identification of word families and also to reconstruction methodology. Wolfenden, following Karlgren, limited his word families to only those where the final consonant had the same place of articulation, but LaPolla shows that this assumption is problematic, and would lead us to miss a large number of words that have the same meaning and form except for the final consonant. LaPolla also argues that instead of trying to account for all the variation found in the families by reconstructing overly complex systems or abstract symbols representing unknown variables, we should reconstruct a simple system (see the next paper for one example) and then try to explain the variation or accept it as is for the time being, if it cannot yet be explained. The paper then goes on to show the statistical preponderance of -*Ø* ~ -*k* variation, and attempts to explain this and other variations. The paper also argues for rigour in talking about variation, that is, variation can only be talked about in the context of regularity: if we want to say that two forms that differ in the final are cognate, we should make sure the other segments involved all correspond regularly, otherwise anything goes.

Our last paper in this section, Gong 1980, is an independent comparison of just the three oldest written languages in Sino-Tibetan, Tibetan, Chinese, and Burmese, with a view towards reconstructing the vowels of Proto-Sino-Tibetan. Prof. Gong's methodology differs from many other scholars, in that a) he reconstructs on the basis of only three languages, and b) he does not treat the finals (rimes) as units in opposition to the initials, as is usually done in Sino-Tibetan studies, but in this paper looks only at the vowels, regardless of the final consonant (or not) of the word. His results are quite straightforward, with clear cognate sets and the proposal of a simple system of four vowels (i, a, u, ə) and two rising diphthongs (ia,

ua). He argues that within Sino-Tibetan (or at least within these three languages) only Chinese maintained *ə, and the vowels e and o in Tibetan are secondary.

References

Benedict, Paul K. 1942. Thai, Kadai, and Indonesian: a new alignment in Southeastern Asia. *American Anthropologist* 44: 576-601.

Benedict, Paul K. 1972. *Sino-Tibetan: A conspectus (Princeton-Cambridge Studies in Chinese Linguistics II,* James A. Matisoff, Contributing Editor). Cambridge: Cambridge University Press.

Benedict, Paul K. 1975. Where It All Began: Memories of Robert Shafer and the 'Sino-Tibetan Linguistics Project', Berkeley, 1939-40. *Linguistics of the Tibeto-Burman Area* 2.1: 81-92.

Benedict, Paul K. 1976. Sino-Tibetan: Another look. *Journal of the American Oriental Society* 96.2: 167-197.

Benedict, Paul K. 1991. Comment on Matisoff's comparison between Greenberg and Benedict. *Linguistics of the Tibeto-Burman Area* 14.2: 169-70.

Bradley, David. 1997. Tibeto-Burman languages and classification. In D. Bradley (ed.), *Tibeto-Burman languages of the Himalayas*, Canberra: Australian National University.

Bradley, David. 2002. The subgrouping of Tibeto-Burman. In C. Beckwith (ed.), *Medieval Tibeto-Burman Languages,* 73-112. Leiden: Brill.

Burling, Robbins 1983. The 'Sal' languages. *Linguistics of the Tibeto-Burman Area* 7.2: 1-32.

Chappell, Hilary. M. (ed.). 2015. *Diversity in Sinitic Languages.* Oxford: Oxford University Press.

Coupe, Alexander R. 2012. Overcounting numeral systems and their relevance to sub-grouping in the Tibeto-Burman languages of Nagaland. *Language and Linguistics* 13.1: 193-220.

DeLancey, Scott. 1987. The Sino-Tibetan languages. In B. Comrie (ed.), *The World's Major Languages*, 799–810. New York: Oxford University Press.

Ebert, Karen H. 1990. On the evidence for the relationship Kiranti-Rung. *Linguistics of the Tibeto-Burman Area* 13.1: 57-78.

Gong Hwang-cherng. 1980. A comparative study of the Chinese, Tibetan, and Burmese vowel systems. *Bulletin of the Institute of History and Philology* 51.3: 455-490.

Greenberg, Joseph H. 1987. *Language in the Americas*. Stanford: Stanford University Press.

Karlgren, Bernhard. 1933. Word families in Chinese. *Bulletin of the Museum of Far Eastern Antiquities* 5: 9-120.

Randy J LaPolla. 1994. Variable finals in Proto-Sino-Tibetan', *Bulletin of the Institute of History and Philology* 65.1: 131-173.

LaPolla, Randy J. 2001. The role of migration and language contact in the development of the Sino-Tibetan language family. In R. M. W. Dixon & A. Y. Aikhenvald (eds.), *Areal diffusion and genetic inheritance: Case Studies in Language Change*, 225-254. Oxford: Oxford University Press.

LaPolla, Randy J. 2006. Sino-Tibetan languages. In Keith Brown (ed.), *Encyclopedia of Language and Linguistics, 2nd Edition*, 393-397. London: Elsevier.

LaPolla, Randy J. 2012. Comments on methodology and evidence in Sino-Tibetan comparative linguistics. *Language and Linguistics* 13 (1), 117-132 (Special issue on Sino-Tibetan Comparative Studies).

LaPolla, Randy J. 2013. Subgrouping in Tibeto-Burman: Can an individual-identifying standard be developed? How do we factor in the history of migrations and language contact? In Balthasar Bickel, Lenore A. Grenoble, David A. Peterson, & Alan Timberlake (eds.), *Language typology and historical contingency*, 463-474. Amsterdam and Philadelphia: John Benjamins Pub Co.

LaPolla, Randy J. 2017. Overview of Sino-Tibetan Morphosyntax. In Graham Thurgood and Randy J. LaPolla (eds.), *The Sino-Tibetan Languages, Second edition*, 40-69. London and New York. Routledge.

LaPolla, Randy J. & John B. Lowe. 1994. *Bibliography of the Papers from the International Conferences on Sino-Tibetan Languages and Linguistics.* STEDT Monograph Series, No. 1, Second edition, including conferences I-XXV, 1994.

Li, Fang-Kuei. 1936-1937. Languages and dialects. *The Chinese year book,* 121-128. Reprinted in *Journal of Chinese Linguistics* 1.1: 1-13, 1973.

Li, Fang-Kuei. 1976. Sino-Thai. *Computational Analyses of Asian & African Languages* (National Inter-University Research Institute of Asian and African languages and cultures, Tokyo) 3: 39-48.

Li, Fang-Kuei. 1978. Siamese khot. *Monumenta Serica* 33: 403-406.

Luo, Yongxian. 2008. Sino-Tai and Tai-Kadai: Another look. In Anthony V. N. Diller, Jerold A. Edmondson & Yongxian Luo (eds.), *The Tai-Kadai languages,* 9-28. London & New York: Routledge.

Matisoff, James A. 1973a. Notes on Fang-Kuei Li's 'Languages and dialects of China'. *Journal of Chinese Linguistics* 1.3: 471-474.

Matisoff, James A. 1973b. The annual Sino-Tibetan Conferences: The first five years, 1968-1972. *Journal of Chinese Linguistics* 1.2: 152-62.

Matisoff, James A. 1990. On megalocomparison. *Language* 66.1: 106-20.

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

Matisoff, James A. 2003. *Handbook of Proto-Tibeto-Burman: System and philosophy of Sino-Tibetan reconstruction.* Berkeley: University of California Press.

Matisoff, James A. 2013. Re-examining the genetic position of Jingpho: Putting flesh on the bones of the Jingpho/Luish relationship. *Linguistics of the Tibeto-Burman Area* 36.2: 15-95.

Nichols, Johanna. 1996. The comparative method as heuristic. In M. Durie & M. Ross (eds.), *The comparative method reviewed,* 39-71. New York & Oxford: Oxford University Press.

Post, Mark W. & Jackson T-S. Sun. 2017. Tani languages. In Graham Thurgood & Randy J. LaPolla (eds.), *The Sino-Tibetan Languages, 2nd Edition,* 322-337. London & New York: Routledge.

Post, Mark W. & Jackson T-S. Sun. 2017. The Tibeto-Burman languages of Northeast India. In Graham Thurgood & Randy J. LaPolla (eds.), *The Sino-Tibetan Languages, 2nd Edition,* 213-242. London & New York: Routledge.

Jean Przyluski and Gordon H. Luce. 1931. The number 'a hundred' in Sino-Tibetan. *Bulletin of the School of Oriental Studies* 6.3: 667-8.

Shafer, Robert. 1955. Classification of the Sino-Tibetan languages. *Word* 11.1: 94-111.

Sun, Jackson T.-S. 1993a. The linguistic position of Tani (Mirish) in Tibeto-Burman: a lexical assessment. *Linguistics of the Tibeto-Burman Area* 16.2: 143-88.

Sun, Jackson T.-S. 1993b. A historical comparative study of the Tani (Mirish) branch of Tibeto-Burman. PhD dissertation, University of California, Berkeley.

Sun, Jackson T.-S. (ed.). 2014. *Phonological Profiles of Little-Studied Tibetic Varieties. Language and Linguistics Monograph Series 55.* Taipei: Institute of Linguistics, Academia Sinica.

Sun, Jackson T.-S. 2017. Tshobdun Rgyalrong. In Graham Thurgood & Randy J. LaPolla (eds.), *The Sino-Tibetan Languages, 2nd Edition,* 557-571. London & New York: Routledge.

Sun, Jackson T.-S. 2000. Parallelisms in the verb morphology of Sidaba rGyalrong and Guanyinqiao in rGyalrongic. *Language & Linguistics* 1.1: 161-190.

Thurgood, Graham. 1984. The "Rung" languages: a major new Tibeto-Burman subgroup. *Proceedings of the 10th Annual Meeting of the Berkeley Linguistics Society,* 338–49. Berkeley: Berkeley Linguistics Society.

Thurgood, Graham. 1985. Pronouns, verb agreement systems, and the subgrouping of Tibeto-Burman. In G. Thurgood, J. A. Matisoff and D. Bradley (eds.), *Linguistics of the Sino-Tibetan Area: The State of the Art. Papers presented to Paul K. Benedict for his 71st birthday (Pacific Linguistics C, 87),* 376–400. Canberra: Australian National University.

Wolfenden, Stuart N. 1937. Concerning the variation of final consonants in the word families of Tibetan, Kachin and Chinese. *Journal of the Royal Asiatic Society* 4: 625-655.