Tibetan enters history as the language of the Yarlung Valley, the cradle of the Tibetan Empire (Takeuchi, 2012b, 4). Together with the troops of this empire, the language colonized the entire Tibetan plateau, extinguishing the languages formerly used across that territory (Takeuchi, 2012b, 6). The use of Tibetan across and beyond the plateau at the height of the empire set the stage for its breakup into the spoken languages of today. Researchers often classify Tibetan languages into groups based on the provinces of Tibet in which they are spoken (Denwood, 1999, 23–36; Tournadre & Dorje, 2003, 28–32), but the identity of such geographic groupings with genetic groupings remains to be demonstrated (see Sun, 2003, 794–797). The monograph series Beiträge zur tibetischen Erzählforschung provides a good point of departure into the vast research on Tibetan dialects.

The Earliest Tibetan Literature

Paleographic analysis indicates that the Tibetan script derives from a late Gupta script as used in Nepal and northern India in the early 7th century (van Schaik, 2011; but also comp. Schuh, 2013). Although an often repeated legend avers that a writing system was developed with the express purpose of recording the Buddhist scriptures being imported from India, there is little doubt that during the empire’s initial expansion, writing was introduced circa 650 CE to facilitate administration (Hill, 2010a, 111n4). A gap of about a century separates the invention of the alphabet and the earliest securely datable extant Tibetan documents. A stele inscription in front of the Potala Palace, known as the Zhol or Ngam lam Sri inscription (post-763 CE), is generally recognized as bearing the oldest sample of Tibetan writing (Hazod, 2010). However, a bell discovered circa 2010 in Dpa’ ris in eastern Tibet may date to the earlier reign of Khri Lde gtsug brtsan (704–755 CE; Lha mchog skyabs, 2011). The capacity to inscribe writing on a cast bell implies significant prior familiarity with a writing system. Although the Old Tibetan Annals, representing the official version of imperial history, recorded as it happened (Dotson, 2009, 10–11), the extant manuscript is probably postimperial (Iwao, 2012).

Two types of material date from the Tibetan Empire: inscriptions (Li Fang-Kuei & Coblin, 1987; Iwao et al., 2009) and documents excavated at the Central Asian forts of Miran and Mazar Tagh (van Schaik, 2013, 119; Takeuchi, 1997–1998). Old Tibetan continued to function as a lingua franca of commerce and administration in the oasis cities of the Silk Road for some decades after the Tibetan Empire lost control of these territories in the 9th century (Uray, 1981; 1988; Takeuchi, 1990, 187–189; 2012a; 2012b, 7–9). Paper manuscripts preserved in the library cave of Dunhuang, sealed during the first half of the 11th century (Imaeda, 2008), generally date from this postimperial period. The documents from Dunhuang include historical texts, official documents, foreign literature in translation, divination

Figure 1: Tree diagram of the Bodish languages


Languages: Tibetan
texts, and both canonical and noncanonical Buddhist texts. The most important collections of these documents are found now in London and Paris, with smaller collections elsewhere. The collections of both the Bibliothèque nationale de France (Lalou, 1939–1961) and the British Library (La Vallée Poussin, 1962; Iwao et al., 2012) are partially catalogued; images and metadata for many are available on the website of the International Dunhuang Project (http://idp.bl.uk/), and transliterations of many are available at Old Tibetan Documents Online (otdo. aa.tufs.ac.jp). J. Dalton and S. van Schaik (2006) more recently recatalogued the tantric manuscripts of the British Library. Y. Imaeda et al. (2007) provide 65 well-known and well-studied Tibetan Dunhuang documents in transliteration with complete bibliographic references.

The Translation of Buddhist Texts

Begun already in the 8th century, the translation of Buddhist literature into the Tibetan tongue continues to this day. For example, Dge ’dun chos ’phel published his translation from Pali of the Dhammapada in 1946 (Chompel, 1946), and in 2010 ’Bri gung skyabs mgon ce shang translated the second section of the Chan-related Leng qie shizi ji (楞伽師資記; Records of the Lankavatara Masters) to augment the missing sections of an existing old Dunhuang version. The flow of texts also goes in the other direction; there is currently an ongoing project to translate Tibetan sūtras found in the Kanjur into Chinese (http://khyentsefoundation.org/2014/03/ddbc/), and the 84000 project (http://84000.co/) aims to translate the entire Kanjur into English. Nonetheless, the preponderance of translation work took place long ago. The bulk of non-tantric material was translated already during the latter part of the Tibetan Empire (c. 760–850 CE), and tantric materials were translated mostly before the decline of Buddhism in India in the 13th century.

Standardization and Language Reform in Translations of the Imperial Period

Important for an understanding of the language of Buddhist texts is the existence of a particular style of “Translationese,” which characterizes most translations into Tibetan, not only from Indic language sources but also from Chinese and other languages. This style of language attempts at times to mechanically reproduce the syntax of its source, at the same time that it creates a vocabulary from Tibetan word-stock with which to render new ideas, terms, and concepts. There was, moreover, a self-consciousness to this process among the Tibetan scholars responsible for translation and editing of Buddhist literature. One example of this self-consciousness is the effort at standardization that was undertaken both to revise existing translations and to set parameters for new translation works. The colophons of revised early Tibetan translations refer to the language reform that resulted in a “new terminology.” For example, the Lalitavistara was translated and corrected “in conformity with the decision relative to the new terminology” (bskad gsar bcad kyis kyang bcas nas gstan la phab pa; see Simonsson, 1957, 223n2; Scherrer-Schaub, 2002, 310n33). What this reform meant in terms of the daily practice of translators and editors remains unclear. The confused tale of reform is a story of three triples: three edicts, three treatises, and three catalogues.

The timing and import of the three edicts are subject to disagreement. For example, Bsod nams rgyal mtshan in the Rgyal rabs gsal ba’i me long (1568) explains that at the time of Khri Gtsug lde brtsan (r. 815–841), translation practices were standardized according to a "new terminology" (skad gsar bcad) following the prescriptions of three edicts (bka’ bcos nas gstan pa gsum du mdzad do; Uray, 1989, 7; Sørensen, 1994), but Rin chen bkra shis in his Li shi gur khang (1536) separates each of the “edicts” from one another in time: the first is from the invention of the alphabet (c. 650 CE) until the reign of Khri Gsrong lde brtsan (742–800[?]), the second during the reign of Khri Gsrong lde brtsan (r. 815–841), and the third from the time of Rin chen bzang po (958–1055) until the time of Dharmaśālaśāhāra (1441–1526; see Taube, 1978, 173–174). Relying on the oldest available primary sources, C. Scherrerr-Schaub (2002) concludes that the dating of the first edict is uncertain, the second edict was issued probably in 783 CE (but possibly in 795), and the third edict appeared in 815 CE. It seems that the third edict was essentially a renewal of the second, and the first edict could be a retrospective invention, part of the public-relations efforts of Khri Gsrong lde brtsan to present the national conversion to Buddhism as a renewal of the efforts of his ancestor Khri Gsrong brtsan (605[?]–649).

Three lexicographical works assisted the authorized work of translating Buddhist texts and revising earlier translation: the Bye brag tu rtogs byed chen
po (Mahāvyutpatti), the Bye brag tu rtogs byed' bring po, and the Bye brag tu rtogs byed chung ngu. The second work is better known under the title Sgra sbyor bam po gnyis pa. The third work is no longer extant. The two extant works were in circulation at least by 814 CE (Uray, 1989; Panglung, 1994: Scherrer-Schaub, 2002; Hermann-Pfandt, 2008).

As the work of translation progressed, it became necessary to keep track of the completed translations. Three catalogues aided in this purpose, Lhan kar ma, Phang thang ma, and Mchims phu ma, reflecting respectively the collections held at the imperial residences of Lhan kar, Phang thang, and Mchims phu. The Mchims phu ma is no longer extant. The exact dating of the catalogues is controversial – G. Halkias (2004) and B. Dotson (2007) date the Lhan kar ma to 824 CE, and the Phang thang ma to 842 CE, whereas A. Hermann-Pfandt (2008) dates them respectively to 812 and 806 CE.

N. Simonsson (1957) remains the seminal study comparing Sanskrit originals with early and later translations of the same text as a window into the ateliers of translators and revisers. However, his detailed study has not yet been followed up by enough subsequent research to yield a coherent overall picture. Because no available Tibetan Buddhist text is securely datable to before 783 CE, even an exhaustive diachronic study of Tibetan translations may fail to shed light on the effects of the reforms associated with the “new terminology.” While later Tibetan writers, such as the aforementioned Rin chen bkra shis, place all of those linguistic features now associated with “Old Tibetan” into the period before the first edict (Taube, 1978, 173–174), this view is not correct. These features are pervasive in Dunhuang texts and other materials that postdate the Tibetan Empire (see Tropper, 2010, 17), and the occasional old feature, such as la stogs pa instead of la sogs pa (generally “et cetera”), is found even in the 18th-century Derge Kanjur (comp. D 341, mdo sde, a, 1385).

The translation of such a large volume of Buddhist works, mostly from Sanskrit, required the redeployment of inherited vocabulary to Buddhist ends and the coinage of new terminology. Many existing terms received new interpretations or new associations. For example, the word mdo originally (and frequently still) refers to a place where roads or rivers meet. This word was then apparently used as a metaphor for a text in which various points were brought together – used in this sense, for example, in the term mjal dam guyi mdo chen po (great treaty accord) in the east face of the Sino-Tibetan treaty inscription of 821/822 CE, which stands in front of the Jokhang Temple in Lhasa. From that metaphorical use, it was only a small step to use mdo as a calque of Sanskrit sūtra, the discourse of the Buddha. However, the word mdo is not a literal translation of sūtra, which means “thread,” “string,” and so forth. In other cases terms were carefully calqued on the Sanskrit original, thus Sanskrit tathāgata (“thus-gone”) became Tibetan de bszhin gshegs pa (“thus gone”). For example, the Sanskrit word arhat, which derives from the root arh- “to be worthy,” is translated in Tibetan as dra bcom pa (“one who conquers the enemy”), utilizing a traditional Indian idea that arhat is formed from the noun ari (enemy) and the root han- (to slay). Although such neologisms may be based on historically incorrect etymological analyses, in the vast majority of cases, it is demonstrable that the etymologies on which the translation equivalents were created are found deployed in Indian Buddhist literature, and thus represent traditional Buddhist interpretations of important terms (Simonsson, 1957, 269–270; Verhagen, 2001, 67–69, 75).

Reading “According to Context”

One often encounters the claim that the elaborate inflectional categories of Sanskrit allow for a greater reliability and precision in the interpretation of Sanskrit Buddhist texts, whereas in Chinese and Tibetan, one must “rely on context” to know who is doing what to whom. It is, however, merely a greater familiarity with Sanskrit grammar as opposed to Tibetan grammar and a comparative lack of research into Tibetan that leads to such an impression. Despite the fact that Tibetan translations from Sanskrit, which make up the vast majority of translated Buddhist literature, often allow a rather close reconstruction of both the word-stock and syntax of the underlying Sanskrit source text, the Tibetan system of case marking is quite different and rather incompatible with that of Sanskrit. Tibetan distinguishes ten cases:

1. absolutive: Ø;
2. agentive: -kis (Ç -gisis, -gis, -yis, -s);
3. genitive: -ki (Ç -gisi, -gi, -i, -yi);
4. allative: -la;
5. locative: -na;
6. terminative: -tu (Ç -du, -ru, -su, -r);
7. ablative: -las;
8. instrumental: -la;
9. partitive: -nyi;
10. negative: -m. 

Languages: Tibetan 919
8. elative: -nas;
9. associative: -dang; and

Tibetan does not mark case on each individual word as does Sanskrit but instead only once at the end of each noun phrase (Gruppenflexion). The incompatibility of the grammatical structures of the two languages has led to confusion both within the Tibetan grammatical tradition and in the study of Tibetan in modern academia. As Tibetan distinguishes ten cases, two more than the eight cases of Sanskrit, one cannot claim that in this regard Tibetan is by nature less precise than Sanskrit.

Tibetan is a so-called ergative language. From the perspective of European grammar, this means that in intransitive clauses, the “subject” is in the absolutive case, but in transitive clauses, the “subject” is in the agentive case and the “object” is in the absolutive case. Thus, the case alignment of Tibetan transitive clauses is superficially similar to passives in European languages (Schmidt, 1839, 61; Schiefner, 1855, 418; von der Gabelentz, 1861, 472; Giraudieu, 1916, iv; Tournadre, 1996; Vollmann, 2008). Restating these case uses from the Tibetan perspective, the absolutive is used for the single nominal argument of an intransitive clause (ngus pa’i mchi ma Ø char bzhiin du ba b za nas: “the tears of [their] weeping fell like rain”; D 341, mdo sde, a, 132a1–2) and the patient of a transitive clause (ngas de dag kun Ø bsad nas: “I killed them all”; D 341, mdo sde, a, 169b1), whereas the agentive is used for the agent of a transitive clause (bcom ldan ’das kyis chos bstan pas: “the Bhagavān taught the dharma”; D 341, mdo sde, a, 153b2). The agentive also marks infinitives: bcom ldan ’das kyis utpal sngon po dang ’dra ba shin tu m no ba’i ral gris sha bcad cing: “The Bhagavān cut flesh with a very sharp sword like a blue utpala lotus”; Bailey, 1951, 42).

A Tibetan verb consists of four monosyllabic stems: present, past, future, and imperative (Hill, 201b). Not all verbs separately distinguish the four stems (e.g. gshegs, “to go,” has only one form); those that do distinguish all four do so in a complicated way involving ablaut, prefixes, and suffixes (e.g. gsod, bsad, gsad, and sod, “to kill”; za, zos/bzas, bsa; and zos, “to eat”); these morphological processes are complex and opaque (Hill, 201b; 2014). Although broadly speaking the names of the four stems do say something about their use, their meaning and function is also poorly understood (Zeisler, 2004; for a study of the verb forms in Sanskrit and Tibetan as used in the Bodhicaryāvatāra, see Taube, 1953–1954).

Although Tibetan lacks person agreement, other means do much the same work in making explicit who is doing what. The modern Tibetan languages are well known for encoding the source of information by which a speaker has come to know of what he or she says (Tournadre, 1996; Hill, 2013c), and such evidential systems have been characteristic of the language from an early period (Takeuchi, 1990, 12; 2012b, 11–12; 2015, 410; Hill, 2013b). Similarly, clause-marking morphemes are not well understood; there are many options, -na, -nas, -te, -cing, -pa-la, -pa-las, -pa-dang, and -pas, all of which mean something like “then” or “and.” These suffixes appear to distinguish whether or not the subject changes from one clause to the next (Zadoks, 2002; Haller, 2009). Increasing scrutiny of Tibetan grammar will no doubt lead to an appreciation for the refinement of the structure of the Tibetan language and yield new insights into and better understanding of Buddhist canonical literature.

The Tibetan language has a refined system of honorific terminology, in which certain vocabulary items replace others when used, for instance, of the Buddha or an honored teacher. This applies equally to nouns, adjectives, and verbs. An example of this is the variety of verbs that all can be translated as “to speak”/“to say”: smra ba/ lab pa conveys a neutral stance with regard to the person who speaks or is spoken to; gsung ba indicates that the speaker is someone who is highly regarded by the person who reports this act; zhu ba marks that the person addressed is considered to be of high standing, while giving lesser importance to the speaker himself. The latter category is often found in the first person. These three different levels of honorific verbs are less systematized in the earliest Tibetan literature but soon after became very common in written and spoken Tibetan alike (Kitamura Hajime, 1975; Takeuchi, 1987; DeLancey, 1998).

In overall sentence structure, there is a tendency for translations to maintain a syntax that attempts fidelity to the original; these in many cases might be called calques, both in their syntax and in the formation of their vocabulary. This can lead to shorter sentences than would be expected in native Tibetan, usually terminated by final particles (e.g. in the case of the copula, yin no). A clear instance of the artificiality of translations is the precise calquing of Sanskrit preverbs in Buddhist neologisms. For example, rab tu byung ba at face value in Tibetan means “to emerge thoroughly,” but is used to mean “to take ordination” with Sanskrit pravrajya as a model (comp. Inaba Šōju, 1954, 101–105; Dimitrov, 2007).
In contrast, original Tibetan compositions tolerate extremely long sentences that make use of extensive subordination and generally eschew final particles. While the vocabulary of Buddhist texts composed in Tibetan generally follows established translation equivalents for technical terms (in contrast to the nonreligious literature of legal documents, contracts, etc., which tends to innovate or borrow terms from neighboring languages such as Chinese, Mongolian, or Manchu), a hallmark of this style is an extensive use of clipping (a common pattern abbreviates phrases with an AB-CD structure as AC, such that *thal rang* stands for *thal ’gyur ba* and *rang rgyud pa*, i.e.*Prāsaṅgika* and *Śvātantrika*), as well as the coining of new terminology for concepts not present in the inherited Indic tradition.

Comparable with the syntactic divergence witnessed in translations versus indigenous compositions, there is a large difference between the syntax of prose versus that of verse or poetry. Due to strict restrictions on the syllable length of lines, extensive compounding in the latter often leads to the omission of relational morphology. An example is *khrag ’thung dam can rgya mtsho* (“blood-drinking-guardian-ocean”; de Jong, 1959, 119) where in prose one might have expected *khrag ’thung ba’i dam can gyi rgya mtsho* (“ocean of guardians who drink blood”). This phenomenon has received scant attention. Another understudied area of Tibetan poetry is metrics (Poucha, 1950, 1954; Vekerdi, 1952). Tibetan does not allow reproduction of the often complex meters of Sanskrit verse, and in place of Sanskrit heavy and light syllables, Tibetan verse utilizes stressed and unstressed syllables. Translators, therefore, although not in an always unified or systematic manner, attempted to reproduce variations in Sanskrit meters by varying the length of the lines of Tibetan verse. An aspect of Tibetan prosody that should not be overlooked is the increasing poetization of Tibetan prose that begins in the second half of the 13th century, under the influence of the Tibetan translation of Daṇḍin’s *Kāvyādarśa* (Mirror of Poetics; see Tucci, 1949, 104; van der Kuijp, 1992), a process in which increased awareness of Indian norms of poetics more or less directly shaped developing Tibetan styles of composition.

**Language Change**

Like all languages, Tibetan continuously changes. As an example, although modern pronunciation of Tibetan is related to orthography in complicated ways, reading out Tibetan in transliteration, with its plethora of consonants, giving value to each element, well approximates the language’s sound during the early imperial period; dialects such as Balti, spoken in areas remote from the political center, provide evidence of such pronunciations by maintaining complex onsets (e.g. Balti *bzhi*, “four”; see Sprigg, 2002, 38). Nonetheless, in a few places, orthography misleads – that is, the entire initial cluster agrees in voicing, and finals are voiceless (Hill, 2001a). Thus, *bsgrub* was pronounced [bsgrup], and *bskal* was pronounced [pskal]. In contrast, a reasonable approximation to the pronunciation of these syllables in the speech of Lhasa today would be [trup L] and [kää H] (where L and H represent “low” and “high” tones). The system of personal pronouns offers a second example of language change (Hill, 2007; 2010d; 2013a). In Early Old Tibetan of the 8th and 9th centuries, the respectful use of the plural *khyed* for a singular (*pluralis majestatis*) is absent, but by the advent of Classical Tibetan one or two centuries later, this usage was in place (Hill, 2013a).

**Grammars and Dictionaries**


The single most impressive work of Tibetan lexicography is the ongoing *Wörterbuch der tibetischen Schriftsprache* published by the Bayerische Akademie der Wissenschaften (Francke et al., 2005ff.). Each entry gives copious citations of original sources precisely cited to page and line number. The use of previous dictionaries is carefully distinguished from the evidence of textual attestations. In addition, thorough reference to previous scholarship is given when relevant. Especially so long as this work remains incomplete, several other resources are also valuable. H.A. Jäschke (1881) wrote a lexicographical work of the highest standard that is still important today. Lokesh Chandra compiled...
a 12-volume Tibetan-Sanskrit dictionary on the basis of canonical Buddhists texts available in both languages (1958–1961), continued with seven supplementary volumes (1992–1994) and a one-volume Sanskrit-Tibetan index (2007). J.S. Negi (1993–2004) compiled another 16-volume Tibetan-Sanskrit dictionary. It includes extensive quotations in addition to citations and makes reference to a larger number of texts than does Lokesh Chandra’s. In addition, there are bilingual indices available for a number of Tibetan translations of Sanskrit Buddhist texts, including the Abhidharmakosabhāṣya (Hirakawa, 1973–1978), Bodhicaryāvatāra (Weller, 1952–1955), Kāśyapaparivarta (Weller, 1933), Mahāyāna-sūtrālankāra (Nagao, 1958–1961), Meghadūta (Champa et al., 2011), Nyāyabindu (Obermiller, 1927–1928), Prasannapadā Madhyamakavṛtti (Yamaguchi, 1974), Yagacārabhūmi (Yokoyama, 1996), Lokākāvatārasūtra (Suzuki, 2000), Sukhāvatīvyūha-sūtra (Inagaki, 1984), and Saddharmapuṇḍarikasūtra (Ejima, 1974), Prasannapadā Madhyamakavṛtti (Nagao, 1958–1961), Mahāyāna-sūtrālankāra (Suzuki, 2000), Sukhāvatīvyūha-sūtra (Inagaki, 1984), and Saddharmapuṇḍarikasūtra (Ejima et al., 1985–1993).Ngag dbang tshul khrims (1997) provides a dictionary of difficult or archaic words, drawing heavily on Tibetan commentarial literature. This work provides attestations and cites the works in which they are found, but it does not specify page and line numbers and has an inadequate bibliography; consequently, these citations are not easily verified. The most widely used Tibetan dictionary is Zhang Yisun’s (1985).

Bibliography

Chomphel, G., Dhammapada/chos kyi tshigs su bcad pa bzhugs, Gangtok, 1946.
Giraudieu, F., Dictionnaire Latino-Tibetanum ad usum alumnum Missionis Thibeti, Hong Kong, 1996.
Hill, N.W., “H’dug as a Testimonial Marker in Classical and Old Tibetan,” HiLi 12/1, 2013b, 1–16.
Hill, N.W., A Lexicon of Tibetan Verb Stems as Reported by the Grammatical Tradition, Munich, 2007.
924 LANGUAGES: TIBETAN


Weller, F., Index to the Tibetan Translation of the Kācyapaparivarta, Cambridge MA, 1933.

Yamaguchi, S., Index to the Prasannapadā Madhyamaka-vṛtti, Kyoto, 1974.

Yokoyama Köitsu, Index to the Yogācārabhūmi: Chinese-Sanskrit-Tibetan, Tokyo, 1996.


Zhang Yisun, Bod rgya tshig mdzod chen mo, Beijing, 1985.

NATHAN W. HILL