



KHOTAN IV. THE KHOTANESE LANGUAGE

Khotan was located along the route that is termed the southern Silk Road, which followed the southern rim of the Tarim Basin (see [TAKLAMAKAN](#)). The language of the country, Khotanese, was a Middle Iranian language of the Eastern Middle Iranian type and was spoken before the advent of Islam in this area circa 1000 CE. Three principal stages of Khotanese are known, which may conveniently be termed Old (OKhot.), Middle (MKhot.), and Late Khotanese (LKhot.).

The indigenous name for the people was *hvatana* [hwadana], attested in the 4th century at the city of Niya, to the east of Khotan, as Khotana (in *kharoṣṭhī* script). The land was *hvatana-kšira* [hwadanə-kšira] (MKhot. *hvataṃ-kšira* [hwadā-kšira], LKhot. *hvaṃ-kšira* [hwā-kšira]), and the language *hvatana* ‘in Khotanese’. The term *hvatana* could be from *hvata* ‘self’ and be a self-reference to the Khotanese as the “(rulers) themselves” (pointed out by Konow, 1935, p. 30 [799]). In Indic, the land is called Gostana (Emmerick, 1968b, pp. 88-90), literally “cow = earth breast” (Tibetan *sa-nu*), a name that was also applied to Kuṇāla, son of [Aśoka](#) (q.v.) and legendary founder of Khotan (see Skjærvø, 1998, with references). The Tibetan name for Khotan was Li-yul ‘the land of Li’, with unexplained “Li.” The older Chinese form was 于阗 *yutian* from older **Hwa(h)den*, and the modern form is 和田 *hetian* (for Persian *ḵotan*).

The three principal stages of Khotanese, Old, Middle, and Late, may be



assigned approximately to the 5th-6th, 7th-8th, and 9th-10th centuries, respectively. Old and Middle Khotanese are represented by manuscripts found in the area of Khotan proper and eastward as far as Endere, east of Niya, while Late Khotanese is the language of the manuscripts found at [Dunhuang](#) (q.v.). Traditionally, in Khotanese studies, only two stages of the language have been distinguished, Old and Late (= Middle and Late) Khotanese, but the main linguistic changes took place between Middle and Late Khotanese.

Grammatical descriptions of the language are found in Leumann, 1912; Konow, 1916, 1932, 1941a, 1949; Dresden, 1955; Bailey, 1958; Emmerick, 1968a, 1989 in *CLI*, pp. 204-29 (q.v. for further details), and 2009. Numerous publications of individual texts contain glossaries. H. W. Bailey's *Prolexis to the Book of Zambasta* (1967) contains useful discussions of select Iranian and non-Iranian words in that text, while his *Dictionary* (1979) contains only words of Iranian descent, leaving out the entire Indic vocabulary. Emmerick and Skjærvø (1982-97) contains studies by several authors of individual words and grammatical forms.

DECIPHERMENT AND NAME OF THE LANGUAGE

Khotanese documents first arrived in the West in the late 19th century, when A. F. Rudolf Hoernle (q.v.) received a number of manuscripts in a “cursive” Indian script, mostly containing legal documents, but also some Buddhist texts in formal script, from British agents in [Kashgar](#) (q.v.) and from [M. Aurel Stein](#) (q.v.), who first traveled to Khotan in 1900-1901 and to Khotan and Dunhuang in 1906-8. Some of these, mostly official documents, were in what Hoernle called “cursive” Indian script; others were Buddhist texts in formal script (see Hoernle, 1897, 1899a, 1899b-1901, 1906). Already in his 1901 article, Hoernle proposed that the language of the documents was an Indo-Iranian dialect exhibiting features connecting it with the Pamir languages (pp. 32-33). In the same article, he identified the dating formula and several personal names. He also correctly determined several numerals. The Buddhist texts in formal script, which Hoernle (followed by Stein, 1907, p. 150) hypothesized was “proto-Tibetan,” but which later turned out to be Old Khotanese, were not deciphered till later. The first to realize that the documents represented two stages of the same language, rather than two different languages, was Ernst Leumann, who, in his report on the problems of decipherment of Tokharian and Khotanese, rejected Hoernle's “proto-Tibetan” hypothesis (1907a, col. 707 = 1907b, p. 657).



Leumann's *Zur nordarischen Sprache und Literatur* (1912) contained several important studies of the orthography and phonology of Old and Late Khotanese, and the identity of the language of the documents from Khotan was established by [Sten Konow](#) (q.v.) at about the same time (Konow, 1914, p. 343). In this article, Konow identified the terms *hvaṃna* and *viśa'* found in many documents with Chinese *Huan-na* 匈奴 'Khotan' and 于阗 *Yuchi* (also, incorrectly, *Weichi*; Tibetan *Bijaya*), the name of the dynasty of Khotan under the Tang.

There was less agreement about what to call the language. Konow at one stage followed Alexander von Staël-Holstein, who suggested that the Kushanas spoke "Old Khotanī" (1908, p. 1370) and thought the term "Tokharian" referred to the language of Khotan (e.g., Konow, 1912, pp. 564-65; 1914, p. 13). Albert von Le Coq was apparently the first to suggest, on the basis of the geography and history of the area, that it was "the lost language of the Saka" (1909, p. 318); and Heinrich Lüders argued that the language of the "Śaka kṣatrapas" showed several close similarities with the language of Khotan, among them the use of the ligature <ys> to spell the voiced z, for which there is no sign in the Brahmi alphabet, and so proposed the name Śaka for the language (1913). It should be kept in mind, however, that little was known about the languages of the area at that time and that features that were as yet known only from Khotanese were later found also in other Iranian languages. Also, none of the "Śaka kṣatrapas" have obviously Khotanese names. (*Ysamotika* cannot be derived "without difficulties" from *zam-* 'earth' and *zamawat-* [=?], as Lüders thought [1913, p. 413]; an apparently related name *Zamōdo* is found in the Bactrian document A [see Sims-Williams, 2000, p. 32]).

Meanwhile, in several publications (e.g., 1908, 1912), Leumann maintained that the language was "North Aryan," a separate branch of Indo-Aryan, but Konow (1912), in his review of Leumann, 1912, conclusively disproved Leumann's theory. Similarly, Johann Kirste (q.v.) argued that the term "arisch" was inappropriate and that, since most of the manuscript remains had been found at Khotan, the language ought to be called Khotanese (1912, p. 395; he did not consider the Dunhuang manuscripts). He was followed in this by Hoernle (letters to Stein of 10 January 1913 and to Miss Lorimer of 4 February 1914 [British Library, India Office Records and Private Papers, MSS Eur D 815]; courtesy of Ursula Sims-Williams).

By 1916, the nature of the language was well established, thanks to the manuscripts of the *Vajracchedikā-prajñā-pāramitā-sūtra* (*Diamond Sutra*) and



Aparimitāyuh-sūtra (q.v.) discovered by Stein, which were edited and translated by Konow in Hoernle, 1916, accompanied by their Sanskrit and Tibetan versions and a linguistic sketch of Khotanese. Here, not wishing to take a stand on the exact linguistic position of the “unknown language,” Konow simply called it Khotanese. Later, he began referring to it as “Saka,” as in his editions of the *Bhadrakalpikā-sūtra* (1929a) and the *Saṅghāṭa-sūtra* (1932), where he also discussed the various proposals in some detail and decided to keep following Lüders in calling the language “Saka.” He argued that it was obviously a Scythian language, and, according to Herodotus (7.64), the Persians called the Scythians Sakas, so the term was appropriate (1932, pp. 2-4). With Ernst Leumann’s publication of the *Book of Zambasta* in 1933, however, it became known that the Khotanese referred to their language as *hvatana* ‘Khotanese’. Konow noted this in his review (1934, p. 6), but, although he pointed out that there must have been “other Saka dialects,” he concluded that the use of “Saka” was more convenient.

By 1935, however, Konow had studied the Tumshuqese documents and thought he found the term *hvaḍana* referring to this language, as well. From this, he drew the tentative conclusion that Hvaḍana was the old name of the original tribe that later split into two groups and that the Hvaḍanas who went to Khotan applied the term to the country after making themselves masters there. The immigration of the original Hvaḍanas, he suggested, might be the end of the southward displacement of the Sakas by the Yuezhi (1935, pp. 30-32 [799-801]). The Tumshuqese word may not be what Konow thought it was, however (it may be an infinitive “to speak”; see Konow, 1935, p. 42, text VII, 6-7). In his grammar from 1941, Konow used “Khotanese Saka,” and, in his 1941b edition of the medical text *Jīvaka-pustaka* and later, he used simply “Khotanese,” reverting to “Khotanese Saka” in his article on the “Oldest Dialect of Khotanese Saka” (1947).

Harold W. Bailey (q.v.) began publishing articles on Khotanese in the 1930s and his edition of all the known Khotanese texts under the title *Khotanese Texts* in 1945. From the second volume (1954) on, however, the title pages have *Indo-Scythian Studies being Khotanese Texts Volume....* This practice was followed by Mark Dresden (q.v.), whose edition of the *Jātakastava* (1955) bears the subtitle *Indo-Scythian (Khotanese) Text*. Bailey used the term “Saka” in his 1958 article for *Handbuch der Orientalistik* (on Khotanese and Tumshuqese) and for the facsimile publications for the [Corpus Inscriptionum Iranicarum](#) [q.v.] (1960-67), which contained both Khotanese and Tumshuqese texts. His



student, R. E. Emmerick (q.v.), entitled his grammar *Saka Grammatical Studies* (1968a; which did not cover Tumshuqese) but otherwise used the term “Khotanese.” The term “Saka” was still used by various authors in the 1970s but then became increasingly rare.

DOCUMENTATION

Manuscript collections. The largest collection of manuscripts from Khotan is in the British Library, London (the older British Museum and India Office Library collections). Smaller collections are found in Paris, St. Petersburg, Stockholm, Munich, Kyoto, Washington, D.C., Harvard and Yale Universities, and various places in China (Turfan, Urumqi, Lüshun, Beijing; see Emmerick, 1992a, pp. 4-5; Duan Qing, 1993; 2006). The largest collections of manuscripts from Dunhuang are in the British Library (the Stein collection) and in the Bibliothèque nationale de France, Paris (the Pelliot collection). There is no information about Khotanese manuscripts left at Dunhuang or brought to Beijing, though a few have surfaced on the antiquities market.

The text corpus. Most of the manuscripts from the Old and Middle Khotanese periods were found in the region of Khotan itself, especially the sites of Khadaliq and [Dandān Öiliq](#) (q.v.) east of the city of Khotan and Mazār-e Tāg north of the city, but some were found as far east as Endere. The Late Khotanese manuscripts, as far as can be ascertained, all come from Dunhuang.

There are religious (Mahayana Buddhist) texts; various kinds of literary texts other than religious ones, such as medical, divinatory, and calendrical texts; economic and legal texts; and private and official letters. On the Buddhist texts, see [BUDDHISM iii. BUDDHIST LITERATURE IN KHOTANESE AND TUMSHUQESE](#) (more recent publications include Canevascini, 1993; Skjærvø, 2004; De Chiara, 2013-14; on secular documents see Kumamoto, 1996a; Emmerick and Vorobeva-Desyatovskaya, 1995; and Skjærvø, 2004, 2016a).

Among the letters (*pīḍaka-*) we find orders (*parau*) from superior to inferior officials; reports (*haṣḍi*), usually from inferior to superior officials or private persons or to religious superiors (*vīñatta- < vijñapti*); legal documents (*pīḍaka-, pādā-*) regarding purchase (*gārya-vāḍa* ‘document of purchase’, *pāra-vastū pīḍakā* ‘promissory note, notice of debt[?]’), land or water leases, adoption agreements, various disputes (*gvāra-*), and other issues (see Skjærvø, 2016a); various kinds of vouchers (*kṣau*, □ and □; see [ČĀV](#)), orders, and receipts for goods; military records or registers of recipients of ropes, grain, flour for



baking (Skjærvø, 2001 [2005]), taxpayers, people eligible for guard and canteen service, age, etc. (see Bailey, 1961; Skjærvø, 2002, *passim*; and Zhang Zhang, 2016).

Among these documents are both originals and drafts or copies. The Middle Khotanese documents are frequently original letters that also contain copies of answers to the same or other letters. Those from Dunhuang (cf. Wen Xin, 2017) are most often drafts or letter templates, occasionally original letters, among them the ornately written letter from the king of Khotan (P 5538a; Bailey, 1968, pls. XXX-XXXIII). Syllabaries are common in manuscripts from all periods.

There are three Chinese-Khotanese lists of phrases and words (Takata; Skjærvø, 2002, pp. 35-36, 44-45, 515), one Sanskrit-Khotanese phrase list (Kumamoto, 1988), and a Turkish-Khotanese word list (Emmerick and Rónatas, 1992).

CHRONOLOGY

Evidence for the earliest Iranian-speaking population in Khotan is provided by the 3rd-century Kharoṣṭhī documents discovered at Niya and neighboring sites, which contain Iranian loanwords (Burrow, 1935; Bailey, 1949, pp. 121-28). In particular, on his second expedition in 1906-8, Aurel Stein found a wooden sales document at the site of Endere dated in the regnal year of *Khotana maharaya rayatiraya hinajha Vijida Siṃha* ‘General Vijida Siṃha, great king, king over kings, of Khotan’, with *hinajha* = Khotanese *hīnāysa-* (pronounced *hīnāza*) ‘army leader, general’ (Stein, 1921, pl. XXXVIII; transcription in Boyer, Rapson, and Senart, p. 249; translation in Burrow, 1940, p. 137; see also Emmerick, 1992a, p. 2 and n. 7 with further references). This is also the earliest attestation of the title of the royal house of Khotan, *Vijida* (from Indian *avijita* ‘unconquered’), known from Tibetan as *Bij’aya*, Khotanese *viśya, viśa*’ (pronounced *vižya, viža*).

Old Khotanese. The Old Khotanese literary texts are all undated, and one must rely mainly on paleography and grammar for a relative chronology. Archeological evidence is also of some assistance. Thus, the site of Farhad Beg Yailaki was abandoned before the sixth century, according to Stein (1921, pp. 1254-56), and the manuscripts found there must be from the 5th century or earlier (see Skjærvø, 2002, pp. 561-62). The earliest manuscripts have been dated to the 5th and 6th centuries also on paleographic grounds (Sander, 1986, p. 167; 2005, pp. 134-35, 140; Canevascini, 1993, pp. xiii-xiv). This, in turn,



agrees with the 5th-century date one may assume for the earliest (undated) translations of the *Saṅghāṭa-sūtra* and the *Suvarṇabhāsottama-sūtra*, based on comparison with the Chinese translations, which are dated (see Canevascini, 1993, p. xii; Skjærvø, 2004, I, p. li).

An early Khotanese text is that woven into a 5th/6th century carpet found at Sampula near Khotan in 2007/8, which says *spāvatā meri sūmā hōḍā* ‘soma was given to general Meri’. Here *spāvatā* (< **spāda-pati*) is the preform of the common *spātā* (Duan Qing, 2020, p. 35).

Middle Khotanese. We have no dated documents from the Old Khotanese period, but several from the Middle Khotanese period, from most of the 8th century, the latest (Hedin 20) being from the 36th regnal year, presumably of Viśa’ Vāhaṃ, that is, probably, 802 CE (Zhang and Rong, 1997). The latest dated document from Dunhuang (Ch.00272) is from the 14th regnal year of the period Tianxing, 999 CE, shortly before the conquest of Khotan by the Muslim [Qarakhanids](#) (q.v.) circa 1000 (Gronke, 1986; Hamilton, 1979, p. 51; the period Tianshou is now thought to be earlier, see Zhang and Rong, 1999).

In this period (7th-10th centuries), Khotan was dominated by the Chinese and the Tibetans, as is evident also from secular documents. In 658, Khotan was officially placed under the protectorate of Anxi at Kucha (安西都护府 ‘Protectorate to Pacify the West’) and became one of the Chinese Four Garrisons (四镇; Chavannes, 1903, pp. 113, n. 2, 118). In 665, the Tibetans attacked Khotan and, in 670, took the Four Garrisons and held them until 692 (see Beckwith, pp. 34-36, 54). There is a reference to the turbulent history of Khotan in the *Book of Zambasta* (15.9), where the author complains about the *heinā khoca u huna ciṃṅga supīya* ‘Red-faces (= Tibetans), Huns (Xiongnu?), Chinese, and Supīyas’, who have destroyed the land of Khotan (on the Supīyas, see Bailey, 1985, pp. 79-81). This reference would seem to postdate the first Tibetan occupation (although the main text of the *Book of Zambasta* is probably earlier). The Tibetans were then driven out, but less than a century later they gradually reconquered the Four Garrisons, Khotan last in 791 or 792 (Chavannes in Stein, 1907, pp. 533-36; Beckwith, pp. 150, n. 35, 155). The end of this second Tibetan occupation is uncertain (Hamilton, 1979, pp. 49-50, is based on Pulleyblank’s reading of the date of the document Hedin 24 [in Bailey, 1961, p. 136], which is [X]□□□, as year “54,” but Zhang and Rong [1997, p. 342] later read the date as year □□□□ ‘Zhengyuan 14’ = 798).

The Middle Khotanese secular documents (legal, economic, private letters), as



well as Chinese documents from the same sites (some documents have the same text in both Khotanese and Chinese) are frequently dated, and the literary texts occasionally give the names of donors known from the non-literary texts (Skjærvø, 1991, p. 270).

Already in his first publications, Hoernle noticed that two Chinese documents from Dandān Öiliq were dated in 768 and 786 (Hoernle, 1901, p. 31). Later, Stein noted that all the Chinese papers from ruin vii at Dandān Öiliq were from 782-89 (1907, p. 277) and that the Chinese coins from the sites of Dandān Öiliq and Rawak (q.v.) were from reigns between 713 and 760 (ibid., p. 283). Later, Sten Konow suggested that the king Väśa' Vāhaṃ of some documents was the same as 吐谷浑 Yuchi Yao in the Chinese sources, who, it was thought, ruled from 756 (Konow, 1914, pp. 349-50; 1929b, esp. pp. 73-76). According to the Chinese sources, when his father, 吐谷浑 Yuchi Sheng, left Khotan (756) to help put down [An Lu-Shan's](#) (q.v.) rebellion, which began in 755, he left his son Yuchi Yao as vice-commissioner (*jiedufushi* 节度副使; in Khotanese he may have been called the *yauvarāya* 'young king'; see Skjærvø, 1991, p. 265), but he never returned, and his son then officially became king in 767 under the name Viśa' Vāhaṃ (see Zhang and Rong, 1987, p. 90; 1997, pp. 346-47; Kumamoto, 1996a, pp. 38-40; Zhang, 2017, pp. 149-50).

Documents that specifically mention Viśa' Vāhaṃ are dated in the years 17 and 20 of his reign only, but Zhang and Rong have argued that a number of documents dated in the years 32-36 (see Skjærvø, 1991, pp. 266-67) also belong to this king's reign. If this is correct, he must have ruled at least until 803, which makes the place in the succession of Viśa' Kīrta, who is known from two texts, one a metrical panegyric, mentioning the "masters" (= Tibetans), unclear (Skjærvø, 1991, p. 266). At present, it seems likely that most of the Middle Khotanese documents are from the reign of Viśa' Vāhaṃ. Two other kings, Viśya Sīhya and Viśa' Dharma, who succeeded one another according to four legal documents written on wooden tablets, may be identical with the kings 吐谷浑 Yuchi Gui, who may have ruled circa 737-46, and Yuchi Sheng (see above). Yet another king, Viśya Vikraṃ, may have ruled 692-706+ (Skjærvø, 1991, p. 260; see also Inokuchi; Skjærvø, 1991, pp. 262-65).

Most of the manuscripts from Dandān Öiliq and adjacent sites are from the 8th century or earlier (Stein, 1907, pp. 266, 283-84, 521; 1921, p. 208; 1933, pp. 68-69), which led Stein to conclude that the site of Dandān Öiliq was abandoned by the end of the 8th century. It is now clear, however, that some of the documents from Khotan are probably from the early 9th century (Hedin



20, see above), and the question of dating is being reinvestigated. One document (British Library, MS Or. 11344/17) suggests that the reason for the lack of further documentation may have been the attack of the Uygurs in 802 (Skjærvø, 2002, p. 115, and 2016a).

Late Khotanese. Many Late Khotanese manuscripts contain dates in the 10th century. The manuscript of the *Vajracchedikā-prajñā-pāramitā-sūtra* (the *Diamond Sutra*) contains a date corresponding to 14 April 941; the Khotanese colophons in Ch. c.001 (a long scroll commissioned by Śāṃ Khīṅā Hvām' containing a miscellany of texts) are all dated in a Hare Year, probably the year 943 (Emmerick, 1992a, p. 22). The *Jātaka-stava* (Stories in praise of the [Buddha's] births; Dresden, 1955, p. 446) and one of the Vajrayāna texts, which contain a date that may correspond to 10 August 971, were written by a certain Cā Kīmā Śānā during the reign of Viśa' Śūra (Bailey, 1961, p. 151; Hamilton, 1979, p. 51).

Hoernle discussed the dates of some of Stein's Dunhuang manuscripts in his 1911 article (pp. 469-71; see also Stein, 1921, pp. 1448-55; Kumamoto, 1982). Later works include Wang Binghua (in the first part of Wang and Duan, 1997) and Kumamoto (1986; see also Kumamoto, 1996a). E. Pulleyblank placed the reigns of Viśa' Saṃbhava (Saṃbhata; Chinese 李晟天 Li Shengtian), Viśa' Śūra, and Viśa' Darma in the 10th century. On the discussions of the dates of other kings and their regnal periods (*Tianxing* 天行 era, Khot. *Thyenā-hīṅā* 950-; 天壽 *Tianshou*, Khot. *Thyaina-sīva* 963- [?]), see Hamilton, 1979, 1984; Kumamoto, 1986, 1996a, 1996b; Wang in Wang and Duan, 1997; Zhang and Rong, 1999.

By these chronological data, it is possible to follow the evolution of the Khotanese language for at least half a millennium.

THE WRITING SYSTEM

Hoernle, who was the first to discuss the scripts of the manuscripts from Chinese Turkestan (see BRĀHMĪ), classified the two types of scripts in the Khotanese manuscripts as upright and cursive Gupta, although he recognized that the term "cursive" was inappropriate, as the letters were not connected (Hoernle, 1911, p. 450). He was aided in the decipherment by syllabaries found in some of the Dunhuang manuscripts, which contained all the basic letters as well as several ligatures and the numerals (ibid., pp. 450-60, with four plates of syllabaries).



The upright, or formal, ductus is found in numerous variants in all periods of the language used for Buddhist literary texts. The cursive ductus also exhibits numerous variations. In both the Middle and Late Khotanese periods, we may distinguish between a formal and a less formal style of the cursive ductus (see Skjærvø, 2002, pp. lxxi-lxxii and pls. 4-8).

The Brāhmī script used for Khotanese contained several non-Indic ligatures used to express special Khotanese sounds, most notably the two series of affricates, alveo-palatal and dental, and the voiced and unvoiced sibilants (see Leumann, 1934). Two vowel signs were invented, one for a central vowel [ə] (or [ɪ]) transcribed as *ä* and one for the corresponding diphthong transcribed as *ei* (for *a* + *ä*, see Emmerick 1998). The aspirated voiced stops of the Brāhmī script (*bh*, *dh*, *gh*) were redundant and were used mostly in Indic words, occasionally in Khotanese ones (e.g., *dhāta*- beside *dāta*- ‘Law, *dharma*’).

Khotanese had several phonemes for which the script had no individual letters; to express these, ligatures and other devices were used, e.g., <tc> for [ts], <ts> for [tsh], <js> for [dz], <ys> for [z]. For the rolled *r*, a special letter was invented, transcribed *rr* (though it is not a ligature of *r* + *r*).

A diacritical mark also was invented, a bowl-like curve placed below the letter (commonly transcribed as an apostrophe) to express what was probably rhotacization (retroflexion) of the vowel. This sign has been interpreted as derived from a letter only found in Tumshuquese, where it represents the sound *ž* (or perhaps *ʎ*?) derived from Old Iranian intervocalic *š* (e.g., *pyežu* ‘listen!’ Khot. *pyū*’; Skjærvø, 1987, pp. 84-85, 90) or derived from <h> and representing a kind of breathing (see on phonology, below).

Of the voiced sibilants, only *z* had a special graph [ys]; the other two were spelled with the corresponding unvoiced *ś*, *ṣ*. In Old Khotanese, the voiced and unvoiced sibilants could be distinguished by writing the voiced single and the unvoiced double, whereas, in Middle and Late Khotanese, the unvoiced was usually written single, while the voiced was marked with the subscript curve (e.g., OKhot., MKhot. LKhot. *ṣavā*-, OKhot. *ṣṣavā*- ‘night’; OKhot. *ṣä* [zə] ‘that’, MKhot., LKhot. *ṣi*’; OKhot. *śäta*- /*žäda*-/ ‘second’, MKhot., LKhot. *śe*’ /*že*/).

In Late Khotanese, a colon <:> is used in Chinese and Turkish after syllables with *h* or *hv* to represent *x*, *ɣ*, and *f*, respectively, e.g., *uhū:ysä* = Turkish *uyuz*, *hvū: śai’ne* ‘wife, (royal) consort’ = Chinese 嫔 *furen* (cf. Emmerick and Pulleyblank, p. 32).



The numerals are the regular Brāhmī ones, with the addition of a sign for ½, probably derived from the letter <hā> for *hālai* ‘half’.

GENERAL CHARACTERISTICS

The use of the Brāhmī script and the numerous Indic loanwords lend the language a strong Indic look, which led Leumann to deny its Iranian appurtenance. The Indic loanwords are from two sources in particular: one, the local Northwestern Middle Indic (Prakrit) language found in the Niya documents and the numerous religious texts found in Afghanistan, which today is commonly termed *Gāndhārī* (q.v.); the other, the Sanskrit of the Buddhist texts copied and read in the area, which may reflect both standard Sanskrit and Sanskrit through the lens of *Gāndhārī*. Often, Old Khotanese has the *Gāndhārī* form and Late Khotanese the Sanskrit form. Examples: OKhot. *Gr̥jakūla-*, LKhot. *Gr̥ddhakūṭa-* (Skt. *Gr̥dhrakūṭa*, a mountain), OKhot. *Nairam̐ña*, LKhot. *Nīram̐jana* (Skt. *Nairañjana*, a river), and OKhot. *lovapāla-*, LKhot. *lokapāla-* ‘world protector’ (= Skt.). See also Degener, 1989b.

The phonology is at first glance of the Indic rather than Iranian type. There is a series of aspirated unvoiced stops (*kh*, etc.) and a series of retroflexes (*ṭ*, etc.). The second feature it shares with several East-Iranian languages, however, and the first represents a tendency to eliminate the Old Iranian unvoiced spirants seen also elsewhere in East-Iranian. To what extent the script might obscure a more Iranian-type phonology, especially in the early stages of the language, is still being investigated. In the later stages, the phonology appears to be close to what the script suggests, which may be a secondary development. On the other hand, Khotanese participates in a series of typical East-Iranian phonological developments (*āy* > *ay*, *č* > *ć*, *ṣw* > *ṣuw*, etc.).

The morphology and syntax are typically East Iranian, with the exception of the preterite of transitive verbs, which is formed by means of a possessive active participle (“I am having-done”), matching the use of “to have” in Sogdian and elsewhere (“I have done”; see below). This construction too is known from Sanskrit and may be an areal phenomenon in Khotanese.

PHONOLOGY

Vowels. All the vowels expressed in the Brāhmī script were phonemes in Old Khotanese (*a*, *ā*, *i*, *ī*, *u*, *ū*, *e*, *o*, *ai*, *au*), with *e* and *o* expressing long [ē, ō]. In addition, the meter suggests that there were short variants [ĕ, ǒ] (Emmerick



and Maggi).

In the oldest stage of the language (and the oldest manuscripts), there were two new vowels, a central vowel [ə] (or [ɪ]) spelled <ä>, which contrasted with [i] (e.g., nom. *balysä* [balzə] ‘Buddha’, gen. *balysi* [balzi]) and a diphthong [aɐ] (or [aɪ]) spelled <ei> and contrasting with <ai> (e.g., nom. *pīsei* ‘teacher’, gen. *pīsai*, see Emmerick 1970, pp. xix-xx; 1998).

There may have been nasal allophones before nasals, as suggested by spellings with otherwise unnecessary *anusvāra* <’> (a simple point above the letter), as in *balysā’nu* ‘of the buddhas’ instead of *balysānu*. It is common to transcribe this “unetymological *anusvāra*” as an ogonek—a subscript “hook” (*balysāṅnu*).

Finally, vowels may have been rhotacized (retroflexed), indicated by a subscript curve in the Brāhmī script (see on the script, above). The rhotacization was caused mainly by a neighboring voiced retroflex *ṛ* and may have become phonemic when the *ṛ* was lost. The rhotacized vowel then caused a following or preceding *n* to become retroflex *ṇ* (see historical phonology, below; cf. Emmerick, 1989, p. 214). In Emmerick and Pulleyblank (pp. 54-55, which see for full discussion), it is suggested that the feature expressed by the subscript curve (which Emmerick, 1992b, pp. 158-69, suggested was derived from the letter *h*) might be “breathiness,” citing, in particular, the verb *ā’* ‘sit’ < **āh*-; but the participle *āṇa-*, which they also cite, with its *ṇ*, points to rhotacization. Since Old Iranian intervocalic *h* usually disappeared without a trace (e.g., *urmaysde* < **ahuramazdāh*), the forms with *ā’* are at any rate hard to explain.

Consonants. All the consonant signs of the Brāhmī script are used (except <jh>, which only occurs in alphabets in Khotanese Brāhmī script), but the aspirated voiced stops <gh, dh, bh> do not correspond to any separate phonemes (see on the script, above). The following description does not necessarily apply to Indic loanwords.

There were four series of stops—velar, retroflex, dental, and labial—and two series of affricates, alveo-palatal and dental, all in three variants: unvoiced, unvoiced aspirated, and voiced: /k, kh, g; ṅ, ṅh, ḍ; t, th, d; p, ph, b; ṭṣ (č), ṭṣh (čh), ḍḍ (j); ts, tsh, dz/. The values of these graphs in the 10th century have been established on the basis of a Chinese version of the *Diamond Sutra* from Dunhuang written in Brāhmī script (Emmerick and Pulleyblank, especially pp. 29-47).



Stops are often written double. In initial position *tt-* is usually written instead of *t-*, and *gg* is used in initial position and often after nasal. Double *ṭṭh* is found in a few words.

Of the retroflex stops, *ṭ* is only found in the group *ṣṭ*, while *ṭh* and *ḍ* are phonemes.

The Khotanese aspirated stops correspond to the Old Iranian fricatives (e.g. *khara-* ‘donkey’, Av. *xara-*; *phārra-* ‘fortune’, OIr. *farnah-*). Evidence from the 10th-century manuscripts renders it likely that these phonemes were actually stops, and this may have been the old state of affairs, as suggested by the aspirated affricates that developed through palatalization (e.g., *tso* [tsho] ‘go!’ < *čyawa; *gvach-* [gwatšh-] ‘be digested’ < *wi-pačya- (Emmerick and Pulleyblank, pp. 32-34).

The two-way opposition in the affricates (unless one regards <kṣ> as <tṣh> with Emmerick and Pulleyblank), probably alveo-palatal vs. dental, [tš] vs. [ts], rather than retroflex vs. alveo-palatal, [tʃ] vs. [tʃ̌], contrasts with the three-way opposition in the sibilants, presumably *s* [s] ~ *ś* [š] ~ *ṣ* [ʃ], *z* ~ *ź* [ž] ~ *ṛ* [ʒ].

The nasals were all allophones of /n/ before the corresponding consonants (*ṅg*, *ṅj*, *mb*, etc.). In addition, the palatal nasal <ñ> was either a single phoneme or a variant of the group /ny/ (cf. *sāña-* ‘means’ < Indic *sājña-*, *ñāp-* = *nyāp-* ‘become known’ < Indic *jñāpaya-*; *nyāttara-* = *ñāttarai* ‘lower’ < *nita-). The retroflex <ṅ> was an allophone of /n/ before or after a rhotacized vowel (see above) and was also used after syllables with *r* according to Sanskrit rules (e.g., *prrahauna* and *prrahaṇa* ‘clothes’).

Of the two *r*’s, one (<r>) was presumably a flap, the other (<rr>) rolled (as in Spanish). The simple *r* is found in both Iranian and Indic words, the *rr* mainly in Iranian words in initial position and intervocalically when representing Old Iranian *-rn-* or *-rśn-* (see below). It is also frequently found in the groups *prr-* and *krr* (from OKhot. on) and *trr*, *ttrr* (especially LKhot.).

HISTORICAL PHONOLOGY

Vowels. Khotanese phonology is characterized by syncope of unstressed initial and interior vowels and palatalization, which affected both vowels (fronting) and consonants, both diachronically and synchronically.

Initial vowels were lost in prefixes (e.g., **apa-* and **upa-* > *pa-*, **abi-* > *bä-*, *ba-*,



and **awa-* > *va-*).

Syncope of internal vowels is seen in words such as *gyasta-*, *jasta-* ‘god’ (cf. Av. *yazata-*, with irregular *-st-*, cf. Tumshuqese *jezda-*, by popular etymology with *jasta-* ‘cleaned, healed?’) and *māsta-* ‘big’ (Av. *masita-*, with fronting of *a* > *ä* before *i*). It is a regular feature of suffixes and endings (e.g., *śśādāti-* ‘goodness’ < *śśāra-* + *-tāti-*, with *r-t-* > *ḍ*; *baḍe* ‘he rides’ < **bara-tai*).

Palatalization of vowels and consonants was caused by a following **i* or **y*, which were frequently lost (e.g., *māsta-* ‘big’). When the phoneme causing the palatalization was lost, the palatalization often remained the only distinguishing feature (e.g., nom. *ūtca* [ūt̪sa] ‘water’, loc. *ūca* [ūt̪ša] < **usačā*, **usačayā*; *hālštā* ‘spear,’ gen.-dat. *hālštā* < **ṛštiš*, **ṛštiyah*). See also morphophonology, below.

The diphthongs **ai* and **au* became *ī* and *ū* (e.g., *hīnā-* ‘army’, cf. Old Persian *hainā-*; *ysarūna-* ‘golden’ < **zara-gauna-*, cf. Persian *zar-gūn*).

Consonants. Khotanese belongs to the Northeast-Iranian dialect group (see iran vi. iranian languages and scripts, in *EIr.* XIII, pp. 344-45, 376) and is characterized by the development of Old Iranian **čw* > *ś* and **jw* > *ž* (e.g., *aśśa-* ‘horse’, Av. *aspa-*, OPers. *asa-*; *biśā* [bižā] ‘tongue’, OPers. *hazān-*, Pers. *zabān*).

Voiced stops had already been dropped between vowels and in some other positions by the time of the earliest manuscripts (e.g., *dai* ‘fire’ < **dāgah*; *pai* ‘foot’ < **pādah*, Pers. *pāy*; *mura-* ‘bird’, Av. *mərəya-*, Pers. *morg*).

Initial *d-* may have become spirantized (as in Sogdian) and thus avoided phonemic merger with the newly voiced intervocalic *t* (e.g., *dāta-* = *ḍāda-*). More likely, perhaps, the result of old intervocalic *t* was sufficiently different from *d* to avoid merger. Apparently, it was a dental flap that was soon lost, leaving a hiatus, even when written (e.g., *tsuta-* [tshuda] ‘gone’, < *čyuta-*, later *tsua-*, *tsva-*).

Whether initial *g-* (*gg-*) and *b-* (which merged with initial *w-*) were spirantized is also moot (e.g., *ggara-*, *gara-* ‘mountain’ [gara or ɣara], *bar-* ‘carry’ [bar- or βar-]). Note that Khotanese *b-* and Indic *v-* (in loanwords) are not confused.

Unvoiced consonants between vowels and in certain groups were in general voiced (e.g., *jsa* [dza] ‘from’ < **hača*; *māta* [māda] ‘mother’; *bārgga-* ‘wolf’, *dā(r)-dda-* ‘third’ < **drida-* < **θrita-*; *tcārba-* [tsārba-] ‘fat’ < **čarpa-*, Pers. *čarb*),



but intervocalic *p* became *v* (e.g., *ttavaa-* ‘fever’ < **tapa-ka*) and intervocalic *k* was lost, notably in the *-ka-* suffix (e.g., *suraa-* ‘clean’ < **suxra-ka-*, cf. Pers. *sork* ‘red’, or rather **subra-ka*, cf. Armenian *sowrb* from Iranian, Vedic *śubhrá* [see Schmitt, p. 446; see “Morphology,” below]). Intervocalic *θ*, however, became *h* (e.g., *rraha-* ‘chariot’, Av. *raθa-*). Initial *xw-* became *hv-* (e.g., *hvatä* ‘by oneself’, Av. *x^vatō*).

The group *-rt-* became retroflex *-ḍ-* (e.g., *muḍa-* ‘dead’, Av. *mərəta-*, Pers. *mord*; *hūḍa-* ‘given’ < **frabrta-*).

As in other East-Iranian languages, the Old Iranian alveo-palatal affricates *č*, *ǰ* became dental affricates *ts*, *dz* (e.g., *tcāta-* [tsāda] ‘a well’, Av. *cāt-*, Bactr. *sado* [tsād]; *jsan-* [dzan] ‘kill’ < **jan-*, Av. *jan-*; *paṃjsa-* [pādza] ‘five’ < *panča*). The voicing of intervocalic *-č-* occurred after the general syncope of unstressed vowels (cf. *pasūste* < **saučatai*, with *-st-* rather than *-zd-*; Sims-Williams, 1983a, p. 359). The group *sč-* also became *tc* [ts] (e.g., *pātco* ‘afterward’ < **pasčām*).

A new alveo-palatal series (*č* [tš] spelled <c, ky> and *ǰ* [dž] <j, gy>) was produced by retention of old *č* and *ǰ* before *i*, *ī*, *y*, which prevented them from becoming *ts*, *dz* (e.g., *cu* [čū] ‘what’ < **čim* = Av. *jin-* ‘destroy’ < **jinā-* = Av.). Already in the Old Khotanese period these merged with palatalized *k* and *g*, which led to alternate spellings <c/ky, j/gy> (e.g., *khārga-* ‘mud’, loc. *khārja*). Old initial *čy-*, however, became *ts* [tsh], while intervocalic *-čy-* became *ch* [tšh] (e.g., *tsāta-* ‘rich’ < **čyāta-*, Av. *šāta-* ‘happy’; *gvach-* ‘be digested’ < **wi-pačya-*).

Before *r* and *t*, the unvoiced fricatives were voiced as in other East-Iranian languages and then developed variously (e.g., *brī* ‘dear’, Av. *friia-*; *drai* ‘three’ < **θrāyah*; *grūs-* ‘call’, Av. *xraosa-*, *gruš-* ‘be called’ < **xrusya-*; *baura-* ‘snow’, Av. *vafra-*; *mara* ‘here’ < **imaθra*; *ttīra-* ‘bitter’ < **taira-* < **tayra-* < **taxra-*; *hauda* ‘seven’, Sogd. *avda* <βt> ‘seven’; *dutar-* [δudar] ‘daughter’, Av. *duxtār-*, Sogd. *δuyda*). Whether initial *br-*, *gr-*, *dr-* were pronounced with fricative β-, δ-, γ- (as in Sogdian) or with stops is moot.

The development of initial **w-* and **y-* was similar to that seen in Modern Persian: **w-* > *b-* (= *b-* or β-?) or *g-* according to the phonetic context, and **y-* to *ǰ-* (*dž*) (e.g., *bāta-* ‘wind’, Av. *vāta-*, Pers. *bād*; *bārgga-* ‘wolf’, Av. *vāhrka-*, Pers. *gorg*; *ggu-*, preverb, Av. *vi-*, Pers. *go-*; *jau* ‘fight’, OIr. *yauda-*; *gyasta-* ‘god’, Av. *yazata-*). Khotanese initial *y-* is found only in Indic words, while *v-* is also found in Khotanese words, where it usually represents older **awa-* (in the



common verb *yan* ‘to do’ < **kr-naw*, *y* is the outcome of old *k* treated as intervocalic).

The oldest Indic loanwords also show these various developments, which provides a chronology (e.g., OInd. *vipāka-* ‘fruition (of acts)’ > *vīvāga-*, later *vīvāta-*, *vīvā*; OInd. *mudrā-* > Khot. *mūrā-* ‘coin’, unit of payment).

Old initial *r-* usually became OKhot. *rr-*, while initial OKhot. *r-* is found primarily in loanwords. Among groups with *r*, note *rn* and *rśn* > *rr* and *rd* > *l* (as in Persian; e.g., *kārra-* ‘deaf’, Av. *karəna-*, Pers. *karr*; *tarraa-* ‘thirsty’ < **tršna-ka-*, Pers. *tešna*; *salī* ‘year’, Av. *sarəda-*, Pers. *sāl*) and **r* > *l* before sibilant (e.g., *pālsuā-* ‘rib’, Av. *parsu-*; *bulysa-* [bulza-] ‘long’, Av. *bərəzant-* ‘high’ [but Pers. *boland* < **bṛdant-* < Indo-Ir. **bṛjant-*]; *kālsta-* ‘sown’, **kṛšta-*, Pers. *kešt*).

Intervocalically, *š* was voiced to *ž*, which merged with intervocalic *g* and *t* (into a flap?) and was lost while causing rhotacization (retroflexion) of adjacent vowels and *n* > *ṇ* (e.g., *gū’* [gūr] ‘ear’, instr.-abl. *gū’ṇa* (< **gūža-*, cf. OPers. *gauša-*); OKhot. *nāšāy-* [nəžā(r)y-] ‘to place’ < **ni-šādaya-*, later *ṇā’y-*, *ṇāy-* [ṇāry-, ṇāy-]); OKhot. *hva’nd-*, MKhot. *hvaṇḍ-* ‘man’ (Tumshuqese *hvazand-*); *pāša-*, *pāga’-*, *pāta’-* [pəžar, pəgar, pədar] ‘strength’ < **pauša-*, later *pa’-* [par]. The development before *m* was similar, e.g., *tcei’man-* ‘eye’ < **čašman-*.

Intervocalic sibilants in loanwords were also voiced (e.g., OInd. *ākāśa-* ‘space’ > *āgāśa-* [āgāža]; OInd. *āsana-* ‘seat’ > *āysana-* [āzana-]; OInd. *āsādaya-* ‘obtain’ > *āysai-* [āzāi-]).

MORPHOLOGY AND MORPHOPHONOLOGY

As in other eastern Middle Iranian languages, the nominal and verbal morphological categories of Khotanese are quite archaic and remained remarkably stable throughout the history of the language. The case system is close to that of Old Persian, and all the moods of the verb are represented. Dual forms survive only in “two” and “both,” and the old past tense forms (imperfect, aorist, perfect) have been lost. The main innovations are (as in Sogdian and Chorasmian, qq.v.) the extended declensions (*ka-* and *kā-*stems) and the transitive past tense.

Palatalization. Endings (and suffixes) containing original *-y-* normally caused palatalization (diachronically and synchronically; see Hitch 1990 on synchronic palatalization).



Palatalization in word formation is seen in nouns in *-ya- (e.g., *kīra-* ‘work’ < **karya-*, *ysīrra-* [zīrra-] ‘gold’, Av. *zaraniia-*, Pers. *zarr*) and verbs in *-aya- (e.g., *ber-* ‘to rain’ < **bāraya-*; *bulj-* [buldž-] ‘to praise’, Av. *bərəjaiia-*, cf. *buljsaā-* [buldzaā-] ‘praise’).

In nouns, synchronic palatalization is seen primarily in the locative singular (endings *-ayā, *iyā); in some ā-stems also in the genitive-dative (= instrumental-ablative; ending *-ayāh); and in i-stems everywhere except the nominative singular (endings *-yam, *-yah, etc.; e.g., a-stems: *dasta-* ‘hand’, loc. sing. *dīšta*, cf. OPers. *dastayā*; *ura-* ‘belly’ [Av. *udara-*], loc. *uīra*; ā-stems: *kanthā-* ‘city’, loc. *kīntha*; *hotā-* ‘power’, gen.-dat. *hvetē*; i-stems: *hūni-* ‘blood’ [Av. *vohuni-*], sing. acc. *hūñu*, gen.-dat. *hūñā*; *mulysdi-* [mulzdi-] ‘compassion’ [Av. *mərəždi-*], gen.-dat. *mulśdā* [mulzdə]; *tcari-* [tsari-] ‘face’, loc. sing. *tcīra*, nom.-acc. pl. *tcīrā*).

In verbs, palatalization occurred in the 2nd and 3rd persons singular indicative (endings *-ahi and *-ati) and 1st, 2nd, 3rd singular optative (endings *-yām, -yāh, -yāt; e.g., *puls-* ‘ask’ [Av. *pərəsa-*], indic. 2nd sing. *pulśā*, 3rd sing. *pulśtā*, opt. sing. 1st *pulśo*, 3rd **pulśā*; *gan-*, *yan-* ‘do’ [cf. OPers. *kunau-*], indic. sing. 2nd *yañā*, 3rd *gīndā*; *haur-* ‘give’ [OPers. *frabara-*], sing. 2nd *herā*, 3rd *heḍā*).

Nouns and adjectives. The most common declensions are the masculine *a*-declension (OIr. *a-* and m. *i*-stems), the feminine *ā-* (OIr. *ā-* and f. *i*-stems) and *i*-declensions (OIr. f. *i-* and *ī-*stems), the masculine and feminine *r-* declensions, and the neuter *n*-declension. Residual declensions include diphthong declensions and masculine *n-* and *nd*-stems, and a few *h*-stems. The secondary stems are primarily *k*-extensions of the old vowel stems (*akā-*, *-ākā-*, *-ikā-*, etc.). The intervocalic *k* was lost before the earliest documents, but is still present as *g* in early loanwords in Middle Indic (Niya documents *jheniḡa*, *jheniya*, Khot. *ysīniya*).

There are several productive noun and adjective formations, from verbs, nouns, or adjectives (see Degener, 1989a). The productive suffix *tāti-* ‘-ness’ (cf. OIr. *-tāt-*) makes nouns from adjectives; in older words, syncope is the rule (e.g., *dīḍāti-* < *dīra-* ‘lowly’), but later the suffix is attached as is to vowel stems (e.g., *śārattāti-* ‘goodness’). The productive suffix *āmatā-* makes action nouns from verbs, and *-auñā-* makes abstract nouns from nouns and adjectives (e.g., *pyūvāmatā-* ‘listening’ [but *tsūmatā-* ‘going’]; *arahaṃdauñā-* ‘arhatship’; *dukhā-ttauñā-* ‘poverty’ < *dukhāta-* [< Indic]; *tsāttauñā-* ‘wealth’ < *tsāta-* ‘rich’).



Adjective formations include the suffixes of relationship and possession *īnaa-* and *ānaa-*, which are frequently used to render Sanskrit compounds (e.g., *śśāndeinei ājāvāṣā ūcīnei ājāvāṣā* ‘the snake of earth [*śśāndaā-*], the snake of water [*ūtca-*], Sanskrit *kṣiti-ūragaś ca salila-ūragaś ca*). Prefixes include the privative *a-*, *hu-* ‘good’, *duṣ-* ‘bad’. Compounds are also relatively common (see Degener, 1987), e.g., *uspurra-vīraa-* ‘whose works (*kīra-*) have been completed’; *yāḍa-śśāḍaa-* ‘who has done good deeds (*śśāḍaa-*)’; *mārā-pātara-* ‘parents’. Older compounds include *uysnaura-* ‘living being’ (< **uzana-* ‘breath’ + *bara-* ‘carrying’), *śśūjāta-* ‘one another’ (cf. *śśau* ‘one’, *śśāta-/zāda-* ‘other’).

Comparative and superlative are formed with the productive suffixes *-(ā)tara-* and *(ā)tama-*. There are a few original forms with syncope (e.g., *māštara-* < **masyah-* + *-tara-* beside *māštara-* ‘bigger’) beside secondary ones (e.g., *tcārbātara-* ‘fattier’, *āṣaṇa-pajsama-jserātara-* ‘who is to be made [*tcera-*] worthier of worship’); *śśāra-* ‘good’ has the suppletive forms *hastara-*, *hastama-* ‘better, best’. The superlative *bryāṇdama* ‘dearest’ mirrors an Indo-Iranian phrase type consisting of a genitive plural and a superlative (cf. Vedic *devānām devātama-* ‘the most god-like of gods’, Av. *daēuuanqm daēuuō.tāma-* ‘the most demon-like of demons’, Bactrian *baganodamo* ‘most god-like, most divine’; see Sims-Williams, 2018).

Personal and possessive pronouns. The nominative forms are 1st sing. *aysu* [azu] < **azam*, 2nd sing. *thu*, apparently from **tuhu* (cf. Sogd. *tuyú*, both perhaps from **tuwu*; see Sims-Williams, 1983b, p. 48); the 1st pl. *buhu*, *muhu* probably has its vowels from 2nd pl. *uhu* (< **yūžam?*), which may have its *h* from *muhu*. For other case forms, see below. Possessive pronouns include sing. 1st *mamānaa-*, 2nd *tvānaa-*, pl. 1st *mānia-*, *mājaa-*, 2nd *umānia-*, *umājaa-*.

Demonstrative pronouns. These include the “unmarked” *ṣ-/tt-* ‘that’ ([*z-/t-*] < **aiša-/aita-*), the near-deictic (reduplicated) *ṣ-ṣ-/tt-t-* ‘this’, and the far-deictic *ṣāra-/ttāra-* ‘yonder’ (LKhot. *ṣū`rā*, gen.-dat. *ttūrye*, etc.). The reduplicated pronoun has forms such as sing. nom. m. *ṣātā*, *ṣei* [zədə, zəə], f. *ṣā*, acc. m. *ttutu* (later *ttū*), f. *ttuto* (later *ttuo*, *tvā*). In Late Khotanese, we also find forms expanded by *-ka* (*ṣai`kā* ‘that’, acc. sing. *ttūkā*, etc.).

Relative, interrogative, and indefinite pronouns. These are formed from the stems *k-* (oblique *kam-*) ‘who’ and *c-* (oblique *tcam-*) ‘which’. “Which?” is *kāma-*. Indefiniteness is expressed variously as *kye ṣā kye* ‘whoever’, *kye/cu ju halcä* ‘who/whatever’, *cerä duru gāvu* ‘however many’, etc. Note also *ye* ‘one’ (cf. Germ. *man*).



Demonstrative-relative pairs include *cerä ... tterä* ‘as many ... so many’, *candu ... ttandu* ‘as much ... so much’, *cītä (cīyā, cī) ... ttītä (ttīyā, ttī)* ‘when ... then’.

Reflexive and reciprocal pronouns. The reflexive pronouns are *hävia-* (*hīvia-*) ‘own’, *hamata-* ‘(one)self’, and *hvatä* ‘by oneself’; *hävia-* is commonly used to express possession (e.g., *jasti hīvī parau* ‘the lord’s command’, *hvaṃdānu hīvyā śandā* ‘the men’s ground’, cf. the use of MPers. *xwēš*). Reciprocity is expressed by *śśūjāta-* ‘one another’ (e.g., *hvāñindā śśūjātāna* ‘they speak with [= say to] one another’).

Adverbs. Neuter forms of adjectives can be used as adverbs (e.g., *śśāru* ‘well’). Common adverbs include adverbs of degree and manner: *bihīyu* ‘extremely’, *kāḍe* ‘very’, *samu* ‘only’ (Skt. *eva*), *thatau* ‘quickly’; of time: *vaysña*, *vaṃña* ‘now’, *īmu* ‘today’, *ysai (ysai)* ‘(very) early’; of place: *mara* ‘here’, *ttara* ‘there [where you are]’, *vara* ‘there’ (OIr. **imaθra*, **aitaθra*, **awaθra*), *ku*, *kuṣṭa* ‘where’, *ttatīka* ‘here’; from Indic: *andumaśu* ‘finally’, *avaśśā* ‘certainly’, *ttatvatu* ‘truly’. The suffix *-lsto* (MKhot., LKhot. *-ṣṭā*) added to the locative of nouns or to adverbs of place expresses direction, e.g., *biśśālsto* ‘to the house (*bisā-*)’, *kṣīruvo’lsto* ‘to the lands’, *hālsto*, *hāṣṭā* ‘thither’, *ciṃgvāṣṭā* ‘to among the Chinese (*ciṃga-*), to China’. The suffix *au* denotes languages, e.g., *hiṃdುವau* ‘in Indian’. Adverbial phrases are common, e.g., *ttu bāḍu ttye scātā* ‘at that time, at that time’, *cu pracai* ‘why?’ *ttāna/ttye pracaina* ‘therefore’.

Numerals. The numerals are of the usual type (similar to Persian), with the exception of *śśau* ‘one’. Teens are formed with the cardinal + **dasam*, with various phonetic developments of the initial *d-*, e.g., *śśūndasu* ‘11’, *dvāsu* ‘12’, *tcahulasu* ‘14’ (with *-l- < -r-d-*). The decades include forms such as *bāstā* ‘20’, *dārsā* ‘30’, *tcahulsā* ‘40’, *haudātā* ‘70’. Compounds of the decades are formed by an infix *pare-* ‘beyond, over’, e.g., *dvāvarebistā* ‘22’ (later *dvārebistā*), *pusparekṣaṣṭā* ‘65’, *hauparepaṃjsāsā* ‘57’. Higher numbers are *satā* ‘100’ and *ysāru* ‘1000’. For “10,000,” *byūrru* (Av. *baēuuar/n-*) is the literary word, while, in the Middle Khotanese documents, *ysā’ca-* ‘thousander’ is used.

The ordinals take the suffix *-ma-* (*tcūrāma-*, *tcūrma-* ‘4th’, *dasāma-* ‘10th’, etc.), with the exception of *paḍauysa-* ‘1st’, *śāta-* [*žāda-*] ‘2nd’, *dā(r)dda-* ‘3rd’, *pūha-* ‘5th’.

Gender. Of the most common noun stems, *a-*stems and most consonant stems are masculine, *ā-* and *i-*stems are feminine, and *n-*stems and a few *a-*stems are neuter. Adjectives form their feminine by turning *a-*stems into *ā-*stems. The



old feminine adjective stems in *-ī* (beside *-ā*) are represented by *māstä*, feminine of *māsta* ‘big’. Adjectives with the suffixes *-ānaa-* and *-īnaa-* (< **-āna-ka-*, **-aina-ka*) have feminine stems *-āṃjā-* and *īṃjā-* (< **-āna-čī-* + *-ā-*, **-aina-čī-* + *ā*; e.g., *myānaa-* ‘middle’, f. *myāṃjā-*; *brītīnaa-* ‘of desire [*brītaā-*]', f. *brītīṃjā-*).

Declension. The case system is similar to that of Old Persian, with syncretism of the genitive-dative and instrumental-ablative singular, but also the nominative-accusative plural. In addition, in the feminine, the genitive-dative is the same as the instrumental-ablative.

Most of the endings represent older forms directly, with **-a* and **-ā* > *-a*; **-ah* (and **-ahya*) and **ai* > *-ä*, *-i*; **-āh* > *-e*; **-am* > *-u*, **-ām* > *-o*. In the *k*-declensions, **-aka* and **-akā* > *-ā*; **-akah* (and **-akahya*) > *-ei*, *-ai*; **-akāh* > *-e*; **-akam* and **-akām* > *-au*; **-ik-ah* and **-ik-am* > *-ī*, **-ik-ā* > *-ya*, etc. (see also Sims-Williams, 1990; Hitch 2015; 2016).

The nominative-accusative plural of masculine *a*-stems ends in *-a*, which cannot be from older **-āh* (> *-e*), but, conceivably, from the neuter ending **-ā* or, perhaps, the accusative plural *aṃh*. Feminine *ā*-stems have forms in *-e* < **-āh*; other feminine stems and consonant stems have *-ä* < **-ah*. See also below.

The masculine instrumental-ablative ending is *-āna* (sometimes syncopated *-na*) from the pronouns. The feminine instrumental-ablative has the genitive-dative ending, usually with the postposition *jsa* (cf. Av. *hacā*, Pers. *az*). In Middle and Late Khotanese, the ending *-na* often functions like a postposition and interchanges with *jsa*. The relative pronoun has *kama jsa* ‘from whom’ (< **kahmāt*), cf. the relative adverb *kū* ‘where’, instr.-abl. *kūṃ jsa*, *kū jsa* ‘wherefrom’.

The vocative singular of *a*-stems ends in *-a*, of feminine *ā*-stems in *-ä* < **-ai*. The vocative plural is identical with the instrumental-ablative plural (without *jsa*).

The genitive-dative plural endings are *-ānu* < **-ānam* of *a*- and *ā*-stems and *-ānu* < **-īnam* of *i*- and consonant stems. Khot. *nu* (cf. Sogd. *-nw*) is from East Iranian **nam*, probably an archaism going back to a short ending **om* which can be reconstructed for Indo-European (Peyrot).

The instrumental-ablative plural ending *-yau* is theoretically from **-aibyām* (or similar form); it usually takes the postposition *jsa* ‘from, with’.

The locative plural is Old Khotanese *-uvo* < **aišuwām* (and similar forms),



Middle Khotanese *-vā*, in some manuscripts *-vau*.

The “pronominal” ending of the nominative-accusative plural masculine **-ai* is common (e.g., *biśśā*, Av. *vīspōi* < **wićwai*). Pronouns and “pronominal” adjectives regularly take their genitive-dative and locative singular endings from the *n*-stems (e.g., from *tta-* ‘that’: m., f. loc. sing. *ttiña*, f. gen.-dat. sing. *ttiñe*; *handara-* ‘other’: m. loc. sing. *handarña*, etc.).

Adjectives in *-ānaa-* and *-īnaa-* with feminine *-āmjā-* and *īmjā-* thematized < **āna-čī* and **aina-čī* also have *-ṃja* in the masculine locative singular < **na-ka-ya* (e.g., *maraṇīṃja mahāsamudro* ‘in the ocean of death [OInd. *maraṇa-*]’).

Many masculine *a*-stems can have plural endings from the *n*-stems (nom.-acc. *ksīrañä* ‘lands’, loc. *ksīrañuvo*, etc.). Singular forms from *n*-stems in the *a*-declension are rare in Old Khotanese (note loc. *ysraṃña* < *ysāra-* ‘heart’), more common later.

Some masculine *a*-stems have the nominative-accusative plural ending *-e*, which may represent the old ending **-āh*, perhaps also **ayah*. This also appears to be the ending of the rare neuter *a*-stems (e.g., *datu* ‘wild animal,’ pl. *date*). Neuter nouns and masculine *a*-stems with *n*-stem endings take adjectives in the feminine nominative-accusative plural (e.g., *puñīṃgye ttiṃmañä* ‘seeds of merit’ < nt. *tīman-*; *kīśśāngye bāysañä* ‘luxuriant woods’ < m. *bāysa-*).

Of the consonant stems, we may note the *r*-stems denoting relatives, which have oblique endings like *i*-stems in the feminine (e.g., nom. *pāte* ‘father’, acc. *pātaru* [later *pye*, *pyarä*], gen.-dat. *pīrā* < **piθrah*; but nom. *māta* ‘mother’, gen.-dat. *merä*; nom. *dūta* ‘daughter’, gen.-dat. *duīrā*, nom.-acc. pl. *dutarä* [later *dvarä*]; etc.).

The masculine *n*-stems *urmazdān-* ‘sun’, *rrāysan-* ‘ruler’ (cf. *rrāysanaunda-* ‘dominant’), and *mulysgyaṣṣaun-* ‘merciful’ (< **mulysdi-*) and the masculine *nd*-stems *rrund-* ‘king’, *hāyaund-* ‘master’, and *hva’nd-* ‘man’, later *hvaṇḍ-* (< **uśiwant-* ‘the aware one’ or **auśawant-* ‘mortal’) from old *nt*-stems have nominative singular in *-e* < **-āh*: *urmaysde* < **ahuramazdāh* (gen.-dat. sing. = nom.-acc. pl. *urmaysdānä*), *rrāyse*, *mulysgyaṣṣe*, *rre*, and *hve*’ and vocative *hīye* (Skjærvø/, 2016b, pp. 407, 418). The secondary stem *pandāa-* ‘path’ has nom. *pande*, acc. *pando* (< OIr. **pantāh*, *pantām*).

Of the old *h*-stems, only the nominative is found, e.g., *mase* ‘the size of’.



The personal pronouns form their cases in various ways: 1st sing. acc. *muho* (Tumshuqese *mvo*, both from **muwo* with *-w-* from 2nd sing. **tuwo?*); 2nd sing. *uho* (ultimately from **θuwām*); 1st pl. *ma* (< **ahma*) or *maha*; and 2nd pl. *uhu*, *uho*. The genitive-dative forms include old genitive and dative forms: 1st sing. *mamä*, 2nd sing. *tvī* (Tumshuqese *tivya*, < **ta/ubyah*), 1st pl. *māvu*, *mānu*, *mā*, 2nd pl. *umāvu*, *umānu*, *umā*. The instrumental-ablative forms are sing. 1st/2nd *muho/uho jsa*, 2nd also *tvī jsa*, 2nd pl. *umyau jsa*. The enclitic forms, including 3rd person, are: gen.-dat. sing. *mä*, *tä*, *yä* and *ī* (< **hai*), pl. *nä*, *-ū* (< **wah*), *nä*; instr.-abl. 3rd sing. *-ī jsa* and *-n jsa* (*-m jsa* or, without the nasal, *-jsa*); 2nd pl. *-ū jsa*, 3rd pl. *-n jsa*.

Of the cardinal numbers, ‘one’ *śśau* is inflected like *tta-* (*śye*, *śśāna*, *śśiñā*); ‘two’ has the old dual forms m. *duva*, f., nt. *dvī*, gen.-dat. *dvīnu*; and ‘three’ *drai*, *draya* has gen.-dat. *draiñu*. The cardinal numerals above ‘three’ are inflected like *i*-stems, e.g., *tcahaura* ‘four’, gen.-dat. *tcuīrnu*, instr.-abl. *tcūryau*; *pañja* ‘five’, gen.-dat. *pañjinu*; *dasau* ‘ten’, gen.-dat. *daśśānu*; *haṣṭātā* ‘eighty’, loc. *haṣṭevo*; *ysāre* ‘thousand’, loc. *yservo*, etc.

Pre- and postpositions. Common prepositions and postpositions include *anau*, *vīna*, *vānau* ‘without’, *patā* ‘before’; *jsa* ‘from’ (< **hačā*), *nuva*, *nuvaiya* ‘behind’ (< **ni-padi-ya*), *vara(ta)* ‘to’ (e.g., in letters), *vaṣṭa* ‘throughout’, *vātā* (*patā*) ‘on’, etc. (< **pati*), *vīrā* ‘on’, etc. (< **upari*), *vaska* ‘for the sake of, in pursuit of’ (< **paskāt*), as well as several of Indic origin, e.g., *kāḍana*, *kāḍāna* ‘on account of, for the sake of’ (< *krtena*), *pracaina* ‘because of’ (instr.-abl. of *praca-* < Indic *pratyaya* ‘cause’), *udiśśā* ‘on account of, with respect to’ (< *ud-diśya*). On the use of cases, see Emmerick, 1965.

The verb. The Khotanese verb has all the moods of Old Iranian (indicative, subjunctive, optative, injunctive, imperative) and active and middle (some verbs are conjugated in both the active and the middle with different meanings; see Canevascini 1991 on the reflexive function of the middle). The past tense or preterite (also called perfect, e.g., in Emmerick, 1968a) is formed from a past stem in the nominative singular and plural plus the enclitic copula. The present perfect is formed with the non-enclitic copula (but 3rd sing. *štā*) and the pluperfect with the preterite of ‘be’ (*vāta-*).

The past stem of intransitive verbs is the past participle (e.g., *āta-* ‘come’ < **āgata-*: past sing. 1st m. *ātā mā*, f. *āta mā* = 1st pl. m., f. *āte mā*, etc.; present perfect 3rd sing. *hāmātā štā* ‘has become’), but of transitive verbs it is from an active participle with the ablauting suffix *-ānd/-āt-* < **ā’ant/-ā’at-* and the



masculine nominative singular ending *-e* < **-āh*, feminine *-ātä* (e.g., *buḍe* ‘he carried’, f. *buḍātä*, pl. *buḍāndä*; 1st sing. m. *buḍaimä*, f. *buḍātä mä*, 1st pl. *buḍāndä mä*, etc.; present perfect 3rd sing. *hvate stä* ‘has said’, 3rd pl. *dätändä indä* ‘they have seen’; see Sims-Williams, 1997, pp. 322-23). This formation of the transitive past tense is reminiscent of the Sanskrit periphrastic perfect formation *bhṛtavān asmi* ‘I carried’.

The verbal system is characterized by a large number of intransitive/passive–transitive pairs (e.g., *vasus-* ‘be purified’ ~ *vasūj-* ‘purify’ < OIr. **awa-suxsa-/saučaya-*, *naštav-* ‘burn’ [intr.] ~ *naštēv-* [tr.] < **niš-tapa-/tāpaya-*). Causatives formed with the suffix *-āñ-* (cf. Pers. *-ān-*) and denominatives formed with the suffix *-ev-* (< Indic *āpaya-*) are rare (e.g., *gvachāñ-* ‘make digest’, *pajsamev-* ‘do homage’ < *pajsama-*).

There are four present conjugations, all from Old Iranian thematic presents, with a few survivals of athematic forms (*īmä* ‘I am’ < *ahmi*, etc., *āste* ‘he sits’ < *āstai*). Most verbs belong to the “regular” conjugation (from OIr. *aya-* stems), which has the 3rd sing. present ending active *-ātä*, middle *-äte*, and past stem in *-äta-*, or the “irregular” conjugation (OIr. *a-* stems), with 3rd sing. present ending act. *-tä*, mid. *-te*, and past stem in *-ta-* and numerous phonological changes. A few stems in *ai-* (from OIr. *-āya-*, *-āwaya-*, Indic *-ādaya-*) have 3rd sing. present ending act. *aitä*, *-aiyä*, mid. *aite*, *aiye*, 3rd pl. act. *-aindä*, mid. *-yäre* (e.g., *dai- dāta-* ‘see’ < **dāya* **dāta-*; *ysai- ysāta-* ‘be born’ < **zāya- zāta-*; *paṭhai-paṭhuta-* ‘burn’ [tr.] < **pari-θāwaya-* **pari-θuta-*; *praysai-* ‘have faith’ < Indic *prasādaya-*). Old Iranian stems in *-da-*, *-d(a)ya-*, *-θ(a)ya-*, *-h(a)ya-* have stems in *-y-/v-* with 3rd sing. present in act. *-ttä*, mid. *-tte* (e.g., *nättä* ‘sits’, 3rd pl. *nändä* < **ni-hīda-*; *saittä* ‘seems’, 3rd pl. *saindä* < **sadaya-*; *dättä* ‘appears’, 3rd pl. *diyäre* < **didya-*; *hvaittä* ‘thrashes’ < **hwahaya-*; *hamätte* ‘changes’ (intr.) < **framīθya-*; *butte* ‘knows’, 3rd pl. *buväre* < **bauda-*). Stems in *ai* can be described as ending in a vowel and taking the endings of the “regular” conjugation with the expected contractions, stems in *-y-/v-* as ending in a vowel and taking the endings of the “irregular” conjugation (see Hitch 2017).

Past stems (past participles) are of the common Iranian types. The “regular” verbs take *-äta-*, “irregular” ones take *-ta-*. After vowels, the *t* remains in OKhot., but is then lost, e.g., *panata-* ‘risen’ (pres. *panam-*), *tsuta-*, MKhot. *tsva-* ‘gone’ (pres. *tsū-*), *dāta-*, *dya-* ‘seen’ (pres. *dai-*), *ysāta-*, *ysāva-*, *ysā-* ‘born’. The *t* remains after *s*, e.g., *basta-* ‘bound’, but is assimilated after other consonants, e.g., *buḍa-* ‘carried’ (< **brta-*, pres. *bar-*), *pyūšta-* ‘heard’ (pres. *pyūš-*), *padanda-* ‘made’ (pres. *padīm-*), *purrda-* ‘vanquished’ (pres. *purr-*), *byauda-* ‘found’ (<



**abi-āfta-*, pres. *byeh-*, *byev-*). OIr. *-axta-* > *-īta-*, *-īya-*, e.g., *sīta-* ‘learned’ (pres. *sāj-*). New types include analogical forms in *-āta-*, *-ānda-*, *-aunda-*, *-autta-*, e.g., *huṣṣ-āta-* ‘grown’ [beside *huṣṭa*], *paysān-* *paysānda-* ‘to recognize’, *brembraunda-* ‘weep’, *ysānāh-* *ysānautta-* ‘to bathe’ [intr.] < **snāf-*, *dukhevdukhautta-* ‘make suffer’ < Indic *duḥkhāpaya-*). There are a few suppletive pairs, e.g., *hīs-* *āta-* ‘come’ (< **ā-isa-* *ā-gata-*).

Old preverbs include most of the common ones (Emmerick, 1968a, pp. 229-44), e.g., **abi-*: *baṣṭ-* ‘establish’ < **abi-štaya-*; **apa-*: *pajsem-* ‘pass (time)’ < **apa-ĵāmaya-* ‘make go away’(?); *ā-*: *āvun-* *aurāta-* ‘bless’ < **ā-frnā-* *ā-frīta-*; **frā-*: *hanaśś-* ‘be destroyed’ < **fra-nasya-*; *haur-* *hūda-* < **fra-bara-* *fra-brta-*; *hāruv-* *hārsta-* ‘grow’ < **frā-rauda-* **frā-rusta-*; **ni-*: *nājsāṣ-* ‘show’ < **ni-čaša-*; **pari-*: *parrīj-* *parrāta-* ‘deliver’ < **pari-raičaya-* **pari-rixta-*; *pati-*: *pyūṣ-*, *pyūv-* ‘listen’ < **pati-gauša-*; **upa-*: *panam-* ‘rise’ < **upa-nama-*; **wi-*: *bāysān-* ‘wake up’ < **wi-zānā-*; *guhay-* *guhasta-* ‘wound’ < **wi-xadaya-* **wi-xasta-*. New preverbs include *tca-* (*tcabalj-* ‘scatter’) with the compound forms *ggujsa-* and *pajsa-* (*ggujsabalj-* ‘overcome, scatter(?)’, *pajsabalj-* ‘beat (a drum)’).

Endings. Most of the endings are directly from Old Iranian (e.g., active singular 2nd *-ä* < **ahi*, 3rd *-tä* < **ati*, plural 1st *mä* < **amahi*, 2nd *ta*, < **-ata*, 3rd *-īndä* < **anti*, but 1st sing. *-īmä* < **-āmi*; middle singular 1st *-e* < **ai*, 2nd *-a* < **-aha*, etc.). The 3rd plural indicative and optative middle have the old *r*-endings (indic. *-āre* < **-ārai* [with secondary *-e*], opt. *-īro*, *-īru* < **īrām*); the optative ending was also extended to the 2nd plural (replacing the descendant of OIr. **adwam* > **īyu*?).

The copula is archaic and has enclitic forms (sing. 1st *īmä*, encl. *mä*; 2nd *-ī*, pl. *mä*, *sta*, *īndä*; opt. 3rd sing. *ya* (soon replaced by *īyä*), subj. 3rd sing. *āya*, pl. *āro*). The 3rd sing. *āstä*, encl. *stä* (negated *nāstä*) denotes existence. The copula is suppletive with optative present *v-* (e.g., sing. 1st *vyo*, 3rd *vya*, pl. 3rd *vīro*), past *vāta-* (Skjærvø, 1981, pp. 461-63).

Moods. The subjunctive, optative, and injunctive are used fairly indiscriminately, the injunctive being the least common (e.g., from different manuscripts of the *Suvarṇabhāsottama-sūtra*: OKhot. *śuru ṇu vātä yan-ä* [opt. act.] *hvāṣṭa nä paysān-da* [inj. mid.] *u aysmū nu vātä yan-āte* [subj. mid.] *u pajsamu nä yan-ä* = (archaizing) MKhot. *śśuru ṇu vātä yan-īyä* [opt. act.] *u hvāṣṭa nä paysān-āte aysmū nu vātä yan-āte pajsama nä yan-īyä* = (standard) MKhot. *śuru ṇä ve yan-āte hvāṣṭa-ṃ paysān-āte āysda-ṃ ni yaṃ-da* [inj. mid.] *u pajsamu ni yan-ä* ‘(he who) may do good to them, may revere them, watch over



them, and do homage to them’).

The optative preterite is used in clauses expressing hypothetical or counterfactual conditions (Skjærvø, 1981, pp. 461-63).

Passive. The passive is formed with the auxiliaries *häm-* and *väta-* ‘become’, also in the case of compound expressions such as *ānatä yan-* ‘protect’, *ānatä häm-* ‘be protected’, *hämurä yan-/häm-* ‘forget/be forgotten’. An animate agent is expressed by the genitive-dative, e.g., *īmu mamä puñīgye ttīmañä prānde kälste hämāre* ‘today, by me, seeds of merit will be scattered (*prānde* < *prān-* < **parākana-*) and sown (*kälste* < **kṛṣṭa-*)’ (Skt. *mayā ... avaruptāni bhaviṣyanti*), an inanimate agent by the instrumental-ablative, e.g., *āchyaū skuta väta* ‘they had been touched by illnesses’ (Skt. *roga-sprṣṭāni*).

Potentialis. Khotanese, like Sogdian, has a potential expressed with the auxiliaries *yan-* ‘do’ (transitive, active), expressing possibility (“can, cannot,” Skt. *śakyam*), and *häm-* ‘become’ (intransitive, passive), expressing passive possibility or anteriority and completion (Sanskrit absolute). Examples: *ne haṃkhäṣṭu yanīndä* ‘they cannot count’, (LKhot.) *haṃkhiṣṭa hime* ‘it can be counted’ (both Sanskrit *śakyam gaṇayitum*); (MKhot.) *cītä hā tsue himāte* ‘when he has gone thither’ (Skt. *upetya*).

Participles and infinitives. There are several verbal adjectives: present participle active *anda(a)-*, middle *-āna(a)*; participles of necessity: in *-ia(a)-*, *-āña(a)-* (e.g., *tcēra-* ‘which ought to be done’ < **čārya-*, *haurāña-* ‘which ought to be given’). On the past participle, see above.

There are two infinitives: one from the present stem with the ending *-ä* and one from the past stem with the ending *-ie* (cf. Av. *-taiiaē°*), e.g., *ne haute biśśä dukha näṣem-ä* ‘I cannot calm all sufferings’; *ne hautāre dātu pyūṣ-ṭe* ‘they cannot listen to the Law’.

Note also the use of a verbal noun in *-āmatā-* or an infinitive plus *kṣam-* ‘please’ to render Sanskrit *-tu-kāma-* ‘wishing to do’, e.g., *huṣṣāmata kṣamīyā* ‘increase might please (him)’ and *huṣāñāte kṣamīyā* ‘it might please (him) to make grow’ (both Skt. *vivaraddhayitu-kāmo bhavet*).

Conjunctions and particles: Conjunctions. Coordinating and disjunctive conjunctions include the common *u* ‘and’, enclitic *rro* ‘and,’ *au* (later *ā*), *vā* (< **uta vā*) and *au vā* ‘or,’ *haḍe*, *hāḍe* ‘but’. Subordinating conjunctions include several made from the relative-interrogative stems *k-* and *c-*, e.g., *ka*, *ko* (< *ka* +



-u) ‘if’, *kū* ‘where’, *kho* ‘when, as’, *cu* (later *ci*, *ca*; *ttāna cu*) ‘because’, *cu mānau* ‘although’, *cu bāḍu* ‘at what time, when?’, *cu pracai[na]* ‘for what reason, why?’, *cītā* ‘when’, *kāmu buro* (... *ttāmu buro*) ‘as long as (... so long)’.

Negations. Statements are negated with *ne*, *nā* (later *ni*, *na*), imperative and exhortations with *ma*. Emphatic negations can be expressed by *ne ne* and *ma ne*, *nā härṣtai*, *hārgyu ne*, *nā gāvu* ‘not at all’. Other negations include *ne/nā* ... *ne/nā* ‘neither ... nor’ and *na-ro* ‘not yet’.

Particles. Common enclitic particles include enclitic *ju* (< *cu*, later *ji*, *ja*), used with personal and demonstrative pronouns, relative pronouns, conjunctions, etc., and *vā* ‘and, but, moreover’ (different from *vā* ‘or’, probably short form of *vātcu*, below, e.g., *ttā ju* ‘for they’, *ce ju* ‘he who’, *ce ju šā cu* ‘whoever’, *kho ju* ‘when’), negations *ne ju* ‘although not, not at all’, *ma ju* ‘but do not let’, *kho rro ju* ‘and just like’, *cīyā vā* ‘and/but when’, *cai* (< *ce* + *-ī*) *rro ju vā pyūṣḍe* ‘and he who does listen to it’. Narrative sequences are often introduced by *pātcu* (encl. *vātcu*, later *pātcā*, *vātcā*), *vā* ‘next, then’ (also *pātcu vā*). The enclitic (emphatic?) particle *mī* is especially common from Middle Khotanese on. The most common emphatic particle is *-ī*.

Directional particles, especially common from Middle Khotanese on, include *hā*, *ttā*, *vā*, which express direction to 3rd, 2nd, and 1st person, respectively, sometimes replacing the personal pronouns. They can be extended with the locative particle *-lsto*, later *-āṣṭā* (e.g., *vālsto* ‘hither [to me/us]’, LKhot. *aysa hvāñū ttāṣṭa* ‘I say to you’).

The vocative may take *hai*, *he* ‘O’.

Questions. Direct and indirect questions are introduced by question words, for instance interrogative pronouns and adverbs (*ce* ‘who?’ *cu* ‘what?’ *cītā* ‘when?’ *cerā* ‘how many?’ *cūḍe* and *ce kāḍāna* ‘why?’ *kho* ‘how?’ etc.). Disjunctive questions take *o ne* ‘or not?’ The imperative can take a particle *ne* (e.g., *dya ne thu* ‘won’t you look?’ > “look!” [Skt. *paśyā-hi*]; cf. Sogd. *LA* = *nē*, see Sims-Williams, 1996, pp. 181-82).

Direct speech. Direct speech is commonly introduced by the particle *se* (later *si*, *sa*), perhaps related to *shyty* in the Aramaic inscriptions of *Aśoka*, which G. Morgenstierne (q.v.) suggested might be derived from **sahyati* ‘is said’ (apud Birkeland, p. 233 n. 1). It is used with words of speaking, thinking, knowing, etc. Note *ttye tta* (or *ttai* < *tta* + *-ī*) *hāmātu se* ‘it occurred to him that’ (Skt.



tasyaitad abhavat). In questions: *ttu ne ne bve se kāmā śā padmagarbhā* ‘I do not know (that): Which one is Padmagarbhā?’ Without *se*: *mahākāśavī tta hve badra crrāmā tvānai horā ttrāmu ...* ‘Mahākāśyapa said to him (-ī): O Badra, as is your gift, so ..’; (LKhot.) *tta hve pvīryau* ‘He said (to them): Listen!’

Word order. Normal prose word order is subject + indirect object + direct object + verb. Fronting of the verb is not uncommon, notably in legal language, e.g., (MKhot.) *nāti mī yagurā ttuā ütca ... hauḍä mī ši’ yagurā tti mūri 2000 500. nāmḍūṃ-m-ūṃ mihi braṃgalä... tti mūri uspurri 2000 500* ‘Yagura received (*nāti* ‘took’) that water ... That Yagura paid (*hauḍä* ‘gave’) those 2500 *mūrās*. We, Braṃgala (etc.) received them (-ūṃ), (i.e.) those 2500 *mūrās* complete’. Fronting can be replaced by *cu (... ṣṭe)* ‘as for’.

STAGES OF THE KHOTANESE LANGUAGE

Traditionally, in Khotanese studies, two stages of the language have been distinguished: Old and Late Khotanese, including in the latter category all texts not in regular Old Khotanese and thus applying the term to both the 8th-century texts from Khotan and the 10th-century texts from Dunhuang. The main linguistic divide occurs between the 8th- and 10th-century texts, however, and it is therefore more useful to distinguish three periods: Old Khotanese (ca. 5th-6th centuries), Middle Khotanese (ca. 7th-8th centuries), and Late Khotanese (9th-10th centuries, to the end of Khotanese literature). See also Emmerick, 1979, 1987, 1989.

Old Khotanese. This is phonetically and morphologically still basically close to the “Old Middle Iranian” type. While intervocalic voiced stops had already been lost, in the oldest strata of Old Khotanese intervocalic *g* and *t* still remained, and the final vowels *-i*, *-ä*, and *-u* were still distinct. Occasionally the final syllable *-tä* was lost after long vowels and diphthongs (e.g., *ākṣūtä* > *ākṣū* ‘begins’). In the late Old Khotanese period, however, *g* and *t*, together with the intervocalic voiced sibilant *ž*, merged into one phoneme, which was frequently written *t* regardless of its origin and was presumably realized as a kind of flap or continuant. The final vowels *-i* and *-ä* were no longer distinct phonemes. Old Khotanese texts were also copied well into the Middle Khotanese period, but these usually give themselves away by Middle Khotanese and pseudo-Old Khotanese forms introduced by the scribe(s). This feature is typical of several parts of the main manuscript of the *Book of Zambasta*, upon which much of Emmerick’s description of “standard” Old Khotanese in his *Saka Grammatical Studies* is based.



Middle Khotanese. Here, the intervocalic “glide” phoneme was completely lost and the ensuing vowel sequences suffered various modifications, although archaizing spellings are occasionally found. Contraction of equal vowels: OKhot. *suhāvatāna-* (Skt. *sukhopadhāna-*) > MKhot. *suhāvāna-*; OKhot. *ttagata-*, *ttatata-* ‘wealth’ > MKhot. *ttata-*; OInd. *ākāśa-* ‘space’ > *āgāśa-* [*āgāza*], later *ātāśa-* [*ādāza*], LKhot. *āvaśa-*, *āśa-* [*āza-*]; unequal vowels: OKhot. *suhāta-* > MKhot. *suhya-*; the endings 3rd sing. present OKhot. *-ātā* > MKhot. *-e*, *-ä* and 3rd sing. preterit OKhot. *-āte* > MKhot. *-ye*. Occasionally, the hiatus is maintained, as in *āe* for OKhot. *ātā* ‘he came’. These developments also provide clues to stress, for instance, the 3rd sing. present *hāmāte* ‘becomes’ became *hāmā*, but the 3rd sing. m. preterite *hāmātā* became *himye*; the adverb *thatau* ‘quickly’ first became *thātau*, then *thyau* (see Maggi, 1990 [1993], p. 182). Archaizing spellings are of the type *hamyetä* for MKhot. *himye* + OKhot. *hāmāte*.

Already in Old Khotanese, there was a tendency for *l* to be lost before sibilants while, apparently, rhotacizing the preceding vowel. This became the rule in Middle Khotanese (e.g., OKhot. *pā’sa-* ‘pig’ < **pālsa-*, Av. *parāsa-*; OKhot. *ggei’s-* ‘turn’ (intr.), MKhot. *ge’s-*; OKhot. *puls-* ‘ask’, MKhot. *pu’s-*).

Final *-u* was generally replaced by *-i* and *-ä* (except in *ttu*), and final *-e* merged with *-i/-ä* and final *-o* with *-u*. Final *-nä* and *-mä* were further weakened to *-ṃ*, notably in the genitive-dative plural ending *-ānu* > *-ānä* and *-āṃ*, and in the nominative-accusative singular, genitive-dative singular, and locative singular of stems in *-ana-* and *-ama-* > *-aṃ*. The original finals reappeared in hiatus (e.g., *nāṃdūṃ-m-ūṃ* ‘we received them’ < *nāṃdä mä* + *-ṃ* < *nä*). The enclitic *-ṃ* after long vowels usually takes the hiatus-filler *-t-* (e.g., *tvā-t-ūṃ* ‘this [acc.] to them’ < *ttuto nä*).

In some Old Khotanese manuscripts used in the Middle Khotanese period, Middle Khotanese forms are indicated on the original words, e.g., OKhot. *tsuātāndä* ‘they went’ (with *-ā* added to *tsu*) for MKhot. *tsuāṃdä*; OKhot. *śiraetete* ‘of goodness’ (with *-e* added to *ra*) for MKhot. **śiretä*; OKhot. *ttyāṃnu* (with *-y-* and *-ṃ* added to *ttā*) for MKhot. *ttyāṃ*; OKhot. *uvpeikṣa-* (with a subscript *pe*) for MKhot. *upekṣa-*.

The accusative singular merged with the nominative, and the endings *-u*, *-o*, and *-au* were replaced by *-i/-ä*, *-a*, and *-ā* respectively. In nouns where the accusative was palatalized, the accusative form is used in the nominative (e.g., *Śaṃdrāma māsta gyaštā* ‘the great goddess Śrī’).



Spellings such as *kṣāṇā* for *kṣuṇā* ‘regnal year’ and *rrāṃdā* for *rruṃdā* ‘king’ are found for the first time.

Typical new grammatical forms include the locative plural in *-vā* (dialectal, for OKhot. *-uvoʔ*), and the feminine accusative singular pronoun OKhot. *ttuto* > *ttuo* > MKhot. *ttuā* and *tvā*. The personal pronouns 1st and 2nd plural are *mahe*, *mihe* and *umä*, *amä* and similar forms. The enclitic 2nd singular pronoun is *-e* (OKhot. *tä*), which (synchronically) replaces any final vowel, rather than combining with it (e.g., OKhot. *trāme tä tce'mañä* ‘thus are your eyes’ > MKhot. *ttrāmä tvī tce'mañä* and *ttrām-e tce'mañä*; replacing *-a*: *tcamna* + *-e* ‘whereby your’ > *tcamn-e*). The feminine stems in *āmatā-* and *-āti-* become *āmā-* and *ā-*stems.

The 1st singular indicative ending *-imä* becomes *-ümä* and *-ūṃ* (probably from the by-form *ämä*), while *-ime* becomes the ending of the 1st singular subjunctive/optative active instead of *-iñä* and *-ä*. The 2nd plural imperative active/middle has the new ending *-yari* (e.g., *ārryari* ‘grind [the flour]!’). In the preterite of intransitive verbs, the 1st singular masculine enclitic copula *mä* becomes *-ṃ* and the ending *-ä mä* > *-ūṃ* (e.g., *ātūṃ*, *āvūṃ* ‘I came’), and, similarly, in the preterite of transitive verbs, the 1st plural masculine in *-ändä mä* > *āmḍūṃ* (e.g., *buḍāṃdūṃ* ‘we carried’). The stem *näjsaṣ-* ‘show’ became *nijsub(-)* (pres. 3rd sing. *nijsuṣḍä*), and the preterite of *haur-* ‘give, pay’ is usually *haudä* (OKhot. *hūde*).

The present participles of *jsā-* ‘go’, *ā-* ‘sit’, and *ṣṭ-* ‘stand’ (*jsāna*, *āṇa*, *ṣṭāna*), which already in Old Khotanese may have been used to modify the verb, have become invariable particles used with verbs and other words, apparently to express ongoing events (e.g., late OKhot. *ttānu āysanānu pīro gyasta balysa dātaimä āṇaṃdā kye āṇa dātu hvāñāre* ‘on those seats, I saw lord buddhas sitting, who are (= were) proclaiming the law’).

Late Khotanese. Here, as it appears from the orthography of the manuscripts, the vowel system has changed radically, with, at least, some loss of both quantity and quality distinctions. Long *ā* (and nasalized *ḡ*) to a large extent merged with *au* and *ū* into a single back rounded phoneme (e.g., *rruṃd-*, *rrāṃd-*, *rraud-* ‘king’), and short and long *ī* merged with *ai* into one front vowel (e.g., *siddham*, *saiddham* ‘welfare!’; see Emmerick, 1979, and in *CLI*, p. 209). Final *-i/-ä* and final nasals were frequently not marked. The Chinese text in Brāhmī script analyzed by Emmerick and Pulleyblank, however, suggests that the changes may not have been so drastic, and that, for instance, *i* and *u* were



still distinct from \bar{i} and \bar{u} (Kumamoto, 1995 [paper from 1991]; Emmerick and Pulleyblank, pp. 45-47). Final syllables were lost in weakly stressed words (e.g. $a < \text{MKhot. aysä 'T, pha 'much' < MKhot. pharä}$).

In syllables with OKhot. $-l-$ before s, z , which lost the l in Middle Khotanese, the vowel was raised in Late Khotanese, e.g., $h\bar{ä}lysd\bar{ä}$ 'present', MKhot. $hi'y\bar{s}da$, LKhot. $haiysda$ (etc.); OKhot. $puls-$ 'ask', MKhot. $pu's-$, LKhot. $pvais-$ (beside $pvai's-$, $pves-$, $pvis-$, etc.).

The fact that the documents/letters from Dunhuang are, many of them, drafts and exercises, no doubt lends the language a more irregular look than it deserves. Even religious texts differ considerably in their orthography; for instance, in the *Bhadracaryā-deśanā* (Asmussen), vocalization is much more careless than that of the *Deśanā-parivarta* of the *Suvarṇabhāsottama-sūtra* (Skjærvø, 2004) contained in the same manuscript.

It seems clear, however, that $-āṃ$ and $-au$ merged in a single phoneme [ɔ], which caused the merger of the genitive-dative plural ending $-āṃ$ and instrumental-ablative $-yau$, as well as locative $-vā$, written variously, e.g., gen.dat. pl. ($-ānu$) $jast-āṃ hvāṇḍ-au$ 'of gods (and) men', loc. pl. ($-vā$) $pr̥r̥yivā ... hvāṇḍ-āṃ$ 'among ghosts (and) men'.

There is some vacillation in certain consonant groups, for instance, $ṭh$ and $kṣ$, which may indicate similar pronunciation, e.g., $śśaṭhā-$, $śaksā-$ 'deceit' (from Sanskrit). The graphs $<ṣṭh>$ and $<ṣc>$ alternate, which may be due to graphic similarity rather than phonetic identity, e.g., $nuṣṭhura-$, $nuṣcura-$ 'harsh' (from Sanskrit).

These developments led to some new morphological devices and even some new grammatical categories; for instance, nouns and adjectives ending in long vowels or diphthongs now formed the nominative-accusative plural with the ending $-ta-$, $-va$ (e.g., nom. sing. $nā$ 'nāga', nom.-acc. pl. $nāta$). OKhot. $dātia-$ 'of the Law', MKhot. $dāyia-$ was replaced by $dāvīa-$ (cf. OKhot. $dātīnaa-$ 'of the Law', MKhot., LKhot. $dāvīnaa-$). Past stems in $-āta-$, which became $ya-$ in Middle Khotanese, could lose their $-y-$ after palatal consonants ($bārāsāte$ 'shone' > $birāsē$).

The OKhot. present stem $hvāñ-$ 'speak' became $hūñ-$, and the past stem $byauda-$ 'found' became $bīd-$. The 1st singular optative active has the endings $-īṃ$, $-īne$ ($īna$, also middle), and $-īme$. In the preterite, the 1st singular masculine



transitive/intransitive has *-eṃ, -ai, -i(ṃ)* (cf. OKhot. *vätämä* ‘I was’ > *vyiṃ, vyi, yai*; OKhot. *yäḍaimä* ‘I did’ > *yideṃ, yuḍai*, etc.), and the 1st plural masculine intransitive has *-aṃdūṃ (-adūṃ, -adū)* in analogy with transitive *āṃdūṃ* (e.g., *hamyadūṃ* ‘we became’, OKhot. *hämäta mä*).

Late Khotanese also has a narrative optative, reminiscent of the “preterital optative” in Old Persian, Avestan, and Sogdian, e.g., *bišä hālā pattaviya* ‘it shone in all directions’ (see Dresden, 1970, pp. 136-39).

Lexicon. The lexicon is basically Iranian and Indic (Sanskrit and Prakrit). From Greek there is OKhot. *satīra-*, later *sera* (from *statēr*) and *drammaa-* (probably from *drachmē*). There are a few Tibetan and Chinese terms (see Skjærvø, 2002, pp. lxxvi-lxxviii), and the documents concerning the Uygurs contain some Turkish words.

Among Zoroastrian terms from pre-Buddhist times are *urmaysde* ‘sun’ (< **ahura-mazdāh*) and *śśāndrāmatā-*, the Buddhist goddess Śrī (< **ćwantā aramati-*, Av. *Spəntā Ārmaiti*, the earth). See CHINESE TURKESTAN ii; Skjærvø, 1998.

BIBLIOGRAPHY

Basic reference works include Emmerick, 1992a, and Kumamoto, 1985, for miscellaneous information and bibliographies. For the early history of Khotan, see also Stein, 1907, and chinese turkestan i. Grammatical descriptions of the language are found in Leumann, 1912; Konow, 1916, 1932, 1941a, 1949; Dresden, 1955; Bailey, 1958; Emmerick, 1968a, 1989, and in *Compendium Linguarum Iranicarum* (CLI), pp. 204-29 (q.v. for further details). Numerous publications of individual texts contain glossaries. Bailey’s Prolexis to the *Book of Zambasta* (1967) contains useful discussions of select Iranian and non-Iranian words in that text, while his *Dictionary* (1979) contains only words of Iranian descent, leaving out the entire Indic vocabulary. Emmerick and Skjærvø (1982, 1987, 1997) contains studies by several authors of individual words and grammatical forms. A bibliography of Bailey’s work up to 1970 was published in *BSOAS* 33, 1970, pp. ix-xiv; the complete bibliography promised by Emmerick (1999, p. 348) has not appeared, but it is found on the website of the Ancient India and Iran Trust, Cambridge, U.K. (www.indiran.org).



Emmerick's bibliography to 2005 is found in M. Macuch, M. Maggi, and W. Sundermann, eds., *Iranian Languages and Texts from Iran and Turan: Ronald E. Emmerick Memorial Volume*, Wiesbaden, 2007, pp. xxi-xlii; M. Schwarz and V. Blažek, "Ronald Eric Emmerick (9.3.1937, Sydney – 31.8.2001, Hamburg)," *Linguistica Brunensia* 60, nos. 1-2, 2012, pp. 265-73 provide a broad thematic index in Czech to Emmerick's books and articles.

J. P. Asmussen, *The Khotanese Bhadracaryādeśanā*, Det Kongelige Danske Videnskabernes Selskab, Hist.-fil. Medd. 39.2, Copenhagen, 1961.

H. W. Bailey, *Khotanese Texts I*, Cambridge, 1945. Idem, "Irano-Indica II," *BSOAS* 13, 1949, pp. 121-39.

Idem, *Khotanese Buddhist Texts*, London, 1951. Idem, *Indo-Scythian Studies being Khotanese Texts II*, Cambridge, 1954.

Idem, "Languages of the Saka," in *HO I, IV: Iranistik 1*, Leiden and Cologne, 1958, pp. 131-54.

Idem, *Indo-Scythian Studies being Khotanese Texts IV*, Cambridge, 1961.

Idem, *Indo-Scythian Studies being Khotanese Texts VI: Prolexis to the Book of Zambasta*, Cambridge, 1967.

Idem, *Saka Documents I-IV*, Corpus Inscr. Iran. II/V, portfolios I-IV, London, 1960-67.

Idem, *Saka Documents: Text Volume*, Corpus Inscr. Iran. II/V, London, 1968.

Idem, *Dictionary of Khotan Saka*, Cambridge, 1979.

Idem, *Khotanese Texts VII*, Cambridge, 1985.

C. I. Beckwith, *The Tibetan Empire in Central Asia*, Princeton, 1987.

H. Birkeland, "Eine aramäische Inschrift aus Afghanistan," *Acta Orientalia* 16/3, 1938, pp. 223-33.

A. M. Boyer, E. J. Rapson, and E. Senart, *Kharoṣṭhī Inscriptions I*, Oxford, 1920-29.

T. Burrow, "Iranian Words in the Kharoṣṭhī Documents from Chinese



Idem, "Legends and Ceremonies: Based on the Observation of the Qu Shu Collection at Xinjiang Lop Museum," in Xiao Li, ed., *Non-Han Literature Along the Silk Road*, Silk Road Research Series, Singapore, 2020, pp. 21-45.

R. E. Emmerick, "Syntax of the Cases in Khotanese," *BSOAS* 29, 1965, pp. 24-33.

Idem, *Saka Grammatical Studies*, Oxford, 1968a. Idem, "Names from Central Asia," *Central Asiatic Journal* 12, 1968b, pp. 88-91.

Idem, "The Vowel Phonemes of Khotanese," in B. Brogyanyi, ed., *Studies in Diachronic, Synchronic, and Typological Linguistics. Festschrift for Oswald Szemerényi*, Amsterdam Studies in the Theory and History of Linguistic Science 2, Amsterdam, 1979, pp. 239-50.

Idem, "The Consonant Phonemes of Khotanese," in *Monumentum Georg Morgenstierne I*, Acta Iranica 21, Leiden, 1981, pp. 185-209.

Idem, "Tibetan Loanwords in Khotanese and Khotanese Loanwords in Tibetan," in G. Gnoli and L. Lanciotti, eds., *Orientalia Iosephi Tucci Memoriae Dedicata*, Rome, 1985, pp. 301-17.

Idem, "The Transition from Old to Late Khotanese," in *Transition Periods in Iranian History. Actes du Symposium de Fribourg-en-Brisgau (22-24 Mai 1985)*, Studia Iranica, Cahier 5, 1987, pp. 33-42.

Idem, "Khotanese and Tumshuqese," in R. Schmitt, ed., *Compendium Linguarum Iranicarum (CLI)*, Wiesbaden, 1989, pp. 204-29.

Idem, *A Guide to the Literature of Khotan: Second Edition Thoroughly Revised and Enlarged*, Studia Philologica Buddhica: Occasional Paper Series 3, Tokyo, 1992a.

Idem, "The Dunhuang Ms. Ch 00120: Its Importance for Reconstructing the Phonological System of Khotanese," in A. Cardona, ed., *Turfan and Tun-huang: The Texts. Encounter of Civilizations on the Silk Route*, Florence, 1992b, pp. 145-70.

Idem, "Khotanese *ei*," in N. Sims-Williams, ed., *Proceedings of the Third European Conference of Iranian Studies Held in Cambridge, 11th to 15th September 1995 I*, Old and Middle Iranian Studies, Wiesbaden, 1998, pp. 93-97.

Idem, "Harold Walter Bailey 1899-1996," *Proceedings of The British*



- Academy* 101: “1998 Lectures and Memoirs,” 1999, pp. 309-49.
- Idem, “Khotanese and Tumshuqese,” in G. Windfuhr, ed., *The Iranian Languages*, London, 2009, pp. 377-415.
- R. E. Emmerick and M. Maggi, “Thoughts on Khotanese *e* and *o*,” in R. E. Emmerick and D. Weber, eds., *Papers in Honour of Prof. D. N. MacKenzie on the Occasion of his 65th Birthday on April 8th, 1991*, Frankfurt, etc., 1991, pp. 67-73.
- R. E. Emmerick and E. G. Pulleyblank, *A Chinese Text in Central Asian Brahmi Script*, Serie Orientale Roma 69, Rome, 1993.
- R. E. Emmerick and A. Róna-Tas, “The Turkish-Khotanese Word-List Revisited,” *Central Asian Journal* 36, 1992, pp. 199-241.
- R. E. Emmerick and P. O. Skjærvø, eds., *Studies in the Vocabulary of Khotanese*, 3 vols, Vienna, 1982-97.
- R. E. Emmerick and M. Vorob’yeva-Desyatovskaya, *Saka Documents Text Volume III: The St. Petersburg Collections*, Corpus Inscr. Iran. 2/V, London, 1995.
- Monika Gronke, “The Arabic Yärkand Documents,” *BSOAS* 49/3, 1986, p. 484.
- J. Hamilton, “Les règnes khotanais entre 851 et 1001,” in M. Soymié, ed., *Contributions aux études de Touen-houang* [I], Geneva, 1979, pp. 49-54.
- Idem, “Sur la chronologie khotanaise au IX^e-X^e siècle,” in M. Soymié, ed., *Contributions aux études de Touen-Houang* III, Publications de l’École française d’Extrême-Orient 135, Paris, 1984, pp. 47-48.
- D. A. Hitch, “Old Khotanese Synchronic Umlaut,” *IJ* 33, 1990, pp. 177-98.
- Idem, “Contracted Diphthongs in Old Khotanese,” *IJ* 58, 2015, pp. 293-324.
- Idem, “Contracted Semivowels in Old Khotanese,” *IJ* 59, 2016, pp. 259-94.
- Idem, “Old Khotanese Type A Stems in -a- and -e-,” *BSOAS* 80, 2017, pp. 491-523.
- A. F. R. Hoernle, letters, British Library, London, India Office Records and Private Papers, MSS Eur D 815 (courtesy of Ursula Sims-Williams).
- Idem, “Three Further Collections of Ancient Manuscripts from Central



Asia,” *JRASB* 66, pt. I, 1897, pp. 213-60, plates VII-XXX; repr. Calcutta, 1897, plates I-XXIV.

Idem, *A Note on The British Collection of Central Asian Antiquities Presented to the XIIth International Congress of Orientalists in Rome, October, 1899*, Oxford, 1899a.

Idem, “A Report on the British Collection of Antiquities from Central Asia,” pt. I, *JRASB* 68, pt. I, 1899b (Extra No., Calcutta, 1899), pt. II, *JRASB* 70, pt. I, 1901 (Extra No. 1, Calcutta, 1902).

Idem, “Ancient Manuscripts from Khotan,” *JRAS*, 1906, pp. 695-98.

Idem, “The ‘Unknown Languages’ of Eastern Turkestan,” pt. 1, *JRAS*, 1910, pp. 834-38, 1283-1300; pt. 2,” *JRAS*, 1911, pp. 449-77.

Idem, *Manuscript Remains of Buddhist Literature Found in Eastern Turkestan*, Oxford, 1916.

Taijun Inokuchi, “Utengo shiryō ni yoru Viśa ōke no keifu to nendai” (Succession and chronology of the Viśa Royal Family, based upon the Khotanī-Saka Documents), *Ryūkoku Daigaku ronshū* 364, 1960, pp. 27-43.

J. Kirste, review of Leumann, 1912, in *Vienna Oriental Journal* (= *WZKM*) 26, 1912, pp. 394-400.

S. Konow, review of Leumann 1912, in *Göttingische gelehrte Anzeigen*, 1912, no. 9, pp. 551-65.

Idem, “Khotan Studies,” *JRAS*, 1914, pp. 339-53.

Idem, ed., the Khotanese texts in Hoernle, 1916, pp. 214-356.

Idem, *Saka Versions of the Bhadrakalpikāsūtra*, Oslo Etnografiske Museum Bulletin 5, Oslo, 1929a.

Idem, “The Dates in Saka Texts from Khotan and Tun-huang,” *Acta Orientalia* 7, 1929b, pp. 66-76.

Idem, *Saka Studies*, Oslo, 1932.

Idem, “The Late Professor Leumann’s Edition of a New Saka Text,” *NTS* 7, 1934,



pp. 5-55.

Idem, “Ein neuer Saka-Dialekt,” *SPAW*, phil.-hist. Kl., 20, Berlin, 1935, pp. 772-823 (also individual pagination).

Idem, *Khotansakische Grammatik*, Leipzig, 1941a. Idem, *A Medical Text in Khotanese*, Oslo, 1941b.

Idem, “The Oldest Dialect of Khotanese Saka,” *NTS* 14, 1947, pp. 156-90.

Idem, “Primer of Khotanese Saka: Grammatical Sketch, Chrestomathy, Vocabulary, Bibliography” [ed. G. Morgenstierne], *NTS* 15, 1949, pp. 1-136.

Hiroshi Kumamoto, “Khotanese Official Documents in the Tenth Century A.D.,” Ph.D. diss., University of Pennsylvania, 1982.

Idem, 〇〇〇〇〇〇〇〇〇〇 “Kōtan-go bunken gaisetsu” (An outline of Khotanese studies), 〇〇〇〇 *Kōza Tonkō* 6: 〇〇〇〇〇〇 *Tonkō ko-go bunken* (Dunhuang foreign language studies) II: 〇〇〇〇〇〇〇〇 *Kōtan-go bunken* (Khotanese studies), Tokyo, 1985, pp. 101-40.

Idem, “Some Problems of the Khotanese Documents,” in R. Schmitt and P. O. Skjærvø, eds., *Studia Grammatica Iranica*, Munich, 1986, pp. 227-44.

Idem, “A Sanskrit-Khotanese Conversation Manual for Central Asian Travelers” [in Japanese], in *Seinan-Ajia kenkyū: Bulletin of the Society for Western and Southern Asiatic Studies*, Kyoto University 28, 1988, pp. 53-82.

Idem, “Did Late Khotanese Have a Three Vowel System?” in B. G. Fragner et al., eds., *Proceedings of the second European Conference of Iranian Studies held in Bamberg, 30th September to 4th October 1991, by the Societas Iranologica Europaea*, Rome, 1995, pp. 383-90.

Idem, “The Khotanese Documents from the Khotan Area, with an Appendix by Saitō, Tatuya” *The Memoirs of the Toyo Bunko* 54, 1996a, pp. 27-64.

Idem, “The Khotanese in Dunhuang,” in A. Cadonna and L. Lanciotti, eds., *Cina e Iran da Alessandro Magno alla dinastia Tang*, Florence, 1996b, pp. 95-96.

A. von Le Coq, “A Short Account of the Origin, Journey, and Results of the First Royal Prussian (Second German) Expedition to Turfan in Chinese Turkistan,” *JRAS*, 1909, pp. 299-322.



E. Leumann, “Zwei orientalische Entzifferungsprobleme,” to have been published in *Abh. der Königlichen Akademie der Wissenschaften, Göttingen*, 1899, but replaced by Leumann, 1900.

Idem, “Über eine von den unbekanntem Literatursprachen Mittelasiens,” *Mémoires de l’Académie impériale des Sciences de St.-Pétersbourg*, 8th ser., Cl. hist.-phil., 4/8, 1900b, pp. 1-28.

Idem, “Zwei mittelasiatische Entzifferungsprobleme,” *Internationale Wochenschrift für Wissenschaft, Kunst und Technik* 1, 1907a, cols. 671-76, 703-10.

Idem, “Über die einheimischen Sprachen von Ostturkestan im frühern Mittelalter Erster Teil: Der Gang der Forschung und ihre augenblickliche Gesamtlage,” *ZDMG* 61, 1907b, pp. 648-58.

Idem, “Über die einheimischen Sprachen von Ostturkestan im frühern Mittelalter. Zweiter Teil: Von der arischen Textsprache,” *ZDMG* 62, 1908, pp. 83-110.

Idem, *Zur nordarischen Sprache und Literatur. Vorbemerkungen und vier Aufsätze mit Glossar*, Strassburg, 1912.

Idem, *Das nordarische Lehrgedicht des Buddhismus. Text und Übersetzung von Ernst Leumann. Aus dem Nachlass herausgegeben von Manu Leumann*, Leipzig, 1933-36.

M. Leumann, *Sakische Handschriftenproben*, Zürich, 1934.

H. Lüders, “Die Śakas und die ‘nordarische’ Sprache,” *SPAW*, phil.-hist. Kl., Berlin, 1913, pp. 406-27.

M. Maggi, “Sull’accento cotanese,” *Atti del Sodalizio Glottologico Milanese* 31, 1990 [1993], pp. 180-87.

M. Peyrot, “On the East Iranian Genitive Plural Ending,” *IJ* 61, 2018, pp. 118-30.

E. G. Pulleyblank, “The Date of the Staël-Holstein Roll,” *Asia Major*, N.S. 4/1, 1954, pp. 90-97.

L. Sander, “Brāhmī Scripts on the Eastern Silk Roads,” *Studien zur Indologie und*



Iranistik 11/12, 1986, pp. 159-92.

Idem, “Remarks on the Formal Brāhmī Script from the Southern Silk Route,” *Bulletin of the Asia Institute* 19, 2005, pp. 133-44.

R. Schmitt, “Armenia and Iran iv. Iranian Influences in Armenian Language 1. General,” *EIr.* II, London and New York, 1987, pp. 445-59.

N. Sims-Williams, review of Emmeric and Skjærvø, 1982, in *BSOAS* 46, 1983a, pp. 358-59.

Idem, “Chotano-Sogdica [I],” *BSOAS* 46, 1983b, pp. 40-51.

Idem, “Chotano-Sogdica II. Aspects of the Development of Nominal Morphology in Khotanese and Sogdian,” in G. Gnoli and A. Panaino, eds., *Proceedings of the First European Conference of Iranian Studies held in Turin, September 7th-11th, 1987 by the Societas Iranologica Europaea*, Part 1: *Old and Middle Iranian Studies*, Rome, 1990, pp. 275-96.

Idem, “On the Historic Present and Injunctive in Sogdian and Choresmian,” *MSS* 56, 1996, pp. 173-89.

Idem, “The Denominal Suffix *ant* and the Formation of the Khotanese Transitive Perfect,” in A. Lubotsky, ed., *Sound Law and Analogy: Papers in Honor of Robert S. P. Beekes on the Occasion of His 60th Birthday*, Amsterdam, 1997, pp. 317-25.

Idem, *Bactrian Documents I*, *Studies in the Khalili Collection III*, CII II/VI, Oxford, 2000.

Idem, “Khotanese *bryāñdama* ‘dearest’ and Its Indo-European Background,” *Quarterly Journal of Language and Inscription* 2.1, 2018, pp. 4-9.

P. O. Skjærvø, “The Old Khotanese Fragment H 147 NS 115 and Remarks on Old Khotanese *hamdärväto*, *patīśu*, *vya* and *ya*,” *BSOAS* 44, 1981, pp. 453-67.

Idem, “On the Tumshuqese Karmavācanā,” *JRAS*, 1987, pp. 77-90.

Idem, “Kings of Khotan in the Eighth Century,” in P. Bernard and F. Grenet, eds., *Histoire et cultes de l’Asie centrale préislamique*, Paris, 1991, pp. 255-78.

Idem, “Eastern Iranian Epic Traditions I. Siyāvaš and Kunāla,” in J. Jasanoff, H.



Craig Melchert, and L. Oliver, eds., *Mír Curad. Studies in Honor of Calvert Watkins*, Innsbruck, 1998, pp. 645-58.

Idem, “An Account Tablet from Eighth-Century Khotan,” *Bulletin of the Asia Institute* 15, 2001 [2005], pp. 1-8.

Idem, *Khotanese Manuscripts from Chinese Turkestan in The British Library: A Complete Catalogue with Texts and Translations*, with Contributions by U. Sims-Williams, Corpus Inscr. Iran. II/V, Texts VI, London, 2002.

Idem, *This Most Excellent Shine of Gold, King of Kings of Sutras: The Khotanese Suvarṇabhāsottamasūtra*, Sources of Oriental Languages and Literatures 60-61, Central Asian Sources 5-6, 2 vols., Cambridge, Mass., 2004.

Idem, “Turks and Turkic in the Khotanese Texts from Khotan and Dunhuang,” in É. Á. Csato et al. eds., *Turks and Iranians Interactions in Language and History: The Gunnar Jarring Memorial Program at the Swedish Collegium for Advanced Study*, Wiesbaden, 2016, pp.13-28.

Idem, “Khotanese Land Purchase Deeds,” in Enrico Morano, Elio Provasi, and Adriano V. Rossi, eds., *Studia Philologica Iranica: Gherardo Gnoli Memorial Volume*, Rome, 2017, pp. 455-67.

Idem, “Justice in Khotan,” in D. Gunkel, Joshua T. Katz, B. Vine, and M. Weiss, eds., *Sahasram Ati Srajas: Indo-Iranian and Indo-European Studies in Honor of Stephanie W. Jamison*, Ann Arbor and New York, 2016b, pp. 406–19.

A. von Staël-Holstein, “Tocharisch und die Sprache II,” *Bulletin de l’Académie impériale des sciences de St.-Pétersbourg*, 1908, pp. 1367-72.

M. A. Stein, *Ancient Khotan. Detailed Report of Archaeological Explorations in Chinese Turkestan*, Oxford, 2 vols., 1907.

Idem, *Serindia. Detailed Report of Explorations in Central Asia and Westernmost China*, 5 vols., Oxford, 1921.

Idem, *On Ancient Central Asian Tracks. Brief Narrative of Three Expeditions in Innermost Asia and North-Western China*, London, 1933.

Tokio Takata, “Khotano-Chinese Bilingual Phrase Books,” in idem, *A Historical Study of the Chinese Language Based on Dunhuang Materials*, Oriental Studies Library 33, Tokyo, 1988, pp. 197-227.

