Encyclopedia of Chinese Language and Linguistics

Volume 3

Men–Ser


Katia Chirkova

Northwestern Medieval Chinese

“Northwestern Medieval Chinese” (NWMC) here refers to the variety (or possibly varieties) of Chinese spoken in and around the Héxī 河西 Corridor (situated in today’s Gānsù Province) in the northwest of the Yellow River in late- and post-Táng times (roughly 9th–12th centuries CE). Connecting the Tarim Basin with Northern China, the corridor constituted an important part of the Northern Silk Route, with Dūnhuáng 敦煌 (or Shāzhōu 沙州) as its most important center. Consequently, the variety of Chinese spoken throughout this area is also known as the Shāzhōu Dialect (e.g., Coblin 1988) or alternatively Héxī Dialect (e.g., Takada 1988a).

Northwestern Medieval Chinese merited our attention for several reasons: For one, this variety of Chinese was a major player in the multilingual and -scriptural environment in Dūnhuáng and Turfan, characterized by textual remains in various languages, recorded in a multitude of different scripts (for an overview of attested combinations see, e.g., the table in Yoshida 2004:25). The sources on Northwestern Medieval Chinese as the first larger corpus of Chinese written in segmental scripts may therefore not
only complement the traditional ones in historical phonology, such as normative rime dictionaries and rime tables. Its study is also crucial, for instance, in order to understand the rationale behind phonetic loan characters frequently met with in Dúnhuáng manuscripts and thus to explain part of the textual variation observed among different witnesses of a given Chinese text.

Secondly, the variety of Chinese spoken in the Táng period capital Cháng’ān 長安 (modern-day Xi’an 西安) likewise “belonged to the great northwestern dialect”, occupying the status of “a somewhat refined version” of it in Takata’s (2004:333) words. A certain distance between these varieties of Chinese notwithstanding, Northwestern Medieval Chinese data is hence of paramount importance in the study of contemporary Chinese transcriptions of foreign languages as well as of the so-called → Chinese reading traditions of Chinese characters.

As a consequence of the nature of the sources at our disposal—often transcriptions of preexisting texts not originally written in Northwestern Medieval Chinese—grammatical and lexical features of colloquial Northwestern Medieval Chinese are comparatively difficult to recover from the texts (see especially Takata 1988a:ch. 4 for a conspectus of its grammar). It is therefore unsurprising that research in this field has been dominated by the level of phonology, which will also be the focus of the present article.

The reconstruction of the phonological features of Northwestern Medieval Chinese has focused on two periods due to the availability of sources, namely approximately the 9th to 10th centuries and the 12th century. In studies of the later period (Post-Shāzhōu in Coblin’s terms), a central role is occupied by a Tangut–Chinese bilingual glossary from the end of the 12th century, as well as Tangut transcriptions of dhāraṇī texts. Our data for the earlier period, on the other hand, chiefly derives from manuscripts among the Dúnhuáng findings containing transcriptions of Chinese in segmental scripts (chiefly Tibetan and Khotanese Brāhmi), as well as from phonetic loans observed in a variety of Chinese language manuscripts that have not necessarily been systematically recorded yet. Sources from Turfan, e.g., in the Sogdian and Uighur scripts, provide evidence for the use of Northwestern Medieval Chinese further to the northwest (also cf. Takata 2004:333).

According to Takata (1987, 1988a, 2000), there are actually two different varieties of Chinese reflected in the Dúnhuáng materials: among these, one variety is deemed to have been based on the language of Cháng’ān, which was also current in Dúnhuáng prior to the period of Tibetan rule (787–848). The other variety was Northwestern Medieval Chinese proper as spoken in medieval Dúnhuáng, which began to prosper after relations with central China had diminished in the northwestern regions and constant exchange with other areas of China had been cut off. The use of this local dialect in writing proliferated after the middle of the 9th century and in manuscript copies of the 10th century, phonetic loans betraying their Northwestern Medieval Chinese basis can be encountered more frequently (see the example of the Platform Sūtra below). During the rule of the Cáo 家 family over the quasi-independent Dúnhuáng region in the 10th century, Northwestern Medieval Chinese is thought to have acquired the role of a “standard language” in Dúnhuáng (Takata 2000).

Takata (2000) has also drawn attention to the heavy influence of Tibetan, not only during the period of the Dúnhuáng occupation, but also afterwards, especially during the 10th century when Dúnhuáng was semi-autonomous and communication to Central China reduced to a minimum. The copying of scriptures was initiated on a large scale by the Tibetans during the first half of the 9th century when bilingual Chinese-Tibetan communities were prospering. Notably, the custom of using the Tibetan script to write Chinese was also retained in the 10th century after the end of Tibetan rule.

1. History of the Study of Northwestern Medieval Chinese

The study of Northwestern Medieval Chinese began during the first decades of the 20th century, when the various expeditions to Central Asia and the subsequent discovery of countless valuable sources at Dúnhuáng, Turfan, and elsewhere provided scholars with manuscripts
featuring transcriptions of Chinese into the Khokhanese Brahmi, Manichean, Sogdian, Tibetan, and Uighur scripts. Up until that time, about the only known specimens of foreign transcriptions of T’ang period Chinese were those found in the bilingual Chinese-Tibetan treaty inscription in front of the Jokhang temple in Lhasa, dated 822. Based on rubbings obtained in Beijing in 1869, the inscription was studied and translated by Bushell (1880:535–538), who already pointed out that some Chinese names (likely following Cháng’ān rather than NWMC pronunciation) are found in transcription in the Tibetan text.

Only a few years had passed since the discoveries made at the beginning of the 20th century before a variety of Chinese exhibiting two rather distinct features was recognized in the growing corpus, even though little more than fragmentary evidence was available at the time: loss of syllable coda -ŋ and the presence of -r, in parallel to → Sino-Korean -l, where mainstream Chinese has -t. Both features were already discovered by Müller (1907) in a Sogdian fragment in Manichaean script (Berlin, Turfan Collection, M 115). Staël-Holstein (1910:140–143) adds Uighur data from Turfan on -r and somewhat later Müller (1919:4–95) himself notes that -ŋ is usually ignored in Uighur transcriptions and also gives a case of -r in a Tibetan transcription.

In the following years, some of the manuscripts brought back from Dunhuang by Aurel Stein (1862–1943) and Paul Pelliot (1878–1945) were found to contain transcriptions of Chinese texts in Tibetan script in substantial amounts, either together with the original Chinese in the form of reading glosses or in a stand-alone form consisting of the transcription alone. The significance of such transcriptions was early noted by Pelliot (1912:388–389) himself, who adduces them as evidence for the loss of final -ŋ in Northwestern Medieval Chinese. When Henri Maspero (1882–1945) wrote his study on the dialect of the T’ang capital Cháng’ān (1920), he took into account a fragment of the Qiānzìwén 千字文 together with a Tibetan transcription among Pelliot’s manuscripts (P.ch.3419), besides Chinese transcriptions of Sanskrit, as well as Sino-Japanese, Sino-Korean, and Sino-Vietnamese character readings. Before long the same manuscript was repeatedly reproduced and studied (e.g., Haneda 1923; Pelliot and Haneda 1926). In the same decade, several further transcriptional sources in the form of manuscripts of Buddhist content collected by Stein and now forming part of the India Office Library were introduced and studied by Thomas and Clauson (1926 [ms. C129], 1927 [ms. C130]); and Thomas et al. (1929 [chiefly ms. C93, but also referring to the “Long Scroll”, on which see below]) (also cf. Miyamoto 1929). It is however Luó Chángpéi (1899–1958), who is to be credited for the first monograph-length attempt at a reconstruction of Northwestern Medieval Chinese (1933), based on these sources in conjunction with a manuscript of Mǎ Rénshòu’s (৵ʠྪ) primer Kāiméng yàoxùn කႆ ࠅ ৅, containing sound glosses in Chinese (P.ch.2578; colophon dated 929). He also already took into account modern northwestern dialects for comparison. Noting that “no one has ever […] tried to reconstruct the entire phonological system”, he expressed his “desire to make a definite endeavor in this direction” (1933: viii).

The corpus was enlarged considerably in the following decades. In Paris, the existence of a significant amount of further sources among the Fonds Pelliot tibétain was noted as the cataloguing of the collection progressed. Numerous new manuscripts (P.t.1 [text 3], 448, 1228, 1230, 1238, 1239, 1253, 1256, 1258, 1262) were thus introduced and studied by Simon (1957, 1958), even before the last of the three volumes of Lalou’s catalogue (1939 [P.t.1–849], 1950 [850–1282], 1961 [1283–2216]) was published. The reproduction of a number of manuscripts in the Choix de documents tibétains conservés à la Bibliothèque nationale edited by MacDonald and Imaeda (1978, 1979) also fostered new studies, e.g., on the Tibetan translations from the Shāngrshū 尚書 and Chūnqiū hòuyǔ 春秋後語 in P.t.986 and P.t.1291 respectively, which contain numerous Chinese names in transcription (see Huáng 1981 and Coblin 1991a, 1991b on the former; the latter’s original was identified in Mǎ 1984). From among Stein’s manuscripts in London, two Tibetan-Chinese word-and-phrase books written entirely in Tibetan script were discovered and introduced by Thomas and Giles (1948) (also cf. Ligeti 1968).
Note that while most manuscripts of relevance so far were basically monolingual, we are now also facing truly bilingual ones. Together with some portions of the so-called Long Scroll (IOL C131) as christened by Simon (1958:335)—which was earlier referred to by Thomas et al. (1929), but has only been made readily available by Takata (1993) (also cf. Coblin 1995)—they are considered as specimens of a more colloquial language.

Notable other contributions to the field in the second half of the 20th century include Csongor (1960), Miller (1967), and others, but principally the various studies by Takata (1981, 1983, 1987, 1988a, 1991, 1993) and Coblin (1988, 1989, 1994). Especially important here is Takata (1988a), the most exhaustive study of Northwestern Medieval Chinese up to date—going beyond studying only its phonology and phonetics in giving due attention to grammatical matters as well. Its comprehensive treatment of all aspects of Northwestern Medieval Chinese phonology will serve as our basis below (cf. however Coblin for a second perspective, and also Emmerick and Pulleyblank 1993:57–66 for a discussion of some of Takata’s views).

Drawing upon his earlier experience with transcriptions of Chinese, Thomas (1937) (also cf. Bailey 1938; Thomas 1938) identified the language of a manuscript in Brāhmi script (IOL C134) as Chinese, the text turning out to be the Jīngāng jīng 金刚经 [Diamond Sūtra]. Numerous studies followed, such as Mizutani (1959), Zhāng (1963), Csongor (1972; but cf. already 1959) and Takata (1988a), eventually culminating in a monograph on the manuscript by Emmerick and Pulleyblank (1993). The latter were also the first to include the fragment missing at the beginning of C134, which had been discovered by Emmerick among Pelliot’s manuscripts (P.ch.5597). An important recent study is Takeuchi (2008), who provides an extensive treatment of the sound glosses in Brāhmi script added to another manuscript of the Jīngāng jīng (Peking University Library, D020), including a comparison with C134/P.ch.5597. Both sources are thought to date from the 10th century (Takata 1988a:40; Takeuchi 2008:170), although C134 may be a copy of an earlier manuscript, thus mixing two chronological layers (Csongor 1972:67–68; Coblin 1999:112).

Studies on Chinese loanwords in Uighur as well as transcriptional sources in Uighur script were carried out among others by Csongor (1952, 1954) and especially Shōgaito (1987, 1995, 1997, 2001a, 2001b, 2003). A central position here is held by the translation of Xuánzàng zhuàn 玄奘传 into Uighur, usually thought to date from the late 10th or early 11th century, but there are also numerous fragments of relevance (see, e.g., Raschmann and Takata 1993; Zieme 1996, 2012; Yoshida 2000; Shōgaito and Yakup 2001; Umemura and Zieme 2015). Another notable group of sources are Chinese reading glosses on Chinese characters, following however a Sino-Uighur reading tradition (Takata 1985, 1990; Shōgaito 1995).

Hamilton (1981) has drawn attention to two manuscripts from Dūnhuáng (P.t.1895A, P.t.1689) containing transcriptions of Chinese numerals in Sogdian script. More importantly, thanks to Yoshida (1994; also cf. 2013) we have a comprehensive treatment of both transcriptions of Chinese observed in Sogdian language texts (including those written in other scripts than Sogdian script, such as the Manichaean script) as well as of Chinese Buddhist texts in Sogdian transcription. Note however that the latter sources are from Turfan rather than Dūnhuáng (Berlin collection: So 14830 and Mainz 160, 624, assumed by Yoshida to date from the first half of the 8th century and the late 10th or 11th century respectively), which also applies to many of the Uighur sources.

Finally, a certain amount of data on the sound system of Northwestern Medieval Chinese may also be gleaned from the corpora of Northwestern Medieval Chinese-based transcriptions of foreign names and words. For one of these, consisting of Chinese transcriptions of Syriac names, see Takahashi (2008, 2013, 2014).

The following is a list of the chief sources featuring transcriptions of Chinese in Tibetan script (the majority of which are conveniently available in Takata 1988a [TT] and Zhōu and Xiè 2006 [ZX]; the sigla in quotation marks follow Csongor and Takata; those of Zhou and Xie usually differ) as well as in Brāhmi script:
Chinese texts, originals together with Tibetan transcription:

“C”—Qiānzìwén 千字文 (P.ch.3419 [P.t.1046]) [TT #1; ZX #1];
“T”—Dàoshēng zhōngzōng jiānjī 大乗中宗見解 (IOL C93 [Ch.80.xi]) [TT #4; ZX #2];
“FPa”—Miàofǎ liánhú jǐng pǔmén pǐn 妙法蓮華經普門品 (P.t.1262) [TT #10; ZX #3, A].

Chinese texts in Tibetan transcription only:

“FP”—Miàofǎ liánhú jǐng pǔmén pǐn 妙法蓮華經普門品 (P.t.1239) [TT #6; ZX #3, B];
“K”—Jīn’gāng bōrě bōluómì jīng 金剛般若波羅蜜經 (IOL C129) [TT #2; ZX #4];
“O” & “Oa”—Edēn jīng 阿彌陀經 (IOL C130 [Ch.77,ii,3]) [TT #3; ZX #5];
“TD”—Tiāndì bāyáng shénzhòu jīng 天地八陽神咒經 (P.t.1258) [TT #5; ZX #6];
“NT”—Nándiāniǎn-qùo Pūtādāmō chánshī guānmén 南天竺國菩提達摩禪師觀門 (P.t.1228) [TT #7; ZX #7];
“DA”—Dàōān fǎshī niànfó zàn 道安法師念佛絃 (P.t.1253) [TT #8; ZX #8];
“P”—Bōrě bōluómìduō xīnjīng 般若波羅蜜多心經 (P.t.448) [TT #9; ZX #9];
“HS”—Hànshí piān 寒食篇 (P.t.1230) [TT #11; ZX #10];
“ZC”—[Zháochāo 雜抄 (TT) / Sānhuáng wǔdì xìng 三皇五帝姓 (ZX)] (P.t.1238) [TT #12; ZX #11];
“99”—Jìjīù biāo 九九表 (P.t.1256) [TT #13; ZX #12];
“99”—“Long Scroll” (IOL C31 [Ch.9,ii,17]) [Takata 1993];

Tibetan texts containing transcriptions of Chinese:

Tāng-Fān huìméng běi 唐蕃會盟碑 (822) [TT #14; ZX #14];
Shāngshū 尚書 (P.t.986) [ZX #16];
Kǒngzǐ Xiàngtuó xiàngwèn shū 孔子項橐相問書 (P.t.992, P.t.1284) [ZX #17];
Chūnqiū hòuyǔ 春秋後語 (P.t.1291) [ZX #18];

[Transcribed words in various medical manuscripts] (A: IOL Tib J 756 [S.t.756]; B: P.t.1057; C: IOL Tib J 1246 [I.O.56,57]; D: P.t.127; E: P.t.1044) [ZX #15].

Chinese texts, originals together with Brāhmī transcriptions:

Jīn’gāng bōrě bōluómì jīng 金剛般若波羅蜜經 (Běijīng University Library, D020) [Takeuchi 2008].

Chinese texts in Brāhmī transcription only:

“Kbr” (also “V”—Jīn’gāng bōrě bōluómì jīng 金剛般若波羅蜜經 (IOL C134, in: Khot S 7 [Ch.00120]; P.ch.5597 [= fragment of beginning]) [Thomas 1937/1938].


A look at Northwestern Medieval Chinese from a different angle, thus complementing the transcriptional data, was made possible through the study of dialect loan characters, mainly starting with the efforts of Shào Róngfēn (1963). An ever-growing number of such cases have been identified in Dūnhuáng manuscripts (see e.g., Dèng and Róng 1999, Anderl 2012).

The tremendous progress made in the field of Tangut studies also opened up the new possibility to study a somewhat later, i.e., Sòng...
Dynasty, variety of Northwestern Medieval Chinese through Tangut-Chinese sources such as the bilingual and bисcriptual glossary Fān-Hàn héshí zhǎngzhōng zhū (The Tangut–Chinese Timely Pearl in the Palm; 1190), as was hinted at already by Hashimoto (1961). The phonological systems of the two languages reflected in this glossary were eventually reconstructed by Gōng Huángchéng (1981, 1989, 1995) and Lǐ Fànwén (1994). More recently, Sŭn Bójün (2007a, 2007b, especially 2010) has contributed significantly to our understanding of the pronunciation of both Tangut and 12th century Northwestern Medieval Chinese through the examination of dhāraṇī. Both rerenderings of preexisting Chinese transcriptions (seen through NWMC looking glasses) into Tangut, as well as new transcriptions in Chinese based on Northwestern Medieval Chinese were produced in the Tangut empire. Most of the texts are datable to the period 1140–1193.

2. Some Phonological Characteristics of 10th and 12th Century Northwestern Medieval Chinese

The system of initials in 10th century Northwestern Medieval Chinese as reconstructed by Takata (1988a:107–109; but retranscribed here into IPA) is as follows:

<table>
<thead>
<tr>
<th>stop type</th>
<th>labial, lab.dent</th>
<th>alveolar</th>
<th>alveo-palatal</th>
<th>retroflex</th>
<th>velar</th>
<th>glottal</th>
</tr>
</thead>
<tbody>
<tr>
<td>stops</td>
<td>-voice, /p/</td>
<td>/t/ /ts/</td>
<td>/ʨ/</td>
<td>/ʈʂ/</td>
<td>/k/</td>
<td>/ʔ/</td>
</tr>
<tr>
<td></td>
<td>-asp /pʰ/</td>
<td>/tʰ/ /ʦʰ/</td>
<td>/ʨʰ/</td>
<td>/ʈʂʰ/</td>
<td>/kʰ/</td>
<td></td>
</tr>
<tr>
<td>fricatives</td>
<td>-voice /f/</td>
<td>/s/</td>
<td>/ɕ/</td>
<td>/ʂ/</td>
<td>/x/</td>
<td></td>
</tr>
<tr>
<td></td>
<td>+voice /v/</td>
<td></td>
<td>/z/</td>
<td>/ʑ/</td>
<td>/j/</td>
<td></td>
</tr>
<tr>
<td>nasals /</td>
<td>/m/ /mb/</td>
<td>/n/ /nd/</td>
<td>/ȵ/ /ȵʣ/</td>
<td>/ŋ/</td>
<td></td>
<td></td>
</tr>
<tr>
<td>prenasalized stops</td>
<td>/l/</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

The most salient features of Northwestern Medieval Chinese in terms of initials undoubtedly include: (1) the labials reflect the result of labiodentalization by which /f/, /v/, and /m/ (from /p-/pʰ-, /b- and /m-) had been introduced; (2) the devoicing of voiced fricatives (z-, y- etc., also /v- < /b-) as well as stops and affricates (b-, d-, dz- etc.); (3) the so-called denasalization of what were originally pure nasals, yielding pre-nasalized voiced consonants, which in a sense fill the gap left behind by /mv- (from /m-) and /ɲʑ- (ɲ-) were already fully denasalized to /v- and /ʑ- at this stage, whereas in the remaining cases the denasalization was only partial. At least sporadically we also find cases in which nothing corresponds to original /ŋ- before -y-, e.g., yuàn 順 (cf. Late Middle Chinese [LMC] ɲyan) and near-homophones are repeatedly attested in Tibetan transcriptions as wen (Takata 1988a:31, 370–371).

Note regarding (2) that Takata posits two competing dialects: one in which the old voiced stops and affricates had merged with their voiceless aspirated counterparts regardless of tone (e.g., /b- > /pʰ-; reflected in “T”, “Kbr”—and now also D020, cf. Takeuchi 2008)—and another in which they had merged with their voiceless unaspirated counterparts (cf. Emmerick and Pulleyblank 1993:61–62 for an alternative view).

The system of initials in 12th century Northwestern Medieval Chinese as reconstructed by both Gōng (1981) and Lǐ (1994) is similar in many
respects. In fact, this variety shows both the development of \( b, \) \( d \)- etc. \( \rightarrow p^h, \) \( th\)-, etc. regardless of tone as in one of the two varieties in the 10th century (as also noted by Takata 2013:101–102) as well as the loss of original \( p\)- before \( y\)- and \( w\)- (see Gōng 1981:73–74 for examples). The most notable difference is the non-distinction of \( tc\)-, \( tw\)-, \( c\)- vs. \( ts\)-, \( ts\)-, \( s\)-, leading to only a single series of shibilants. Also note that Lī (1994:261–262) has \( ν\)- where Gōng (1981:70–71) assumes \( w\)-. No separate initial \( ?\)- is reconstructed anymore, but notably some syllables with zero initial from former \( ?\)- appear to have acquired a new (non-phonemic) onset: either \( [ŋ\-] \) or \( [ɣ\-] \) in Gōng’s (1981:60–61) view. This may also explain the aberrant use, e.g., of \( è\)  backButton for LMC \( ?at\) for Skt. \( ga\) (Sūn 2007a:317–318).

As for the finals, we observe sharp differences between the varieties of Northwestern Medieval Chinese dating from the 10th and 12th centuries respectively. We will concentrate on some noteworthy developments here (Takata 1988a; Coblin 1988; Gōng 1989, 1995; Lī 1994 may be consulted for a fuller picture). Among the first characteristics of 10th century Northwestern Medieval Chinese that were noticed around 1900 is the loss of final \( -ŋ\) in certain rime groups together with the lenition of original \( -t\), which almost universally appears as \( -r\). The former phenomenon apparently started from the traditional rime groups \( dâŋ\) backButton and \( gêng\) (Taka backButton—a where it is especially salient. Next was the \( tông\) backButton group, Takata’s \( *(i)\-\text{ŋ} \), which still has \( -ŋ\) according to the Tibetan transcriptions, but frequently drops the coda consonant in renderings in Khotanese Brāhmī script (Takata 1988a:178, especially 1988b; Takeuchi 2008:183–184). Note that in the latter case, \( anusvāra\) \( <m\) is often, although not consistently, used in all three groups mentioned so far to indicate nasализation of the preceding vowel, i.e., here \( -u\) from \( -uy\) etc. (Nasalized allophones of vowels in syllables with nasal finals are likewise indicated by the presence of \( anusvāra\).) In sharp contrast to this, \( -ŋ\) is retained in the rime groups \( jiâng\) backButton, \( -\text{ŋ} \), as well as \( zēng\) backButton, \( *(i\-\text{ŋ})\-\text{ŋ}.\)

The remaining nasal finals \( -m\) and \( -n\), as well as the earlier stop finals \( -p\) and \( -k\) are generally preserved in Tibetan transcriptions. Brāhmī renderings again differ insofar as they show signs of lenition in the form of fricative articulation for the other stops as well, frequently rendering \( -k\) as \( -hā\) and sometimes also \( -p\) as \( -hvā\) (Takata 1988b; Emmerick and Pulleyblank 1993:42). Further evidence for lenition may be gleaned from Uighur transcriptions, which usually have \( -p\)-, \( -r\) and \( -k\-\text{q/-q̈} \), the latter being interpreted as \( [y]\) (see, e.g., Shōgaito 1987:78). Note that the order in which lenition occurred in Northwestern Medieval Chinese is thus \( -t\-\rightarrow -k\-\rightarrow -p\).

Finally, the sources in Brāhmī script provide us with evidence for the development of an “apical vowel” or rather syllabic \( [z]\) in place of \( [i]\) after (non-retroflex) sibilant initials. Compare renderings involving \( <\text{ys}>\) \( [z]\), such as \( siysi\), \( siysâ\) for \( si\) \( \text{四}\) or \( tsisy\) for \( ci\) \( \text{次}\) (cf. Takata 1988a:129–131; Emmerick and Pulleyblank 1993:45, 48–49). This is reminiscent of Korean transcriptions of Chinese from the 15th century onwards, rendering e.g., \( si\) \( \text{四}\) as \( suz\) \( \frac{\Delta}{\Delta}\) \([\text{siz}]\).

A detailed picture of the tonal system of Northwestern Medieval Chinese is lacking (cf. Takata 1988a:182–185), but it appears that the traditional \( shâng\) \( \text{上}\) and \( qù\) \( \text{去}\) tones had merged after voiced obstruent initials, as already pointed out by Shào (1963 [2009:260–264]).

In the 12th-century variety of Northwestern Medieval Chinese, the final stage in the reduction of coda consonants is reached. Thus, all former oral stop codas are lost entirely (Gōng 1989 [2002:285–296]), or at least had merged into \( -ʔ\) (Lī 1994:325–326). Concerning the nasal codas, several diverging views have been presented so far: Gōng (1989 [2002:296–323]) holds not only that all velar nasal codas are lost as segments—leading to open syllables, at least part of which he reconstructs with nasalized vowels \( -(\text{u}<\text{ŋ}, -\text{d}<\text{ŋ}\text{ et c.}, \text{but, e.g., } -\text{jij}<\text{-iey})\) — but also that former \( -n\) vs. \( -m\) are both lost as segments, generally leaving a trace in the form of nasalization of the preceding vowel. Lī (1994:329–332) on the other hand mostly has \( -n\) for older \( -m\), whereas for older \( -n\) and \( -ŋ\) he sometimes reconstructs them still as \( -n/-ŋ\), but mostly as nasal vowels; nasality is retained without exception here. Transcriptions of \( dhârânî\) may shed some light on the issue—however,
even if former -n is for instance found to render Sanskrit -n and -ṅ (Sün 2007a:313, 2007b:18) phonetic details beyond the retention of nasality in some form or another are difficult to specify.

3. Dialect Loans as a Source for the Study of Northwestern Medieval Chinese

Besides the various types of transcriptions of Chinese in non-Chinese scripts, phonetic loans as attested in a large variety of Dunhuang materials (usually “non-canonical” writings such as treatises, sermon transcripts, popular narratives, etc.) are another important witness of Northwestern Medieval Chinese. Consider for instance the numerous textual variants observed between the different Dunhuang manuscripts of the well-known Liù-zǔ tánjīng [Platform Sūtra of the Sixth Patriarch] as assembled in the appendix to Anderl (2012), some of which clearly have a Northwestern Medieval Chinese background. While all the manuscripts show both “regular” loans and dialect-based variations, they are most commonly seen in the 10th century Stein copy (Or.8210/S.5475; abbreviated as “S.” in the following, whereas “D.” refers to Dunhuang Museum, Dunbó 77). Characteristic types of loans (given below in the form “loan character → conventional character”) with parallels in the transcriptional sources include for instance the following. (The reconstructed NWMC forms in parentheses are Takata’s who uses superscript and subscript numbers to indicate the traditional four tones in the yín 陰 and yáng 阳 registers respectively, with subscript “2–3” referring to the merged shàng and qù tones after voiced obstruent initials; cf. Takata 1988a:303.)

a. What were originally open syllables and ones with final -ŋ are frequently equated, e.g., ti 髀 (*tʰiɛi1) → tīng 聽 (*tʰiɛi1) [S], ding 定 (*tiei1) → dì 弟 (*tiei2–3) [S], lì 禮 (*tiei2) → līng 令 (*liɛŋ3) [S] etc.;

b. *[i] and *[y] are commonly confounded (also cf. Takata 1988a:18): yì 義 (*ŋgi1) → yū 語 (*ŋgy2) [S] and yū 語 → yī 議 (*ŋgi3) [D, S], qī 起 (*kʰiɛŋ) → qū 去 (*kʰiɛŋ) [S] and qū 去 → qī 起 [D, S]. zhī 之 (*tʰiɛŋ) → zhū 諸 (*tʰiɛŋ) [S] etc.; and
c. *[iu] and *[ʷi] are sometimes confounded (also cf. Takata 1988a:131–132): xuī 須 (*siu1) → suī 雖 (*sʰiɛŋ) [D, S] and suī 雖 → xuī 須 [D] etc.

The abundant use of such loans especially in manuscripts dating from (late-)Táng times testifies to the important role “orality” plays here. Recording the vocalization of a text—note that the Platform Sūtra was originally the transcript of a sermon—was apparently more important for some scribes than standard orthographic usage of Chinese characters, whereas in other cases the lack of adequate training or the need to transcribe quickly rather than accurately may have yielded the same results.

Bibliography


Deng Wěnkùan 聶文範 and Róng Xīnkuān 融新闗, Dūnhuáng běn Chánjí lùjiào 南京 《敦煌本禅籍路教》 [Chapter scriptures preserved at the Dūnhuáng Museum, collected and collated], Nánjīng 南京: Jiāngsā guǐ 江蘇古籍出版社, 1999.


Haneda Tōru 羽田透, "Kan-Ban taisen Senjimono-no dankan 漢蕃對音千字文的斷簡" [A fragment of the Qianzwen with Tibetan transcriptions], Tōyō gakuhō 東洋學報 13/3, 1923, 84–104.


Mǎ Míngdá 马鸣达, "P·T·1291-hào Dūnhuáng zàngwén wěnshǔ yǐjiě dingwǔ P·T·1291 那煌藏文文书解析订误" [A revision of the annotated translation of the Tibetan document P·T·1291 from Dūnhuáng], Dūnhuángxué jìkān 敦煌学辑刊 6, 1984, 14–24.


Takeuchi Yasunori 武内庸則, “Burafumou moji-de onchū-o fushita kanbun kyōten-ni tsuite: Hokudai 他山の研究成果について通著-の分類について”。


Takeuchi Yasunori 武内庸則, “Burafumou moji-de onchū-o fushita kanbun kyōten-ni tsuite: Hokudai
Notions of “Chinese”

In general usage, the word “Chinese” has many different meanings. In a general sense, it can be used to refer to the territories, peoples, and languages of the People's Republic of China or of other states that have controlled all or parts of China. More specifically, it is sometimes used to refer to people who today identify as part of the Han 漢 ethnic group or who are retroactively identified with that group, to any or all of the Sinitic languages, past or present (that is, languages belonging to the grouping composed of the Sinitic languages, past or present (that is, those with a part or any of the Sinitic languages, past or present (that is, those with a part or any of the Sinitic languages, past or present (that is, those with a part or any of the Sinitic languages, past or present (that is, languages belonging to the grouping composed in part of Old Chinese and all its descendents), to the sinographic writing system, and to Modern Standard Mandarin in particular, among other possible references (+Names for the Chinese Language). In addition to the above, the English word “Chinese” is used as a translation for a number of different contemporary Sinic terms, which include Hanyu 漢語 ‘Han language’, Hanzu 漢族 ‘Han people’, Zhongguo 中国ese Language). In addition to the above, the English word “Chinese” is used as a translation for a number of different contemporary Sinic terms, which include Hanyu 漢語 ‘Han language’, Hanzu 漢族 ‘Han people’, Zhongguo 中国

Christopher Anderl & Sven Osterkamp