How to reconstruct Old Chinese

André-Georges Haudricourt

Non-final version (November 9th, 2017). In preparation for: Haudricourt, André-Georges. Evolution of languages and techniques*. (Ed.) Martine Mazaudon, Boyd Michailovsky & Alexis Michaud. (Trends in Linguistics. Studies and Monographs [TiLSM] 270). Berlin: De Gruyter Mouton.*

How to reconstruct Old Chinese (1954)

Originally published as: Comment reconstruire le chinois archaïque, *Word* 10(2/3). 351–364 (1954). Reprinted (with additions) [1] in *Problèmes de phonologie diachronique*: 161-182.

translated by Guillaume Jacques

Abstract

[The reconstruction of Chinese raises special difficulties due to the nature of the writing system, which only provides indirect evidence concerning pronunciation. The reconstructions proposed by Karlgren and Maspero are revisited systematically. Comparative data from other East Asian languages are adduced, with special attention to early loans from Chinese into Vietnamese and Tai languages, which provide decisive evidence on Old Chinese, i.e. Chinese as it was spoken before the Tang dynasty.

The paper argues against reconstructing a contrast between two series of voiced stops (plain vs. aspirated) in Old Chinese. It provides evidence for the reconstruction of a labiovelar series in Old Chinese, and, taking as a model the development of tonal oppositions from syllable finals in Vietnamese, proposes to reconstruct an Old Chinese derivational suffix *s to account for a series of tonal alternations in Middle Chinese.]

[1. Introduction]

[1.1 Historical linguistics in East Asian and European languages]

The linguistic history of Chinese bears some analogy to that of Romance: Mandarin, Cantonese, Wu (from the lower reaches of the Yangtze River) and Min (from Fujian) correspond to the Romance languages; Classical Chinese, the written language of Korea, Japan and Vietnam, played the

1. [The 1972 reprint includes handwritten Chinese characters for all examples in the text (except in §1.2). These had been omitted in the original publication.]

same role as Medieval Latin, the written language of Germanic, Hungarian and West Slavic countries; early loanwords from Chinese into Miao [Hmong], Thai, Vietnamese and Khmer correspond to the ancient Latin loanwords in Germanic, Albanian, Basque and Berber.[2]

There are nonetheless two important differences. First, Latin, an inflectional language, can be classified in a straightforward way as a member of the Indo-European family, to which Germanic, Slavic and Albanian also belong but Basque and Berber do not. But Chinese, a language with invariable words [i.e. lacking inflectional morphology], is more difficult to classify: it belongs to the Tibeto-Burman group, but is it as close to the Tai languages[3] as people have long believed? In this type of language, relatedness can only be assessed on the basis of vocabulary, and the number of Chinese loanwords has often been underestimated.[4] In addition, the concept of "mixed languages" was accepted: when language C had vocabulary in common both with language A and with language B, it was said that language C originated in a mixture of a dialect of A and a dialect of B.[5]

It seems now that the best hypothesis is that the languages A, B, and C all belong to the same language family. In East Asia, languages would all be linked in a chain as one large family, Tibeto-Burman and Chinese being related to Miao [Hmong] and Austroasiatic, these two to Tai, and the latter to Malayo-Polynesian.

The second difference lies in the writing system: Latin is written alphabetically, and its spelling encapsulates phonetic information that was passed down unchanged through the ages, while the early loanwords from

2. The best recent overview on this subject is Forrest (1948).
3. [In the translation, the national language of Thailand is referred to as "Thai", and the name "Tai" refers to what Haudricourt called "Common Thai", or "Thai proper" (Haudricourt 1956, 1970, in this volume), a group comprising Thai, Lao, Shan, etc., but excluding the Zhuang (a.k.a. Dioi) group of languages. It is a lower level group than what comes to mind nowadays when using the terms Tai or Proto-Tai.]
4. [Haudricourt here questions the view that Tai and Chinese constitute a sub-family within Tibeto-Burman (or "Sino-Tibetan") (Maspero 1920:22n1). Haudricourt's proposed chain of relations (next paragraph) excludes Tai from Tibeto-Burman entirely, following Benedict (1942); this view has gained widespread acceptance, although the Tai-Chinese genetic relationship is still defended by some linguists, especially in China (Luo 2008).]
5. A case in point is H. Maspero's statement (1912: 117) that "Proto-Vietnamese was born from the fusion of a Mon-Khmer dialect, a Tai dialect, and maybe even a third, yet unknown, language."

Latin followed the considerable phonetic evolution of the languages of Europe. These early loanwords are therefore quite eroded in comparison to recent loanwords that retain the Latin graphic form. Chinese, on the other hand, is written ideographically, and hence defenceless against phonetic evolution. The official language as it was spoken varied from century to century, either abruptly – as when the location of the capital city changed and a different dialect came to be looked up to as the standard – or gradually, following the phonetic evolution of the language. In fact, since the capital city was located in the North, in an area which was subjected to invasions by Turks, Mongols, Tungus and Manchus, who founded dynasties and became assimilated after a period of bilingualism, the language of the capital city evolved much more rapidly than the languages of Southern China and Southeast Asia. The vulgar pronunciation of Chinese words from the oldest layers of borrowings preserves the original pronunciation better than the more recent learned borrowings as represented by the official readings of the ideographic characters. Analogous cases are exceptional in Europe, *e.g.* in German, where the word *Kaiser* [ˈkʰaɪzɐ] 'emperor', an early borrowing from Latin *Caesar*, preserves the ancient pronunciation better than the proper name *Cäsar* [ˈtsɛzɐ, a later, learned borrowing of the same word].

[1.2 Resources for the reconstruction of Chinese]

For the reconstruction of the pronunciation of earlier stages of Chinese, we first have at our disposal the tools of comparative linguistics: comparing Chinese dialects with one another, using early Chinese loanwords into other languages, and assessing to what extent the reconstructions are in keeping with the facts observed in neighbouring languages that are genealogically related to Chinese.

[Another resource is the analysis of the writing system.] Most characters are phonograms, which is to say that they work as rebuses. For instance, the ideogram for the word **pa** 巴[6] *bā* 'boa' is used as a phonetic component in the characters 把 *bǎ* 'to grasp', 靶 *bǎ* 'target', 耙 *bà* 'harrow', 疤 *bā* 'scar', etc. However, this writing system goes back to nearly two thousand BC, and some words whose pronunciation used to be similar at that period are now very different. If we could distinguish with confidence between

6. [See the Appendix for a summary of Pinyin and other transcriptions used in the translation.]

phonogram and ideogram, this would provide valuable clues concerning the original pronunciation. Unfortunately, this is not the case. For instance, the character 肥 *féi* 'fat' contains the ideogram for 'boa'; it is difficult to know whether this indicates the idea of fatness that this kind of snake evokes, the pronunciation of a word beginning by a bilabial followed by an open vowel, or both.

It is apparently only in the third century AD that people started to compare the pronunciation of characters by means of the method known as *fǎnqiè* (反切): the pronunciation of a character is indicated by two others, the first having the same initial consonant as the target character, and the second the same rhyme. For example, the pronunciation of 'boa', **pa**, is represented by characters pronounced **pǎk** and **ka**.[7] Basing themselves on *fǎnqiè*, on Japanese and Vietnamese readings of Chinese characters,[8] and especially on a rhyme dictionary, compiled in 601, but which until recently was known only in a 1008 edition,[9] Henri Maspero and Bernhard Karlgren attempted to reconstruct the pronunciation of Middle Chinese, the language of the Tang period [7th to 9th centuries].[10] Maspero and Karlgren were not in agreement on everything, but Karlgren's reconstruction is the more advanced and the more systematic of the two, and it is the one which has been adopted by a majority of sinologists.[11] Moreover, going backwards in time from this Middle Chinese reconstruction and using the phonograms and the rhymes in the *Book of Odes* [11th–8th centuries BC], Karlgren also

7. [The *fǎnqiè* formula for this character is 巴伯加切 |boa uncle add cut|, which translates as: "'boa' cuts up into 'uncle' and 'add'", i.e. "its pronunciation is composed of the initial of 'uncle' and the rhyme of 'add'".]
8. [Local systems of pronouncing Chinese characters became established in medieval times in neighboring countries which adopted Chinese writing; these readings, called Sino-Japanese, Sino-Vietnamese, etc., evolved independently of the pronunciations in Chinese, and are used in reconstructing older Chinese pronunciation.]
9. An edition of this dictionary [the *Qièyùn*] dated 706 has been discovered in Peking in recent years; an article by Wang Lianceng 王聯曾 on this topic is in press in *T'oung-Pao* 通报 [this article was published several years after Haudricourt's article as Wang 1957].
10. [Karlgren called this language, of the Sui (581-618) and Tang (681-907) periods, Ancient Chinese.]
11. Maspero's two works on Middle Chinese are Maspero (1916) and Maspero (1920). Those by Karlgren on the same topic are Karlgren (1915–1926) and Karlgren's *Analytic Dictionary* (1923). In the latter work the characters are numbered; we cite these numbers as "AA...". [This dictionary is arranged by phonetic series: e.g. the series 'boa' cited above is AA683.]

attempted to reconstruct the pronunciation of the epoch of Confucius, the Zhou dynasty [11th to 3rd centuries BC].[12] Karlgren's two reconstructions are linked; we will examine first the initials and then the finals.

[2. Initials]

[2.1 Arguments for Maspero's three OC series vs. Karlgren's four]

Maspero and Karlgren agree in reconstructing three series of stops in Middle Chinese: a series of plain voiceless stops, *p, t, k*, a series of voiceless aspirated stops, *ph, th, kh*, and a third series, on whose phonetic reconstruction they do not agree. For Maspero these were plain voiced stops, *b, d, g*; for Karlgren they were aspirated voiced stops, *bh, dh, gh*.
The first two series are preserved as such in all dialects. The third remained distinct only in Wu dialects as stops with voiced aspiration. Maspero showed, using Chinese transcriptions of Sanskrit words, that the voiced stops became aspirated only toward the end of the Tang dynasty (8th–10th centuries), and that plain voiced stops have to be reconstructed before the aspirated stage (Maspero 1920: 27). A change from voiced stops to aspirated voiceless stops is also found in other languages of Southeast Asia: in Thai[13] and Karen (Haudricourt 1946, 1953 [this volume]), voiced stops became plain or aspirated voiceless stops depending on the dialect.
One might argue that the divergence between Maspero and Karlgren is negligible, insofar as, in the following set of correlations :

p	*t*	*k*
ph	*th*	*kh*
b(h)	*d(h)*	*g(h)*

12. The reconstruction of all characters of this period is found in Karlgren's *Grammatica Serica* (1940). We refer to the numbering in this publication as "GS". [This reconstructed stage, which Karlgren, Maspero, and Haudricourt call "Archaic Chinese", is referred to as Old Chinese (OC) in the present translation.]
13. It is proven in Thai by the use of an Indic alphabet, where the creators of the script could choose between plain voiced and aspirated voiced stops. [The unaspirated voiced letters *b*, *d* were chosen to represent the Old Thai *voiced series. See Haudricourt 1948, 1949a.]

only voicing is distinctive for the third series, so that the choice of either interpretation (plain voiced or voiced aspirated) would only be a difference of opinion on the realization of these phonemes.

But in fact, this detail is actually of great importance in Karlgren's reconstruction. Karlgren was ahead of his time in adopting a phonological approach and viewing the absence of plain voiced stops in his table as a structural gap that needed to be filled.[14] He reconstructs the following symmetrical system for the Zhou period:

***p**	***t**	***k**
***ph**	***th**	***kh**
***b**	***d**	***g**
***bh**	***dh**	***gh**

Karlgren believes that in the Tang period, ***b *d *g** were lost and ***bh *dh *gh** were preserved.[15] The reconstruction of this additional voiced series is motivated by the fact that some characters for words without initial stops are used as phonetic components to transcribe words that have initial stops, as in the examples in Table 1a.[16]

14. On the notion of structural gap-filling, see Martinet (1939, 1952).
15. To begin with, we can remark how unlikely it is, from the phonetic point of view, that aspirated voiced stops should be more resistant to sound change than plain voiced ones.
16. We follow Karlgren's transcriptions, with the aspiration sign (‘) replaced by h. [Haudricourt's original note goes on to explain and disambiguate the different meanings of apostrophe in his original transcriptions; these have all been changed in the present translation. In particular, Karlgren's half-circle notations for the tones *píng*, *shăng*, and *qù* have been restored; Haudricourt used opening and closing single quotation marks for the latter two, and left the *píng* tone unmarked (see Appendix). Haudricourt's remarks on Vietnamese transcription have been moved from the present note to a note introducing Table 2.]

[Table 1a. Characters for words without an initial stop used as phonetics in characters for words with an initial stop.[17]]

AA	character	K-MC	[B-MC]	phonetic in:	K-MC	[B-MC]
264	用 *yòng*	i̯ʷongˀ	yowngH	筩 *tǒng*	ᶜdhung	duwng
				通 *tōng*	ᶜthung	thuwng
				桶 *tǒng*	ʿthung	thuwngX
				誦 *sòng*	zi̯ʷongˀ	zjowngH
292	匀 *yún*	ᶜi̯uěn	ywin	均 *jūn*	ᶜki̯uěn	kjwin
249	又 *yòu*	ji̯ə̯uˀ	hjuwH	求 *qiú*	ᶜghji̯ə̯u	gjuw
				灰 *huī*	ᶜxuâi	xwoj
239	炎 *yán*	ᶜji̯äm	hjem	談 *tán*	ᶜdhâm	dam
				毯 *tǎn*	ʿthâm	thamX
184	台 *tái*	ᶜi	yi	治 *zhì*	dʹhiˀ	driH
				治 *zhì*[18]	ᶜdʹhi	dri
				始 *shǐ*	ʿśi	syiX
				胎 *tāi*	ᶜthâi	thoj

Conversely, some characters for words with a stop initial are used as phonetics for words without an initial stop, as in Table 1b.

[Table 1b. Characters for words with an initial stop used as phonetics in characters for words without an initial stop.]

AA	character:	K-MC	[B-MC]	phonetic in:	[gloss]	K-MC	[B-MC]
425	谷 *gǔ* ‘valley’	kuk	kuwk	欲 *yù*	to desire	i̯ʷok	yowk
				裕 *yù*	opulent	i̯uˀ	yuH
1006	多 *duō* ‘many’	ᶜtâ	ta	爹 *diē*	dad	ᶜtʹi̯a	trjæ
				移 *yí*	to move	ᶜie̯	ye

17. [“AA” numbers in the Table refer to articles in Karlgren’s *Analytic Dictionary* (1923). Each article typically contains a “phonetic series” consisting of a “title character” and the list of the compound characters in which the title character has been incorporated as a “phonetic”, that is, as an indication of pronunciation. For each character cited, the dictionary gives a gloss, modern Mandarin and Cantonese pronunciations, and Karlgren’s Middle Chinese transcription (“K-MC”). We have added Baxter’s (1992) Middle Chinese transcriptions (“B-MC”).]
18. [This Chinese character had two readings: ‘to heal’ (B-MC *driH*), and ‘to govern’ (B-MC *dri*). In Modern Chinese, the second reading has become generalized.]

Karlgren (1923: 18-27) concludes that this "zero" consonant points to the loss of a voiced stop, ***g** or ***d**, so that the pre-Tang dynasty forms should be reconstructed as OC ***di̯ʷong** for MC *i̯ʷong* (AA 264 用 *yòng*), OC ***gi̯uĕn** for MC *ɦi̯uĕn* (AA 292 匀 *yún*), etc., yielding a better fit of the relationship of the phonetic element to the characters in which it is used. The examples in Table 1a would then read for Karlgren:

AA 264: ***di̯ʷong** used as phonetic in *dhung*, *thung*, *zi̯ʷong*
AA 292: ***gi̯uĕ** used as phonetic in *ki̯uĕn*
etc.

Concerning affricates, Karlgren asserts that *z* comes from ***dz**, and *ź* [ʑ] from ***dź** [dʑ], which is plausible. However, he proceeds to deduce from this the existence of unaspirated voiced phonemes ***dz** and ***dź** distinct from aspirated voiced ***dzh** and ***dźh**; and he reconstructs earlier ***z** and ***ź** for the "zero" initial of some characters used as phonetic elements. So, in the Tang period, the relationship of the pronunciation of the phonetic element to the character in which it is used would be as in Table 1c [a "zero-initial" phonetic used to write a word with a sibilant initial], but at a pre-Tang period it would be as shown in Table 1d.

[Table 1c. Tang period: Characters for words with "zero-initial" in the Tang period used as phonetics in words with affricated or fricative initials.]

AA	character	K-MC	[B-MC]	phonetic in	K-MC	[B-MC]
211	羊 *yáng* 'sheep'	ɦi̯ang	yang	祥 *xiáng*	zi̯ang	zjang
187	亦 *yì* 'also'	i̯äk	yek	跡 *jì*	tsi̯äk	tsjek
				夜 *yè*	i̯aʾ	yæH

[Table 1d. Karlgren's pre-Tang reconstructions of the examples of Table 1c.]

AA	character	K-OC	phonetic in	[gloss]	K-OC
211	羊 *yáng* 'sheep'	zi̯ang	祥 *xiáng*	auspicious	dzi̯ang
187	亦 *yì* 'also'	zi̯äk	跡 *jì*	trace	tsi̯äk
			夜 *yè*	night	zi̯ag

To question, following Maspero, the antiquity of the voiced aspirate series *bh*, *dh*, *gh*, *dzh*, *dźh*... [and hence the existence of an opposition between two voiced series], we must reexamine the entire reconstruction.

Karlgren worked in a rigorous way, basing himself on the postulates that phonetic laws are fully regular and that structural gaps tend to become filled; he treated the data as if the Chinese language had had a linear development, free from any dialectal influences, over two or three millennia.

[2.1.1 Absence of a contrast voiced fricatives vs. voiced affricates]

Let us first examine the distinction between voiced affricates and voiced fricatives. Based on the initials found in the dictionaries, Karlgren distinguishes [in Middle Chinese] *dzh* and *z*, *dźh* and *ź* [dzʱ and z, dʑʱ and ʑ], but he does not recognize the existence of a voiced retroflex fricative *ẓ* [ʐ], only the voiced aspirated affricate *ḍẓh* [ɖʐʱ].

[2.1.1.1 Karlgren's MC retroflex affricate ḍẓh]

Table 2 shows the reflexes of Karlgren's MC initial *ḍẓh* [ɖʐʱ] in Beijing Mandarin, in Cantonese, and in Sino-Vietnamese.[19,20]

19. In Vietnamese we follow the official orthography (as does Emeneau 1951). Orthographic *ư* and *ơ* (with the "horn") represent central vowels: ɯ and ə, respectively. Note that *b* and *đ* are preglottalized consonants [IPA ɓ , ɗ] (see Haudricourt 1950), while *d* and *g* are fricatives [ð, ɣ].
20. [Sino-Vietnamese refers to the reading of Chinese characters by officials and literati in Vietnam in the second millenium. When the Red River delta acquired political independence from China in the 10th century, becoming the Đại Việt, the variety of Middle Chinese spoken in the area by Chinese settlers, and also by an important fraction of the city dwellers, became isolated. Its entire phonological system was eventually absorbed into the Austroasiatic (Viet-Muong) ancestor of Vietnamese, spoken in the Red River delta (Nguyễn Tài Cẩn 1979; Ferlus 1992: 114). This is how Vietnamese came to encapsulate a complete set of readings for Chinese characters. Vietnamese and Sino-Vietnamese words are transcribed using the same system because the Sino-Vietnamese layer is now an integral part of the language and has undergone the same phonetic changes as the rest of the system since the 10th century AD. The IPA equivalents added in this translation follow Ferlus's analysis of rhymes, set out in the appendix to "The short vowels of Vietnamese" (this volume), and Haudricourt's analysis of initials as set out in Haudricourt (1949b). Note that this constitutes an interpretation of the state of the language in the 17th century, as reflected in Rhodes (1651); it is close to the present-day dialects, but none of the modern dialects makes all the distinctions of written Vietnamese.]

[Table 2. Reflexes in Mandarin (Mand.), Sino-Vietnamese (SV) and Cantonese (Cant.) of Karlgren's MC words with a retroflex affricate initial *dẓʰ*.]

AA	[gloss]		[Pinyin]	K-MC	[B-MC]	SV	Cant.	Mandarin
1096	firewood	柴	*chái*	˂dẓhai	dzrɛɨ	*sài*	ṣāi	tʂʰai
1024	jackal	豺	*chái*	˂dẓhai	dzrɛj	*sài*	tshāi	tʂʰai
1170	slander	讒	*chán*	˂dẓham	dzrɛm	*sàm*	tshām	tʂʰan
1259	weak, unfit	孱	*chán*	˂dẓhi̯än	dzrɛn	*sàn*	ṣān	tʂʰan~tʂʰuan
1062f	bed	床	*chuáng*	˂dẓhi̯ang	dzrjang	*sàng*	tshōng	tʂʰuaŋ
386	peak	岑	*cén*	˂dẓhi̯əm	dzrim	*sầm*	ṣăm	tsʰən[21]
1088	to worry	愁	*chóu*	˂dẓhi̯o̯u	dzrjuw	*sầu, xầu*	ṣău	tʂʰou
1114	to worship	崇	*chóng*	˂dẓhi̯ung	dzrjuwng	*sùng*	ṣung	tʂʰoŋ
834	food	饌	*zhuàn*	dẓhwan˃	dzrjwenH	*soạn*	tsān	tʂuan
877	scholar	士	*shì*	ˁdẓhi	dzriX	*sĩ*	ṣi	ʂʅ
889	to serve	事	*shì*	dẓhi˃	dzriH	*sự*	ṣi	ʂʅ

We observe that the reflex of this initial is regular in Vietnamese, where we find a fricative *s* [ʂ], and almost regular in the Beijing dialect, where we find an affricate [Pinyin *ch* [tʂʰ] or *zh* [tʂ]], with two exceptions [Pinyin *sh* [ʂ]]. The Mandarin affricate is aspirated on the *píng* tone [marked as "˂"] and unaspirated on the other tones [cf. **tʂuan** 'food'], which is the general rule for Mandarin reflexes of Middle Chinese voiced stops. In Cantonese, on the other hand, we find two reflexes, as if there had existed a MC *ẓ* distinct from *dẓh*. But the phonological contrast between the first two words, Cantonese **ṣāi** and **tshāi** is obviously not enough to prove that the initials of these two etyma have always been distinct in Chinese. In fact, written evidence shows that this distinction did not exist in the Tang period, and that the variable reflexes in Table 2 are simply due to borrowing between dialects: the indigenous evolution in Cantonese was *dẓh* > **ṣ**, as in 'firewood'; words like **tshāi** 'jackal' were borrowed from the dialects of the central or northern capital cities, where the evolution had been *dẓh* > **tsh**.

[2.1.1.2 Karlgren's MC palatal affricates and fricatives]

Should the *dźh* vs. *ź* contrast of the Tang period be projected back to an earlier stage? This is doubtful, as this contrast is not found in modern dialects, as shown in Table 3. [22]

21. [Haudricourt gives the pronunciation tʂʰən for this word.]

22. [Baxter's transcriptions of K-MC *ź* as an affricate and *dźh* as a fricative in the table reflect the theory, advanced by Lu 1947 and now widely accepted, that the phonetic values of these initial categories were inverted in traditional Chinese historical phonology, as followed by Karlgren and Haudricourt. The error apparently originated in the late Middle Chinese rhyme tables (see Baxter 1992: 52-53).]

[Table 3. Reflexes in Beijing Mandarin, Cantonese, and Sino-Vietnamese (SV) of MC *ź* and *dźh*.]

AA	[character]	[meaning]	K-MC	[B-MC]	SV	Cantonese	Mandarin
856bd	常裳 *cháng*	often; skirt	꜀źi̯ang	dzyang	*thường* [t^{h}ɯəŋ^{2}]	ṣöng [ʃœŋ^{11}]	tʂ^{h}aŋ
857b	償 *cháng*	to compensate	꜀źi̯ang	dzyang	*thường* [t^{h}ɯəŋ^{2}]	ṣöng [ʃœŋ^{11}]	tʂ^{h}aŋ
1196	晨 *chén*	morning	꜀źi̯ĕn	dzyin	*thần* [t^{h}ɤ̆n^{2}]	ṣăn [ʃɐn^{11}]	tʂ^{h}ən
1203	丞 *chéng*	assistant	꜀źi̯əng	dzying	*thừa* [t^{h}ɯə^{2}], *thằng* [t^{h}ăŋ^{2}]	ṣing [ʃiŋ^{11}]	tʂ^{h}əŋ
1144g	純 *chún*	pure	꜀źi̯uen	dzywin	*thuần* [t^{hw}ɤ̆n^{2}]	ṣun [ʃøn^{11}]	tʂ^{h}un
1267	垂 *chuí*	to droop	꜀źwi̯ĕ	dzywe	*thùy* [t^{hw}i^{2}]	ṣui [ʃøy^{11}]	tʂ^{h}uei
1247hi	殊 *shū*	special	꜀źi̯u	dzyu	*thù* [t^{h}u^{2}]	ṣu [ʃy^{11}]	ʂu
1129	社 *shè*	society	꜂źi̯a	dzyæX	*xã* [ɕa^{6}]	ṣe [ʃɛ^{23}]	ʂə
228b, 1187r	墅 *shù*, 曙 *shǔ*	villa, bright	꜂źi̯^{w}o, ꜂źi̯^{w}o	dzyoX, dzyoH	*thự* [t^{h}ɯ^{4}]	ṣü [ʃy^{23}]	ʂu
897, 1265n	受 *shòu*	to receive	꜂źi̯əu	dzyuwX	*thụ* [t^{h}u^{4}]	ṣau [ʃɐu^{22}]	ʂou
1220	植 *zhí*	to plant	źi̯ək	dzyik	*thực* [t^{h}ɯk^{8}]	tsik [tsik^{2}]	tʂʅ
860d	芍 *sháo*	peony	źi̯ak	dzyak	*thược* [t^{h}ɯək^{8}]	tsök [tsœk^{2}]	ʂao
1207	乘 *chéng*	to ride (a vehicle)	꜀dźhi̯əng	zying	*thừa* [t^{h}ɯə^{2}]	ṣing [ʃiŋ^{11}]	tʂ^{h}əŋ
232d	船 *chuán*	boat	꜀dźhi̯^{w}än	zywen	*thuyền* [t^{hw}iən^{2}]	ṣün [ʃyn^{11}]	tʂ^{h}uan
868e	神 *shén*	spirit, god	꜀dźhi̯ĕn	zyin	*thần* [t^{h}ɤ̆n^{2}]	ṣan [ʃɐn^{11}]	ʂən
865	射 *shè*	to shoot	dźhi̯a꜄	zyæH	*xạ* [ɕa^{4}]	ṣe [ʃɛ^{22}]	ʂə
1261	順 *shùn*	to obey	꜀dźhi̯uen	zywin	*thuận* [t^{hw}ɤ̆n^{4}]	ṣun [ʃøn^{22}]	ʂun
862	舌 *shé*	tongue	dźhi̯ät	zyet	*thiệt* [t^{h}iət^{8}]	ṣit [ʃit^{2}]	ʂə
445	實 *shí*	true	dźhi̯ĕt	zyit	*thật* [t^{h}ɤ̆t^{8}], *thực* [t^{h}ɯk^{8}]	ṣat [ʃɐt^{2}]	ʂʅ
891	食 *shí*	to eat	dźhi̯ək	zyik	*thực* [t^{h}ɯk^{8}]	ṣik [ʃik^{2}]	ʂʅ
632	繩 *shéng*	rope	꜀dźhi̯əng	zying	*thằng* [t^{h}ăŋ^{2}]	ṣing [ʃiŋ^{11}]	ʂəŋ

It appears from Table 3 that Beijing Mandarin has exactly the same reflex for MC *ź* as for *dźh*. It cannot be argued that the Mandarin treatment [of MC *ź*] reflects an old ***dź** distinct from ***dźh**, since its reflex is aspirated in the *píng* tone[23] in the same way as the reflexes of the stops reconstructed by Karlgren as *bh*, *dh*, *gh*.

Sino-Vietnamese, like Cantonese, shows that *dźh* was treated as the voiced counterpart of *ś*. In Sino-Vietnamese, MC *ś* [ɕ] becomes *th* [tʰ][24] [with upper register tone, parallel to the change of *dźh* to SV *th* [tʰ] with low register tone shown in Table 3]. By contrast, the voiceless Middle Chinese affricates *tś* and *tśh* [tɕ and tɕʰ], respectively, became SV *ch* [c] and *x* [ɕ].[25] Maspero (1920:15) cites a sixth-century author, Yan Zhitui 颜之推, who reported that Southerners confused the following pairs of words

źi̯äk (石 AA 883 [B-MC: *dzyek*]) and *dźhi̯äk* (射 AA 865 [B-MC: *zyæH*])

źi̯ę (是 AA 890 [B-MC: *dzyeX*]) and *dźhi̯ę* (舐 GS867f [B-MC: *zyeX*]).

[2.1.1.3 Karlgren's MC dental affricates and fricatives]

The same author indicates that Southerners confuse *dzhi̯en* (錢 AA 1072e [B-MC: *dzjen*]) and *zi̯än* (涎 AA 235d [B-MC: *zjen*]). This is indeed what happens in Cantonese (Table 4).

[Table 4. Karlgren's MC *z* realized as an affricate in Cantonese.]

AA	character	[gloss]	K-MC	[B-MC]	Cantonese
211ij	祥 *xiáng*	auspicious	ᶜzi̯ang	zjang	tshöng [tsʰœŋ[11]]
810c	祠 *cí*	temple	ᶜzi	zi	tshi [tsʰi[11]]
787	像 *xiàng*	portrait	ᶜzi̯ang	zjangX	tsöng [tsœŋ[22]]
811	寺 *sì*	monastery	ziˀ	zijH	tsi [tsi[22]]
253k	袖 *xiù*	sleeve	zi̯əuˀ	zjuwH	tsau [tsɐu[22]]
775, 779	夕席 *xī / xí*	evening; seat	zi̯äk	zjek	tsik [tsik[2]]
781	習 *xí*	practice	zi̯əp	zip	tsap [tsap[2]]

23. [Karlgren's MC tone marks were omitted from Table 3 in the original article; they have been restored here, including the low open hook for *píng* tone, which Haudricourt always left unmarked.]
24. [This sound change is part of a larger chain shift that occurred in Vietnamese and affected the Sino-Vietnamese layer as well as the native vocabulary: Proto-Vietic *t and *d changed to ɗ, and *s to t, cf. Ferlus 1982.]
25. On this point, we must correct Maspero (1912: 26, 53), who uses doubtful readings of rare characters.

In order to explain all of these facts, a deeper study of the various dialects is required. It would involve examining them from the point of view of linguistic geography and conducting a historical study of the various capital cities and administrative centers whose dialects may have had a leading influence on the other dialects. In any case, it is far from obvious that the contrast between voiced affricates and fricatives goes back to the period of Old Chinese; it can hardly serve as serious evidence for the existence of the contrasts ***dz** vs. ***dzh** and ***dź** vs. ***dźh** in the oldest reconstructible form of the language.[26]

[2.1.2 The argument from "zero" initial phonetic elements: comparative evidence]

Let us now examine Karlgren's second argument: the "zero" initial consonants of phonetic elements for which Karlgren proposes older initial ***b *d *g *z**. The only argument for Karlgren's "gap-filling" hypothesis is the equilibrium of the reconstructed system. If we look for evidence in languages that are genealogically related to Chinese, such as the Tibeto-Burman group, we observe that Karlgren's ***g-** regularly corresponds to **w-**, **sw-** and **rw-**, and that his ***d-** and ***z-** correspond to **j-** (in Lushai, Tibeto-Burman ***y-** [j] regularly changes to **z-**: see Benedict 1948:205).[27]
If we turn to languages that have early loans from Chinese, we find phonemes that are not stops. For instance, Thai, Dioi (Li 1945[28]) and the northern Austroasiatic[29] languages Lamet (Izikowitz 1951: 171) and Khmu (Roux and Tran 1927: 184) have borrowed the twelve names of the Chinese

26. Martin (1953: 13) mentions that *dźh* [dʑʱ] was not counted among the initials in the Tang period. We could conclude that *ź* and *dźh* [ʑ and dʑʱ] were not yet differentiated at that time. [Neither Karlgren's nor Haudricourt's analysis of the opposition between *ź* and *dźh* [ʑ and dʑʱ] is currently accepted. However, Haudricourt's argument concerns the justification or not of a general opposition between a voiced series and a voiced aspirate series, and whether the fricatives/affricate orders contribute to the evidence.]
27. But Benedict does not dare to criticize Karlgren, and hypothesizes that d and g appeared later in Chinese. Shafer (1940, 1944, 1950) does not take Karlgren's reconstructions of Old Chinese into account.
28. Li formulates clear criticisms of Karlgren in this article.
29. [The original article has "Austronesian" here, clearly unintentionally.]

calendric cycle.[30] Two of these names have initials that are of interest to us (see Table 5).

[Table 5. Names of the third and tenth characters of the twelve-year cycle.]

char.	GS	K-OC	AA	K-MC	[B-MC]	Sino-Viet.	Thai	Dioi	Khmu	Lamet
寅 *yín*	450	*di̯ər	283	ʻi̯ĕn	yin	*dần* [ðɤ̆n^{2}]	ñi	ñien	ñi	ñi
酉 *yǒu*	1096	*zi̯og	258	ʻi̯əu	yuwX	*dậu* [ðɤ̆^{w4}]	rau	ru	rau	rau

It is possible that Thai **ñ-** [ɲ] comes from an older **j-**, but **r-** is ancient in Thai.

An early loanword is apparently found in Vietnamese for 'left-over, extra': Vietnamese *thừa* [tʰɯə^{2}], from Chinese 餘 *yú* (AA 1322b ◌*i*^{w}*o*; GS 821 *dio*; Sino-Vietnamese *dư* [ðɯ^{1}]).

With a low register tone, Vietnamese **th** can only come from an earlier **ź* [ʑ]. [This entails that this word must have had a **ź* initial in the donor Chinese dialect, not a stop.]

Even Karlgren's reconstruction for the Tang period, which is relatively late, does not seem to be correct, since it is contradicted by evidence from Sino-Vietnamese, which only came into existence in the 10th century. This can be shown by the data in Table 6.

[Table 6. Sino-Vietnamese examples contradicting Karlgren's reconstructions.]

AA n°	character	[gloss]	K-MC	[B-MC]	Sino-Vietnamese
1348c	越 *yuè*	overstep, get over	ji̯^{w}ɒt	hjwot	*việt* [viət^{8}]
1345b	園 *yuán,*	garden	◌ji̯^{w}ɒn	hjwon,	*viên* [viən^{1}],
1345e	遠 *yuǎn*	far	ʻji̯^{w}ɒn	hjwonX	*viễn* [viən^{6}]
118b	域 *yù*	domain	ji̯^{w}ək	hwik	*vực* [vɯk^{8}]
524y	位 *wèi*	position	j^{w}iˀ	hwijH	*vị* [vi^{4}]
183	矣 *yǐ*	particle	ʻji	hiX	*hĩ* [hi^{6}]
249	又 *yòu*	again	ji̯əuˀ	hjuwH	*hựu* [hɯ^{w} ^{4}]
1267e	郵 *yóu*	mail	◌ji̯əu	hjuw	*bưu* [ɓɯ^{w} ^{1}]
239	炎 *yán*	burning hot	◌ji̯äm	hjem	*viêm* [viəm^{1}]
1138de	悅 *yuè* 閱 *yuè*	pleased to pass through	i̯^{w}ät	ywet	*duyệt* [ðʷiət^{8}]
232b	沿 *yán*	along	◌i̯^{w}än	ywen	*duyên* [ðʷiən^{1}]

30. [These are also called the "Twelve Earthly Branches".]

425bc	浴 *yù* 欲 *yù*	bath to desire	i̯ʷok	yowk	*dục* [ðuk^{8}]
1265Agh	惟 *wéi* 維 *wéi*	only, sole to bind	ʿʷi	ywij	*duy* [ðʷi^{1}]
181l	以 *yǐ*	using	ʿi	yiX	*dĩ* [ði^{6}]
259c	遊 *yóu*	to travel	ʿi̯əu	yuw	*du* [ðu^{1}]
1327g	愈 *yù*	to heal, to get better	ʿi̯u	yuX	*dũ* [ðu^{6}]
247	閻 *yán*	surname	ʿi̯äm	yem	*diêm* [ðiəm^{1}]

It is true that Vietnamese *d-* can be the reflex of an earlier ***j-** (Karlgren's *j̯-*) as well as of an earlier ***ð-** [and therefore Karlgren's reconstruction of a ***j** for these items in Table 6 would be acceptable]. On the other hand, Karlgren's *ji̯-* is obviously incorrect, and one has to reconstruct ***ɣ-**, following the work of the Chinese scholars Chao (1941) and Lo (1951).[31] All these pieces of evidence converge to show that the consonants that ought to be reconstructed for these "zero" initials are not stops, but voiced fricatives. Karlgren's second argument vanishes and instead of an Old Chinese system such as:

***t**	***th**	***dh**	***d**
***tś**	***tśh**	***dźh**	***dź**
***k**	***kh**	***gh**	***g**

it is more plausible to reconstruct the following:[32]

***t**	***th**	***d**	***ð, j**
***tś**	***tśh**	***dź**	***ź, r**
***k**	***kh**	***g**	***ɣ, w**

31. Note, however, that one should not conclude that this *ɣ goes back to a more ancient *g, as Kennedy (1952) does.
32. [The solution to this problem was found by Yakhontov (1960b) five years after the publication of Haudricourt's article. For details, see the comments appended to this article.]

[2.2 Initial consonant clusters]

Another point of divergence between Maspero and Karlgren is the reconstruction of initial consonant clusters. When one finds characters using the same phonetic element which transcribe both words in **l-** and words in **k-**, the two scholars agree in interpreting the **k-** as the reflex of old ***kl-**. For Maspero, these are words with initial ***l-,** to which a ***k-** prefix has been added. For Karlgren, the words with **l-** in Modern Chinese originate from words with a ***gl-** initial.

Maspero had an advantage in that he made use of Tai evidence[33] (see Maspero 1930). The study of early loanwords from Chinese is indeed the only method which makes it possible to solve this problem. For instance, Karlgren interprets the character 剝 *bāo* 'to peel' (AA 574f *påk*) as an ideogram formed from [the two co-significs] 彔 *lù* 'to carve' (AA 574 *luk*) and 刀 *dāo* 'knife' (AA 975 *tâu*), whereas Maspero (1930:321) interprets 剝 *bāo* 'to peel' as a phonogram [containing AA 574 *luk* as a phonetic] and reconstructs ***plɔk** in Old Chinese by comparison with the Thai word **plɯək** 'tree bark'. It seems that, in this particular instance, Maspero was mistaken, since 'to peel' is also found as ***pɔk** in Proto-Tai, and *bóc* [**ɓɔk^{7}**] in Vietnamese.[34]

The complex initials ***kl-** and ***pl-** were present in Vietnamese until a relatively late period, but to this day we have not found either of these initials in any word of Chinese origin. Tai evidence is more revealing. Some examples are shown in Table 7.

---

33. [Maspero indicates : "The comparisons with the Thai languages are made on the sole basis of Siamese, because it is the only language of the group (apart from Ahom, too poorly known to use) which preserved the initial consonant clusters *kl, pl,* etc. This method, totally indefensible from the theoretical point of view, [...] was the only one practically possible." "Siamese is given in a transliteration of the written form." (1930 : 321n1)]
34. On the basis of a collection of comparisons of this kind, Wulff (1934) believed that he had demonstrated the existence of infixes in Chinese and in Tai.

[Table 7. Four examples where Karlgren mistakenly reconstructed consonant clusters (K-OC column).]

gloss	Proto-Tai	character	AA	K-MC	GS	K-OC	[B-MC]
to change	plienʾ	變 *biàn*	590j	pi̯ᵂänʾ	178o	pli̯an	pjenH
bee	ʿphɯng	蜂 *fēng*	31e	⸢phi̯ᵂong	1197s	pʰi̯ung	phjowng
cyclic character 2/12	ʿplaw	丑 *chǒu*	1242	ʿtʿhi̯əu	1076	tʰni̯ôg	trhjuwX
cyclic character 4/12	ʿhmaw	卯 *mǎo*	602	ʿmaw	1114	mlôg	mæwX

Karlgren and Maspero agree on the first example. The second, from Maspero, and the last two, from Li (1945), show how little basis there is for Karlgren's reconstructions.

It is probably necessary to distinguish clusters in **-l-** from clusters in **-r-**, since data from Tai languages and from Vietnamese distinguish between **r** and **l** in old borrowings of words that have **l** initial in Modern Chinese or in Sino-Vietnamese[35] (Table 8).

[Table 8. Early Chinese borrowings into Vietnamese and Tai, pointing to the presence of r distinct from l in the donor Chinese dialect.]

number	character	[gloss]	K-OC>MC	Vietnamese	Tai	[B-MC]
GS 1193	龍 *lóng*	dragon	li̯ung > li̯wong	*rồng* [roŋ^{2}]		ljowng
GS 185f	闌 *lán*	enclosure	glân > lân	*ràn* [ran^{2}]		lan
GS 185i	練 *liàn*	to train	glian > lien	*rèn* [rɛn^{2}]		lenH
AA 552	簾 *lián*	screen	gli̯am > li̯äm	*rèm* [rɛm^{2}]		ljem
AA 576	轆 *lù*	pulley	(MC) luk	*rọc* [rɔk^{8}]	rɔk (Thai)	luwk
GS 1032	六 *liù*	six	li̯ôk > li̯uk		*hrŏk (Proto-Tai)	ljuwk

When ***r** was preceded by a stop, we find *s* [ʂ] in Vietnamese. For instance the name of the squirrel, **prɔk** in Bahnar, Khmu, and Siamese, became *sóc* [**ʂɔk**^{7}] in Vietnamese. The Chinese etyma had a prefix in the cases shown in Table 9.

35. [The forms in Sino-Vietnamese, a later layer of borrowings than the forms in Table 8, are *long* for 'dragon', *lan* for 'enclosure', *luyện* for 'to train', *liên* for 'a blind', and *lộc* for 'pulley'.]

[Table 9. Chinese borrowings into Vietnamese, pointing to the presence of a prefix preceding *r in the donor Chinese dialect.]

number	character	[gloss]	K-MC	[B-MC]	Vietnamese
GS 928	力 *lì*	strength	liə̯k	lik	*sức* [ʂɯk^{7}]
AA 550	蠟 *là*	wax	lâp	lap	*sáp* [ʂap^{7}]
AA 551	蓮 *lián*	lotus	lien	ljen	*sen* [ʂɛn^{1}]

Only the study of *fǎnqiè* spellings that predate the Tang period could indicate whether Old Chinese distinguished ***l*** from ***r***, as most Sino-Tibetan languages did.

[2.3 An uvular series in Old Chinese]

Proto-Tai distinguishes the velar stops ***k*** and ***g*** from the uvulars ***q*** and ***ɢ***, and some Chinese loanwords display a uvular, *e.g.* ***ɢam*** 'gold' and ***qaŋ*** 'pig iron', raising the issue of whether Old Chinese too had such a contrast. Loanwords never constitute decisive proof, but they provide valuable clues that should not be overlooked.

[3. The reconstruction of rhymes]

Karlgren's reconstructions of vowels raise even more serious difficulties.

[3.1 Medials]

We know that the syllabic structure of Modern Chinese allows for a medial phoneme: a semi-vowel ***j***, ***w*** or ***ɥ*** placed between the initial consonant and the main vowel. When they started their reconstructions, Maspero and Karlgren believed that the ***-w-*** of Modern Chinese could be traced back to the *hékǒu* [i.e. rounded] category of the rhyme tables[36], and the ***-j-*** to a difference of "division" (*děng*).[37] Maspero analyzed *ẅ* [***ɥ***] as a combination

36. [In traditional Chinese phonology, rounded rhymes are referred to as *hékǒu* 合口'closed mouth', and unrounded rhymes as *kāikǒu* 开口 'open mouth'.]
37. [On the four divisions děng 等 of Middle Chinese, see Branner (2006), Ferlus (2009) and references therein.]

of **w** plus **j** whereas Karlgren analyzed it as **j** plus **w**. Karlgren distinguished between *w* and *u*, and between *i̯* [**-j-**] and *i*, depending on whether this phoneme belonged to the rhyme or not. In 1915 Karlgren interpreted one of the *děng* distinctions as a palatalization contrast: *kji̯ən* (AA 384; 巾 *jīn* [B-MC: *kin*] 'kerchief') vs. *ki̯ən* (AA 385; 斤 *jīn* [B-MC: *kɨn*] 'ax'), whereas in 1923 he interpreted it as a contrast in vowel quality: *ki̯ĕn* (AA 384; 巾 *jīn*) vs. *ki̯ən* (AA 385; 斤 *jīn*). However, he did not pursue this line of thought consistently, and he does not hesitate to reconstruct most of these medial phonemes back to Old Chinese.

In contrast, Maspero explains the genesis of these medials as diphthongizations, on the evidence of loanwords in Thai,[38] as in the following three examples:

character	GS no. OC > MC	Proto-Tai
兼 *jiān* 'to gather'	GS 627a **kliam** > *kiem* [B-MC: *kem*]	***kɛm**
天 *tiān* 'sky'	GS 361a **thien** > *thien* [B-MC: *then*]	***tʰɛn**
凡 *fán* 'all'	GS 625a **bhi̯wam** > *bhi̯wɒm* [B-MC: *bjwom*]	***ᶜbrɔm**

A diphthong ***ie** existed in Proto-Tai; the first two examples therefore testify to a change ***ɛ** > **ie**[39] in Chinese. Loanwords into Vietnamese, although they are more recent, illustrate the same change (cf. Vietnamese *rèn* 'to train' in Table 8 and Vietnamese *sen* 'lotus' in Table 9). In the third example, we observe both the change ***ɔ > *uo > wɒ** and the change from *-**r**- to -*i̯*- [**-j-**] that we have already seen in the word for 'bee' (Table 7).

Another origin for the medial -*w*- could be found in the breaking of earlier labiovelars into velar+*w*. It seems that scholars have overlooked the fact that some rhymes[40] in the *Analytic Dictionary* only appear with velar initials (*k, kh, g, x, ng*),[41] for instance -*iʷei*, -ʷɒng, -*iʷäng*, -ʷâk, -*iʷet*, etc.[42]

38. Of course, Maspero viewed Thai as genealogically related to Chinese, but if one agrees with Benedict (1939) that the words at issue are loanwords, their value as evidence to reconstruct Old Chinese is all the greater.
39. [The direction of this change, reversed ("ɛ < ie") in the original article, was corrected by the author in the 1972 reprint.]
40. [In the framework of traditional Chinese phonology, each initial consonant, and each rhyme (vowel + coda) of Middle Chinese has a specific name

If we accepet Maspero's hypothesis [of a diphthongization], Karlgren's distinction between *-w-* and *-u-*, based on the rhymes, might simply be a distinction between a labial phoneme and a labial feature. Sino-Vietnamese offers an argument. In the version of the rhyme table dated 706, there are two **-u-** [*hékǒu*] rhymes : *-uậi* and *-uən*, that are distinguished from their corresponding unrounded rhymes. These became *-ôi* [**-oj**] and *-ôn* [**-on**] in Sino-Vietnamese while the corresponding unrounded rhymes became Sino-Vietnamese *-ơi* [**-ɤj**] and *-ăn* [**-ăn**]. In contrast, those rhymes in **-u-** that were distinguished only in later rhyme books, such as *-uâ* and *-i̯uən*, as well as the **-w-** rhymes *-ʷa* and *-ʷai*, became rhymes with a full labial phoneme [rather than a labial feature] in Sino-Vietnamese, respectively *-oa*, *-uân*, *-oa*, and *-oai* [**-ʷa**, **-ʷɤ̆n**, **-ʷa** and **-ʷaj**].[43]

Besides Tai and Sino-Vietnamese, which Maspero had already made use of, we will now turn to early Chinese loanwords in Vietnamese (predating the Sino-Vietnamese layer), which until now have been almost completely ignored.[44] Although they are less ancient than loanwords into Tai, they are

---

indicated by a single character. The traditional name of the rhyme has been provided in a note or between parentheses in this section for reference.]

41. [This is a very important observation, which was fully exploited only five years later by Yakhontov (1960a). Most reconstruction systems now include a series of labiovelars.]
42. [Respectively: rhyme 齊 *qí*, reconstructed by Baxter as B-MC: -wej, 唐 *táng* B-MC: wang, 清 *qīng* B-MC: jweng, 鐸 *duó* B-MC: wak, 屑 *xiè* B-MC: wet].
43. [For ease of reference, here is a summary in table form of the rhymes mentioned, with the addition of the names traditionally given to the rhymes in Chinese phonology: a character that exemplifies the rhyme.]

source	K-MC	[B-MC]	rhyme name		Sino-Vietnamese	
early rhyme books	-uậi	-woj	灰	*huī*	*-ôi*	[-oj]
	-uən	-won	魂	*hún*	*-ôn*	[-on]
	-ậi	-oj	咍	*hāi*	*-ơi*	[-ɤj]
	-ən	-on	痕	*hén*	*-ăn*	[-ăn]
late rhyme books	-uâ	-wa	戈	*hū (gē)*	*-oa*	[-ʷa]
	-i̯uən	-jwin	諄	*zhūn*	*-uân*	[-ʷɤ̆n]
	-ʷa	-wæ	麻	*má*	*-oa*	[-ʷa]
	-ʷai	-wɛɨ	佳	*jiā*	*-oai*	[-ʷaj]

44. Originally, Maspero did not admit that any borrowings from Chinese in Vietnamese were earlier than Sino-Vietnamese. He acknowledges the existence of some in a short article (Maspero 1916), and mentions a few others

more numerous, which allows for the establishment of conclusive correspondences with Chinese.

Let us begin by examining loanwords ending in the diphthongs *-ia*, *-ưa* and *-ua* [**iə**, **ɯə** and **uə**] in Vietnamese, comparing them to Karlgren's reconstructions of the corresponding Chinese words (Table 10).

[Table 10. Chinese loanwords in Vietnamese ending in one of the diphthongs *-ia*, *-ưa* and *-ua*, and Karlgren's reconstructions of the Chinese words.]

number	char.	K-OC	K-MC	[B-MC]	Vietnamese	[gloss]
GS 23f	離 *lí*	lia	˂ljie	lje	*lìa* [liə^{2}]	to leave
GS 2ru	義 *yì,*	ngia	ngjieˀ	ngjeH	*nghĩa* [ŋiə^{6}]	justice
	儀 *yí*			ngje		
GS 358j	紫 *zǐ*	tsiar	ˈtsie	tsjeX	*tía* [tiə^{3}]	purple
AA 890b	匙 *chí*		˂zie	dzye	*thìa* [t^{h}iə^{2}]	spoon
	shi				*chìa* [ciə^{2}]	key[45]
GS 25a	皮 *pí*	bhia	˂bhjie	bje	*bìa* [ɓiə^{2}]	cover
GS 874d	碑 *bēi*	pie̯g	˂pjie	pje	*bia* [ɓiə^{1}]	stele
GS 49s′	鋸 *jù*	kio	kiwoˀ	kjoH	*cưa* [kɯə^{1}]	saw
AA 579j	驢 *lǘ*		˂liwo	ljo	*lừa* [lɯə^{2}]	donkey
GS 84g	貯 *zhù*	tio	ˈtiwo	trjoX	*chứa* [cɯə^{3}]	to contain
GS 82l	餘 *yú*	dio	˂iwo	yo	*thừa* [t^{h}ɯə^{2}]	rest, left-over
GS 60i	許 *xǔ*	xio	ˈxiwo	xjoX	*hứa* [hɯə^{3}]	to allow
GS 46z	助 *zhù*	dzhio	dzhiwoˀ	dzrjoH	*chữa* [cɯə^{6}]	to cure
GS 60p	禦 *yù*	ngio	ngiwoˀ	ngjoH	*ngừa* [ŋɯə^{2}]	to prevent
GS 102h	斧 *fǔ*	piwo	ˈpiu	pjuX	*búa* [ɓuə^{3}]	hammer, axe
GS 136c	府 *fǔ*	piu	ˈpiu	pjuX	*(chợ) búa* [ɓuə^{3}]	market
GS 136m	符 *fú*	bhiu	˂bhiu	bju	*bùa* [ɓuə^{2}]	amulet
GS 1001	婦 *fù*	bhiŭg	ˈbhiəu	bjuwX	*(goá) bụa*	widow[46]

in Maspero (1920: 61, 93n2). Since then no work has been published on this topic. [Haudricourt was apparently unaware of Wang Li's substantive study, published in 1948, which identifies dozens of older Sino-Vietnamese loanwords, including many of the same items as in Haudricourt's article, and relevant phonological patterns. After Haudricourt's article, few studies have been devoted to older borrowings from Chinese into Vietnamese. Michel Ferlus (p.c.) mentions Nguyễn Khắc-Kham (1971) and David Tryon Ray (1979).]

45. [Haudricourt translated this word as *cure-oreille* 'ear-pick'.]
46. [By itself, 婦 *fù* simply means 'wife; married woman'. The term 'widow' requires a compound with 寡 *guǎ* 'scant, few; widowed', as 寡婦 *guǎfù*. The

					[ɓuə^4]	
GS 101f	扶 *fú*	bhi̯wo	ʿbhi̯u	bju	*vùa* [vuə^2]	to support
GS 129	主 *zhǔ*	t̂i̯u	ʻtśi̯u	tsyuX	*chúa* [cuə^3]	master
GS 129d	註 *zhù*	t̂i̯u	tśi̯uˀ	trjuH	*chua* [cuə^1]	to make a note
GS 1109j	務 *wù*	mi̯og	mi̯uˀ	mjuH	*mua* [muə^1]	to get[47]
GS 1028e	繡 *xiù*	si̯ôg	si̯ə̯uˀ	sjuwH	*thùa* [t^{h}uə^2]	to embroider

For the *-ia* rhyme [**iə**], there is a good correspondence between the Vietnamese borrowings and Karlgren's reconstruction of a medial ***-i-**. Likewise for the second rhyme, *-ưa* [**ɯə**], if we take into account the fact that Vietnamese *ư* [**ɯ**] shares its lip position with **i** and its tongue position with **w**. However, the third diphthong, *-ua* [**-uə**], is problematic [since Karlgren reconstructs a **j**-like sound whereas the Vietnamese evidence points to a **w**-like sound]. The parallelism between these three medials is also found in Sino-Vietnamese, where they appear respectively as *–i*, *–ư*, and *–u* [**i**, **ɯ**, **u**]. It seems that Karlgren made a mistake in chronology, reconstructing in the case of the third rhyme a more recent form than for the first two rhymes.[48] I would propose reconstructing ***-ie**, ***-i^{w}ə** and ***-wu**[49] respectively.

Karlgren could object that the Chinese words in Vietnamese do not belong to the dialect that he reconstructs, but to Wu, a southern dialect, to which he attributes the oldest Japanese system for reading Chinese characters, the Go-on 吳音 [Pinyin: *wúyīn*].[50] It is indeed very likely that this dialect was

---

early Vietnamese loan *goá bụa* [ɣ^{w}a^3 ɓuə^4] 'widow' has a Sino-Vietnamese (i.e. later) doublet: *quả phụ* [k^{w}a^5 fu^4], with the same meaning.]

47. [In both Old Chinese and Modern Chinese, 務 *wù* is associated with two sets of meanings: 'strive, exert oneself, be necessary' and 'task, affair', but not 'to get'. The meaning in Vietnamese is 'to buy'.]
48. [That is, Karlgren correctly reconstructed an older form in Old Chinese for the first two rhymes, but in the case of the third rhyme, he projected back into Old Chinese a form belonging to a later period.]
49. [The original *Word* article has *wo*; the 1972 reprint has *wu*.]
50. Karlgren (1940: 71–72) admits that his reconstruction does not account well for the ancient Wu dialect ("Go" is the Japanese pronunciation of 吳 *Wú*), but concludes that this Wu dialect was a mixture. Is it plausible that the first system of Chinese readings to come to Japan in the 6th century was more mixed than the dialect of the capital in the 8th century? Demiéville (1946) said: "Hypothesizing the existence of several Wu dialects to explain the incoherence in the Go-on readings is entirely ad hoc and utterly questionable. It has no support whatsoever in what we know from Japanese sources concerning the historical conditions in which Go-on was formed."

different from that of the capital city during the Tang period, but then it is difficult to understand why Karlgren ignored it when he reconstructed Old Chinese and, in the majority of cases, projected the final *-i̯u* back into the earliest period.

[3.2 Back vowels]

Another rhyme, whose reflex in Sino-Vietnamese is *ô* [IPA: **o**], is found [as Vietnamese *o* (IPA ɔ)] in many early loanwords, as shown in Table 11a.

[Table 11a. Karlgren's OC *o/â realized as Vietnamese open *o* in early loans]

GS	character	K-OC	K-MC	[B-MC]	Vietnamese	[gloss]
74e[51]	庫 *kù*	kho	khuoˀ	khuH	*kho* [kʰɔ¹][52]	warehouse
49u	苦 *kǔ*	kho	ʿkhuo	khuX	*khó* [kʰɔ³]	difficult
60	午 *wǔ*	ngo	ʿnguo	nguX	*ngọ* [ŋɔ⁴]	7th Earthly Branch
63	兔 *tù*	tho	thuoˀ	thuH	*thỏ* [tʰɔ⁵]	hare
801b	渡 *dù*	dhâg	dhuoˀ	duH	*đò* [ɗɔ²]	ferry
53	戶 *hù*	gho	ʿɣuo	huX	*họ* [hɔ⁴]	family
94z	弩 *nǔ*	no	ʿnuo	nuX	*nỏ* [nɔ⁵]	crossbow
792j	訴 *sù*	sâg	suoˀ	suH	*tỏ* [tɔ⁵]	to express
69l	爐 *lú*	lo	luo	lu	*lò* [lɔ²]	oven
766t	露 *lòu*	ʿglâg	luoˀ	luH	*lõ* [lɔ⁶]	to reveal

However, beside these reflexes with **ɔ**, there are also some in **a**, as shown in Table 11b.

[Table 11b. Karlgren's OC *o/â realized as a in early loans in Vietnamese and other languages.]

GS	char.	K-OC	K-MC	[B-MC]	Viet.	Proto-Tai	[gloss]
802f	墓 *mù*	mâg	muoˀ	muH	*mả* [ma⁵]		tomb
58	五 *wǔ*	ngo	ʿnguo	nguX		*ʿhŋa	five
94z	弩 *nǔ*	no	ʿnuo	nuX	*ná* [na³]	*ʿhna (cf. Khmer ស្នា *sna*)	crossbow
60	午 *wǔ*	ngo	ʿnguo	nguX		*saŋa (cf. Khmu sŋa)	7th Earthly Branch

51. [Corrected in translation from GS75e.]
52. [In present-day Vietnamese, orthographic *kh* is pronounced x.]

Sino-Tibetan comparison shows that this rhyme comes from an older ***-a**.[53] It seems that this rhyme was pronounced as **-a** in Old Chinese and into the Han dynasty [2nd century B.C. – 2nd century A.D] (the crossbow seems to have been invented and diffused at this period), and that this **a** was gradually raised to a mid-open **ɔ** in the 6th to 7th centuries (early loans into Vietnamese), then through mid-close **o** in the 9th century (Sino-Vietnamese), finally becoming **-u** in Modern Chinese.

This brings into question the entire reconstructed vowel system of Old Chinese. For instance, basing themselves on the fact that 馬 *mă* [B-MC: *mæX*] 'horse' rhymes with **-o** words in the *Book of Odes*, Maspero and Karlgren concluded that this word should be reconstructed with an open **-ɔ** or an **-å** [**ɒ**]; however, since at the stage of Old Chinese the alleged rhyme **-o** was in fact **-a**, there is no reason not to reconstruct an **-a** vowel for 'horse'. Additional support for this reconstruction may be found in Japanese *uma* and Korean *mal* 'horse'. The final consonant, preserved in Korean, explains why the **-a** was maintained in this word [in Chinese], and did not become raised.[54]

Sino-Tibetan comparisons show that some vowels must have become more open, such as 三 *sān* 'three': Tibetan *gsum*, GS 648 ***səm**>*sâm*.[55]

Karlgren reconstructs the rhyme **-ək** for both Middle and Old Chinese. However, the reflexes of this rhyme are different in Sino-Vietnamese (first example in Table 12) and in earlier loans.[56]

---

53. For instance, Tibetan *lnga* 'five', Burmese *nga*.
54. [Haudricourt's ideas about *-a have been widely accepted, except for this last sentence. It is now generally accepted that 馬 'horse' [B-MC: *mæX*] should be reconstructed with a medial *-r- in Old Chinese; it is the presence of this medial which prevented the rhyme from changing to *-u*.]
55. [Haudricourt is alluding here to the fact that the numeral 'three' in Chinese has an irregular development. In the *Book of Odes*, it rhymes as *-ɨm; the expected outcome in Middle Chinese is *som*, but Baxter's Middle Chinese reconstruction (based on strong evidence) is *sam*, similar to Karlgren's *sâm*). No satisfactory explanation has been found to this day for this irregularity.]
56. [Haudricourt's argument is that the vowel in Sino-Vietnamese readings is ă whereas an earlier stratum of loanwords has ɯ. This is confirmed by examining the Sino-Vietnamese readings of characters in Table 12, as compared to the pronunciation of early loanwords. In Sino-Vietnamese, the character 特 *tè* reads *đặc* [dăk^{8}] (and has the meaning 'special', not 'male'; see next footnote), and 墨 *mò* 'ink' reads *mặc* [măk^{8}], both with the vowel ă. The word *đức* [dɯk^{7}] 'virtue' (德 *dé*) is considered by Haudricourt as an early loanword, on the basis of its ɯ vowel; if it belonged to the Sino-Vietnamese layer, one would expect the vowel ă. Interestingly, *đức* [dɯk^{7}] serves as the

[Table 12. Correspondences between words reconstructed by Karlgren with the rhyme ***-ək*** and words with two different rhymes in two different layers of Vietnamese borrowings. SV=Sino-Vietnamese.]

GS	char.	K-OC	K-MC	[B-MC]	Vietnamese	[gloss]
905	得 *dé*	tək	tək	tok	SV *đắc* [dăk^{7}]	to obtain
919k	德 *dé*	tək	tək	tok	*đức* [ɗɯk^{7}]	virtue
961h′	特 *tè*	dhək	dhək	dok	*đực* [ɗɯk^{8}]	male (animal)[57]
904c	墨 *mò*	mək	mək	mok	*mực* [mɯk^{8}]	ink

This could support the hypothesis of the following evolution: ***u*** > ***ɯ*** > ***ə*** > ***a***.

The preceding observations help to improve the symmetry of the vowel system that is reconstructed for Old Chinese; Karlgren had reconstructed five back vowels, ***u***, ***o***, ***ɔ***, ***ɒ***, ***ɑ***[58] and no front vowel in open syllables.

[3.3 Central and front vowels in open syllables]

Karlgren believed that front vowels in Old Chinese were always followed by a consonant. However, early Vietnamese loanwords show that one of the

reading of the corresponding character (德 *dé*) in the system referred to as 'Sino-Vietnamese'; in light of Haudricourt's argument, it can be hypothesized that the early Chinese loanword for 'virtue' was carried over into the new (Sino-Vietnamese) system for reading Chinese characters. On the topic of interferences between different layers of Chinese borrowings in Vietnamese, which created exceptional readings within the Sino-Vietnamese system, see Nguyễn Tài Cẩn 1979. A further glimpse into the complex overlay of Chinese loanwords into Vietnamese is offered by Haudricourt's first example, 得 *dé* 'to obtain': prior to its Sino-Vietnamese reading as *đắc* [dăk^{7}], this word had been borrowed with yet another vowel (neither ă nor ɯ), as *đ*ư*ợc* [ɗɯək^{8}].]

57. [The gloss that Haudricourt cites represents the oldest sense of this etymon, i.e. 'bull, male (animal)', attested both in Old Chinese and as an early Vietnamese loanword. In Modern Chinese, the meaning of 特 *tè* is 'special, exceptional' (as also in Sino-Vietnamese, where 特 reads *đặc* [dăk^{8}]). The sense development from 'male (animal)' to 'special' is well documented from the end of the Old Chinese period onward.]
58. [Haudricourt's original mentions five vowels but only lists the first four, as: u, ô, o, å.]

-i rhymes [the Middle Chinese rhyme 之 *zhī*] originated in a central vowel[59] (see Table 13).

[Table 13. Early Chinese loanwords in Vietnamese suggesting a central vowel in Old Chinese.]

GS n°	character	K-OC	K-MC	[B-MC]	Vietnamese	[gloss]
956	疑 *yí*	ngi̯əg	ᵓngji	ngi	*ngờ* [ŋɤ²]	suspect
976h	似 *sì*	dzi̯əg	ˁzi	ziX	*tợ* [tɤ⁴]	as, like
963	市 *shì*	d̂i̯əg	ˁźi	dzyiX	*chợ* [cɤ⁴]	market
971	事 *shì*	dẓhi̯əg	ḍẓhiˀ	dzriH	*thờ* [tʰɤ²]	to serve
972j	詞 *cí*	dzi̯əg	ᵓzi	zi	*tờ* [tɤ²]	document
974	絲 *sī*	si̯əg	ᵓsi	si	*tơ* [tɤ¹]	silk
961z	時 *shí*	d̂i̯əg	ᵓźi	dzyi	*giờ* [ʒɤ²]	hour, time
961d′	詩 *shī*	śi̯əg	ᵓśi	syi	*thơ* [tʰɤ¹]	verse, poem
952v	旗 *qí*	ghi̯əg	ᵓghji	gi	*cờ* [kɤ²]	flag

These Vietnamese reflexes agree with those found in Proto-Tai for three names of the twelve-year cycle (Li 1945), as shown in Table 14.

[Table 14. Proto-Tai comparanda for three names of Earthly Branches.]

GS n°	char.	K-OC	K-MC	[B-MC]	Proto-Tai	[gloss]
937	亥 *hài*	ghəg	ˁɣâi	hojX	* ˁgəɯ	12th Earthly Branch
964	子 *zǐ*	tsi̯əg	ˁtsi	tsiX	* ˁčəɯ	1st Earthly Branch
967	巳 *sì*	dzi̯əg	ˁzi	ziX	* ˁsəɯ[60]	6th Earthly Branch

The other rhyme in **–i** [the Middle Chinese rhyme 脂 *zhī*][61] was probably a front vowel, contrary to what is reconstructed by Karlgren (see Table 15).

[Table 15. Examples of probable Old Chinese front vowels.]

character	GS n°	K-OC	K-MC	[B-MC]	Vietnamese	gloss
几 *jī*	602	ki̯ɛr	kji	kijX	*ghế* [ɣe³]	chair
眉 *méi*	567	mi̯ər	ˁmjʷi	Mij	*mầy* [mɤ̆j²][62]	eyebrow

59. [In Baxter's system, the reconstruction for the 之 *zhī* rhyme is OC *-ɨ.]
60. The correspondence Tai *s- for Chinese z- is expected, as Proto-Tai lacked z-, which appeared later from the cluster *dr-. Another case in point is GS 686 [*d̂i̯əp > ź i̯əp], Proto-Tai *sip 'ten'.
61. [In Baxter's system, the reconstruction for the 脂 *zhī* rhyme is OC *-ij.]

屍 *shī*	561	śi̯ər	śi	Syij	*thây* [t^{h}ɤ̆j^{1}]	corpse

Karlgren reconstructs an ***-r*** when an open syllable in Middle Chinese is written using a phonetic element which is also used to write words with a final -*n. Vietnamese does not provide any evidence on this aspect of the reconstruction, because Proto-Austroasiatic ***-r*** and ***-l*** both became **-j** or zero in Vietnamese.
In some cases, the Chinese loanword had a final **-j** [not found in Middle Chinese], as in the examples in Table 16.

[Table 16. Examples of Chinese loanwords in Vietnamese providing evidence of a final -j in the donor dialect.]

char.	[gloss]	GS n°	K-OC	K-MC	[B-MC]	Vietnamese
蛾 *é*	moth	2q	ngâ	ᶜngâ	nga	*ngài* [ŋaj^{2}]
摩 *mó*	to grind	17e	mwâ	ᶜmuâ	ma	*mài* [maj^{2}]
瓦 *wǎ*	tile	20	ngwa	ʿng^{w}a	ngwæX	*ngói* [ŋɔj^{3}]
移 *yí*	to change	3q	dia	ᶜie̯	ye	*dời* [ðɤj^{2}]
騎 *qí*	to ride	lu	ghia	ᶜghji̯e̯	gje	*cưỡi* [kɯəj^{6}][63]
寄 *jì*	to send, to give	lx	kia	kji̯e̯ˀ	kjeH	*gởi* [ɣɤj^{5}], *gửi* [ɣɯj^{5}]

It is possible that this final **-j** represents the trace of a final ***-l*** or ***-r*** which Karlgren did not reconstruct.[64]

[3.4 Karlgren's codas in ***-g***]

A more complex problem arises with the codas in ***-g*** which Karlgren sprinkled throughout his Old Chinese reconstructions. Karlgren and

62. Vietnamese *-ây* [ɤ̆j] represents an earlier diphthong *-ei. For instance AA 776 *siei* becomes *tây* [tɤ̆j] 'West'.
63. [This early loanword is written as *cỡi cữi* by Haudricourt ([kɤ^{6}] and [kɯj^{6}]); both forms are now rare, the usual form being *cưỡi* [kɯəj^{6}]. The Sino-Vietnamese form for this word is *kị* [ki^{4}].]
64. [The rhyme 歌 *gē* is now universally reconstructed as *-aj for the Old Chinese stage, in keeping with Haudricourt's insight.]

Maspero noticed that the four tones of Middle Chinese[65] are not evenly distributed among sets of words written with the same phonetic element [i.e. the words that make up phonetic series]. The *píng* and *shăng* tones on the one hand, and the *rù* tone on the other, are never found in the same series. This is not surprising, since the *rù* tone is the only one [of the three] that occurs on words ending in *-p*, *-t* or *-k*, while the *píng* and *shăng* tones occur on open syllables, or syllables ending in sonorant consonants. But the *qù* tone appears in series of both of these categories. Words in the *qù* tone whose phonetic element is a *rù* tone character never have nasal finals. Examples are provided in Table 17.

[Table 17. Examples of *qù* tone words whose phonetic element is a *rù* tone character.]

example					char. used as phonetic		
char.	[gloss]	AA n°	K-MC	[B-MC]		K-MC	[B-MC]
裕 *yù*	wealth, opulence	425	i̯uˀ	yuH	谷	kuk	kuwk
夜 *yè*	night	187	i̯aˀ	yæH	亦	i̯äk	yek
懿 *yì*	beautiful, exemplary	201	ʔiˀ	ʔijH	壹	ʔi̯ĕt	ʔjit

Karlgren and Maspero agree that in these cases the words with *qù* tone must have had a final stop in Old Chinese, and they reconstruct a final voiced stop; the three words in Table 17 are reconstructed as ***iug**, ***iag** and ***id**.

Karlgren goes further. Noticing that in the oldest poetry book, the *Book of Odes*, *qù* tone words which belong to *rù* tone phonetic series (reconstructed with a final ***-g**) rhyme perfectly with *qù* tone words which belong to *píng* or *shăng* series, he generalized the reconstruction of ***-g** to all words belonging to these series. The result of this choice was to suppress all *-j* and *-w* finals in his Old Chinese system. The word 'cat' [貓 *máo*, K-MC *mi̯äo*, B-MC *mæw*], which sounds so clearly onomatopoeic, becomes Old Chinese ***miog** (GS 1159c)! Even from the point of view of the rhyme patterns of the *Book of Odes*, the resulting reconstruction is not satisfactory, as Karlgren is led to conclude that his ***-o** and ***-âg** rhymed with each other, and to postulate a dialectal evolution in the most archaic of all classical texts (Karlgren 1940: 31).

65. In the Tang period, according to the rhyme dictionaries, Chinese only had four tones. The loss of the voicing contrast doubled the number of tones from the 9th century on. See Martin (1953: 10–13) for a clear overview.

Vietnamese data allow for the formulation of a different hypothesis. The oldest layer of Chinese loanwords into Vietnamese shows a correspondence between the Chinese *qù* tone and the Vietnamese *hỏi* and *ngã* tones (Haudricourt 1954), as recapitulated in Table 18.

[Table 18. Examples illustrating the correspondence between the Chinese *qù* tone and the Vietnamese *hỏi* and *ngã* tones [tones 5 and 6; etymological C1 and C2].]

char.	[gloss]	GS n°	K-OC	K-MC	[B-MC]	Vietnamese
寄 *jì*	to send, give	1x	kia	kjiᵉ̯ᴴ	kjeH	*gởi* [ɣɤj^{5}]
義 *yì*	justice	2ru	ngia	ngjiᵉ̯ᴴ	ngjeH	*nghĩa* [ŋiə^{6}]
露 *lù*	dew	766t′	glâg	luoᴴ	luH	*lõ* [lɔ^{6}]
訴 *sù*	to express	792j	sâg	suoᴴ	suH	*tỏ* [tɔ^{5}]
墓 *mù*	grave, tomb	802f	mâg	muoᴴ	muH	*mả* [ma^{5}]

Maspero (1912: 102) showed long ago that these Vietnamese tones come from a final fricative ***-h** originating in a yet earlier ***-s**.

[3.5 Reconstructing a suffix *-s in Old Chinese]

Let us now postulate the existence of a suffix ***-s** in Old Chinese. This suffix could be added to any word: to words with a vocalic or nasal rhyme (the *píng/shăng* series), or with a stop-final rhyme (the *rù* series). In the latter case, the finals ***-ks**, ***-ts**, ***-ps** changed to ***-gs**, ***-ds** and ***-bs**, and then to ***-js,** ***-ws** and ***-s**. It seems that rhymes in ***-ks** became vocalized in ***-ws** / ***-js** as early as Old Chinese, as they rhyme in the *Book of Odes* with rhymes ending in ***ws** and ***js** which result from the addition of the ***-s** suffix to words with ***w** and ***j** finals.

The ***-ts** finals, for their part, survived longer. One of the characters of the 12-year cycle [viz. the eighth, 未 *wèi*], hence belonging to the earliest layer of loanwords into Tai (and the only one of these names to be in the *qù* tone) is reconstructed by Karlgren as GS 531 ***mi̯wəd** > *mjwi̯ᴴ* [B-MC: *mjɨjH*]. This word is found in Tai, Khmu and Lamet as **mot**. According to our hypothesis, we would reconstruct ***mwəts** or ***mots**. The lenition of the ***-ts** finals must have occurred shortly before the Tang period, as during that

period we find a whole series of finals that only appear in the *qù* (not *píng* or *shăng*) tone: AA *-âi, -uâi , -iʷäi, -iʷei*, etc.[66]

The function of this ***-s*** suffix still needs to be determined. In my opinion we should avoid the excessively "mentalist" approach of Karlgren, when he comments:

> We have the word pair âk 'bad' and âg 'to hate'. Here the voiced final -g indicates a verb. But we also had the word pair dhâk 'to measure' and dhâg 'a measure'. Here the voiced final indicated a noun. (Karlgren 1949: 96)

The best way to resolve this issue is not to start out by trying to determine which "part of speech" a given word would belong to according to its translation into a European language. What matters is to identify the word's basic meaning [and to distinguish it from the derived meaning]. Thus, in my opinion, the pairs of words mentioned by Karlgren in the above citation participate in a system of derivation, as set out in Table 19.

[Table 19. Reconstructed OC word-pairs showing basic (underived) and *s-suffixed forms]

basic		with suffixed *-s	
*âk	bad, evil	*âks	to hate
*xâu	good	*xâus	to love
*dâk	to measure	*dâks	a measure
*ṣi	to send	*ṣis	envoy

Karlgren's pair 'evil/to hate' is thus parallel to the pair 'good/to love', and his pair 'to measure/a measure' parallel to 'send/envoy'. By not taking [Middle Chinese] tones into account in his reconstruction [of Old Chinese], he fails to distinguish 'good' from 'to love' (GS 1044 OC ***xôg***) or 'to send' from 'envoy' (GS 975n OC ***ṣl̂iəg***), proposing a single Old Chinese reconstruction for each of these pairs.[67]

66. [B-MC: -ajH (rhyme category: 泰 *tài*), -wajH (夬 *guài*), -jejH (祭 *jì*), and -jojH (廢 *fèi*).]

67. [In GS, Karlgren omitted tone marks from his MC and Mandarin transcriptions, on the grounds that he had not used the MC tones in reconstructing OC (Karlgren 1940:15). MC reconstructions of the words cited appear with tone marks in AA (Karlgren 1923). Karlgren's transcriptions of these words are listed in the following table.

character	gloss	GS	K-OC>K-MC	AA	K-MC	B-MC
惡 *è*	bad	805h	ˑâk > ˑâk	209	ˑaˀ < *ˑag	ak

惡	*wù*	to hate	805h	·âg > ·uo	209	·uoʾ < *·uog	ʔuH
好	*hǎo*	good	1044a	xôg > xâu	1089	˂xâu	xawX
好	*hào*	to love	1044a	xôg > xâu	1089	xâuʾ	xawH
度	*duó*	to measure	801a	dhâk > dhâk	1128	dhâk	dak
度	*dù*	a measure	801a	dhâg > dhuo	1128	dhuoʾ < *d^huog	duH
使	*shǐ*	to send	975n	slị̂əg > ṣi	526	˂ṣi	sriX
使	*shǐ*	envoy	975n	slị̂əg > ṣi	526	ṣiʾ	sriH

Karlgren restored the tonal notation of MC and Mandarin in the revised edition of GS (Karlgren 1957) but did not use them in reconstructing OC. He was convinced that OC had had tones, but believed that they could not be reconstructed on existing evidence. See Comment.]

Comments

This article is one of the most important of Haudricourt's contributions to linguistic science. Together with his article on the origin of tones in Vietnamese, also published in 1954 ("The origin of tones in Vietnamese", this volume), it provides the foundation for the theory of tonogenesis in East Asia. The application of this theory to Chinese not only explains the origin of tones, but clarifies several rhyming patterns found in the *Book of Odes* (the oldest rhymed text in Chinese), where words having a final stop rhyme with words in departing tone, (去聲 *qùshēng*) according to their Middle Chinese pronunciation. Haudricourt's theory that the departing tone comes from *-s (see "The origin of tones in Vietnamese", this volume) explains this phenomenon: the words with departing tone rhyming with words in final stop should be reconstructed with a final cluster *-ks, *-ts or *-ps. Besides, Haudricourt's theory of tonogenesis allows for the reconstruction of several *-s suffixes (in particular a nominalizing suffix) which can be shown to be cognate with those found in conservative languages of the Sino-Tibetan family, such as Tibetan.

This theory has been accepted by nearly all specialists in Chinese historical phonology, apart from some scholars in China who hold fast to the traditional views of Chinese philology.

A second major finding reported in this article consists in the hypothesis that there were labiovelars in Old Chinese. This idea was used later by Yakhontov (1960a) and Pulleyblank (1962) to revise the reconstruction of the Old Chinese vowel system, and is the basis for the six-vowel system that is common to the recent systems of Starostin (1989), Baxter (1992) and Zhengzhang (2003).

On the other hand, two of the ideas proposed by Haudricourt have been disproved since the time of publication.

First, in his discussion of the Middle Chinese initials *dzy- and *zy-, he argues that the contrast is only due to dialect mixture. However, we now know that (i) the phonetic reconstruction of these initials in Middle Chinese by Karlgren was erroneous: it was based on a wrong interpretation of the philological evidence; and (ii) these two initials have entirely distinct origins in Old Chinese, and reflect a genuine contrast, though this contrast is poorly preserved in modern dialects.

Second, his hypothesis concerning the existence of fricatives, which he set up to solve the issues raised by Karlgren's reconstruction of voiced stops in Old Chinese, is now known to be a dead end. Yakhontov (1959) and Pulleyblank (1962) put forward another interpretation, which is now accepted by most scholars: that Karlgren's ***d-** should instead be reconstructed as a lateral. (A clear account of the necessity of reconstructing laterals in Old Chinese is provided by Baxter 1992: 196–199.)

Guillaume Jacques

Non-final version (November 9th, 2017). In preparation for: Haudricourt, André-Georges. Evolution of languages and techniques*. (Ed.) Martine Mazaudon, Boyd Michailovsky & Alexis Michaud. (Trends in Linguistics. Studies and Monographs [TiLSM] 270). Berlin: De Gruyter Mouton.*

Appendix: Transcriptions used in the translation

Haudricourt's general practice is to cite forms as they appear in his sources, but a few exceptions, some motivated by typographical constraints, are mentioned below. In his own reconstructions, unlike Karlgren, he uses *y* for yod [IPA **j**]. We have systematically revised Haudricourt's *y* to *j* in his own reconstructions and in a few citations. However, *y* does appear with the value of IPA **j** in Pinyin transcriptions of Mandarin and in Baxter's reconstructions (see below). In transcriptions in this translation, *j* is never used for IPA **ʤ**, **ʣ** (etc.), and the IPA front rounded vowel **y** does not appear. Haudricourt's phonetic transcription δ has been replaced by its IPA equivalent **ð**.

Karlgren's transcriptions

Haudricourt cites character numbers and their reconstructed pronunciations from Karlgren's *Analytic Dictionary* (AA, 1923) and *Grammata Serica* (GS, 1940[68]) (two separate numberings). Each of these dictionaries gives three pronunciations for each character:
AA: modern Mandarin/Cantonese/Ancient Chinese
GS: Archaic Chinese/Ancient Chinese/modern Mandarin.
Note that in the translation, "Old Chinese" (OC) is used for "Archaic Chinese" and "Middle Chinese" (MC) for "Ancient Chinese".
Haudricourt cites OC and MC forms using Karlgren's transcriptions, except that he replaces the apostrophes used to indicate aspiration of stops by *h*, thus: *kh, gh, tsh…* instead of *k', g', t's…* (AA) or *k', g', ts'…* (GS) (IPA: **kʰ**, **gʱ**, **tsʰ**…). We have followed him in this, and in writing *x* for Karlgren's *χ* (AA only) (IPA: **x**). Karlgren's half-circle notations (see below) for the tones *píng*, *shǎng*, and *qù* have been restored; Haudricourt used opening and closing single quotation marks for the latter two, and left the *píng* tone unmarked. Haudricourt or his typesetters made a few other minor changes

68. Present-day readers are more likely to be familiar with the revised edition of this work (GSR, Karlgren 1957), which has the same character numbering but slightly different transcription conventions.

or simplifications without incidence for the linguistic argument; we have restored Karlgren's notations without comment in many cases.
In the introductions to his dictionaries Karlgren explained the conventions adopted for the notation of his Middle and Old Chinese reconstructions in articulatory terms. We can summarize some of the possibly unfamiliar conventions as follows (aspirated consonants are not listed):

	AA	GS	IPA
dental affricates	ts, dz		ʦ, ʣ
supradental (retroflex) stops, fricatives	t, ḍ, ṣ, ẓ		ʈ, ɖ, ʂ, ʐ
supradental (retroflex) affricates	ʦ̣, ʣ̣	ʦ̣, ʣ̣	ʈʂ, ɖʐ
palatal stops	t′, d′	t̂, d̂	c, ɟ
palatal affricates	t′ś, d′ź	tś, dź	ʨ, ʥ
palatal fricatives, nasal	ś, ź, ń		ɕ, ʑ, ɲ
yod ("sonant prepalatal fricative")	j		j
velar fricatives, nasal	χ, ɣ, ng	x, ɣ, ng	x, ɣ, ŋ
glottal initial	˙ (dot)	˙	ʔ

Concerning vowels, here are explanations provided in AA (pp. 6-7), with IPA equivalents.

symbol	Karlgren's explanation	interpretation in IPA
i, e, ä, ö, o, u	as for instance in German	i, e, ɛ, ø, ɔ, u
a	*a* "aigu", an open *a*, French p*a*tte	a
â	*â* "grave", French p*â*te (not considering the length of the vowel)	ɑ
å	something like Engl. l*aw*	ɒ
ɒ	Engl. b*u*t	ʌ
ə	Germ. Gab*e*	ə
ŭ	a very open *u*, with the tongue position of *ə* and half narrow labialisation	ɵ
ʷ	a subordinate labial vowel in Anc. Chin.	ʷ

The difference between back and front "a"s, which is symbolized in the IPA by script [ɑ] vs typographic [a], was marked by diacritics both by Karlgren and by Haudricourt, distinguishing *â* from *a* in Karlgren and in

Haudricourt 1954, and distinguishing **â** from **a**, in the 1972 reprint and the present translation.

For Middle Chinese tones in AA, Karlgren used half-circles before and after the syllable: *píng* (平) tone *ˌba*; *shǎng* (上) tone *ˈba*; *qù* (去) tone *ba˺*. No tone mark is used (or needed) on syllables with stop-finals; the latter suffice to identify the *rù* (入) tone. No tones are reconstructed for OC.

Haudricourt often cites Karlgren's OC and MC forms together, with a GS identifier, as [OC form] > [MC form]. Karlgren omitted tonal notation from the MC reconstructions in GS. In cases where the MC forms have tone markings, we assume that these are in fact from AA. Note that no asterisks appear in Haudricourt's article; the present translation follows both Karlgren and Baxter in marking OC, but not MC forms, with asterisks.

Baxter's transcriptions

Reconstructions/transcriptions of Middle Chinese from Baxter 1992, the first complete system published since Karlgren, have been added to the translation to give a point of comparison with more recent work. Although both systems are based on the same written sources, their objectives, and their views concerning the principal source for MC, the *Qièyùn*, are somewhat different. Where Karlgren claimed to present a strictly phonetic transcription of the dialect spoken in the Tang capital (1957:4), Baxter (1992: 26-31) aimed at a phonologically structured representation of the distinctions made in the *Qièyùn*, which many believe were somewhat more numerous than would have been found in any single dialect, and of whose precise phonetic values he claimed less certainty. The following illustrate Baxter's transcription (which was designed to be typeable on an American keyboard[69]) with approximate IPA equivalents:

aspiration: *h* (*e.g. ph, th*, etc.: IPA **pʰ**, **tʰ**, etc.)
palatal sibilants, affricates: *Cy* (*e.g.* sy, *zy*, *tsy*; IPA **ɕ**, **ʑ**, **tɕ**)
retroflex stops, sibilants, affricates: *Cr* (*tr, sr, tsr*; IPA **ʈ**, **ʂ**, **ʈʂ**)
yod: initial: *y*; medial: *j*; IPA: **j**.

69. Karlgren (1954:366) famously characterized phonemic analysis, and notations adapted to the typewriter, as a "craze".

laryngeal: *x* (voiceless), *h* (voiced); IPA x ~ h, ɣ ~ ɦ.
tones: capital letters at the end of open syllables: *píng* (none); *shǎng* –X, *qù* –H; *rù* (stopped) tone is indicated by syllable-final *p*, *t*, *k*.

Pinyin (Mandarin)

For modern Mandarin, Haudricourt used the now obsolete French transcription of Demiéville 1953, which we have abandoned. In the translation, we have systematically provided Mandarin transcriptions, in the standard Pinyin romanization, alongside all Chinese characters cited, as an aid to identification and reference. But because the correspondence of Pinyin to standard phonetic categories is not transparent, we have also provided IPA segmental equivalents (see below) for the few words whose Mandarin pronunciation is under discussion.
Pinyin can be recognized by the fact that it appears in italics, without an asterisk, and has tone marks over the vowels (macron for Mandarin tone 1, acute accent for tone 2, haček for tone 3, grave accent for tone 4). In Pinyin, voiceless unaspirated stops (IPA: p, t, k, ts, tʂ, tɕ) are represented by *b, d, g, z, zh, j*, and voiceless aspirated stops (IPA: pʰ, tʰ, kʰ, tsʰ, tʂʰ, tɕʰ) by *p, t, k, c, ch, q*. Coronal fricatives (IPA: s, z, ɕ, ʂ, ʐ) are represented by *s, z, x, sh, zh*. Yod is represented by *y* (IPA: j).
For modern Cantonese, we have followed Haudricourt in using Karlgren's transcriptions (but with h for aspiration as above).

The editors

References

Benedict, Paul K. 1939. Semantic differentiation in Indo-Chinese. *Harvard Journal of Asiatic Studies* 4. 213–229.

Benedict, Paul K. 1948. Archaic Chinese g and d. *Harvard Journal of Asiatic Studies* 9. 197–206.

Chao, Yuen-ren. 1941. Distinctions within Ancient Chinese. *Harvard Journal of Asiatic Studies* 5. 223–227.

Demiéville, Paul. 1946. Review of *The Chinese Language* (Forrest 1948), *Bulletin de la Société de Linguistique de Paris* 43(2) (book reviews). 267–281. [Editors' notes: the publication date indicated on the *Bulletin* volume in which this review was published (1946) is earlier than that of the reviewed book (1948); this is a consequence of the disruption of the publication of the *Bulletin de la Société de Linguistique de Paris* during the war and its aftermath.]

Demiéville, Paul. 1953. *Matériaux pour l'enseignement élémentaire du chinois* [Materials for elementary Chinese teaching]. Paris: Adrien Maisonneuve.

Emeneau, Murray Barnson. 1951. *Studies in Vietnamese Grammar*. Berkeley: University of California Press.

Forrest, Robert A.D. 1948. *The Chinese Language*. "The Great Languages" series. London: Faber and Faber.

Haudricourt, André-Georges. 1946. Restitution du karen commun [The reconstruction of Proto-Karen]. *Bulletin de la Société Linguistique de Paris* 42. 103–111.

Haudricourt, André-Georges. 1948. Les phonèmes et le vocabulaire du thai commun [The phonemes and vocabulary of Common Thai]. *Journal Asiatique* 236. 197–238.

Haudricourt, André-Georges. 1949a. La conservation de la sonorité des occlusives du thai commun dans le dialecte de Cao-bang [The preservation of voicing of Common Thai stops in the Tho dialect of Cao-bang]. *Actes du XXI^e^ Congrès des Orientalistes*. Paris. 251–252.

Haudricourt, André-Georges. 1950. Les consonnes préglottalisées en Indochine [Preglottalized consonants in Southeast Asia]. *Bulletin de la Société Linguistique de Paris* 46. 172–182.

Haudricourt, André-Georges. 1953. A propos de la restitution du karen commun [More on the reconstruction of Proto-Karen]. *Bulletin de la Société Linguistique de Paris* 49. 129–132.

Haudricourt, André-Georges. 1954. De l'origine des tons en vietnamien [The origin of tones in Vietnamese]. *Journal Asiatique* 242. 69–82.

Izikowitz, Karl Gustav. 1951. *Lamet Hill Peasants in French Indochina*. Göteborg: Ethnografiska Museet.

Karlgren, Bernhard. 1915–1926. *Études sur la phonologie chinoise* [Studies in Chinese phonology]. Archives d'études orientales 15(1–4). Leiden: Brill.

Karlgren, Bernhard. 1923. *Analytic Dictionary of Chinese and Sino-Japanese*. Paris: Geuthner.

Karlgren, Bernhard. 1940. Grammata Serica: Script and Phonetics in Chinese and Sino-Japanese. *Bulletin of the Museum of Far Eastern Antiquities*12. 1–471.

Karlgren, Bernhard. 1949. *The Chinese language: an Essay on its Nature and History*. New York: Ronald Press.

Kennedy, George A. 1952. Voiced gutturals in Tangsic. *Language* 28. 457–464.

Li Fang-kuei. 1945. Some old Chinese loanwords in the Tai languages. *Harvard Journal of Asiatic Studies* 8. 333–342.

Lo Chʻang-pʻei 罗常培 (Luo Changpei). 1951. Evidence for amending B. Karlgren's Ancient Chinese -j- to -ɣ-. *Harvard Journal of Asiatic Studies* 14. 285–290.

Martin, Samuel Elmo. 1953. The phonemes of Ancient Chinese, Supplement to the *Journal of the American Oriental Society* 16.

Martinet, André. 1939. Rôle de la corrélation dans la phonologie diachronique [On the role of correlations in diachronic phonology]. *Travaux du Cercle Linguistique de Prague* 8. 273–288.

Martinet, André. 1952. Function, structure and sound change. *Word* 8. 1–32.

Maspero, Henri. 1912. Etudes sur la phonétique historique de la langue annamite : les initiales [Studies in Annamite historical phonetics: initial consonants]. *Bulletin de l'École Française d'Extrême-Orient* 12. 1–127.

Maspero, Henri. 1916. Quelques mots annamite d'origine chinoise [Some Annamite words of Chinese origin]. *Bulletin de l'École Française d'Extrême-Orient* 16. 35–39.

Maspero, Henri. 1920. Le dialecte de Tch'ang-Ngan sous les T'ang [The Chang'An dialect during the Tang dynasty]. *Bulletin de l'École Française d'Extrême-Orient* 20. 1–124.

Maspero, Henri. 1930. Préfixes et dérivation en chinois archaïque [Prefixes and derivation in Old Chinese]. *Mémoires de la Société de Linguistique de Paris* 23. 313–327.

Roux, Henri & Tran Vân-Chu. 1927. Les Tsa Khmu [The Tsa Khmu people]. *Bulletin de l'Ecole française d'Extrême-Orient* 27. 169–222.

Shafer, Robert. 1940. The vocalism of Sino-Tibetan. *Journal of the American Oriental Society* 60(3). 302–337.

Shafer, Robert. 1944. Problems in Sino-Tibetan phonetics. *Journal of the American Oriental Society* 64(3). 137–143.

Shafer, Robert. 1950. The initials of Sino-Tibetan. *Journal of the American Oriental Society* 70(2). 96–103.

Wulff, Kurt. 1934. *Chinesisch und Tai: sprachvergleichende Untersuchungen* [Chinese and Tai: comparative linguistic studies]. Det Kongelige Danske Videnskabernes Selskab. Historik-filologiske Meddelelser 20, 3. Copenhagen: Levin & Munksgaard.

Additional references (mentioned in the critical apparatus)

Baxter, William H. 1992. *A Handbook of Old Chinese phonology*. Berlin: Mouton de Gruyter.

Branner, David P. (ed.). 2006. *The Chinese rime tables, linguistic philosophy and historical-comparative phonology*. Amsterdam: John Benjamins.

Ferlus, Michel. 1982. Spirantisation des obstruantes médiales et formation du système consonantique du vietnamien [The spirantization of medial obstruents in Vietnamese, and the development of the Vietnamese consonant system]. *Cahiers de linguistique Asie Orientale* 11(1). 83–106.

Ferlus, Michel. 1992. Histoire abrégée de l'évolution des consonnes initiales du vietnamien et du sino-vietnamien [A short history of the evolution of Vietnamese and Sino-Vietnamese initial consonants]. *Mon-Khmer Studies* 20. 111–125.

Ferlus, Michel. 2009. What were the four divisions of Middle Chinese? *Diachronica* 26(2). 184–213.

Haudricourt, André-Georges. 1949b. L'origine des particularités de l'alphabet vietnamien. *Dân Việt-Nam* 3. 61–68. Translated as: The origin of the peculiarities of the Vietnamese alphabet. *Mon-Khmer Studies* 39. 89–104 (2010).

Karlgren, Bernhard. 1954. Compendium of Phonetics in Ancient and Archaic Chinese. *Bulletin of the Museum of Far Eastern Antiquities* 26. 211–367.

Karlgren, Bernhard. 1957. Grammata Serica Recensa: Script and Phonetics in Chinese and Sino-Japanese. *Bulletin of the Museum of Far Eastern Antiquities* 29. 1–332.

Lu, Zhiwei 陸志韋. 1947. *Gǔyīn Shuōlüè* 古音說略 [Phonology of Ancient Chinese]. 北平：哈佛燕京學社华书局 (Beiping : Hafo Yanjing xue she).

Luo, Yongxian. 2008. Sino-Tai and Tai-Kadai: another look. In A. Diller, J.A. Edmondson & Y. Luo (eds.), *The Tai-Kadai languages*. 9–28.. Routledge, London.

Nguyễn Khắc-Kham. 1971. Influence of Old Chinese on the Vietnamese language. *Area and Culture Studies* 21. 153–181. Tokyo University of Foreign Studies.

Nguyễn Tài Cẩn. 1979. *Nguồn gốc và quá trình hình thành cách đọc Hán Việt* [Origin and development of Sino-Vietnamese]. Hà Nội: Nxb Khoa học Xã hội. Reprinted by Nxb Đại học Quốc gia Hà Nội (2002).

Pulleyblank, Edwin. 1962. The consonantal system of Old Chinese. *Asia Major* 9. 58–144, 206–265.

Pulleyblank, Edwin. 1984. *Middle Chinese: a Study in Historical Phonology*. Vancouver: University of British Columbia Press.

Ray, David Tryon. 1979. Sources of Middle Chinese phonology: a prolegomenon to the study of Vietnamized Chinese. M.A., Southern Illinois University (270 p).

Rhodes, Alexandre de. 1651. *Dictionarium Annamiticum Lusitanum et Latinum* [Annamite-Portuguese-Latin Dictionary]. Rome.

Starostin, Sergej A. (Старостин, Сергей Анатольевич). 1989. Реконструкция древнекитайской фонологической системы *Rekonstrukcija drevsnekitajskoj fonologicheskoj sistemy* [Reconstruction of the Old Chinese phonological system]. Moscow: Nauka.

Wang Li 王力. 1948. Hànyǔ Yuèyǔ Yánjiū 汉语越语研究 [Research on Chinese and Vietnamese]. *Lǐngnán xuébào* 岭南学报 9(1). 1–96.

Wang Lien-Tseng 王聯曾. 1957. Un dictionnaire phonologique des T'ang, le *Ts'ie yun corrigé et complété* de Wang Jen-hiu [A phonological dictionary of the Tang dynasty: Wang Jen-hiu's *Corrected and improved Qie-Yun*]. *T'oung Pao* 45. 51–150.

Yakhontov, Sergej E. (Яхонтов, Сергей Евгеньевич). 1959. Фонетика китайского языка I тысячелетия до н. э. (система финалей). Fonetika kitajskogo jazyka I tysjacheletjia do n.e. (sistema finalej) [Chinese phonetics in the 1st millennium BCE: the system of finals]. *Проблемы востоковедения* / *Problemy Vostokovedenija* 2. 137–147.

Yakhontov, Sergej E. (Яхонтов, Сергей Евгеньевич). 1960a. Фонетика китайского языка I тысячелетия до н. э. (лабиализованные гласные). Fonetika kitajskogo jazyka I tysjacheletjia do n.e. (labializovannye glasnye) [Chinese phonetics in the 1st millennium BCE: labialized vowels]. *Проблемы востоковедения* / *Problemy Vostokovedenija* 6. 102–115.

Yakhontov, Sergej E. (Яхонтов, Сергей Евгеньевич). 1960b. Consonantal combinations in Archaic Chinese. In *Papers presented by the USSR delegation at the 25th International Congress of Orientalists*. Moscow: Oriental literature publishing house.

Zhengzhang Shangfang 郑张尚芳. 2003. 上古音系 *Shànggǔ yīnxì* [The sound system of Old Chinese]. Shanghai: Shanghai Jiaoyu Chubanshe 上海教育出版社.