

Chinese loans in Old Vietnamese with a sesquisyllabic phonology¹

While consonant clusters, taken broadly to include presyllables, are commonly hypothesized for Old Chinese, little direct evidence is available for establishing the early forms of specific words. This essay examines a hitherto overlooked source: Old Vietnamese, a language substantially attested in a single document, which writes certain words, monosyllabic in modern Vietnamese, in an orthography suggesting sesquisyllabic phonology. For a number of words loaned from Chinese, Old Vietnamese provides the only testimony of the form of the Vietic borrowing. The small list of currently known sesquisyllabic words of Chinese origin attested in this document includes examples of both words with a secure initial Chinese cluster and words with plausible Vietic-internal prefixation.

Keywords: Old Chinese language, Old Vietnamese language, historical reconstruction, sesquisyllabic words, prefixal morphology.

1. Context

Old Vietnamese² is a Vietic language of which substantial attestation is limited to a single document, the 佛說大報父母恩重經 *Phật thuyết Đại báo phụ mẫu ân trọng kinh* (*Fóshuō Dàbào fùmǔ ěnzhòngjīng*, “Sūtra explained by the Buddha on the Great Repayment of the Heavy Debt to Parents”, henceforth *Đại báo*). The language of this document preserves Proto-Vietic sesquisyllabic phonology. In Old Vietnamese, we find words of Chinese origin like 鏡 **s-kuwŋ* ‘mirror’, borrowed from Chinese 鏡 **kiaŋH*.³ This paper examines whether words like

¹ I would like to acknowledge the generous support of the European Research Council for supporting this research, under the auspices of ‘Beyond Boundaries: Religion, Region, Language and the State’ (ERC SynergyProject 609823 ASIA). This paper would have remained a mere idea, were it not for Huáng Shìqí and Masaaki Shimizu who took the trouble to send or borrow me the necessary literature. Nathan Hill gave precious advice and comments throughout the drafting of this essay. Finally, I would like to thank the anonymous reviewers, Laurent Sagart, Masaaki Shimizu, Đào Huy Linh, Guillaume Jacques, Alexis Michaud, Mark Alves, Nguyễn Minh-Châu, Wolfgang Behr, Trần Trọng Dương and Jonathan Smith for their comments which have considerably improved this essay.

² On the term *Old Vietnamese* see §1.3. In this essay, Old Vietnamese is transcribed with reconstructed initial, in roman type, and projected Modern Vietnamese rime in italic type. For example, **plòi* ‘heaven’ denotes that I reconstruct Old Vietnamese initial **pl-*, and that the Modern Vietnamese cognate of the word is *trời*.

³ Chinese (and Sino-Vietnamese) are transcribed in several different ways depending on the context. In proper names and terms of cultural nature, the Chinese is rendered in pinyin, in Sino-Vietnamese or both, depending on the specific context to which the term pertains.

Chinese sources of loanwords under consideration are transcribed in a modified version of Baxter’s Middle Chinese transcription (1992). Most notably, the transcription of this paper uses medial *-i-* instead of *-j-*, and the vowel *o* is written *ɔ*. Other sounds are transcribed in an IPA-like fashion: among Baxter’s alternative orthographies for vowels, *æ*, *ɛ*, *ɨ* are preferred to *ae*, *ea*, *+*. Retroflex stops are written *t̪*, *ʈh*, *q̪*, *ŋ̪*, retroflex sibilants *t̪s̪*, *t̪ʂh̪*, *d̪z̪*, *ʂ̪*, *z̪*, palatal sibilants *tɕ*, *tɕh*, *dɕ*, *ɲ*, *ɕ*, *ʐ*. Finally, we have *j* for *y*, *ɲ* for *ng*, *ʔ* for ‘ and *ɣ* for *h*.

Finally, Chữ Nôm orthography is transcribed first in Sino-Vietnamese followed by Middle Chinese transcription: 破散 (*phá tán* < **phaH sanX*). Unencoded Chữ Nôm characters are represented with Unicode ideographic

this support the reconstruction of initial consonant clusters in Old Chinese. More specifically, it investigates whether Chinese loans in Old Vietnamese that exhibit sesquisyllabic phonology confirm or challenge Baxter and Sagart's reconstruction of Old Chinese (2014), which employs data from Vietic languages (notably Vietnamese and Rục) to support Old Chinese clusters.

This section introduces the context of the paper, beginning with a general introduction (§1.1) to the difficulties of reconstructing Old Chinese initial consonant clusters. §1.2 discusses the importance of Vietic languages, which offer an important source of evidence for them in cases where they might have disappeared without a trace. The source *Đại bảo* is introduced in §1.3, and the phonetic interpretation of its sesquisyllabic orthography in §1.4.

1.1. Initial consonant clusters in Old Chinese

While Middle Chinese and contemporary Chinese dialects have a simple syllabic canon, scholars have long hypothesized that Old Chinese has a more complex phonotactics, especially on the left periphery of the syllable. Scholars from Henri Maspero (1930) reconstruct initial consonant clusters like *pr- or *sm-, while recent reconstructions (Pan Wuyun 2000, Baxter and Sagart 2014) also hypothesize a sesquisyllabic phonology with initial minor syllables. Baxter and Sagart (2014), for example, reconstruct Old Chinese forms such as *k.teʔ for 紙 *tʰeX “paper” and *mə.lat for 舌 *ʒet “tongue”.

In this paper, *par abus de langage*, we call both initial clusters like *kt- and the sequence of an initial minor syllable followed by the initial consonant like *mal- ‘consonant clusters’. I follow the convention of Baxter and Sagart (2014) in referring to the first components, such as *k- or *mə-, as *preinitials*, and distinguish them as consonantal (*k-) and syllabic (*mə-).

While various kinds of evidence point to consonant clusters in Old Chinese, it is very difficult for scholars to agree on the clusters to reconstruct for precise etymons. One of the reasons for this uncertainty is that, under the type of evolution that languages of China and Mainland Southeast Asia usually undergo, different phonotactic types of clusters have almost intrinsically different rates of survival in descendants. A case in point is Old Tibetan, in which most consonants can take preinitials *s-* and *d-*. However, the existence and identity of preinitials is much better preserved before sonorants than stops: in a typical modern Kham dialect (for example the dialects of Derge and Batang, cf. Skal·bzang 'Gyur·med and Skal·bzang Dbyangs-can, 2002), Old Tibetan *ng-*, *sng-* and *dng-* remain distinct as $_ŋ^4$, $_ŋ̃$ and $_ŋ$, but *k-*, *sk-* and *dk-* all merged into $_k$. If one were reconstructing Old Tibetan without the benefit of orthographic forms, it is likely that *ŋ-, *sŋ- and *dŋ- would be reconstructed correctly, but *k-, *sk- and *dk- would be very hard to distinguish.

This asymmetry in the preservation of clusters before obstruents and sonorants explains the curious situation of Old Chinese reconstructions, where there is a clear gradient to the levels of consensus among different reconstruction of different phonotactic types of OC clusters. For example, recent reconstructions of Old Chinese agree on either one of two theories about *sm- type initial consonants (see Mei 2012, Sagart & Baxter 2012). Given that there is an implicational hierarchy to the effect that languages with *sm- type initial consonants usually have *sk- type initial consonants (Goad 2011), Old Chinese likely had *sk- type initial consonants as well. However, authors cannot agree on which particular OC words have *sk- type consonants (see Gong & Lai 2017 for a brief account).

description characters such as 𠂇 or 𠂈. For example, the sequence “𠂇目它” represents a character composed horizontally of 目 and 它.

⁴ $_$ designates the high tone, and $_$ the low tone, in tonal modern dialects of Tibetan.

In order to better understand both the phonological structure of Old Chinese and the reconstruction of individual Old Chinese words, we need to search for more direct evidence of consonant clusters.

1.2. Does Vietic evidence support Old Chinese clusters?

From the very beginning of research on Vietnamese words of Chinese origin, it has been noted that certain words borrowed from Chinese exhibit lenition of initial consonants (Maspero 1912: 19–39, Wáng Lì 1948: 71). An example is Chinese 劍 **kiæmH* ‘sword’, which is borrowed as Vietnamese 劍 *guom* [ɣ-]. This contrasts with the absence of initial lenition in the Sino-Vietnamese pronunciation of the word, *kiếm*.

Vietnamese belongs to the Vietic branch of Austroasiatic languages. Modern Vietic languages exhibit the whole typological spectrum from monosyllabic Chinese-like Việt-Mường languages⁵ to completely sesquisyllabic languages such as the Chứt (Rục-Sách) cluster. Haudricourt (1965) first explained Vietnamese spirant initials like *v-*, *d-* [z- < ʔ-], *g-* [ɣ-] and *r-* as reflecting a process of lenition caused by lost Proto-Vietic presyllables. According to this theory (cf. also Thompson 1976: 1131–1133, Ferlus 1976, 1982), Proto-Vietic presyllables, still preserved almost intact in conservative Vietic languages like Rục, disappeared in all modern Việt-Mường languages. Their former presence triggered lenition in Mainstream (Hanoi-Saigon) Vietnamese, but not in Mường varieties nor the “Haut-Annam” (Maspero 1912) or “heterodox” dialects of Vietnamese, such as Vinh (Ferlus 1991) or Quảng Bình (Michaud, Ferlus and Nguyễn 2015).

Table 1: Lenition in Vietnamese

Proto-Vietic	Vietnamese	Mường ⁴	Rục ⁶
* <i>p-</i> ‘four’	<i>bốn</i> ([b-])	<i>pon</i> ᵛᵛ	<i>pó:n</i>
* <i>CVp-</i> ‘lime’	<i>vôi</i>	<i>pol</i> ᵛᵛ	<i>kəpu:l</i>
* <i>s-</i> ‘hand’	<i>tay</i>	<i>thăj</i> ᵛᵛ	<i>sɪ:</i>
* <i>CVs-</i> ‘snake’	<i>rắn</i>	<i>thăj</i> ᵛᵛ	<i>pəsí:n</i>

As predicted by this theory of the origin of Vietnamese lenition, for Chinese words just as for native words, conservative Vietic languages have a cognate with a prefix: Vietnamese 劍 *guom* [ɣ-] ‘sword’ is cognate to Rục *təkiam*. Pulleyblank (1981: 281–286), the first focused treatment of Chinese loans in Vietnamese showing lenition, considers the possibility that the preinitial was present in the original Chinese form ‘of great interest’ but ‘not easy to pursue further without additional information to enable one to determine the kind of clusters involved’.⁷

⁵ This essay follows the terminology of Hayes (1992), now in general usage. *Vietic* designates the larger group, which includes notably Chứt (Rục-Sách). *Việt-Mường*, a term which formerly often comprises all Vietic languages, designates only the languages spoken by the ethnic Vietnamese and Mường groups. For the terminological questions, cf. Michaud, Ferlus and Nguyễn (2015: 126).

⁶ Rục forms are cited from Nguyễn, Trần and Ferlus (1988). ‘Mường’ designate Khen Mường (Hòa Bình province), transcribed by Milton and Muriel Barker, and cited from Thompson (1976).

⁷ Maspero (1912: 21–23 et passim) deems Vietnamese words with lenition to be a ‘*dérivé récemment formé*’ of their counterparts without lenition in Mường and Vinh-type Vietnamese dialects. Wáng Lì (1948) includes lenition in what he calls the process of *Vietnamization* (越化, Việt hoá), which Nguyễn Tài Căn (1979) defines as “development under a different path than that of Sino-Vietnamese readings” (diễn biến theo một con đường khác với cách đọc Hán Việt).

Baxter and Sagart (2014) were the first to systematically use loanword material in Vietic languages to reconstruct consonant clusters in Old Chinese. By comparing the Vietic data with Chinese loans in the Kra-Dai language Lakkia, Baxter and Sagart (2014: 36–37, 93–97) give a compelling argument that preinitials in these words do come from the Chinese source. Their examples are 紙 **tɕeX* ‘paper’, borrowed as Vietnamese *giấy* [z-], Rục *kacáj*, Lakkia *khjei*³, and 賊 **dzak* ‘bandit’, borrowed as Vietnamese *giặc* [z-], Rục *kacák*, Lakkia *kjak*⁸. The oldest sizeable set of Chinese loans in Vietic and Lakkia, just as their counterparts in Tai and Kam-Sui, all result from the same event of massive borrowing which follows the Qín-Hàn conquest of the region corresponding to modern Guǎngdōng, Guǎngxī and Vietnam. For a Chinese word to receive a Vietic prefix and be then borrowed into Lakkia would be extremely implausible. Even if the number of correspondences is limited, we have here an important argument for the Chinese origin of the preinitials.

In the system of Baxter and Sagart (2014), Vietic preinitials are understood as reflecting genuine Old Chinese preinitials, thus 紙 **tɕeX* ‘paper’, Vietic **k-* < **k.teʔ*; 賊 **dzak* ‘bandit’, Vietic **k-* < **k.dzək*.

While there are a great number of words of Chinese origin with a softened initial in Vietnamese, only a few of them can be found in conservative Vietic languages as Rục. In order to further evaluate Baxter and Sagart's use of Vietic data in Old Chinese reconstruction, we need additional sources that attest to Chinese loanwords with preinitials.

1.3. Old Vietnamese and *Đại báo*

We find many cases where a word of Chinese origin shows a lenited modern reflex in Vietnamese, but is not attested in conservative Vietic languages such as Rục and Thavưng. We know that the Proto-Việt-Mường form must have contained a preinitial, but it is unclear which.

The preponderance of older Chinese loans in Vietnamese and Việt-Mường but not other Vietic languages is not surprising given the linguistic history of Vietic. As shown by Ferlus' work (2010) on the Khmer names of the duodecimal year cycle (地支 *dìzhī*, *Địa Chi* ‘earthly branches’), the Vietic branch of Austroasiatic had already split into subbranches by the time the early layer of Chinese words arrived. Chinese words entered the ancestor of today's Việt-Mường, and were then borrowed into other languages. It is not surprising that lexical Sinicization was followed by phonological Sinicization: the group that originally received the Chinese loans, which also contains the largest number of them, was then monosyllabized under the weight of Chinese influence. Middle Vietnamese, as attested in Alexandre de Rhodes' dictionary (1651), had already lost all sesquisyllabic preinitials, and retained only a few relictual initial clusters such as *bl-*.

This paper continues Baxter and Sagart's work (2014) on Vietic borrowings in Old Chinese by examining a remarkable document, which greatly improves our knowledge of Việt-Mường historical phonology. The document, 佛說大報父母恩重經 *Phật thuyết Đại báo phụ mẫu ân trọng kinh* (“Sūtra explained by the Buddha on the Great Repayment of the Heavy Debt to Parents”, henceforth *Đại báo*), is held in the *Société asiatique*, Paris. It is a version of a popular Chinese apochyphon more commonly known under the title 父母恩重難報經 *Fùmǔ Ēnzhòng Nánbào jīng*, *Phụ mẫu ân trọng nan báo kinh* (“Sūtra on the Difficulty of Repaying the Heavy Debt to Parents”), in which the Chinese text is accompanied by a vernacular translation (called 解音

Maspero, Wáng Lì and Nguyễn Tài Cẩn essentially postulate a non-Neogrammarian unconditional split, the condition of which is convincingly explained by Haudricourt, Thompson and Ferlus as Proto-Vietic presyllables. As a terminological convenience, we can understand *Vietnamization* as morphological prefixation within Proto-Vietic or Proto-Việt-Mường, which developed into Vietnamese lenition.

giải âm in Vietnam) in a rudimentary form of Chữ Nôm, where vernacular words are written with Chinese characters and modified versions thereof.

The language of the vernacular translation in *Đại báo* is clearly Việt-Mường. I propose to call this language Old Vietnamese, as the language reflected in *Đại báo* shows an important archaism: certain words that are monosyllabic in Modern Vietnamese are written with two Chinese characters (*digraphic orthography*), the first of which corresponding to a preinitial in conservative Vietic languages and proto-Vietic: *rắn* ‘snake’, Middle Vietnamese *rắn* ‘cobra; anguis’ (Rhodes 1651: 636), is written 破散 (*phá tán* < **phaH sanX*).

The traditional periodization schemes of Vietnamese distinguish a period *Archaic Vietnamese* (Nguyễn Đình-Hoà, 2009) or *Old Vietnamese* (*vietnamien ancien*, Ferlus, 2010). I believe that calling the language of the *Đại báo* is justified by the fact that there is a fundamental divide between, on the one hand, Archaic Vietnamese as traditionally defined as well as the language of *Đại báo*, and, on the other hand, Middle Vietnamese of *Annán Yiyǔ* (安南譯語, *An Nam Dịch Ngữ*) and Rhodes (1651). For example, ‘snake’ was borrowed from Archaic Vietnamese to Old Khmer as *masāñ* (Ferlus 2010: 9). The Archaic Vietnamese form clearly sides with the form in the *Đại Báo* **p-sǎn* 破散, and is distinguished from Middle and Modern Vietnamese *rắn*. Old Vietnamese, as defined in this essay, can be regarded as the last stage, uniquely attested, of Archaic/Old Vietnamese as hitherto defined in the literature.

Apart from the *Đại báo*, an extensive text, Old Vietnamese is attested in a much smaller scale in two sources: remnants in Modern Chữ Nôm usage (cf. §3.1) and proper names attested in Chinese-language inscriptions, most importantly the Hộ Thành Mountain (護城山, Ninh Bình province) inscription (Shimizu, Lê & Momoki 2005).

Đại báo was first brought to scholarly attention when a copy of it was sent to the Hán Nôm Institute in 1979; Nguyễn Ngọc San (1982) wrote about the digraphic orthography of *Đại báo*, which he interpreted as clusters and preglottalized consonants. Shimizu Masaaki (1996) made the first systematic study of *Đại báo*’s digraphic orthography, in which he collected a corpus of 24 glyph-word pairs exhibiting cluster orthography, as well as 47 glyph-word pairs exhibiting sesquisyllabic orthography. Hoàng Thị Ngộ’s candidate thesis (1996), later published as Hoàng (1999), is the first and only transcription of the text into modern Vietnamese.

Đại báo shows a great number of cases of digraphic orthography, where a word, monosyllabic in Modern Vietnamese, is written with two Chinese characters (*two-character orthography*), or in a composite character made of two different characters (*composite-character orthography*). Following Shimizu (1996), we classify Old Vietnamese digraphic orthography into two phonotactic classes:

- Consonant-liquid cluster orthography: The word transcribed is strictly monosyllabic and has a consonant-liquid initial cluster: CRVC; the first character transcribes the initial consonant C; the second character transcribes the medial RVC.

For example, the word *trời* ‘heaven’, MViet *blời* ‘*ceo; cælum*’ (Rhodes 1651: 45), is written 巴例 (*ba lệ* < **pa lieiH*) in *Đại báo*. I reconstruct Old Vietnamese **plòi*, the first character 巴 *ba* < **pa* transcribes the initial consonant **p*, the character 例 *lệ* < **lieiH* transcribes the remaining part of the syllable **lòi*.

- Sesquisyllabic orthography⁸: the word transcribed is sesquisyllabic C-CVC. The first character transcribes the presyllable C-; the second character transcribes the initial and the rime CVC.

⁸ This type is called *disyllabic construction* (双音节构造 *sō-onsetsu kōzō*) in Shimizu (1996).

For example, the word *rắn* ‘snake’, MViet *rắn* ‘cobra; anguis’ (Rhodes 1651: 636), is written 破散 (*phá tản* < **phaH sanX*). I reconstruct Old Vietnamese **p-sǎn*, cf. Rục *pəsí:n*. The first character 破 *phá* < **phaH* transcribes the preinitial **p-*, and the second character 散 *tản* < **sanX* transcribes the remaining part **sǎn*.

In this study, we focus our attention on the second kind of digraphic orthography, which relies on Vietic and Old Vietnamese sources. The first kind of digraphic orthography, which transcribe consonant clusters of the TR- type, will be deferred to another discussion along with other sources of evidence.

Concerning the dating of the *Đại Báo*, the document held in the *Société asiatique* itself was printed at the initiative of Trịnh Quán 鄭欖 around 1730 (Shimizu 1996: 84). The text conspicuously avoids the character 利 *lợi*, the name of the first emperor (c. 1384–1433) of the Lê dynasty (1428–1788). Mainly based on this taboo, Shimizu (1996, 2015) dates the text to the beginning of the Lê dynasty (XVth century). On the other hand, Nguyễn Tài Cẩn (2008) suggests that the text is likely to be copied from a Trần-dynasty (1225–1400) precedent, by pointing to a number of less conspicuous taboo practices in the text⁹.

A number of texts are dated, with more or less certainty, to the Trần dynasty. Trần Trọng Dương (2011) pointed out that, compared to known texts dated to the Trần dynasty, *Đại Báo* contains at least an order of magnitude more cases of two-character orthography. For example, compared with the text 指南玉音解義 *Chỉ nam ngọc âm giải nghĩa*, which shows the most cases of orthography in two characters apart from the *Đại Báo*, there are 16 cases of two-character orthography for a total length of 15000 characters; in the *Đại Báo*, on the other hand, Trần Trọng Dương counts 103 cases of two-character orthography for a total length of 4942 characters. Independently of Trần Trọng Dương's work, I checked the list of words showing sesquisyllabic orthography in Early Vietnamese texts given in Trần and Nguyễn (2007). Among the 62 examples, 50 are attested only in *Đại báo*. The result also shows that there is a qualitative difference between the language of *Đại báo* and that of later texts. As Nguyễn Quang Hồng (2008: 127–144) and Trần Trọng Dương (2011) observe, the text should be dated before Trần dynasty, likely to XIIth century.

1.4. The reconstruction of Old Vietnamese preinitials in this paper

The reconstruction of Old Vietnamese preinitials in this paper roughly follows that of Shimizu (1996). I reconstruct the following Old Vietnamese preinitials: **p-*, **t-*, **k-*, **ɕ-*, **s-*, ?-. We note that the list of Old Vietnamese preinitials coincides with the more numerous preinitials in Michel Ferlus's reconstruction of Proto-Vietic (2007): **p-*, **t-*, **k-*, **c-*, **s-*, **ʔa*. In the following list, I provide the spellers and corresponding Proto-Vietic initial for each Old Vietnamese preinitial.

Old Vietnamese **p-* < Proto-Vietic **p-*

Spellers: 波 (*ba* < **pa*), 巴 (*ba* < **pæ*), 破 (*phá* < **phaH*)

Example: **p-sǎn* ‘snake’ 破散 (*phá tản* < **phaH sanX*) = *rắn*, Middle Vietnamese *rắn* ‘cobra; anguis’ (Rhodes 1651: 636)

Vietic cognates: Rục *pəsí:n* ‘snake’

Old Vietnamese **t-*, cf. Proto-Vietic **t-*

Speller: 多 (*đa* < **ta*)

⁹ Trần Trọng Dương (2010) disagrees with Nguyễn Tài Cẩn's analysis, and takes the taboo characters under question to be normal graph variants at the period.

Example: *t-mai ‘shoulder’ 多埋 (*đa mai* < **ta mɛj*) = *vai*, Middle Vietnamese *bai* ‘ombros: humerus’ (Rhodes 1651: 65)

Vietic cognates: Thavung ʔap̚iːn / ʔap̚iːŋ ‘shoulder’

Old Vietnamese *k- < Proto-Vietic *k-

Spellers: 可 (*khả* < **khaX*), 𠂔 (*cá* < **kaH*)¹⁰

Example: *k-pó ‘remember’ 可汝 (*khả như* < **khaX nɿX*) = *nhớ*, Middle Vietnamese *dó* or *nhó* ‘lembarse; recordor’ (Rhodes 1651: 175)

Vietic cognates: Rục *kəp̚á*: ‘to remember’

Old Vietnamese *ɕ- < Proto-Vietic *c-

Spellers: 車 (*xa* < **tɕhæ*), 舍 (*xả* < **ɕæX*)

Example: *ɕ-mǎng ‘hear’ 車莽 (*xa mǎng* < **tɕhæ maŋX*), 舍莽 (*xả mǎng* < **ɕæX maŋX*) = *mǎng*, obsolete in Modern Vietnamese, but cf. Kiều 1.535 嚙信掣妥驚惶 *Mǎng tin xiết nỗi kinh hoàng* ‘How he was scared after hearing the news!’ Middle Vietnamese cf. *mǎng tin* ‘fama; fama’ (Rhodes 1651: 450)

Vietic cognates: Rục *caməŋ* ‘to hear, listen’

Note: The Chinese character 車 has two alternative readings: MC **kiɿ* = Mandarin *jū*, Sino-Vietnamese *cư*; MC **tɕhæ* = Mandarin *chē*, Sino-Vietnamese *xa*. Nguyễn Hữu Vinh et al. (2009: 764) and Trần Trọng Dương (2012) read 車 as *cư* < **kiɿ*, reconstructing *k-mǎng. This reading is to be rejected considering the alternative orthography and Vietic cognates.

Old Vietnamese *s-, cf. Proto-Vietic *s-

Speller: 司 (*tư* < **si*)¹¹

Example: *s-pui ‘merry, joyful’ 司盃 (*tư bôi* < **si pwi*) = Vietnamese *vui*, Middle Vietnamese *bui* ‘allegre; hilaris’ (Rhodes 1651: 74)

Vietic cognates: Rục *tupuːj* ‘merry, happy’

Old Vietnamese *ʔ-, comparable to Proto-Vietic *ʔa

Speller: 阿 (*a* < *a*)

Example: *a-pôi ‘early’ 阿盃 (*a bôi* < **a pwi*) = Vietnamese *vội* ‘hasty’, Middle Vietnamese *bội* ‘cousa apressada; properus’ (Rhodes 1651: 71).

(No known cognates in conservative Vietic languages)

Shimizu (1996) reconstructs *b- (or *p-), *ph-, *d- (or *t-), *l-, *ś-, *kh-, *s- and *ʔ-. My reconstruction of Old Vietnamese preinitials differs from his treatment in the following respects:

- Voiceless unaspirated stops, akin to other Vietic languages and reconstructed Proto-Vietic, are preferred to implosives.
- I do not admit the difference between *p- and *ph-. Although two different sets of spellers are used with bilabial stops, with unaspirated stops (波 *ba* < **pa* and 巴 *ba* < **pæ*) and with aspirated stops (破 *phá* < **phaH*), the same behaviour is seen in *k-, with

¹⁰ This character, a simplified form of 箇 (*cá* < **kaH*) ‘one’, is a graphical variant of other simplified forms of 箇 in currency in East Asia, such as Simplified Chinese 个 (*gè*) or Japanese ケ (*ka*). All derive from one half of the bamboo component 𥵹 in 箇.

¹¹ Nguyễn Ngọc San (1982) and Trần Trọng Dương (p.c.) propose that the speller 司 (*tư* < **si*) should be read as *t- instead of *s-. This is unlikely, since the sound change *t* < **s* in syllable initial did not happen before lenition, as *đ* < **t* gives Middle Vietnamese *d* [ð] in lenition, but *t* < **s* gives *r*.

unaspirated speller 𠂔 (*cá* < **kaH*) as well as aspirated speller 𠂔 (*khả* < **khaX*). In both cases, there is no pressing evidence to support the distinction, which does not conform to usual Austroasiatic phonological patterns, as in those of Khmer or reconstructed proto-Vietic.

- I do not consider the word 𠂔打 (*la đả* < **la tæŋX*) ‘rock’, which would be **l-tá* according to Shimizu (1996) and other studies, as a case of an Old Vietnamese sesquisyllabic word. It does not cause lenition in Modern Vietnamese: *đá* not **dá*; it survived in forms of Vietnamese later than Old Vietnamese: we find 羅砢 (first character *la* < **la*, second character Nôm with 多 *đa* < **ta* as phonetic component) in 國音詩集 *Quốc âm thi tập*; more strikingly, we find Middle Vietnamese *là đá* ‘*lagea de pedra, lapis planus*’ (Rhodes 1651: 390).

2. Some sesquisyllabic Old Vietnamese words borrowed from Chinese

In this section, I study some words of Chinese origin from *Đại báo*. I collect Old Vietnamese words in two-character orthography from two studies of the document (Shimizu 1996, Trần and Nguyễn 2007), and select the words given in the two studies for which the Chinese origin appears secure to me.

The *quốc ngữ* reading of the text of *Đại báo* is based on Hoàng Thị Ngọc (1999), who provides the only complete transcription so far available. I report all cases where my judgment differs from that of Hoàng Thị Ngọc (1999), Shimizu (1996) or Trần and Nguyễn (2007).

Three examples will be discussed not in this section, but in subsequent sections. Section 3 discusses 𠂔 **s-kuong* ‘mirror’, which bears important consequences on several aspects concerning Old Chinese reconstruction in general and Baxter and Sagart’s reconstruction (2014) in particular. Section 4 treats two words with Old Vietnamese preinitial *ʔ-, which could be a case of Vietic-internal prefixation.

2.1. 巴拭 **p-ɕíc* ‘to spread, to apply’ < 拭 **ɕik* ‘to wipe’

(1) 巴拭 仍 粉

**p-ɕíc* *những* *phấn*

spread PL powder

‘apply different kinds of powder’ (*Đại báo* 8a-1)

Original: (女人在世) 濃塗脂粉 ‘(Women live in this world,) apply themselves thickly with rouge and powder...’

I follow Trần and Nguyễn (2007, #37) in reading 巴拭 (*ba-thức* < **pa-ɕik*) as *xúc* ‘to anoint, to rub, to apply’, Middle Vietnamese *xúc* ‘*untar, ungir; ungo*’ (Rhodes 1651: 896). Hoàng Thị Ngọc (1999: 147) has *súc* ‘force’, likely merely an alternative spelling of the same word.

The Old Vietnamese reconstructs to **p-ɕíc*. The Old Vietnamese preinitial **p-* fails to cause lenition in **ɕ* and yields *x-* ([s]) in Modern Vietnamese. This behaviour is probably regular, and observed in **k-ɕa* > *xa* ‘far’ (§ 2.3) and *ʔ-*ɕuóng* > *xuóng* ‘chant’ (§ 4).

The Chinese original of 巴拭 **p-ɕíc* is 拭 **ɕik* ‘to wipe’ < OC (Baxter-Sagart) **ɕək*. The Old Vietnamese form supports an initial cluster **pɿ-* or **pəɿ-* in Old Chinese.

Under the Baxter-Sagart system, a syllabic preinitial is lost before **ɿ-* in pre-Middle Chinese, while a non-syllabic preinitial prevails over **ɿ-*: cf. 𣪠 **thwat* < **ɿʔot* < **mə-ɿʔot* ‘peel off’; 𣪠 **phen* < **phʰen* < **[pɿ]ʰen* ‘frank words’. The preinitial needs to be syllabic **pə-* under the

Baxter-Sagart system. Old Vietnamese data support a revised Baxter-Sagart reconstruction of 拭 **ɕik* ‘to wipe’ as **pə.lək*.

2.2. 多边 **t-pen* ‘edge’ < 邊 **pen* ‘edge, side’

(2) 庄 故 多边 揆

*chăng có *t-pen côi*

NEG have edge border

‘does not have edges or borders’ (Đại báo 44b-3)

Original: (所有) 無邊 (塵刹佛) ‘(all the Buddhas of lands as countless as specks of dust) without borders’

I read 多边 (*đa-biên* < *ta-pen*) as Modern Vietnamese *ven* ‘(river)bank; near’, cf. Middle Vietnamese *uen soũ* ‘borda do rio; ripa fluminis, melius’ (Rhodes 1651: 865)¹². My reading differs from Trần and Nguyễn (2007, #79) and Hoàng Thị Ngọc (1999: 147), both reading *biên*, the Sino-Vietnamese reading of the same etymon 邊 **pen* ‘edge, side’. Compared with *biên*, or other words deriving from the Chinese etymon, such as *bên* ‘side, edge’, the reading *ven* is preferable in order to account for the presyllable, which triggers lenition: Old Vietnamese **t-* should cause lenition of initial **p* into Vietnamese *v*. Also note the locution *ven côi* survived in the glossing register of Modern Vietnamese: in Thiều Chửu’s *Hán-Việt tự điển* (1942) the Chinese character 陲 **dʒwe* ‘frontier, border’ is glossed as *ven côi*.

The Old Vietnamese is reconstructed as **t-pen*. As is explained above, Old Vietnamese **t-p-* yields a lenited initial in Modern Vietnamese *v-*.

The Chinese original is 邊 **pen* ‘edge, side’, Baxter-Sagart **p^he[n]*. The rime correspondence, where Chinese *e*-like vowel is rendered as Vietnamese *e* [ɛ] puts the word among the oldest borrowings from Chinese to Vietnamese, cf. forms preserving Old Chinese *r-*: *sen* ‘lotus’ < **kr-*, borrowed from Chinese 蓮 **len*, Baxter-Sagart **k.[r]^he[n]*; *rèm* ‘curtain’, borrowed from Chinese 簾 **liem*, Baxter-Sagart **rem*.

Chiang Chia-lu (2011: 106) proposes another etymology for *ven*, namely from Chinese 緣/沿 ‘edge’ **jwen* < Baxter-Sagart **lon*. As she herself recognizes, it is highly improbable for Middle Chinese *j-* < Old Chinese **l-* to be rendered in Vietnamese by *v-*, an anomaly she explains as a late borrowing from an unspecified Southern Chinese dialect. Analysing the word as coming from 邊 **pen* ‘edge, side’, which is furthermore supported by the Chữ Nôm orthography 边, 邊 and 𠂔𠂔边, avoids these formal problems.

The Old Vietnamese form **t-pen* for the Chinese word 邊 **pen* ‘edge, side’ supports an initial cluster **tp-* or **təp-* in Old Chinese.

Under the Baxter-Sagart system, syllabic and non-syllabic preinitials exist with the same consonantism, which has different treatments in Late Old and Middle Chinese. A syllabic preinitial is lost before a voiceless stop, cf. 九 **kiuwX* < **[k]u?* ‘nine’, which could derive from **tə.ku?* (Baxter and Sagart 2014: 155). On the other hand, a non-syllabic alveolar preinitial prevails over a grave initial: cf. 𦵏 **tɕuwX* < **tu?* < **[t.p]ə?* ‘peel off’. The preinitial in 邊 **pen* needs to be syllabic **tə-* under the Baxter-Sagart system. Old Vietnamese supports a revised reconstruction under Baxter-Sagart system as 邊 **tə.p^he[n]*.

¹² Guillaume Jacques (p.c.) points to a difficulty in this etymology: the word is spelled *uen* in Rhodes, with the letter *v-* (probably [w-]) instead of *h* ([β-]), the predicted outcome of a *p-* with lenition. However, as Haudricourt (1974) notes, there is a fluctuation between letters *h* and *v* in Rhodes. For example, *bai* ‘ombros; humerus’ (Rhodes 1651: 65), annotated ‘alij vai’.

2.3. 可耶, 可賒, 可車 *k-ɕa < 賒 *ɕæ ‘far’

(3) 多 可耶

đi *k-ɕa

go far

‘go far’ (Đại báo 15b-1)

Original: 遠行 (憶念恩) ‘(the moral debt of parents missing you when you) travel far’

(4) 盎 那 強 可賒

áng nà cànɣ *k-ɕa

father mother more far

‘the parents become still more distant’ (Đại báo 36b-5)

Original: (及長大爲子索妻得他女子) 父母轉疎 ‘(Parents look for a wife for their son; after they obtained another woman,) the parents become on the contrary distant from their son.’

(5) 當 召 庄 可車

đáng chịu chǎng *k-ɕa

will suffer NEG far

‘will suffer it; it’s not far’ (Đại báo 43a-2)

Original: (三塗苦報) 將受非遙 ‘(We will suffer the bitter repayment in the three ways;) we will suffer it not far away’

The orthographies 可耶 (*khả da* < **khaX jæ*), 可賒 (*khả xa* < **khaX ɕæ*) and 可車 (*khả xa* < **khaX tɕhæ*) transparently spell the Modern Vietnamese word *xa* ‘far’, Middle Vietnamese *xa* ‘longe; distans’ (Rhodes 1651: 879), as Shimizu (1996), Trần and Nguyễn (2007, #90, #91) and Hoàng Thị Ngọc (1999: 153 et passim) observe.

The Old Vietnamese form of this word reconstructs to **k-ɕa*. The Old Vietnamese preinitial **k-* fails to cause lenition in **ɕ-* and yields *x-* [s] in Modern Vietnamese. This behaviour is probably regular and observed in **p-ɕiuc* ‘to spread, to apply’ > *xúc* ‘to spread, to apply’ (§ 2.1) and **?-ɕuóng* > *xuóng* ‘chant’ (§ 4). The orthography 可耶 (*khả da* < **khaX jæ*) might indicate an alternative pronunciation with lenition **k-ɕa*, which did not survive into Modern Vietnamese.

The Chinese original is 賒 **ɕæ* ‘far’¹³. The Chinese character, which contains the component 貝 ‘cowrie shell, wealth’, is generally used to write a word **ɕæ* that has the sense ‘buy or sell on credit, defer payment’. This economic sense must be original, as it is the only one attested in pre-Hàn and Hàn texts such as *Rites of Zhou* 周禮 *Zhōulǐ* and *Book of Han* 漢書 *Hànshū* (fascicles 24, 91 and 99), as well as the definition given in the *Shuōwén* 說文.

The sense ‘far’ is attested much later. Its first occurrence according to *Hànyǔ Dà Cídiǎn* (Luo Zhufeng et al. 1993) is in *Bàopǔzi* 抱樸子 (Book of the Master Who Embraces Simplicity), a text ascribed to 葛洪 Gě Hóng (283–343), who led his whole life in 句容 Jùróng, in the 江左 Jiāngzuǒ (Lower Yangtse Basin) region. The subsequent attestations are from Southern Dynasties literature.

The word 賒 **ɕæ* ‘far’ is likely to derive from 賒 **ɕæ* ‘to defer payment’, via the following semantic development ‘defer payment’ > ‘delay, postpone’¹⁴ > ‘late’¹⁵ > ‘far’. The word is likely

¹³ The Southern provenance of the Chinese word suggests another possibility, namely that the Chinese word is borrowed from Vietnamese. This possibility can be dismissed, as the original Vietic word for ‘far’ is Vietnamese *ngái* - reduced to a secondary synonym of *xa*, as in the synonymic locution *xa ngái* - and *Rục cənáj*.

¹⁴ Cf. Táo Qián 陶潛’s poem 和胡西曹示顧賊曹 (~ 400 CE): 悠悠待秋稼, 寥落將賒遲 ‘I’m slowly waiting for the autumn harvest, but the crops look meagre and poor; the harvest will be delayed long.’

a Southern dialect word which entered the literary language during the period of the cultural dominance of the Jiāngzuǒ region over the rest of China (roughly 317–619).

We finally note that the word 賒 **ɕæ* has a third sense in Literary Chinese, namely ‘many’, which might have been borrowed into Vietnamese as *xe* (Trần Trọng Dương, 2012b).

The late date and derived sense suggest that Old Vietnamese **k-ɕa* is borrowed from a Southern dialect form that corresponds with Middle Chinese *ɕæ*. Barring unexpected later prefixation, however, the **k* in the source dialect to the Vietnamese form can be projected back in Old Chinese, and supports an initial cluster **k̚l-* or **k̚l-* in Old Chinese. Under the Baxter-Sagart system, the Old Vietnamese data suggest that the reconstruction of 賒 **ɕæ*, currently **l̚A*, should be revised to **k̚l̚A*.

2.4. 𠂔正 **k-ciěng* < 井 **tsienX* ‘well’

- (6) 工 𠂔正 工 法 准 𠂔目它
*trong *k-ciěng trong bếp chôn đá*
 in well in stove place stone

‘to wells, to stoves or the milling-places’ (Đại báo 35a-3,4)

Original: (行來東西鄰里) 井竈碓磨 ‘(when the children) go to neighbours east and west, or to wells, stoves, pestles or mills’

I follow Trần and Nguyễn (2007: #45) in reading 𠂔正 (*cá-chính* < **kaH-tɕienH*) as *giếng* ‘well’, Middle Vietnamese *gyiếng* ‘*poço de agoa pera beber; puteus aquae ad potum*’ (Rhodes 1651: 283). Hoàng Thị Ngọc (1999: 165) reads *chiếng*, with the voiceless initial *ch-* based on the spelling 正 *chính* < **tɕienH*. This reading should be rejected for two reasons. First, no word *chiếng* ‘well’ exists in Modern Vietnamese, a fact indirectly admitted by Hoàng herself, who explicitly glosses this form as ‘*giếng*’. Second, Old Vietnamese **k-ciěng* regularly results in modern *giếng* by lenition.

The Old Vietnamese reconstructs to **k-ciěng*. Old Vietnamese **k-c-* yields lenited initial in Modern Vietnamese *gi-*. Compare for example 賊 **dzak* ‘bandit’, where Vietnamese *giặc* derives by lenition from a form akin to Rục *kac/ik*.

The Chinese original is clearly 井 *tsienX* ‘well’. The Old Vietnamese form supports an initial cluster **kts-* or **kəts-* in Old Chinese. Specifically, under the Baxter-Sagart system, the current reconstruction **C.tsenʔ* can be safely refined to **k.tsenʔ*.

2.5. 𠂔𠂔針 **k-kǎm??* <? 箴, 鍼, 針 **tɕim* ‘pin, needle’

- (7) 恒 肝 脾 𠂔 𠂔𠂔針
*dút gan ruột làm *k-kǎm??*
 break liver intestines make pin

‘breaks liver and intestines into pin-sized pieces’ (Đại báo 41b-5)

Original: 寸斷肝腸 ‘(the work of bringing children up) breaks the liver and intestines (of parents) into inch-sized pieces’

As is the case for ‘edge’, we have again the situation where we need to choose a reading among different Vietnamese words meaning ‘pin, needle’, with similar pronunciation, and which can all be written with Chữ Nôm characters based on the Chinese characters 箴, 鍼, 針 **tɕim*.

¹⁵ Compare the locution 賒促 **ɕæ-tshio̯k* ‘later or sooner, long or short’: 今古既異, 賒促不同 “Today is different from long ago; (naturally, whether the period of mourning is) long or short is different” (宋書 Book of Sòng, fascicle 15 禮制二 = 通典 Tōngdiǎn, fascicle 82).

Judging from dictionaries of later Chữ Nôm usage, the major possibilities are *găm*, *kim* and Sino-Vietnamese *châm*. Another word, *ghim*, is usually written 金 in later Chữ Nôm, but the orthography 鍼, 針 is not unimaginable.

Trần Trọng Dương (p.c.) suggests that the character 𠂔針 should be read as *trăm* ‘hundred’, with *làm trăm* interpreted as ‘make into a hundred pieces’. This is an unlikely interpretation, as *tr-* does not show any sign of being confounded with *ch-* until very late. In ‘modern’ Chữ Nôm, *trăm* is written in 𣎵 as 林 (*lâm* < **lim*) or with characters such as 𣎵 with the phonogram 林.

The Vietnamese words concerned has been hitherto analysed as deriving from Chinese 箴, 針 **tɕim* ‘pin, needle’, Baxter-Sagart **t.[k]əm*. However, Baxter and Sagart (2017) propose that the words *kim* and *ghim* derive rather from Chinese 金 **kim* ‘metal, bronze’, Baxter-Sagart **k(r)[ə]m*.

Trần and Nguyễn (2007: #45) interpret the composite character 𠂔針 as a case of digraphic orthography, composed of 𠂔 (*cá* < **kaH*) and 針 (*châm* < **tɕim*). Trần and Nguyễn's reading *châm* does not conform with their orthographical interpretation. If their orthographical interpretation is correct, the Modern Vietnamese reading would be one showing lenition, namely *găm* (or *ghim*), which would derive from Old Vietnamese **k-kăm*.

However, this reconstruction **k-kăm* is problematic on both internal and external grounds. Internally, sesquisyllabic languages typically do not admit the homorganic stops as the preinitial and as the initial: Ferlus' Proto-Vietic (2007) contains no syllables with **k-k-* or **t-ɕ-*. Externally, the Baxter-Sagart reconstruction **t.[k]əm* is supported by Lakkia *them*¹.

Another possible interpretation of the orthography is an early case of the *dấu cá* (the *cá* 𠂔 sign, cf. Nguyễn Quang Hồng, 2012), an orthographic device that indicates that the character is a specifically Nôm character, not to be read *à la chinoise*. If 𠂔針 is a case of *dấu cá*, all the non-Sino-Vietnamese readings, namely *găm*, *kim* (and probably *ghim*) are possible, with no implications on Old Chinese preinitials.

3. 𣎵 **s-kuong* < 鏡 **kiæŋH* ‘mirror’:

Old Vietnamese preinitials and Old Chinese morphology

In this section, I examine the Old Vietnamese word 𣎵 **s-kuong*, borrowed from Chinese 鏡 **kiæŋH* ‘mirror’, Baxter-Sagart **C.qraŋʔ-s*, which touches on several questions of interest in the reconstruction of Old Chinese. The Old Vietnamese form is reconstructed in §3.1. The reconstructed Vietnamese form enables the Baxter-Sagart form to be refined to **s.qraŋʔ-s*, which, however, reveals an internal inconsistency in the Baxter-Sagart reconstruction of Old Chinese, discussed in §3.2. The unique value of this loanword is discussed in §3.3, as, regardless of the phonological interpretation, this word offers an example of an Old Chinese **s*-stop cluster which is supported both by morphology with well-known Tibeto-Burman correlates as well as directly transcribed with *s-* in foreign borrowing.

3.1. Reconstruction of the Old Vietnamese form

(8) 𣎵 朗 補 工 架

**s-kuong* *sáng* *bỏ* *trong* *giá*

mirror clear put in stand

‘A clear mirror is put on a stand.’ (Đại báo 12a-5)

Original: 明鏡掩粧臺 ‘A clear mirror covers the dressing table.’

I agree with Hoàng Thị Ngọc (1999: 151) and Trần and Nguyễn (2007: #115) in reading 𪛗 (司+姜, *tu-cuong* < **si-kian*) as *guong* ‘mirror’, Middle Vietnamese *guong* ‘*espelho*; *speculum*’ (Rhodes 1651: 301). This reading is further supported by the fact that the glyph 𪛗 survived into later Chữ Nôm usage, spelling the word *guong* ‘mirror’.

The Old Vietnamese reconstructs to **s-kuong*. Old Vietnamese **s-k-* corresponds with lenited Modern Vietnamese *g-* [y-].

This word is borrowed from Chinese 鏡 **kiæŋH* ‘mirror’. The correspondence between Chinese *qù* 去 and Vietnamese *bằng* 平 (*ngang* and *huyền*) tones is characteristic of older loans (cf. Alves 2016: 271). The Vietnamese data supports a cluster **s(ə)k-* or **s(ə)q-* in Old Chinese.

3.2. Can we reconstruct Baxter-Sagart **s.qraŋʔ-s*?

On account of the Vietnamese initial lenition, Baxter and Sagart (2014) reconstruct **C.qraŋʔ-s* for this word. The Old Vietnamese data enables the refinement of this reconstruction into **s.qraŋʔ-s*.

This emendation, as attractive as it seems, is not compatible with the Baxter-Sagart reconstruction of Old Chinese as it currently stands. Initial **s.q(ʔ)-* is supposed to yield Middle Chinese *s-*, while **s.q(ʔ)r-* should yield Middle Chinese *tʃ-* (Baxter and Sagart 2014: 137–138). The amended reconstruction **s.qraŋʔ-s* would have become Middle Chinese **tʃiæŋH* and not *kiæŋH*.

Further examination reveals that there is a more general problem with **C.q-* clusters in the Baxter-Sagart reconstruction of Old Chinese. Baxter-Sagart **C-* stands for a preinitial that “cannot be identified because it has been lost in all the pronunciations under consideration” (Baxter and Sagart 2014: 168). According to this definition, whenever there is a **C-* in Baxter-Sagart, it can only stand for one of the preinitials reconstructed in the Baxter-Sagart OC phonology: **N-*, **m-*, **s-*, **p-*, **t-*, **k-*:

- **N.q-* and **m.q-* give Middle Chinese *j-* < **ɟ*;
- **s.q-* gives Middle Chinese *s-*;
- **p.q-* has no examples but should give Middle Chinese *p-*, cf. **p.k-* which gives *p-* and **p.q^h-* which gives *ph-*;
- **t.q-* gives Middle Chinese *tʃ-* < **t-*;
- **k.q-* probably did not exist, but would be, indeed, the only one that could give Middle Chinese *k-*.

Baxter-Sagart **C.q-* cannot give Middle Chinese *k-*, except in the unlikely case where **C* is **k*.

On the other hand, Baxter and Sagart (2014) explicitly reconstruct **C.q-* in order to account for the word 鏡 *kiæŋH* ‘mirror’. The velar initial in Middle Chinese *kiæŋH* and Vietnamese *guong* is explained by a sound change by which “that uvulars shifted to velars after a tightly attached nonnasal preinitial: **C*” (Baxter and Sagart 2014: 101).

Once we become cognizant of the fact that **C* stands for any initial in Baxter-Sagart’s notation, we notice that **C.q-* clusters are reconstructed twice, with different Middle Chinese reflexes. They are reconstructed once with specific preinitials, for which cases the preinitial prevails, and once in **C*, for which cases the Middle Chinese reflex is fronted *k-*. For a hypothetical Old Chinese **s.qraŋʔ-s*, the first reconstruction would yield Middle Chinese *tʃiæŋH*, while the second would give *kiæŋH*. How the two classes should be reconstructed in a Baxter-Sagart-like system remains a problem to be solved.

3.3. 鏡 **kiæŋH* ‘mirror’ as an example of Old Chinese **s*-stop cluster

Baxter and Sagart (2014: 101, 168) reconstruct an initial **q*- and a *qù*-tone (**-s*) derived from *shǎng*-tone (**-ʔ*) for 鏡 **kiæŋH* < **s.qraŋʔ-s* ‘mirror’, as they consider this word as being in the same family as 景, 影 **ʔiæŋX* < **qraŋʔ* ‘shadow’. The refinement of the reconstruction into **s.qraŋʔ-s* enables us to also refine its morphological derivation. Independently of the specific forms reconstructed, 鏡 **kiæŋH* ‘mirror’ is one of the strongest examples of **s*-stop clusters in Old Chinese.

A morphological prefix **s*- in Old Chinese has been hypothesized since Conrady (1896). As the recent argument between Mei (2012) and Sagart & Baxter (2012) shows, there is broad agreement on the fact that the function of **s*- would correspond to that in other Sino-Tibetan languages. However, due to different suppositions on the Middle Chinese reflexes of **s*-, there are few words where scholars can actually agree on the reconstruction of **s*-.

One of the well-attested functions of **s*- is that of deriving an instrumental noun. Rgyalrongic languages have a *sA*- prefix¹⁶, that notably derives instrumental nouns from the verb roots. In Zbu Rgyalrong, the prefix *se*- derives from *ke-twé?* ‘to open’ the instrumental noun *se-thwé?* ‘key’. The Zbu form belongs to a relic formation (Sun 2004: 293–294, Gong 2018: 211) which shows the age of prefix. Similarly, in written Tibetan, *s*- can derive instrumental nouns from verb roots: from *nod* ‘to receive’, *snod* ‘container, vessel’; from *nyan* ‘to listen’, *snyan* ‘ear (honorific)’ (cf. Zhāng Jichuān 2009: 246, Gong Xun 2017: 154–155, Jacques to appear). Baxter and Sagart (2014: 101) reconstruct **s*- with an instrumental value. A good example (Sagart & Baxter 2012: 50–51) is 鋤 **dzin* ‘hoe’, reconstructed as **s-l<r>a*, derived from 除 **dʒin* < **l<r>a* ‘remove’, as a hoe is an instrument of removal.

Based on the Old Vietnamese data and the Sino-Tibetan comparative evidence, we can reconstruct the noun 鏡 **kiæŋH* < **s.qraŋʔ-s* ‘mirror’ as derived from the verb 映 **ʔiæŋH* < **qraŋʔ-s* ‘to reflect’, which is itself a denominal from 景, 影 **ʔiæŋX* < **qraŋʔ* ‘shadow > reflection’. It is relevant to note that the formation has a parallel in the Late Middle Chinese/Early Mandarin period. During the Sòng dynasty, when the syllable *kiæŋH* was taboo, being the personal name of the grandfather of the first emperor 趙敬 Zhào Jìng (MC **dʒieuX kiæŋH*), mirrors were then called 照子 *zhàozǐ*, MC **tɕieuH-tsiX*. This formation is entirely parallel to what is proposed here for 鏡 **kiæŋH* < **s.qraŋʔ-s*, as 子 **zǐ*, MC *tsiX* was the dominant instrumental noun suffix, added to 照 *zhào*, MC **tɕieuH* ‘to shine on’, later ‘to reflect’.

The word 鏡 **kiæŋH* for ‘mirror’ is a Warring States period (475–221 BC) neologism replacing original 鑑 **kæmH*, coined inside Warring States period Chinese and not inherited from an older state of the language. It can be regarded as one of the safest examples both of a **s*-stop cluster in Old Chinese and an example of the instrumental value of the **s*- prefix in Old Chinese.

4. **ʔ*:- template-filling dummy prefix?

We find two words of Chinese origin showing the preinitial **ʔ*-, written 阿 (*a < a*):

阿路 **ʔ-lò* ‘kiln, oven’ < 爐 **lu* ‘oven’

(9) 阿路 堵 毒 坡律

**ʔ-lò dǒ dōc sôt*

kiln red poison burning

‘red kiln and burning hot poison’ (Đại Báo 29a-3,4)

Original: 炎爐熱毒 ‘kiln of fire and burning poison’

¹⁶ This prefix is, to my knowledge, first described in Lín Xiàngróng (1993: 162), cf. also Sun 1998: 142, Yán-mùchū 2005, Jacques 2008: 332–333, Lai 2017: 158 etc.

Both Trần and Nguyễn (2007, #6) and Hoàng Thị Ngọc (1999: 161) read 阿路 (*a-lộ* < **a-luH*) as *lò* ‘kiln, oven’, Middle Vietnamese *lò* ‘*fogão, forno; fornax*’ (Rhodes 1651: 417). This word is borrowed from Chinese 爐 **lu* ‘oven’ < Baxter-Sagart **[r]a*. Note that the word is not from the earliest layer of Chinese loans in Vietnamese, as the initial Middle Chinese **l-* < Old Chinese **r-*, is reflected as *l-* in this word but *r-* in the earliest Chinese loans in Vietnamese.

阿唱 **ʔ-ɣuóng* ‘to chant’ < 唱 **tɕhaŋH* ‘to chant’

(10) 省 阿唱 浪
*tiếng *ʔ-ɣuóng rằng*
 sound chant this

‘They chant as follows’ (Đại Báo 22b-4,5)

Original: 高聲唱言 ‘They loudly chant as follows’

I follow Shimizu (1996) in reading the Sino-Vietnamese *xuóng* ‘to chant’. Hoàng Thị Ngọc (1999: 158) reads *hét* ‘to cry’, usually written phonetically as 喝 (*hát* < **hat*) or 嗽 (*hiết* < **hiət*). This reading has no argument in its favour except that *xuóng* is infrequent in contemporary Vietnamese. Concerning this point, cf. *xuóng ten*, ‘*nomear alguém como por rol; vocare nomen alicuius tanquam è catalogo*’ (Rhodes 1651: 900), which should be emended into *xuóng tên* ‘chant the name-list’. This word is borrowed from Chinese 唱 **tɕhaŋH* ‘to chant’ < Baxter-Sagart **mə-tʰaŋ-s*.

Both borrowings show the preinitial **ʔ-*, which has not been hypothesized for Old Chinese. We also note that they show signs of rather late borrowing from Chinese: for **ʔ-ɣuóng* ‘chant’, the phonological correspondence is of the newest type, and completely coincides with the Sino-Vietnamese pronunciation. We further notice that in the So dialect of Thavung (Suwilai 1996), monosyllabic Thai-Lao verbs are borrowed with a prefix *ʔa-*: *lâ:k* ‘to drag along the ground’ is adapted as *ʔalâ:k*. For **ʔ-ɣuóng* ‘to chant’, at least, it is attractive to suppose a similar prefix **ʔ-*, a dummy prefix that fills a sesquisyllabic phonological template. The problem of the prefix **ʔ-* in Old Vietnamese and other Vietic languages deserves further study.

5. Conclusion

This paper is a preliminary examination of Old Vietnamese loans from Chinese attested with sesquisyllabic orthography in *Phật Thuyết Đại Báo Phụ Mẫu Ân Trọng Kinh*.

Old Vietnamese data is one of the most eloquent pieces of evidence that argues for the presence and determines the identity of **C₁(ə)C₂-* initial clusters in Old Chinese. It enables the proposal of a few changes to Baxter and Sagart’s reconstruction of Old Chinese.

- In two cases, where Baxter and Sagart reconstruct an unknown preinitial **C-* based on lenition in Modern Vietnamese, Old Vietnamese enables the identification of the unknown preinitial: 井 *tsienX* ‘well’, Baxter-Sagart **C.tsənʔ*, modified to **k.tsənʔ*; 鏡 *kiəŋH* ‘mirror’, Baxter-Sagart **C.qraŋʔ-s*, modified to **s.qraŋʔ-s*.
- In two cases, where Baxter and Sagart reconstruct Old Chinese forms without preinitials, Old Vietnamese suggests that a preinitial likely existed in Old Chinese: 拭 *ɕik* ‘to wipe’, Baxter-Sagart **lək*, modified to **pə.lək*; 邊 *pen* ‘edge, side’, Baxter-Sagart **pʰe[n]*, modified to **tə.pʰe[n]*.

The Old Vietnamese word 阿路 **ʔ-lò* ‘kiln, oven’ and 阿唱 **ʔ-ɣuóng* ‘to chant’ are possible examples of a prefix **ʔ-*, internal in Vietic, which is also attested in the conservative Vietic language So Thavung (Suwilai 1996).

The word 鏡 **s-kuəŋ* < 鏡 *kiəŋH* ‘mirror’, which permits a revised Baxter-Sagart reconstruction as **s.qraŋʔ-s*, is important in two different ways. First, it points to the fact, hitherto ignored, that **C.q-* clusters are reconstructed twice in the Baxter-Sagart system, with different reflexes in Middle Chinese and other sources. Second, this loan in Old Vietnamese provides one of the strongest examples both of an Old Chinese initial consonant cluster of the **s*-stop type, and of the Old Chinese prefix **s-*, which derives here an instrumental noun from a verb.

References

- Alves, Mark J. 2016. Identifying Early Sino-Vietnamese Vocabulary via Linguistic, Historical, Archaeological, and Ethnological Data. *Bulletin of Chinese Linguistics* 9: 264–295.
- Baxter, William H. 1992. *A Handbook of Old Chinese Phonology*. Berlin: Mouton de Gruyter.
- Baxter, William H., Laurent Sagart. 2014. *Old Chinese: A new reconstruction*. Oxford University Press.
- Baxter, William H., Laurent Sagart. 2017. Old Chinese reconstruction, a response to Schuessler. *Diachronica* 34(4): 559–576.
- Chiang, Chia-lu. 2011. *Yuènnán hànziyīn de lìshǐ céngcì yánjiū* (越南漢字音的歷史層次研究). Doctor thesis, National Taiwan Normal University.
- Ferlus, Michel. 1976. Du nouveau sur la spirantisation ancienne en vietnamien. *Bulletin de la Société de Linguistique de Paris* 71: 305–312.
- Ferlus, Michel. 1982. Spirantisation des obstruantes médiales et formation du système consonantique du vietnamien. *Cahiers de linguistique — Asie orientale* 11: 83–106.
- Ferlus, Michel. 1991. Le dialecte vietnamien de Vinh. *24th International Conference on Sino-Tibetan Languages and Linguistics, October 1991, Bangkok*. Available online at: <https://halshs.archives-ouvertes.fr/halshs-00922737>.
- Ferlus, Michel. 2007. *Lexique de racines Proto Viet-Muong (Proto-Vietic Lexicon)*. Ms., accessed from the *SEAlang Mon-Khmer Etymological Dictionary* <<http://sealang.net/monkhmer/dictionary/>>.
- Ferlus, Michel. 2010. Le cycle khmer des douze animaux : histoire d'un contact ancien entre Vietnam et Cambodge. *Cahiers de linguistique - Asie Orientale* 39: 3–19.
- Goad, Heather. 2011. The Representation of sC Clusters. In: Marc van Oostendorp, Colin Ewen, Elizabeth Hume, Keren Rice (eds.). *The Blackwell Companion to Phonology*: 898–923. Oxford: Blackwell Publishing.
- Gong, Xun, Yunfan Lai. 2017. Consonant clusters. In: Rint Sybesma et al. (ed.). *Encyclopedia of Chinese language and linguistics*: 665–672. Leiden: Brill.
- Gong, Xun. 2017. Jiāróng-yǔzǔ yǔyán yǔ Shàngǔ Hànyǔ de Xíngtàixué (嘉絨語組語言與上古漢語的形態學). In: Park Haeree, Chen Shaoxuan (eds.). *Gǔwénzì yǔ hànyǔ lìshǐ bǐjiào yīnyùnxué* (古文字與漢語歷史比較音韻學) [*Paleography and Chinese Historical and Comparative Phonology*]: 132–154. Shanghai: Fudan Daxue Chubanshe.
- Gong, Xun. 2018. *Le gyälrong zbu, une langue tibéto-birmane de Chine du Sud-ouest. Une étude descriptive, typologique et comparative*. Doctoral thesis, INALCO.
- Haudricourt, André-Georges. 1965. Les mutations consonantiques des occlusives initiales en môn-khmer. *Bulletin de la Société de Linguistique de Paris* 60(1): 160–172.
- Haudricourt, André-Georges. 1974. Hai Chữ B, trong cuốn từ điển của A-léch-xan-đo Rôt. *Ngôn Ngữ* 1974(4): 37–38.
- Hayes, La Vaughn H. 1992. Vietic and Việt-Mường: a new subgrouping in Mon-Khmer. *Mon-Khmer Studies* 21: 211–228.
- Hoàng, Thị Ngo. 1996. *Nghiên cứu về chữ Nôm và tiếng Việt qua bản giải âm Phật thuyết đại báo phụ mẫu ân trọng kinh*. Candidate Thesis. Hanoi: Viện nghiên cứu Hán Nôm.
- Hoàng, Thị Ngo. 1999. *Chữ Nôm và tiếng Việt qua bản giải âm Phật thuyết Đại báo phụ mẫu ân trọng kinh*. Hanoi: Nxb Khoa học Xã hội.
- Jacques, Guillaume (Xiàng, Bólin). 2008. *Jiāróngyǔ yánjiū* (嘉絨語研究). Beijing: Minzu chubanshe.
- Jacques, Guillaume. Forthc. Fossil nominalization prefixes in Tibetan and Chinese. To appear in: *Bulletin of Chinese Linguistics*.
- Lai, Yunfan. 2017. *Grammaire du khroskyabs de Wobzi*. PhD thesis, Université Sorbonne Nouvelle Paris 3.
- Lin, Xiangrong. 1993. *Jiāróngyǔ yánjiū* (嘉絨語研究). Chengdu: Sichuan minzu chubanshe.
- Luo, Zhufeng et al. 1993. *Hànyǔ dà cídiǎn* (漢語大詞典). Shanghai: Shanghai cishu chubanshe.

- Maspero, Henri. 1912. Études sur la phonétique historique de la langue annamite. *Bulletin de l'Ecole française d'Extrême-Orient* 12: 1–124.
- Maspero, Henri. 1930. Préfixes et dérivation en chinois archaïque. *Mémoires de la Société de Linguistique de Paris* 23(5): 313–327.
- Mei, Tsu-lin. 2012. The Causative *s- and Nominalizing *-s in Old Chinese and Related Matters in Proto-Sino-Tibetan. *Language and Linguistics* 13(1): 1–28.
- Michaud, Alexis, Michel Ferlus, Minh-Châu Nguyễn. 2015. Strata of standardization: the Phong Nha dialect of Vietnamese (Quảng Bình Province) in historical perspective. *Linguistics of the Tibeto-Burman Area* 38(1): 124–162.
- Nguyễn, Ngọc San. 1982. Góp vài ý kiến về âm đầu tiếng Việt cổ qua cứ liệu Nôm trong sách «Phật thuyết đại báo phụ mẫu ân trọng kinh». *Ngôn ngữ* 1982(3): 34–41.
- Nguyễn, Quang Hồng. 2001. *Khái luận văn tự học chữ Nôm*. Hanoi: NXB Giáo dục.
- Nguyễn, Tá Nhí. 1987. Lối đánh dấu cá trong chữ Nôm. *Tạp chí Hán Nôm* 1987(1): 35–38.
- Nguyễn, Tài Cẩn. 1979. *Nguồn gốc và quá trình hình thành cách đọc Hán Việt*. Hanoi: NXB Khoa học Xã hội.
- Nguyễn, Tài Cẩn. 2008. *Một số vấn đề về ngành Nôm học*. Ms. (Hội nghị Nôm học, April 2008, Temple University).
- Nguyễn, Phú Phong, Trí Dôi Trần, Michel Ferlus. 1988. *Lexique Vietnamien-Ruc-Français: parler d'une minorité ethnique des montagnes de Quảng Bình, Vietnam avec notes ethnographiques et introduction linguistique*. Paris: Sudestasia.
- Nguyễn, Hữu Vinh, Đặng Thế Kiệt, Nguyễn Doãn Vượng, Lê Văn Đăng, Nguyễn Văn Sâm, Nguyễn Ngọc Bích, Trần Uyên Thi. 2009. *Từ Điển Chữ Nôm Trích Dẫn: Dictionary of Nôm characters with excerpts*. Westminster: Viện Việt-Học.
- Pan, Wuyun. 2000. *Hànyǔ lìshǐ yīnyùnxué* (汉语历史音韵学). Shanghai: Shanghai jiaoyu chubanshe.
- Pulleyblank, Edwin G. 1981. Some notes on Chinese historical phonology. *Bulletin de l'École française d'Extrême-Orient* 69: 277–288.
- De Rhodes, Alexandre. 1651. *Dictionarium annamiticum, lusitanum et latinum*. Rome: Sacrae Congregationis de Propaganda Fide. Copy held at the Bayerische Staatsbibliothek, accessed at: <http://mdz-nbn-resolving.de/urn:nbn:de:bvb:12-bsb10521857-2>.
- Sagart, Laurent, William H. Baxter. 2012. Reconstructing the *s- Prefix in Old Chinese. *Language and Linguistics* 13(1): 29–59.
- Shimizu, Masaaki. 1996. Kanbun-Chunomubun taiyaku Bussetsu Daihō Bumo Onjūgyō ni miru chunomu ni tsuite (漢文=字喃文対訳『仏説大報父母恩重經』に見る字喃について). *Ningen-kankyōgaku* 5: 83–104.
- Shimizu, Masaaki, Lê Thị Liên, Shiro Momoki. 2005. A Trace of Disyllabicity of Vietnamese in the 14th Century - Chu Nom Characters Contained in the Inscription of Hộ Thành Mountain. *Kōbeshi Gaikokugo Daigaku Gaikokugaku kenkyū* 64: 17–49.
- Shimizu, Masaaki. 2015. A reconstruction of Ancient Vietnamese initials using Chữ Nôm materials. *NINJAL research papers* 9: 135–158.
- Skal-bzang 'Gyur-med (Gésāng Jūmiǎn), Skal-bzang Dbyangs-can (Gésāng Yāngjīng). 2002. *Zàngyǔ Fāngyán Gàilùn* (藏语方言概论). Beijing: Minzu chubanshe.
- Sun, Jackson T.-S. 1998. Nominal morphology in Caodeng rGyalrong. *Bulletin of the Institute of History and Philology, Academia Sinica* 69(1): 103–149.
- Sun, Jackson T.-S. 2004. Verb-stem variations in Showu rGyalrong. In: Lin Ying-chin et al. (eds.). *Studies on Sino-Tibetan Languages: Papers in Honor of Professor Hwang-Cherng Gong on His Seventieth Birthday*: 269–296. Taipei: Academia Sinica.
- Suwilal, Premsrirat. 1996. Phonological characteristics of So (Thavung), a Vietic language of Thailand. *Mon-Khmer Studies* 26: 161–178.
- Suwilal, Premsrirat. 2000. *So (Thavung) Preliminary Dictionary*. Salaya & Melbourne: Institute of Language and Culture for Rural Development, Mahidol University and the University of Melbourne.
- Thiều Chửu. 1942. *Hán-Việt Tự Điển*. Hanoi: Đuốc Tuệ xuất bản.
- Thompson, Laurence C. 1976. Proto-Viet-Muong Phonology. In: Philip N. Jenner and Laurence C. Thompson and Stanley Starosta (eds.). *Austroasiatic Studies*: 1113–1204. Honolulu: The University Press of Hawaii.
- Trần, Trọng Dương. 2010. Về dấu vết chữ Nôm kỳ húy trong sách 'Phật thuyết'. Ms. (Tạp chí Văn hóa Nghệ An 7).
- Trần, Trọng Dương. 2011. *Phật thuyết có phải dịch phẩm Nôm của thế kỉ XII? Ngôn ngữ* 2011(4): 31–47.
- Trần, Trọng Dương. 2012a. *Một số từ gốc Hán có cấu trúc CCVC qua ngữ liệu thơ Nôm trong Quốc âm thi tập*. Presentation made at the Hội thảo kỷ niệm 1 năm ngày mất GS Nguyễn Tài Cẩn.

- Trần Trọng Dương. 2012b. Từ nguyên của XE và các điệp thức của nó. In: *Thông báo Hán Nôm học năm 2010–2011*: 557–562. Hanoi: Thế giới.
- Trần, Uyên Thi, Hữu Vinh Nguyễn. 2007. *Ai vẽ được, ai xóa được? Dấu vết âm việt cổ: từ song tiết và phụ âm kép*. Presentation made at the Hội Nghị Quốc Tế về Tiếng Việt, Viện Việt-Học, Westminster.
- Wang, Li. 1948. Hànyuèyǔ yánjiū (漢越語研究). *Lingnan xuebao* 9(1): 1–96.
- Yánmùchū. 2005. Shìlùn Jiāróngyǔ zhōng sa de yǔfǎ gōngnéng jí xíngshì (试论嘉戎语中sa 的语法功能及形式). *Aba Shifan Gaodeng Zhuanke Xuexiao Xuebao* 22(1): 91–92.
- Zhang, Jichuan. 2009. *Zàngyǔ cízú yánjiū: gǔdài zàngzú rúhé fēngfù fāzhǎn tāmen de cíhuì* (藏语词族研究: 古代藏族如何丰富发展他们的词汇). Beijing: Shehuikexue Wenxian Chubanshe.

Гун Сюнь. Китайские заимствования с полуторасложной структурой в старовьетнамском языке

В настоящее время практически не остается сомнений в том, что в древнекитайском языке присутствовали начальные сочетания согласных, включавшие в себя пресиллабы; однако прямых указаний на то, какие конкретные формы реконструировать для тех или иных слов, сохранилось очень немного. В данной статье вводится в научный обиход новый источник, ранее остававшийся неисследованным — старовьетнамский язык, достаточно подробно зафиксированный в одном сохранившемся тексте, где ряд слов, которые в современном вьетнамском языке являются односложными, записан в орфографии, предполагающей для них изначальный полуторасложный характер. Для некоторых слов, заимствованных из китайского, только варианты их записи в старовьетнамской орфографии позволяют определить изначальную форму, в которой они были заимствованы. Небольшой список известных на сегодняшний день полуторасложных слов, зафиксированных в старовьетнамском тексте и имеющих китайское происхождение, включает как лексемы, для которых можно надежно реконструировать в китайском сочетании согласных, так и слова, где вероятна префиксация уже на вьетнамской почве.

Ключевые слова: древнекитайский язык, старовьетнамский язык, историческая реконструкция, полуторасложные слова, префиксальная морфология.