What Are Cognates and What Are Variants in Chinese Word Families?

Axel Schuessler

Wartburg College

In Chinese, some words are obviously genetically related to others and can be grouped into word families. But there is only limited agreement as to the morphology and the range of sound alternations in word families. Generally, investigators agree that change of tone and change in voicing of initial consonants have been common morphological devices in Proto-Chinese or Old Chinese (OC). On the other hand, variants of words are common in Chinese, consider, for example, the distinctions between colloquial vs. literary layers. Between recognized morphological changes and variants lies a large gray area. This paper will consider these issues from the angle of variants and attempts to identify individual cases of cognate set which are not the result of morphological derivation, as for example the two words for ‘crow’: wú / MC ùo < OC *ʔa and yā / ?a < *ʔa. Does yā really derive from OC *ʔa? This paper suggests that wú and yā are merely phonetic variants of the same word (OC *ʔa), where wú followed the ‘regular’ phonological development, and yā represents an archaism which preserves the original vowel, perhaps due to sound symbolism.

This paper also discusses, among others, cases in which two Chinese words can only be understood as variants of the same stem when one goes outside of the Chinese language. For example, tân / tʰaŋm / ‘dark’ and gàn / kʰəm / ‘purple’ are variants of an area word *klám ‘dark, dark red’ etc. There is apparently no internal Chinese morphological mechanism which would allow us to derive one from the other, they reflect two different strata of borrowing in Proto-Chinese or OC.

Chinese word families

It has long been recognized that the Chinese vocabulary can to some extent be grouped into families of related words. For the purposes of this
paper, ‘allofams’ (members of the same word family) can be connected in one of two ways: either they are derived by some morphological mechanism (as English ‘song, sing, sang, unsung’, or ‘free, freedom’), or they are variants of the same original etymon, as English ‘shirt’ and ‘skirt’. For Chinese this distinction is normally only made in individual cases studies. Sorting these out is important for etymological studies.

Investigators have differed significantly over the range of sound alternations within a word family (wf). At a minimum, there is general agreement on two types of sound changes which had a morphological role in word derivation. These are (1) change in tone, and (2) change in the manner of articulation of the initial consonant. Both have parallels in TB languages, particularly Written Tibetan (WT) and Kuki-Chin languages, and are therefore of ancient Sino-Tibetan (ST) provenance. For example:

(1) 
zhī 織 / tśjāk / *tak ‘to weave’ ≠ zhī / tāːtʃ ‘*tak-h < *tak-s) ‘weaved material’

Compare:
WT ‘tḥag- (< *N-tak), btags (< *b-tak-s) ‘to weave’ ≠ ṭhtags (< *tak-s) ‘weaved material’

The change in the manner of articulation includes this often cited example:

(2) jiān 見 / kien ‘to see’ ≠ xiān 現 / yan ‘to be seen, appear’

Note
WT ‘gebs-pa, bkab, dkab, kʰob ‘to cover’

Some broad meanings or functions can be associated with these morphological changes: quēshēng (tone C) originally reversed the direction of attention which usually translates into intransitives or passives in Western languages (Schuessler 1985), while voicing tends to change verbs to intransitives. Apart from these generally accepted morphological changes, individual investigators have proposed other sound alternations.

Among the features which are widely recognize are alternations between MC x- and m- or η- as in hēi 黑 vs. mò 黑; the alternation between MC l- and ʂ- as in lǐ 里 vs. shǐ 史; alternations between s and ts-
As to finals, a stop consonant can on occasion alternate with a homorganic nasal, as in lüè 彈 vs. liàng 'robber', although compelling examples are not abundant.

Many scholars allow for additional interchanges. In general, there are two avenues towards assembling cognate sets. One is to start with words with identical or associated meaning and secondarily explain their phonological differences. In order not to go off in all directions, one ought to keep to a given rime and allow the initials to vary, or keep to one category of initials and then allow for variations in finals. Wang Li gives examples for both approaches; same initial but different rimes (such as the negations with initial m-), as well as same rime but different initials (such as rime -ă ‘bright’) (1958:542-545). Looking for cognates with the same rime is the objective of those who are interested in identifying prefixes or pre-initials in wf; for example, lùò 落 ‘to fall down (incl. leaves from a tree)’ and tuò 落 ‘fallen leaves’ have incompatible initials (MC l- vs. th-). But if we assume an OC *rák vs. *hrak, these words look like a plausible set (Sagart 1999:125ff; Sagart’s OC reconstructions are different, though).

The other approach, keeping to the same type of initial but allow the rimes to vary, is exemplified in a set proposed by Pulleyblank (1973:121), e.g. róu 肉 / ńţjœu 軟 / ńţjwâń B / nuò 聲 / ńţju, ńţjwâń B, nuán C / ruò 弱 / ńţjak all meaning ‘soft’, but he has apparently reservations about including items with just any final, therefore he leaves rèn nèn, nu nèn (p.571) is distinct from róu 肉 / ńţjœu 軟 / ńţjwâń B / nén, nûn nèn (p.236). As long as we do not know more about OC morphology, we cannot tell if the this wf is due to morphological derivation, or to convergence in which the initial n- would be associated with ‘soft’, just as in English, words with initial gl- typically suggest ‘gliding, glossy’ etc. Alternatively, some of these words for ‘soft’ might be variants of some sort.

In his Word Families (1933), Karlgren allows for a rather broad pattern. A word family could have a final of the type -K, -T or -P etc. in
conjunction with the initial consonant type K-, T-, N- or P- etc. where T-
includes any acute initial consonant, i.e. any which is not a guttural or
labial. Thus his wf with items 242-262 (1933:69), for example, has a root
T-K and includes the following words (Karlgren’s ‘archaic Chinese’, i.e.
OC; in parentheses Baxter’s (with minor amendments)):

(3) yang / *dianj (*lanj) ‘light’ × zhao / *tioj (*tao) ‘bright’ ×
zhou / *tioj (*trukh or *rtukh) ‘day time’ × xing / *sienj
(*s’enj) ‘star’

In this proposed wf, the OC initials, as understood today (Baxter), are
*l-, *d- / *t-, *s-; the vowels are *a, *au, *e, *u; the finals are *-1, *-k, *-V
(vowel). The TB cognate for yang is *lanj (e.g. WB lajn ‘be bright’, o-lanj
‘light’), zhou is clearly cognate to WT gdugs (< *g-duk-s) ‘mid-day, noon’.
These two TB items are certainly not related. Therefore, Karlgren’s
phonological parameters are much too broad, we need to set up narrower
ones for wfs.

But how can we determine these parameters? So far, the narrowest
and safest frame work for a wf is to be content with tonal variations and
those in those within a manner series of the OC initial, including MC ji- <
*l ~ MC tO < *lh- etc. The next step out of this safe core would be to
accept as a wf variation initial variations like presence / absence of pre-
initial *s- which is usually recognized as a morphological relic. Work
recently carried out may help us identify more prefixes (Baxter and Sagart

A step beyond this would be recognition of sound alternations whose
morphology is not clear to us, but which includes has a few striking examples
which make a morphological connection of some kind difficult to brush
away, as perhaps ling / bing ‘ice’, ling / pinging ‘ oppress,
maltreat’. We could not persuasively identify a morphological role of the
assumed labial pre-initial here; this situation is reminiscent of IE roots
extensions whose morphological function cannot be pinpointed, but which
are nevertheless accepted as facts, as in IE *gel- ‘ball’ (e.g. > Lat. galla ‘gall-
nut’) ~ *gel-g (e.g. > Engl. ‘cling’) ~ gel-t (e.g. > Engl. ‘child’) ~ gel-b (e.g.
> ‘globes’) (Pokorny 1994:357). Here we are gradually moving into a gray
area of uncertainty, and we may be prepared to encounter variants within wf which are not the result of morphological change.

Variants are common occurrences in dialects, i.e. 白 colloquial vs. 文 literary forms, such as Mandarin col. 他 ‘he, she, it’ vs. lit. tuō ‘other’. These are lexically two different words but historically one and the same etymon, there was no ablaut morphology that derived one from the other. This phenomenon is so ubiquitous in China that one might expect this to have occurred already in ancient and archaic times. For illustration, the wf ‘strong’ strongly suggests to me that it includes OC variants:

(4) qiáng 强 / gian ‘be strong’ [Shi]; 強 [Meng]. Wang Li 1982:341 includes many other words in this wf. Perh. the wf→gāng

* qiáng 强 / gian ‘make an effort, compel’ [Meng]
* qìng 勢 / gian ‘strong, powerful’ [Zuo]
* jìng 劲 / kian ‘strong’ [Zuo] (Wang Li 1982:341)
* háng 行 / yian ‘strong, vigorous’ [Lunyu]
* gēng 棍 / kian ‘strong’ [Chuci]

I find it difficult to conceive of a morphological system which derives words from a root with the help of minute shades of vowel timbre, all meaning the same thing, more or less. It seems that the original etyma are qiáng and qiăng because the former has survived in colloquial layers of modern dialects: qiáng in Mand., also other dialects, e.g. Min: Xiamen col. kiuA2 (lit. kio1A2); qiăng is a morphological derivation from qiáng by way of shaŋshe (tone B) which derives active verbs from verbs; another minimal pair of this rare tonal pattern is:

(5) zhăng 長 / tjajB / *traj ‘to grow tall’ [Meng], ‘increase, elder’ [Shi]; < zhăng 张 / tjaj / *traj ‘make long, to string a bow’ [Shi], ‘string an instrument’ [Guoce], ‘stretch, extend’ [Lao]. QYS tj- has probably more than one source, therefore *tr- is not a certainty. <> WB tajB ‘tighten, become tense, stiff’, WT ṭhaj-po ‘hardy, strong, tense’.
Since it is often difficult to decide whether we are dealing with a morphological relationship or not, I will in the remainder of this paper look at this problem from the other end and try to identify variants within wfs. i.e. items whose differences are not morphological in origin:

1. Grammatical variants: due to syntactic position
2. QYS div. II vocalism
3. QYS homophones with different OC rimes
4. Vocalic variation in OC
5. Shang sheng vs. -k
6. ST *řj- ~ *ř-j- initials and their Chinese doublets
7. ‘Beard’ ~ ‘snout’
8. ‘Black’

1. Grammatical words

In some grammatical words, the origin of the vowel change can be inferred from the word’s syntactic use. In OC, če / *neiʔ ‘you’ can occur in all sentence positions: pre-verbally as subject or topic, post-verbally as object, before nouns as possessive ‘your’. In Early Zhou texts, rǔ / *ruʔ appears in topic and object position, nǎi / *naiʔ ‘your’ is found only in the adjectival position immediately preceding a noun. Nǎi is therefore an unstressed proclitic in which the vowel was reduced to a neutral central *ə. The exact parallel to this is the relationship between yú / jiwo / *laʔ (or *jaʔ) ‘I, we, me, us’ and the possessive yí / ji / *la ‘my, our’. The ‘ablaut’ in these words does therefore not have the meaning ‘possessive’, nor is it evidence for ‘cases’, but is the result of a non-morphemic development. Incidentally, this is mutatis mutandis also the relationship between wó 我 and wú 吾: wó, occurs in all positions in the sentence, while wú is typically (always?) either the subject / topic or possessive; in other words, wú is always proclitically linked to the following constituent, cases are not involved at all.
2. Variants with MC div. II vocalism

We find in Shijing several onomatopoetic words for ‘bang!’:

(6) bāng 彈 / pēŋ ‘bang!’
    pánɡ ~ pénɡ (馬 + 彈) / bānɡ ~ bēnɡ ‘bang!’
    bānɡ 彤 / pānɡ ‘bang!’

According to current theories on OC phonology, the forms with the MC rimes -ŋ should be reconstructed in OC with a medial ‘r’, thus *praŋ ~ *braŋ as opposed to *p’raŋ ~ *b’aŋ (the symbol ‘ indicates a later MC div. I/IV syllable, later div. III with medial ‘j’ are unmarked). I believe it unlikely that a sound like ‘bang!’ had an r-cluster in OC or any language, for that matter. I suggest that the forms which yielded a later MC dic. II type vowel simply preserved the ‘a’-color of the original word, while the regular development moved the ‘a’-sound back more to ‘$’ (p$1), in southern dialects even to o / o. This alerts us to the strong possibility that not all div. II syllables derive from OC r-clusters.

Once this suspicion is aroused, other possible words and doublets come to mind where a div. II vowel was preserved for reasons of sound symbolism. Generally, onomatopoetic words do not always follow regular phonological developments (H. H. Hock 1986:50, thus Middle English pīpen (to chirp of birds) is peep [pīp], not the expected [pajp]). Examples:

(7) yā 聞 / ?a8 / *?a? (not *?ra?) ‘mute’ <> TB *(m-)a ‘mute’
(8) è 聞 / ?æk / *?ak (not *?rak) ‘to laugh’
(9) ài 聞 / ?aiC ‘choke’
(10) ài 乞 / ?aiC ‘belch’
(11) hé 聞 / xap ‘laugh’
(12) kēng 今 / kʰæŋ ‘sound of metallic instrument’
(13) zhēng ㆜ / tɛŋ ‘sound of beating’

This is true also for:
(14) wū 鳥 / uo / *?a ‘crow, raven’
    ~ yā 鴉 / a / *?a (not *?ra) ‘crow, raven’ [Zhuangzi], is a later variant of wū without medial *-r-, according to Pulleyblank (AM n.s. 9.1, 1962:103)

Common grammatical words often have archaic features (Demiéville, Stimson), thus we find already in Shijing an interrogative in two variants:

(15) hú 胡 / yuo / *g’a ‘what, where’ vs.
    xiá 遠 / ya / *g’a (not *gra) ‘how, why’

We cannot be absolutely certain about such character doublets because the different graphs may have been agreed upon only in Han time and reflect some Han period opinion on what the Shijing ought to have sounded like. In Western Zhou versions of the Songs, only a *g’a might have existed, the distinction having been introduced later.

Along these lines, we may speculate that, among others, Mandarin final 阿 might be the same word as hú 胡 / yuo < *a (or *fia, if you prefer).

One might now wonder about the nature of div. II vocalism in sets like these:

(16) bāng 邦 / pān ‘country’ [Shi]
    ※ fēng 封 / pjwoŋ / *pəŋ, Min: Xiam lit. hoŋ A1, col. pəŋ A1; ‘mound, tumulus, raise a mound, boundary embankment, fief’ [Bl, Shi]. <> WT pʰuŋ-po ‘heap’ ※ spuŋ ‘a heap’, spuŋ-pa ‘to heap’
(17) pēng 碰 / bēng [Zìhuì], bān [Duan Yucai] ‘to hit, run into / meet unexpectedly’ is perhaps a recent col. form of fēng 迦 (Wang Li 1982:390)
(17a) mò 寞 / māk [Zhuang] ~ / māk 嘘 [Chuci] ‘quiet, still’; Chuci is a late OC (Han) text which might contain dialectal features
What Are Cognates and What Are Variants in Chinese Word Families? / 233

3. QYS homophones with different OC finals

The following MC homophone sets reflect variations in reconstructed OC vowel:

(18) jī 蹴 / tsjäk / *tsak (*tsjak?) ‘footprint, track’ [Shujing] (written with radical 162).
    ~ jī 蹴 / tsjäk / *tsek ‘footprint, track’ [BI, Shijing].
(19) ji 隤 / tsjäk / *tsak (*tsjak?) ‘walk reverently’ [Lunyu].
    ~ ji 隤 / tsjäk / *tsek ‘walk with small steps’ [Shijing].
(20) jí 蹱 / dzjäk / *dzak (*dsjak?) ‘trample’ [Liji].
    ~ jí 蹱 / dzjäk / *dzek ‘trample, oppress’ [Zhuangzi].
(21) zhì 転 / tiC / *trits ‘to slip, trip’ [Shi]. <> WT ’dred-pa ‘to slip, slide, glide’; Kanauri *bret. This etymon is apparently distinct from the syn. → dié 跌.
    ~ zhì 転 / tiC / *trts ‘to stumble’ [Zuo]; Duan Yucai considers this just a variant of 蹴.

Each of these sets might in reality reflect only one word. The OC finals *-ak and *-ek (div. III) had merged by late Han, therefore the distinction in the received texts may simply be graphic. Alternatively, we might assume OC variants, but in any event, these are almost certainly not morphological cognates.

A clear case of a graph misleadingly suggesting a the wrong OC rime is:

(22) zhī 紙 / tśieB / LHan tśiaiB ‘paper’ [Hou Hanshu].

Bodman (1980:184) relates this to Viet. giâý, PVM *k-caj? (Ferlus) ‘paper’ which, like the PMin form *ťśiuI8, presupposes an OC rime *-ai rather than the *-e which would be required for OC because of the phonetic element. When this word was committed to writing, the rimes OC *-ai and *e had already merged in some dialects, a process which is observed already in late Zhou texts (Pulleyblank 1962:216). No doublets are involved here.
Variants or cognates, however, make up this nearly homophonous pair:

(23) zhì 至 / tśiC / *tji(t)s ‘to come to, get to, arrive at, reach to’ [BI, Shi].
<> WT mēbi-bal, mēbis ‘come, go, exist’, WB ceC ‘come, arrive’.
★ zhí 底 / tśiB / *tji ‘to come to, bring about, effect, accomplish, achieve’ [Shi]; ‘come to’ [Chuci]; ‘to bring about, establish, settle’ [Shi].

Here, MC tśiB and tśiC appear to be parallel to the WT cognates mēbi ~ mēbis.

4. Vocalic variation in OC

Apart from these cases, there are clearly OC/QYS vocalic alternations in wfs. We have already seen that Karlgren allowed any vowel within a wf. Pulleyblank maintains a theory by which OC had only two vowels *a and a which alternate in wfs (Pulleyblank 2000:33). Examples are:

(25) tán 談 / dâm / *d’am ‘to speak’. <> WT gdam-pa ‘to advise, give council’.
★ tán 談 / dâm / *d’am ‘to speak’.
(26) yī 螈 / 1jjieB 3 / *ŋaj ~ yī 搏 / *ŋaj

However, wfs usually have only one major OC vowel: *a, e, i, o, u or *, or one diphthong like *ai, *au. Obvious cases of vowel alternations within a wf are not often found, therefore the pair like ‘speak’ and ‘ant’ above may well be variants of some sort.

In the following group no. 27 (if indeed consisting of cognates), many different vowels are represented:

(27) zhū 濰 / tśjwoB / *taʔ ‘islet’ [Shi].
~ zhū 濰 / tśiB / *taʔ ‘islet’ [Shi].
Note also chī 時 / di ‘islet’ [Shi], and zhōu 江 / tśjau/ *tu ‘island in a river’ [BI, Shi].
What Are Cognates and What Are Variants in Chinese Word Families?

(28) cāng 蒼 / QYS tsʰāŋ ‘green, azure’ [Shi].
    ~ qīng 青 / QYS tsʰięb, OC *tsh’ěŋ ‘green, blue’ [Shi]; PMin *tsʰaŋ ‘raw’ (Norman identifies the PMin form with 青, Nakajima with 生). The doublet with vowel ‘a’ seems to forshadow the Min and southern developments.

(29) tāng 鴻 / tʰāŋ / *th’aŋ ‘sound of drum’ [Shi].
    ~ tiān 天 / dien / *d’in ‘sound of drum’ [Shi some versions].

(30) tíng 庭 / die / *deŋ ‘courtyard, court of a palace, mansion, temple’ [OB, Shi].
    ~ cháng 場 / djaŋ < *diaŋ? ‘vegetable garden, a (pounded) threshing area’ [Shi 154, 7 etc.].

The PMin form for tíng would be *diaŋ which looks like earlier forms for cháng. The initials of the two words do not match; yet retroflex d- in the QYS is non-phonemic (one could instead write djaŋ), and by default, the only way to render a syllable diaŋ in the QYS would be djaŋ. But this is rather speculative.

5. Variants with MC shàngshēng vs. final *-k:

Variants are also suspected as the origins of consonantal alternations in the final. We assume here, for the sake of argument, that the MC shàngshēng (tone B) reflects and OC final glottal stop:

(31) gè 格 / kɐŋ / *kraŋ ‘to come, go to’ [OB, BI, Shujing], originally written 各.
    ~ jiǎ 假 / kɐŋ / *kraŋ? ‘to come, go to’ [Shijing].

Here we may have a trace of genuine dialect differences because the Shijing has only jiǎ, other texts only gé. However, because jiǎ is not a rime word in Shijing, it is impossible to decide whether this character was writing a variant form of gé, or was simply borrowed to write it (or vice-versa). Note also the possible cognate WT ’gro-ba ‘to go’ which can come from either *Ngro or *Ngra, without final -k. As the next examples suggest,
it might just as well be the form in final -k which is secondary.

The first two sets nos. 32 and 33 are in fact cognates. Additional wf connections clearly favor an original final *-ʔ over *-k. This -k may have come about as some sort of backward formation since -ʔh and -kh, whatever their phonetic nature, may have overlapped phonetically:

(32) shè 舍 / šja / *hlaʔ ‘to let off, bestow, grant set aside, leave’
[Shijing], ‘give, bestow’ [Zuo]; ‘give up, let go’ 舍 [Guoyu].
Note Middle Mon salah ‘to give away, disburse’, Late Mon ‘give up, free’.
~ shì 許 / šjäk / *hlak ‘put away; let go, detach’ [Zuo].

(33) shè 舍 / šja / *hlaʔ ‘put down, deposit’ [Zuo]; ‘let go’ [Shi],
‘let off, liberate’ [Zuo], ‘pardon’ [Shu].
~ shè 詡 / šja / *hlakh ‘let go’ [Shi], ‘let off, liberate’ [Zuo],
‘pardon’ [Shu].

(34) lǐ 理 (2) / li / *raʔ ‘cut jade according to its veins’ [Guoce].
在美国 preferences prefer the veins of stone 詡 [Zhouli], ‘split according to the veins’ (stone) 詡 [Zhouli] >
‘engrave’ 詡 [Li]; ‘space between fingers’ 詡 [Yi] > ‘a tenth’ 詡 [Li].

(35) yù 餘 / njwo / *ŋaʔ ‘to defend, object, oppose, prevent’ [BL, Shi].

(36) wù 午 / nuo / *ŋaʔ ‘resist’ [Li].

(37) bō 博 / pük / *p’ak ‘be wide’ [BL, Shi].

(38) bō 趿 / puã / *p’aiʔ ‘to walk lame’ [Yi]. -- Archaic forms are preserved in southern dialects, e.g. CMin *pai ‘lame’. <> TB *paj ≈ *baj ‘lame, limp, oblique’, *baj ‘left’ (side) (Matisoff 1995:42) > WB phai ‘go aside, put aside’ ≈ phai ‘go aside, get out of the way’ ≈ pai ‘put aside, away, reject’ ≈ bai ‘left side’.
≈ bi 騙 / pjiäk 4 / *pek (< **pajk?) ‘to walk lame’ [Li]
Obviously, doublets do not always have completely identical meanings. But that is not to be expected, or else both would not have survived, just like Engl. ‘shirt’ and ‘skirt’ are not homonyms.

Then, there are alternations with -k and MC pingsheng (tone A), i.e. OC open syllables. These cases appear to be different in so far as their meanings and usages are not quite as close as in the preceding examples, hence we may have some kind of morphological final -k root extension or suffix:

(39) yí 移 / jie / *laj ‘to transfer, move’ [Shu], ‘change, alter’ [Meng].
\begin{itemize}
\item TB *laj ‘change, exchange’ (Matisoff 1995:42) > WB laiB ‘change, exchange; empty’ (contents of one vessel into another, also hläiB) \(\neq\) hläiB-phäi ‘exchange’; JP lai\(^1\) ‘to change’, gä\(^1\)-läi\(^1\) ‘change, exh.’, mä\(^2\)-läi\(^2\) ‘change, substitute’; Tiddim Chin läi\(^2\) ‘change’, lei ‘buy’, Dimasa salai ‘alter, change, exh.’, Garo sre ‘change, exh.’. (STC p.64).
\item \(\neq\) yì 易 / jiäk / *lek (< **lajk) ‘to change’ [OB, Shi]. → Tai: S. lëk\(^2\)-L *dl- ‘to change, exchange’; KS *hlik\(^7\) ‘exchange’.
\end{itemize}

(40) su 肅鱐 / sjuk / *suk < **sruk? ‘shrivel’ [Shijing], ‘contract, shut’ (as flowers), ‘slice of dried fish’ 鳝 [Liji]. \(\Rightarrow\) PTai *hru\(\uparrow\)otD1 ‘to shrink, contract’; Be 33sut ‘shrink’.
\begin{itemize}
\item \(\sim\) suo 縮 / sjuk / *sruk ‘to draw back, shrink’ [Huainanzi].
\item xiū 脩 / sju / *su < **sru ‘dry up’ (of plants), ‘shrink’ [Shijing], ‘dried meat’ [Lunyu].
\end{itemize}

6. Variants which can only be identified through foreign languages

So far, we have been looking at wfs and variants within Chinese. Many ostensibly unrelated OC words turn out to be variants, though, when we go outside of the Chinese language.

A set of items which are close in meaning and rime, but alternate with MC initials lj- and ji- can be connected etymologically only by going beyond
OC to ST. As a preliminary, let us consider yán ‘salt’ for a moment:

(41) yán 盐 / jiâm / **r-jam > *jam ‘salt’

which is puzzlingly alone in a xiesheng series with initials like MC l- and MC k- < *r- and *kr-. The WT cognate is rgjam ‘salt’ < *r-jam (in such configurations, r- was felt a pre-initial and an epenthetic -g- emerged, see Li Fang-kuei 1959; otherwise *rj- > WT z-). If the Ch. word also derived from ST *r-jam, then the simplest explanation would be an early OC **r-jam with loss of the pre-initial. If this be indeed the Chinese fate of ST *r-j-, then we have further cases which not only have a WT cognate in rgj-, but also an OC doublet in *r(j)- > lj-:

(42) ST *r-ju → yóu 游 / jiouu / *ju < **r-ju ‘to flow, roam about’.
    ← WT rgju-ba < *r-ju ‘to walk, move, wander, range’ ≠ rgjuin ‘the flow, current’.
    → liú 流 / ljou / *rju ‘to flow, float’.

(43) ST *r-jut → yú 湖 / jiuet / *jut ‘well-rope’, also MC kjuet.
    ← WT rgjùd < *r-jut ‘string, cord’.
    → lù 縵 / ljue / *rjut ‘rope’.

(44) ST *r-jut → yù 逾 / jiuet / *jut ‘following, then’ vs.
    → lù 律 / ljue / *rjut ‘follow a model’. These are perhaps the same words as in the previous set.

(45) ST *r-jam → yán 刃 / jiâmB / *jam? ‘pointed, sharp’ (of plowshare) 刃 [Shi]; ‘sharp, pierce’ 刃 [Li], ‘cut’ [Xun].
    → lián 磨 / ljäm / *rjam ‘sharp, keen’ (of soldier) 磨 [Hanfei]; ‘sickle’ 磨 [Mo]. The graph seems to refer to a whetstone (‘sharpener’).

Incidentally, this way we can also account for words of ST provenance like:
What Are Cognates and What Are Variants in Chinese Word Families?

What Are Cognates and What Are Variants in Chinese Word Families?

(46) yè 液 / jiäk / *jak < **r-jak ‘fluid’ <> TB *rjak ‘grease, juice’
(STC #204), from ST *r-jak ~ *rjak.

Evidently, QYS ji- never derived from OC *r-, only from *l- and *j-.

This is a convenient place for a brief digression on ‘night’ and ‘evening’:

(47) yè 夜 / jiaC / *jakh ‘night’ <> TB *ja ‘night’ (STC #417)
(48) xì 夕 / zjäk / *s-jak ‘evening, night’ <> TB *r-jak ~ *s-rjak
‘day/night’ (STC #203)

These words have been much discussed. ‘Night’ yè 夜 / jiaC is so
close to TB *ja ‘night’ in every way (except for a possible OC final -k
which has, however, solid parallels as a Chinese innovation) that, all else
being equal, these two items form a cognate set. Assuming this, zjäk < *s-
jak (simplified from St **s-rjak?) ‘evening’ should be related to TB *r-jak
~ *s-rjak ‘day/night’ (i.e. 24 hr period). This is a better fit than WT zla
‘moon’. While ‘day’ and ‘sun’ often are expressed by the same word, as in
OC, because one does not exist without the other, ‘night’ is not closely
correlated with ‘moon’ because there are nights without moonlight and
days when the moon is visible. Hence WT zla ‘moon’ is a poor candidate in
the first place, the graphic representation ‘moon’ for mnemonic purposes
notwithstanding.

7. ‘Beard’ ~ ‘snout’

The two words hui / xjw{/OC *huats 顪 ‘beard’ and 喙 ‘snout’ go
back to one etymon:

(49) 顪 ‘beard of chin’ [Zhuang], GY also: ‘jia 顪 jaw, cheek’. This
appears to be a loan from Tai: S. (Siamese) nual³L < *hn-
‘beard’. This in turn appears to be of Lolo-Burmese provenance:
WB hnut ‘mouth, womb’ (< TB *snot). The change from
‘mouth’ to ‘beard’ might have been occasioned by LB words for
‘beard’ which literally mean ‘mouth hair’: Achang 55 _noncei-mui,
Leqi 5²nuat-3³m, etc. WB and some related languages do not use the etymon TB *snot, but *mut ‘mouth’ instead. This may explain why the Kam-Sui languages have either mut or nut ‘beard’ (Thurgood 1988).

(50) 喙 ‘snout’ [Zuo], ‘mouth’ [SW] (this graph has additional readings which might originally have belonged to the meaning ‘to pant’). With ‘beard’ having ultimately the LB ‘mouth’ as origin, it appears now that 喙 is the very same etymon, but with its original meaning; it has presumably entered Ch. directly from a TB language instead through the Kam-Tai filter.

(51) A variant from the same ST root is perhaps xu 須鬚 / sju, OC *sno ‘beard of chin’ 須鬚, related to TB *sno(w) ~ *sno(t) (Benedict LTBA 1976:16) > WT snod ‘vessel’, bu-snod ‘womb’, Motuo Monpa no-waŋ ‘mouth’, Kanauri sto ‘face’, Zhzh snu ‘vessel’ (< Tib.?), WB hnut (see hui above) ‘mouth, womb’; Pwo, Sgaw Karen no? ‘mouth’ (STC p.144f); Achang 55ŋot-3¹mui ‘beard’ (‘mouth hair’), Leqi 55nua-3³m. Here the same shift from ‘mouth’ to ‘beard’ has occurred in Chinese, facilitated perhaps by such forms as Jiarong t₃nos ‘lips’ → t₃nos mo rŒ ‘beard’. On the other hand, the source of Ch. could be Mon-Khmer languages instead: Khasi ḫu?’ hair’ (ӷ is infix), Khm /s꼬/, Palaungic *s- > h-: Ta-ang hu?, which in turn is reminiscent of Mand. hũ ṭ⁴h ‘beard’, probably a coincidence.

Hui in its meanings ‘beard’ and ‘snout’ seem to be loans which had entered PCh. by two different avenues. Also, if xu had been directly inherited from ST, we would expect a voiceless nasal in OC (QYS š- or th-). In any case, Ch. pre-initial s- hardly ever corresponds to s- in TB languages, although it is the same ST morpheme. Prefix *s- had a somewhat limited productive period in PCh. and OC, so words with this morpheme are almost all Ch. creations. Or alternatively, the word xu ‘beard’ might also be a loan, but from ‘northern TB’ (preservation of s-prefix in a cluster as in WT, Jiarong et al.), while hui is from ‘southern TB’ (all this happened about 3000 or more years ago!).
8. ‘Black’

Not only TB relations show historical connection between two or more Chinese words, but also other language families. The following is a widespread area word:


(53) gàn / kənˈ / *kləmˈ ‘purple’ [Lun] <> PTai *kləmˈ¹ ‘dark red, purple, dark, black’ (Li 1977:221-222); AN ɗəm, kələm ‘be dark’ ≠ laməm ‘be gloomy’ (‘düster sein’).

Both tǎn and gàn apparently go back to the same foreign etymon **kləm, in tǎn the voiceless initial consonant devoiced the lateral and was then lost, in gàn the initial *k- survived into the QYS and beyond.

As to the phonology of these forms, note these parallels for QYS th- corresponding to foreign voiceless *l or a cluster with *l:

táo / tāu ‘to pour water, wash’ <> PHmong ?luəˈ ‘to pour’.

ti / tʰiɛk ‘distant, remove’ tʰiɛk ‘distant, keep at a distance’ (GSR 850k). Above, we have already seen more cases of OC *-ek from *-aj+k. <> KS *klaiˈ¹ ‘far’, PTai: S. kəlaiˈ¹ ‘far’.

tiè / tiɛt ‘iron’ <> WT lčags < *hljaks ‘iron’.

<> Tai: S. lekˈ¹S < *hl- ‘iron’, KS *klətˈ ‘iron’; PVM *k̡hæc ‘iron’ (Bodman 1980:103).

tūn / tʰən ‘to swallow’.

<> PT *kl-: S. kliiˈ¹ ‘to swallow’, Ahom k(l)en, Wuming klwan; KS ?dun.

As to MC k- (in div. I/IV) from OC *kl-, note:

gàn / kəm ‘be sweet’, probably from **kləm < **kləm <> TB *kləm ‘sweet’ (STC p.75 n.231).
jiān จัน / kiem(ก) / *klem(h) ‘to combine’, ‘grasp, hold together, all-embracing’.
<> WT glem-pa ‘to squeeze, crush, squash’.
gū 改 / kã<sup>C</sup> ‘to change’ <> PTai *kla<sup>A1</sup> ‘to pass by, change into’.
gāo 合 / kāu<sup>B</sup>, kuok / *kluk(h) ‘to announce, inform’ <> Tai: S. klaau<sup>B1</sup> < *kl- ‘to say, declare’ (Li 1976:46).
gè 胳 / kãk / *klak ‘armpit, armpit seam’ GSR 766e <> PMonic *knlak ‘armpit’, LitMon knak.
gōng ㄍ / kuŋ / *kloŋ ‘palace’ <> Mon gloŋ ‘citadel, palace’.
gōng ㄍ / kuŋ / *kloŋ ‘clan head, uncle’.
<> Tai: S. luŋ<sup>A2</sup> < *l- ‘parent’s elder brother, uncle’.
gōu 犬 / kǒu<sup>B</sup> / *kloʔ ‘dog’ <> HM *klu<sup>2</sup> (Purnell); AA: Written Mon kluiv.
gǔ ㄍ / kuo<sup>B</sup> / *klaʔ ‘drum’.
<> PTai *kloŋ<sup>A1</sup> ‘drum’ (Li F. 1976:40), PHLai *l- (laŋ<sup>A4</sup>).
gǔ 鹹 / kuo<sup>B</sup> / *klaʔ ‘salt, salty marsh’ <> PTai *klɔ<sup>A1</sup> ‘salt’.

Concluding remarks

The Chinese lexicon can be grouped into wfs. Some morphological derivational devices are well known. In many instances, however, members of a wf are mere variants of an earlier etymon, be it due to position in a sentence, be it due to archaisms, or a split into colloquial vs. literary layers. In some cases, the historical connection between Chinese words can be established only by going outside of the Chinese language. In between these identifiable types of relationships, there is a large gray area in which words are suspected of being members of a wf, but what their exact relationships are is so far difficult or impossible to determine. Recently, the issue of OC morphology has been addressed by Baxter (Baxter and Sagart 1998), Sagart (1999) and Pulleyblank (2000).

It also emerges that the concept ‘word family’ has fuzzy edges. Do tàn
What Are Cognates and What Are Variants in Chinese Word Families?

黮 and 甘 belong to the same Chinese word family? We should perhaps point out that they may be cognates, but they are not members of the same Chinese wf because no internal Chinese morphology seems to be involved, just as English ‘head’ and ‘capital’ derive from the same IE root, but cannot be related through internal English morphology.
References


什麼是漢語詞族中的同源與異讀？

Axel Schuessler
Wartburg College

在漢語之中，儘管有些詞明顯的屬同一詞族，但是對於同族詞都有哪些構詞行為和音韻變化，卻少有定論。一般都認為聲調轉換和清濁轉換是上古漢語同源詞間常見的構詞手段。不過異讀詞，例如文白異讀，卻也相當普遍。如何分辨同源詞和異讀詞，就成了一項有待澄清的問題。本文的目的，就是想從異讀詞的角度提供進一步的思考。以“鳥” MC t'uo < OC *ʔa、“鴉”ʔa < *ʔa 爲例，這兩個語詞所指相同，他們是否果真具有構詞關係、抑或僅屬不同層次的異讀，值得深究。本文論證兩詞均來自 OC *ʔa，其中 “鳥” MC t'uo < OC *ʔa 爲常用詞，“鴉”ʔa < *ʔa 則為受形符影響而具存古特徵的異讀詞。本文同時也討論了藉由域外借詞才能判定的異讀詞，例如 tân 筆 / tʰàmⁿ 與 gân 染 / kʰàmⁿ “purple”等。