

## A PRINCIPLE IN THE PHONETIC COMPOUNDS OF THE CHINESE SCRIPT

By BERNHARD KARLGREN

In a dictionary published recently<sup>1</sup> I endeavoured to show that the phonetic compounds (諧聲) of the Chinese script were not composed so casually and with such lack of phonetic strictness as has been always supposed hitherto, and I established certain general phonetic rules which, though subject to many exceptions (the characters having been invented during centuries and by many authors), can be said to be fundamentally true. Now, thanks to a remark made by a reviewer, I have arrived at a new observation through which the methods followed by the character inventors can be shown to be even more strict in a certain respect than I imagined when writing my dictionary.

One of the main results of my introductory investigation in the dictionary was the reconstruction of certain initial consonants which must have existed in Archaic Chinese (say 200 B.C.) but were lost in Ancient Chinese (6th century A. D.). As 甬  $\dot{x}^{wong}$  is phonetic in 甬  $d'ung$ , 通  $t'ung$  etc.—a long series of words with *dental* initials (none with gutturals), and as 勻  $\dot{ju}^n$  is phonetic in various words with *guttural* initials, 均  $kju^en$  etc. (none with dentals), I concluded that in Archaic Chinese these words had initials which were dropped before the time of Ancient Chinese: a *d*- in 甬  $\dot{x}^{wong}$ , a *g*- in 勻  $\dot{ju}^n$ . As a significant fact I considered the regular existence, in the member lacking initial consonant in Ancient Chinese, of an  $\dot{x}$ - or a  $\dot{j}$ - (崇  $j\dot{x}am$  phonetic in 談  $d'am$  etc.), which phonemes I believe to be traces of the lost consonants, or possibly the phonetic cause of their disappearance. In certain cases—e. g. 羊  $\dot{y}ang$ , phonetic in 詳  $\dot{z}iang$ —I reconstructed, not a *d*- but a *z*- (羊  $\dot{y}ang < z$ - phonetic in 詳  $\dot{z}iang < dz$ -) for reasons for which I refer to the introduction of the dictionary.

Now Mr. Arthur Waley, in a very interesting review of my work, makes the following remark:<sup>2</sup>

<sup>1</sup> Bernhard Karlgren, *Analytic dictionary of Chinese and Sino-Japanese*, Paris, Paul Geuthner, 1923.

<sup>2</sup> *Bulletin of the School of Oriental Studies*, vol. III, p. 364.

"It is, however, very remarkable that (at any rate, in all the more typical series) it is the phonetic itself (i. e. the simple, uncompounded form) which lacks the initial consonant and therefore originally had an initial unaspirated sonant (*b, d, g*, etc.). I do not think that Dr. Karlgren's theory fully explains this phenomenon".

I had noticed the fact myself, but I did not pay much attention to it, as it could not invalidate my main argument: it could not be due to some law to the purport that a word without consonant in Arch. Chinese ( $\dot{x}^{wong}$ ) could serve as phonetic in a word with consonant ( $t'ung$ ), for this would leave the most important phenomenon unexplained—that cases like 甬  $\dot{x}^{wong}$ , 崇  $j\dot{x}am$  have only *dental* derivatives, cases like 勻  $\dot{ju}^n$ , 爰  $j\dot{y}^un$  only *guttural* ones. However, as Mr. Waley says, my theory does not explain the phenomenon that a  $\dot{y}a < d$ - is frequently phonetic in  $d'a$ ,  $t'a$  etc., while a  $d'a$  or a  $t'a$  seldom occurs as phonetic in an  $\dot{y}a < d$ -. I have to find the solution of this problem.

In order to do so some statistics will be necessary. For this purpose the 6,000 characters of my dictionary are not sufficient, but I work with a stock of words about equal to that of Giles' dictionary—say 12,000 (of the 13,848 entries of this work many are only cross references and rare variants), which are practically all the really current words in Ancient Chinese.

In the following discussion I leave out entirely the cases where explosives and fricatives<sup>1</sup> or nasals interchange in the phonetic series, e. g. 乞  $k'it$  phonetic in 迄  $xit$ , 更  $kwng$  phonetic in 硬  $ngvng$ —though far from rare they are nevertheless exceptions. Thus, to start with, we deal exclusively with explosives and affricates.

The principle I seek to establish is this:

*The authors of the characters freely use a word with a weaker initial consonant as phonetic in a word with a stronger one, but seldom vice versa.*

We cannot, of course, expect the old script creators to have been over-sensitive in this respect, and we therefore get only rough distinctions. I have not been able to discover any difference in function between  $g'$ - and  $k$ - etc., between  $d'$ - and  $t$ , which interchange quite

<sup>1</sup> But of course in this discussion I count Ancient  $\gamma$  among the explosives, and take it together with  $g'$  ( $g'j\dot{a}n$  etc.) in the statistics, because I proved in my dictionary that an Ancient  $ya$  etc. came from an Arch.  $g'a$  etc.

freely. Both are stronger than  $g^-$ ,  $d^-$ , both are weaker than  $k^-$ ,  $t^-$ . Thus we get a scale like this:

Weak	Middle	Strong
	$\alpha$ $\beta$	
$g^-$	$g^-$ and $k^-$	$k^-$
$(b)^1$	$b^-$ and $p^-$	$p^-$
$d^-$	$d^-$ and $t^-$	$t^-$
$dz^-$	$dz^-$ and $ts^-$	$ts^-$
$d'^2$	$d'^-$ and $t'^-$	$t'^-$
	$d'z^-$ and $t's^-$	$t's^-$
	$d'tz^-$ and $t's'$	$t's'$

If my rule is true, we shall find that

1. *Weak* words will serve as phonetics in *weak*, *middle* and *strong* words. Thus a  $(d)j\alpha$  (e. g. 也  $j\alpha < d^-$ ) will do in  $d'a$  and  $ta$ , and also in  $t'a$ .

2. *Middle* words will serve in *middle* and *strong* words:  $d'a$ ,  $ta$  will do in  $d'a$ ,  $ta$ , and also in  $t'a$ , but not in  $(d)j\alpha$  ( $j\alpha < d^-$ ).

3. *Strong* words will serve in *strong* words only:  $t'a$  will do in  $t'a$ , but not in  $d'a$ ,  $ta$ , nor in  $(d)j\alpha$  ( $j\alpha < d^-$ ).

If this rule holds good, it is easily seen that Mr. Waleys observation is explained: a 兩  $j^wong$  is phonetic in 簡  $d'ung$  and 通  $t'ung$  according to point 1, but a  $d'ung$  or a  $t'ung$  cannot be phonetic in  $j^wong$ , according to points 2 and 3. Let us therefore test it.

The positive statements in the three points above need no support of statistics. That an Ancient Chin.  $j^-$  or  $j\tilde{k}^-$  is very often phonetic in words with  $j^-$ ,  $j\tilde{k}^-$ ,  $d^-$ ,  $t^-$ ,  $t'^-$  etc. has been fully illustrated in my dictionary e. g. 兩  $j^wong$  phonetic in 簡  $j^wong$ , 簡  $d'ung$ , 通  $t'ung$ ; 易  $j\tilde{a}ng$  in 楊  $j\tilde{a}ng$ , 腸  $d'j\tilde{a}ng$ , 遑  $d'\tilde{a}ng$ , 趨  $t'\tilde{a}ng$ ; 况  $j\tilde{a}m$  in 忱  $j\tilde{a}m$  ( $< d^-$ ), 沈  $d'j\tilde{a}m$ , 枕  $t'j\tilde{a}m$  ( $< t^-$ ), 耽  $t\tilde{a}m$ ; 炎  $j\tilde{a}m$  in 談  $d'\tilde{a}m$ , 毯  $t'\tilde{a}m$  etc.

That an Ancient Chinese  $d^-$  is phonetic in  $d^-$ ,  $t^-$ , and  $t'^-$  promiscuously, can be seen in any page of my dictionary, e. g. 定  $d'ieng$  phonetic in 錠  $tieng$ , 弟  $d'iei$  in 第  $d'iei$ , 制  $t'iei$ , 喬  $g'j\tilde{a}u$  in 輪  $g'j\tilde{a}u$ , 嬌  $k'j\tilde{a}u$ , 躑  $k'j\tilde{a}u$  etc.

Equally, that an Ancient Chin.  $t^-$  is phonetic in  $d^-$ ,  $t^-$ , and  $t'^-$ , e. g. 旦  $t\tilde{a}n$  phonetic in 怛  $t\tilde{a}n$ , 但  $d'\tilde{a}n$ , 坦  $t'\tilde{a}n$ .

<sup>1</sup> Uncertain if it existed in Archaic Chinese.

<sup>2</sup>  $d^-$  became Ancient Chin.  $z^-$ , e. g. 市 Arch.  $d'i >$  Auc.  $z_i$ .

The important point is whether the negative statements are true. Thus there are two questions:

a) Is it true that as a rule neither words with middle consonant initials ( $b^-$ ,  $d^-$ ,  $g^-$ ,  $p^-$ ,  $t^-$ ,  $k^-$  etc.) nor words with strong initials ( $p'^-$ ,  $t'^-$ ,  $k'^-$  etc.) occur as phonetics in words with weak initials ( $j^-$ ,  $j\tilde{k}^-$  derived from Arch.  $d^-$ ,  $g^-$ ) — in other words, is Mr. Waleys observation correct?

b) Is it true that words with strong initials ( $p'^-$ ,  $t'^-$ ,  $k'^-$  etc.) seldom occur as phonetics in words with middle initials ( $b^-$ ,  $d^-$ ,  $g^-$ , etc. and  $p^-$ ,  $t^-$ ,  $k^-$  etc.)?

The statistics out of about 12,000 characters will give us sufficiently reliable answers to these questions.

a) The exceptions to Mr. Waleys rule are comparatively very few. I here give a list of them:

	Middle $\alpha$ :	
室 $g^-$	is phonetic in 1 word	} with Anc. $j^-$ or $j\tilde{k}^-$ ( $<$ Arch. $d^-$ , $g^-$ )
丿 $g^-$	" 1 "	
𠂔 $\gamma^- < g^-$	" 5 words	
完 $\gamma^- < g^-$	" 1 word	
𠂔 $\gamma^- < g^-$	" 2 words	
盜 $\gamma^- < g^-$	" 1 word	
翟 $d^-$ or $d'^-$	" 2 words	
朕 $d^-$	" 1 word	
虫 蟲 $d'^-$	" 1 "	

Example: 𠂔  $\gamma\tilde{a}m < g^-$  phonetic in 𠂔  $j\tilde{a}m < g^-$ .

	Middle $\beta$ :	
軍 $k^-$	is phonetic in 4 words	} with Anc. Chin. $j^-$ or $j\tilde{k}^-$ ( $<$ Arch. $g^-$ , $d^-$ )
均 $k^-$	" 1 word	
監 $k^-$	" 1 "	
貴 $k^-$	" 1 "	
谷 $k^-$	" 6 words	
多 $t^-$	" 2 "	
詹 $t's^- < t'^-$	" 1 word	
占 $t's^- < t'^-$	" 1 "	
佳 $t's^- < t'^-$	" 5 words	
粥 $t's^- > t'^-$	" 1 word	

Example: 谷  $k\tilde{u}k$  phonetic in 𠂔  $j^w\tilde{o}k < g^-$ .

Strong:

頃  $k'$  is phonetic in 2 words } with Anc.  $\dot{z}$  (< Arch.  $g$ -,  $d$ -)  
 彖  $t'$  " 3 " }

Example: 彖 *tuân* phonetic in 緣  $\dot{z}'\ddot{u}n < d$

This harvest is sufficiently small to allow us to state that Mr. Waley's rule is fundamentally correct. The exceptions are not more numerous than can be expected in view of the unsystematic genesis of the characters (many authors and long periods of creation). There is, consequently, a tendency not to spell weak-initialled words with middle- or strong-initialled phonetics, strong enough to be called a leading principle in the composition of the characters.

b) While there are hundreds of cases of strong words as initials in strong ones, e. g. 康  $k'àng$  in 糠  $k'àng$  (空 10 strong derivatives, 忝 7, 充 6, 夸 8, 匡 8, 楚 6, 口 5, 秋 12, 采 8, 翳 15, 豈 6, 受 7, 區 7, 此 10, 昌 7, 去 8, 殷 5, 倉 18, 丕 10, 羞 10, 參 13, 婁 7, 又 6, 春 6 etc.), the cases of strong-initialled words occurring as phonetics in middle-initialled ones, e. g. 乞  $k'jət$  phon. in 紇  $juət < g'$  and 訖  $k'jət$ , are sufficiently limited in number to allow us to speak of them as exceptions:

	Derivates in			Derivates in	
	$b'$ -, $d'$ -, $g'$ -, $dz'$ - etc.	$p$ -, $t$ -, $k$ -, $ts$ - etc.		$b'$ -, $d'$ -, $g'$ -, $dz'$ - etc.	$p$ -, $t$ -, $k$ -, $ts$ - etc.
受		2	溥	1	
崔	2		浦	1	
取	1	10	普		1
秋	1	6	川	1	
妾		2	彖	4	
東		1	夸	1	
息		5	殷	1	5
差	1	3	口		1
函		2	邵		1
察		1	万	1	
此	2	8	刖		2
七		1	去	1	1
戚		2	乞	1	2
	1		气	3	2
孚	4	1	豈		2
尊	3	6	由		1

In these statistics I have left out derivatives with the following phonetics, which had alternative readings with strong and middle initials and which therefore serve just as well in strong as in middle-initialled derivatives 番 ( $p'$  and  $b'$ ), 亢 ( $k'$  and  $k$ ), 开 ( $k'$  and  $k$ ), 可 ( $k'$  and  $k$  [original form of 哥]), 高 ( $k'$  and  $k$ ), 潘 ( $p'$  and  $b'$ ), 辟 ( $p'$  and  $p$ ), 票 ( $p'$  and  $p$ ), 大 ( $t'$  and  $d'$ ), 湯 ( $t'$  and  $\dot{z}$ ), 王 ( $t'$  and  $d'$ ), 它 ( $t'$  and  $d'$ ), 替 ( $ts'$  and  $ts$ ), 眞 ( $ts'$  and  $ts$ ), 且 ( $ts'$  and  $ts$ ), 青 ( $ts'$  and  $ts$ ), 次 ( $ts'$ ,  $ts$ , and  $dz'$ ), 卒 ( $ts'$  and  $ts$ ), 竣 ( $ts'$  and  $ts$ ), 士 ( $t'$ ,  $t$ ,  $d'$ ), 芻 ( $ts'$  and  $ts$ ) 鬲 ( $p'$  and  $b'$ ) 困 ( $k'$  and  $g'$ ).

Further I have left out 3 groups which would have considerably increased the number of exceptions but which cannot be relied upon because they reveal some Arch. compound initial:

出  $t's'juət$  phonetic in 屈 讎  $k'juət$  etc.

僉  $ts'jäm$  phonetic in 檢  $kjəm$ , 斂  $ljäm$  etc.

音 Arch.  $p'zu$  and  $t'zu$ .

Perhaps one ought to set apart also the phonetic 專 with its sub-group 溥, as it is phonetically very irregular (see Anal. Dict. p. 50), which would reduce our 96 exceptions to 86.

It is easy to observe that in no less than 50 out of these 96 (or 86) cases of strong-initialled words as phonetics in middle-initialled ones, we have to do with *affricates*:  $ts'$ -:  $ts$ -, and (less often)  $ts'$ -:  $dz'$ - etc. Now this is quite interesting, for there is acoustically less difference between  $ts'$ -:  $ts$  than between  $t'$ -:  $t$ -, which fact might explain the carelessness of the character inventors in regard to these cases.

To sum up: the table above shows here, as in the case a, that there was a marked tendency *not* to use words with *strong* initials in words with *middle* initials, but only in such with *strong* ones. And so my general thesis may be said to be proved: an aversion from the use of words with stronger initials as phonetics in words with weaker ones is a leading principle with the character inventors, while the opposite, weaker words as phonetics in stronger ones, is admissible and very frequent.

This phenomenon, observed in the explosive-affricate group of initials, can be found also in another group. While cases like

<sup>1</sup> Cf. certain modern dialects in Kiang-su where the distinction  $ts'$ -:  $ts$ - seems to be disappearing.

羊 *iang* < *z*-<sup>1</sup> phonetic in 詳 *ziang* < *dz*- are quite frequent, I have found only two cases of the opposite type:

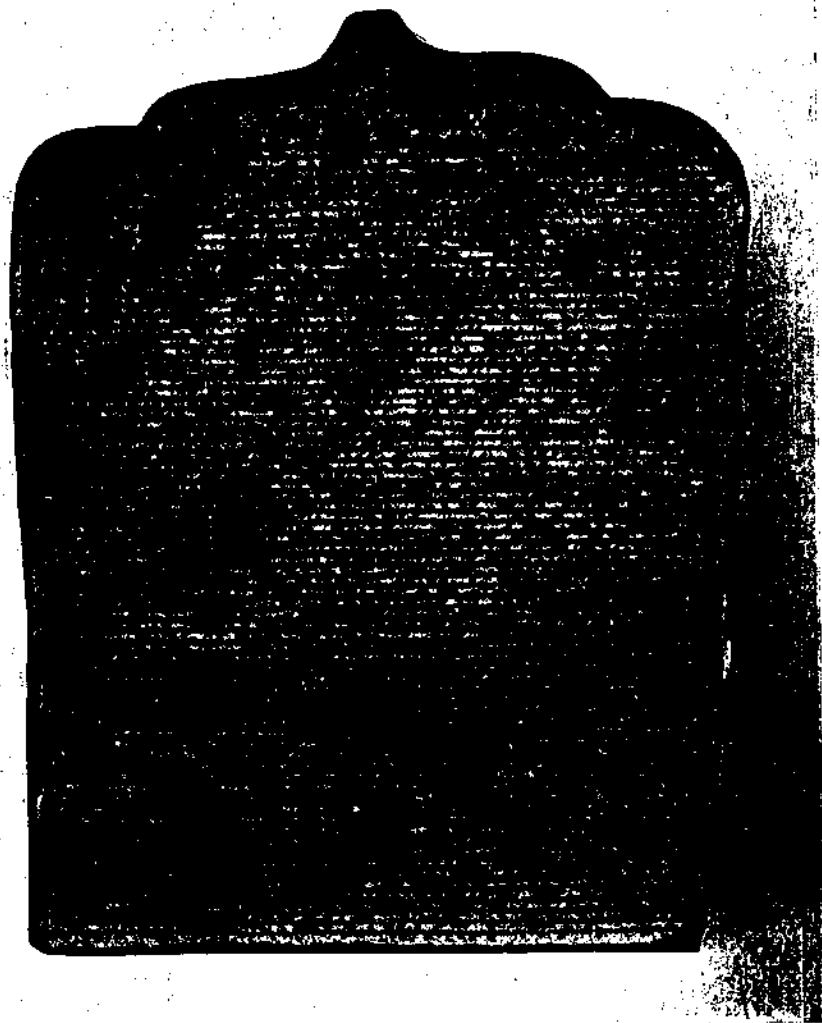
齋 *dz'zi* phonetic in 3 words *zi* < *z*-

巳 *zi* < *dz*- phonetic in 1 word (圮) *i* < *zi*.

Therefore, here again, the same tendency is evident: the weaker words (*z*)*ia* are admissible in the stronger ones (*d*)*zia*, but not vice versa.

The laws governing the choice of phonetics are thus found to be still stricter than even I had imagined. I think it quite possible that a widened knowledge of Archaic Chinese will let us see consistency and method also in many of the numerous cases which now appear to us as fanciful exceptions to the fundamental rules.

<sup>1</sup> Arch. *z*- and not *d*-, see Anal. Dict. p. 26; cf. Mencius, Book III, chap. III, 10: 庠者養也校者教也 — “*ziang* (< *d*-) means *iang* (< *z*-), *iau* (< *g*'-) means *kau* . . .”



Eine der Marmorplatten in Mandalay (Birma), die den buddhistischen Kanton enthalten.