

**Beyond Rhyme:
The Development of the *Digital Etymological Dictionary of Old Chinese*
and Examples of its Use in Textual Analyses from the Western Zhou to the Tang**

Jeffrey R. Tharsen, University of Chicago
Creel-Luce Conference April 23-25, 2010

I. Introduction and Methodology

Chinese readers have known for millennia that the sounds of the language(s) of their forebears were different from their current pronunciations of the characters. Sound glosses (*shēng xùn* 聲訊) in early dictionaries like the *Er ya* 《爾雅》, *Fang yan* 《方言》 and *Shuowen jiezi* 《說文解字》 were the first attempts at transmitting the phonetic values of a graph, complemented in the Western Han 西漢 dynasty by the creation of the *dú ruò* 讀若 method and in the Eastern Han 東漢 by the *fǎnqiè* 反切 system to give the pronunciation of at least the initials and finals of a specific graph. Since then, a wide variety of methods have been developed to help give readers a rough approximation of how the original sounds and rhyme schemes in the text may have functioned, including the use of *dú yīn* 讀音 and *yǔ yīn* 語音. Most recently, the application of the Western phonological methodology for reconstruction of the sounds of extinct languages, in particular the concepts of the allofam (word family) and etonym (word root) as evidenced through related Sinitic languages and in concert with Chinese *xiéshēng* 諧聲 series has become instrumental in the analysis of ancient Chinese phonetics.

The methodology outlined in this study represents a logical next step in the development of tools which can be employed to help us understand the tapestry of sounds and how these ancient texts may have been read throughout their early history, by allowing any text to be

parsed into its possible phonetic readings and then read, using the International Phonetic Alphabet, in what modern phonologists have reconstructed as most approximating the pronunciations of the Chinese language of various early time periods.

For current examples of how computer technology can help make research into the study of ancient Chinese more efficient, for paleography Richard Sears's chineseetymology.org¹ or Chen Wei's 陳偉 database of Warring States character forms on bsm.org.cn; for phonological data, you may have used [eastling.org](http://www.eastling.org)'s database (www.eastling.org/OC/oldage.aspx); and for text searches and philological studies, you may have used the Chinese University of Hong Kong's extensive and well-edited database (chant.org) or Donald Sturgeon's collection of digitized texts and Chinese-English dictionary (chinese.dsturgeon.net). However useful, all of these sites have a basic flaw in that they are only able to return data for a single graph at a time (although Sturgeon's modern Chinese-English dictionary does permit long text strings as input).

Phonology and the phonetics of a text, however, are systems wherein graphs reside within a matrix of sound and grammatical constructs, and I believe the proper next step is to test reconstructed pronunciations not simply in graph-by-graph analyses but on a large scale against texts known to have particular phonological characteristics (rhyme, alliteration, puns and wordplay) to see how and where patterns arise, with the goals of both increasing our understanding and appreciation of the texts themselves and hopefully also showing where our current knowledge of ancient Chinese phonemes and phonic systems may be deficient.

Therefore, for the past several months I have been developing a digital metadictionary and relational database incorporating the phonetic data from the two earliest Chinese rhyme dictionaries: Lu Fayān's 陸法言 *Qièyùn* 《切韻》 from 601 C.E. and the *Guǎngyùn* 《廣韻》,

¹ For cursive forms, see also Stanislas Millot's *Dictionnaire des Formes Cursives des Caractères Chinois* online at www.goulnik.com/cursives/.

presented at the court of Emperor Zhenzong 真宗趙恆 of the Northern Song dynasty in 1011, along with Axel Schuessler's *ABC Etymological Dictionary of Old Chinese*, which contains reconstructed pronunciations in IPA for the Song 宋, Tang 唐 and Late Han periods, and a “Minimal Old Chinese” (OCM) reconstruction roughly corresponding to the period from ca.1000-200 B.C.E. as reconstructed by William Baxter (and when Schuessler and Baxter diverge on a reading, Schuessler's pronunciation is given under “OCM” and Baxter's under the heading “OCB”). OCM, according to Schuessler, “incorporates those features on which there is broad agreement among investigators today...OCM is eliminating much that is highly hypothetical in others' (and my former) proposals, is on firmer ground, and appears relatively straightforward.” Clearly, there remains much which we do not know about the language of the ancient Chinese, and despite Schuessler's assurances, both the reconstructions provided and thus the results of the current study must be viewed as preliminary and still relatively hypothetical; hopefully, as our knowledge of these systems and the sources of data for refining them continue to grow, our ability to more exactly understand the language of these early periods in Chinese history will expand in turn.²

My methodology for construction of the database and its interface was as follows: I converted Schuessler's dictionary (which I have continued to update following his recent addenda and corrigenda) to unicode encoding (UTF-8), including not just the phonetic data but also his allofam-based structure³ along with his comments on the etymology of each graph, and

² See Schuessler, *Minimal Old Chinese and Later Han Chinese*, p.ix-x. In future versions of this database, *fǎnqiè* 反切 data from the *Shuowen jiezi* 《說文解字》 and the Three Kingdoms-period *Erya Yinyi* 《爾雅音義》 may also be worth including, but as the methodology used by the compilers of these sources is still unclear (and their value continues to be debated by paleophonologists), I have not included them in the current version of the Digital EDOC.

³ “Allofams” are groups or “families” of words similar in meaning and sound; the best current example of these groups in the Chinese context is the structure of Bernhard Karlgren's 1957 *Grammata Serica Recensa*.

then proofed each entry in the database against the printed version.⁴ Once proofed, the data was placed into a single large Microsoft Access table and I built a relational database linking the 5,281 entries in the table with the data from the 16,917 entries in the *Qièyùn* and the 25,334 entries in the *Guǎngyùn*.⁵

There are currently three main interfaces which I have built for use with the database:

- 1) A full text search, where all the results for a single search value are returned using any of the listed database fields, and where the user can then iterate through all the results of the search query. This interface can search for graphs, keywords or even phonetic values (in IPA format). As all the returned data are in unicode, the results of the search can be copied and pasted into virtually any application (e.g. Microsoft Word, Excel). Full allofam (word family) data as laid out in Schuessler's dictionary for each graph can also be returned with this interface, though this part of the application is still in development.
- 2) A lookup function for a string of characters (with a maximum limit of 1024 characters), where the text in the search box is parsed and data is returned for each graph according to the database fields chosen using the checkboxes. Unfortunately, as the formatting options for the text display area are extremely limited, some of the data columns do not line up perfectly, so I recommend copying and pasting the data into Excel or a similar spreadsheet application for ease of reading and analysis.
- 3) A text file converter, which reads an input file (in either unicode UTF-8 or ANSI text format) and outputs a tab-delimited unicode text file with the full data from all the database fields chosen

⁴ There may still be rare errata in the phonetic data within the comments field; any errors are solely the fault of the developer and I would greatly appreciate it if users would be so kind as to inform me should they encounter errors in any of the *Digital EDOC*.

⁵ Since any version of any text, digital or otherwise, contains its own idiosyncracies, I would recommend double-checking with a printed version before publishing findings resulting from use of the *Digital EDOC*. Also, all citations should be to printed versions of the aforementioned dictionaries. Page numbers for checking *Guǎngyùn* data can be found online in the MDBG Chinese-English dictionary: www.mdbg.net/chindict/chindict.php.

by the user. As this interface places no limitations on the size of the input file, entire texts can be converted to graph-by-graph recordsets of phonological data. For examples of the full output from this interface, see the Appendices to this paper. This application currently runs at about three seconds per graph. Again, although the output is unicode text and any word processor can read it, as the output file is tab-delimited, I would recommend using Excel to read and manipulate the data.

These functions are nothing that a determined scholar could not do by looking up each word in each of the dictionaries and writing down the values, in much the same way that scholars in the Qing 清 dynasty spent years compiling massive concordances and dictionaries by hand, but the use of digital tools and media has made the process far more efficient.

Each of the following examples is worthy of a much fuller treatment than is given here, but I wanted to share some of my preliminary findings and demonstrate the wide range of applications the use of modern database tools can provide.

II. Five Preliminary Examples of Applications of the *Digital EDOC*

I have long had the intent to lay out the reconstructed pronunciations for the entire *Shijing* 《詩經》 in order to perform a through analysis of the more subtle phonetic structures which lie within the well-documented rhyme schemes. The current study is not that analysis. It is, however, short discussions of five examples of subtle phonetic structures within rhyming texts from the earliest poems of the *Shijing* through Tang-period regulated verse, the research into all of which was made much more efficient by using the *Digital EDOC*. I was not looking for rhymes or rhyming per se (in the case of the *Shijing*, the dominant rhyme schemes have been well-documented for centuries), but attempted rather to discern the more subtle ways that parallel phonemes and phonetic structures may have been used by the

ancient Chinese in the hopes that these structures will help inform our knowledge of how texts from this formative period of Chinese literary style and form might have been composed, read, transmitted and/or performed.

II.1 〈周頌·桓〉“*Huán*” : A Poem Featuring a “Proto-Rhyming” Phonetic Structure

The “Hymns of Zhou” 〈周頌〉 are generally considered to be the oldest poems in the *Shijing*. According to Wang Li’s 王力 1980 *Shijing yundu* 《詩經韻讀》, the poem “*Huán*” 〈桓〉, like many of the poems in the section, contains no rhymes at all. William Baxter, in his 1990 *Handbook of Old Chinese Phonology*, indicates an AB-AB rhyme scheme, with *yáng*-group (陽部) rhymes on *wáng* 王 and *fāng* 方 and *zhī*-group (之部) rhymes on *shì* 士 and *zhī* 之. However, a closer look at the phonetic structure of the poem reveals that neither of the observations by two of the foremost linguists of early Chinese in the modern era should be considered completely correct; I believe Li is right in that there is no formal rhyme scheme akin to the standard tetrasyllabic rhyme structure found in much of the corpus, and Baxter is correct in noting the four rhyme words, but following an analysis of the phonetics based on the reconstructions published in Schuessler’s *ABC Etymological Dictionary of Old Chinese*, it seems what we have here is in fact what I would term a “proto-rhyming” couplet frame rhyme structure of AB-CD-Cd-Ab (where the lowercase letters represent near phonetic matches or “cross rhymes”; see Chart II.1 below).

OCM			OCM		
綏	sūi	*snui			
萬	wàn3	*mans			
邦	bāng	*prōŋ			
，			于	yú	*wa
婁	lóu1	*róʔ	以	yǐ2	*ləʔ
豐	fēng6	*phuŋ	四	sì1	*sisʔ
年	nián	*nîn	方	fāng1	*paŋ
，			，		
天	tiān	*thîn	克	kè1	*khôk
命	mìng	*mrenh	定	dìng2	*têŋ
匪	fěi1	*peiʔ	厥	jué6	*kot
解	xiè	*grêʔ	家	jiā7	*krâ
。			。		
桓	huán1	*wân	於	yú	*ʔa
桓	huán1	*wân	昭	zhāo3	*kiauw
武	wǔ4	*maʔ	于	yú	*wa
王	wáng2	*waŋ	天	tiān	*thîn
！			，		
保	bǎo1	*pûʔ	皇	huáng1	*wân
有	yǒu2	*wəʔ	以	yǐ2	*ləʔ
厥	jué6	*kot	間	jiàn	*krêns
士	shì1	*dzrəʔ	之	zhī1	*tə
，			！		

Chart II.1 : Reconstructed Phonetics and Proto-rhyming in “*Huán*” (周頌·桓)

The reason the repetitive phonetics in “*Huán*” are not a true rhyme scheme is because although the near-rhyme pairs *xiè* 解 (in the *zhī* rhyme-group 支部) and *zhī* 之 (in the *zhī* rhyme-group 之部) and *shì* 士 (in the *zhī* rhyme-group 之部) and *jiā* 家 (in the *yú* rhyme-group 魚部) are phonetically quite similar, and although Li provides two examples of *zhī* rhyme-group (之部)

and *yú* rhyme-group (魚部) rhyming as “combining rhymes” (*héyùn* 合韻) in the *Shijing*,⁶ [-e] and [-ə] are distinct vowels which do not rhyme in the poems of the *Shijing*. Therefore, I propose that while the phonetics do demonstrate a regular, repeating phonic structure to the poem, a structure which might be foreshadowing the regular tetrasyllabic rhyme schemes found in later poetry, at this early stage the system may have been in its nascent stages and yet to be codified.

II.2 Parallel Phonetics in “*Guān jū*” 〈關雎〉

“*Guān jū*” 〈關雎〉 is very likely the most thoroughly discussed poem in Chinese history, and to think one could add anything to the extensive discussions and commentaries on the poem from the past two millenia is perhaps wantonly hubristic. However, in my preliminary analysis of the phonetic structures of the poems of the *Shijing*, “*Guān jū*” stands out as having a remarkable number of parallel phonemic structures, so much so that one has to wonder if it might have been placed at the head of the corpus not only because of the intricacy of its layers of meaning within its finely-wrought rhyming structure but also because of its large amount of phonic repetition and the overall phonic harmony which becomes apparent when read aloud in the reconstructed pronunciations. (See Chart II.2 below.)

⁶ The two poems in which Li notes a “combining rhyme” (*héyùn* 合韻) between -ə (之部) and -a (魚部) are 〈庸. 〉 and 〈小雅. 巷伯〉; he lists no poems as containing an -e (支部) and -ə (之部) cross-rhyme.

OCM	OCB	OCM	OCB	OCM	OCB	OCM	OCB	OCM	OCB	
關	*krôn	*kron	參	*ts ^h rəm	求	*gu	*grju	參	*ts ^h rəm	
關	*krôn	*kron	差	*tshrai	之	*tə	差	*tshrai	參	*ts ^h rəm
雝	*tsha	苻	*grāŋ?	不	*pə	*pjə	苻	*grāŋ?	苻	*grāŋ?
鳩	*ku	*k(r)ju	菜	*ts ^h əh	得	*tək	菜	*ts ^h əh	菜	*ts ^h əh
、		、		、		、		、		
在	*dzə?	左	*tsāi?	寤	*ŋāh	左	*tsāi?	左	*tsāi?	
河	*gāi	右	*wə?, *wəh	寐	*mi(t)s	右	*wə?, *wəh	右	*wə?, *wəh	
之	*tə	流	*r(i)u	思	*sə	采	*tshə?	采	*tshə?	
洲	*tu	之	*tə	服	*bək	*bjək	之	*tə	之	*tə
。		。		。		。		。		
窈	*ʔiŋ?	窈	*ʔiŋ?	悠	*liu	*ljiw	窈	*ʔiŋ?	窈	*ʔiŋ?
窈	*diŋ?	窈	*diŋ?	哉	*tsə	窈	*diŋ?	窈	*diŋ?	
淑	*diuk	淑	*diuk	悠	*liu	*ljiw	淑	*diuk	淑	*diuk
女	*nra?	女	*nra?	哉	*tsə	女	*nra?	女	*nra?	
、		、		、		、		、		
君	*kun	寤	*ŋāh	輶	*thran?	琴	*gəm	鍾	*toŋ	
子	*tsə?	寐	*mi(t)s	轉	*tron	瑟	*srit	鼓	*kâ?	
好	*hŋ?	求	*gu	*grju	反	*pan?	友	*wə?	*wʝə?	
速	*gu	之	*tə	側	*tsrək	之	*tə	之	*tə	
。		。		。		。		。		

Key: rhyme phonetic correlation

Chart II.2 : Reconstructed Phonetics and Parallel Phonemics in “Guān jū” 關雝

I have marked the well-known rhyme pattern with boxes around the rhyme words, but there are several further instances of parallel or repetitive phonic structures which deserve mention. The first of these parallels is found in the third stanza: while *dé* 得 and *fú* 服 are known to make up part of the poem’s rhyme scheme, they are actually only half of a two-graph rhyming couplet, as in the reconstruction *bù* 不 and *sī* 思 complement them perfectly, and indicate a double-rhyme: *pə-*tək* ~ *sə-*bək*.

Secondly, the graph *cāi* 菜 (*ts^həh) should be viewed as part of the repetitive pattern of ending the first, second and fourth lines of the nearly identical second, fourth and fifth stanzas with the final vowel [*-ə]; although *zhī* 之 is not considered to be a rhyming graphs here (the “true rhyme” comes on the preceding graph), the power of the phonic repetition cannot be denied.

That the final graph in the third line of each of these nearly identical stanzas is nǚ 女 (*nraʔ), not a true rhyme but as discussed above in Section II.1, a *héyùn* “combining rhyme” with a similar phonic effect ([a] is an open front unrounded vowel and [ə] is an open mid unrounded vowel); we can draw the conclusion that these graphs were chosen specifically to create a series with phonically harmonic final phonemes: *-əh / *-ə / *-a / *-ə.

Other subtle phonetic parallels within the poem include *mei* 寐 (*mi(t)s) and *sè* 瑟 (*srit) in the fourth line, second position in the second and fourth stanzas, and *yōu* 悠 (*liu) in the third stanza to parallel *yǎo* 窈 (*ʔiúʔ) in the other four. Less striking than the thrice repeated phonic series outlined above but also worth pointing out is the *-a *-u / *-ə *-u / *-ə *-u binomial repetition in the first stanza. Finally, there seems to be an intentional repeated use of the *gē*-rhyme group (歌部, *-ai) in the first (*zuǒ* 左) or second position (*hé* 河, *cē* 差) within a line of four graphs; while noteworthy for its phonic effect ([ai] is a particularly strong diaphoneme), this is perhaps an example of phoneme mirroring, as *cān cē* 參差 (*ts^hrəm *tshrai) is always followed by *zuǒ yòu* 左右 (*tsâiʔ *wəʔ) in the same position in the next line.⁷

II.3 The Phonetic Structure of the “San De”

Moving into the Warring States period, the text entitled “*Sān dé*” 〈參德〉 was published in the fifth volume of the *Shanghai Bowuguan cang Zhanguo Chu zhushu* 《上海博物

⁷ If one were to read the standard Chinese locution “to the left and right” (*zuǒ yòu* 左右) as “to the right and left” (*yòu zuǒ* 右左), there would be a perfect 之部~歌部 parallel in the first two graphs of the second line throughout all but the middle stanza; this makes sense semantically but not culturally, as the phrase *yòu zuǒ* 右左 does not seem to appear anywhere in the early Chinese canon.

館藏戰國楚竹書》 and contains several features which particularly lend it to detailed analysis of its phonetic structures. First, the text clearly rhymes in regular patterns throughout, often in tetrasyllabic meter, and employs a wide variety of rhyme schemes and patterns. Secondly, it contains the most “punctuation” marks (it is still not clear exactly what role these marks played in the composition or reading of the text) of any Warring States-period manuscript yet discovered, and these symbols often delineate semantic and phonic divisions in the text, though this is not always the case (as will be seen below; the symbols are marked by a “■” in Chart II.3).

Having analyzed the text in great detail for its phonetic structures, one particular feature stands out: there seem to be sections within the text which *do not* rhyme interspersed among sections which clearly do rhyme in regular patterns. One can then separate out these non-rhyming lines, which also seem to be somewhat different semantically from the rhymed sections, and analyze their content as separate from the verse sections. While it is still unclear what function the non-rhyming sections performed in the text vis-à-vis the rhyming sections (and I have as yet only completed the initial analysis of exactly how all these sections are laid out within the text and what the differences in semantics exist between divisions), I’d like to present a cursory example of this phenomenon and what a complete phonological analysis can show in such a text, using the strips numbered 1 and 2 from the Shanghai Museum publication.

	OCM	OCB	Rhyme Group		OCM	OCB	Rhyme Group		OCM	OCB	Rhyme Group		OCM	OCB	Rhyme Group		OCM	OCB	Rhyme Group	
1	天	tian	*thîn	真	供	gōng	*koŋ(h)	東	時	shí	*dɔ	之	■							
1	地	dì	*lɾjajs ?	*lɾjajs ?	歌	供	gōng	*koŋ(h)	東	材	caí	*dzə̌	之	■						
1	民	mín	*mîn	真	供	gōng	*koŋ(h)	東	力	lì	*rək	*C-rək	職	■						
1	明	míng	*mraŋ	*mrjaŋ	元	王	wáng	*waŋ	陽	無	wú	*ma	魚		思	sì	*deʔ	之		
1	是	shì	*deʔ	支	謂	wèi	*wjəts	物	參	sān	*tshəm	侵		德	dé	*tək	職			
1	卉	huì	*hwəiʔ	*hwəih	微	木	mù	*mók	屋	須	xū	*ts'io	*sno	侯	時	sì	*dɔ	之		
1	而	ér	*nə	之	後	hòu	*gəʔ / fiəʔ	侯	奮	fèn	*pəns	文	■							
1	天	tian	*thîn	真	惡	è	*ʔəkh	鐸	如	rú	*na(h)	魚		欣	xīn	*hən	文	■		
1	平	píng	*beŋ	耕	旦	dàn	*tāns	元	毋	wú	*mə	之		哭	kū	*khók	屋			
1	晦	huì	*hməʔ	*hma(k)ʔ(s)	職	毋	wú	*mə ?	之	歌	gē	*kái	歌	■						
1	弦	xián	*gîn	真	望	wàng	*waŋ	*wjaŋ	陽	齋	zhāi	*tsrí	脂	宿	sù	*sāh	*sɲa	魚		
								(齊	qí	*dzí	脂	素	sù	*sāh	魚)			
1	是	shì	*deʔ	支	謂	wèi	*wjəts	物												
1	順	shùn	*m-luns	文	天	tian	*thîn	真	之	zhī	*tə	之		常	cháng	*daŋ / *djaŋ	陽	■		
2	敬	jìng	*kreŋh	耕	者	zhě	*taʔ	鐸	得	dé	*tək	職		之	zhī	*tə	之	■		
2	怠	dài	*ləʔ	之	者	zhě	*taʔ	鐸	失	shī	*lhit	質		之	zhī	*tə	之	■		
2	是	shì	*deʔ	支	謂	wèi	*wjəts	物	天	tian	*thîn	真		常	cháng	*daŋ / *djaŋ	陽	■		
2	天	tian	*thîn	真	神	shén	*m-lin	真	之	zhī	*tə	之		X						
2	X				X				X					X						

Chart II.3 The “*Sān dé*”: Reconstructed Pronunciations and Rhyming / Non-rhyming Sections in Strips 1 – 2

As can be seen in the chart above, the initial six lines contain a * $-ə$ (之部) and * $-ək$ (職部) rhyming pattern, followed by a * $-en$ (文部) couplet. In the next eleven graphs, there seems to be no rhyming pattern at all; this section is followed by the phrase “this is what is called” (*shì wèi* 是謂) which occurs repeatedly throughout the text, and then the next four lines seem to rhyme in a *yáng* rhyme-group (陽部) frame around a *zhī* rhyme-group (之部) couplet in perfect tetrasyllabic meter. A rough translation of the non-rhyming section might be:

平旦毋哭	At level dawn, do not weep
晦毋歌■	On the last day of the month, do not sing.
弦望齋宿	At the half-moon and full moon, be abstinent when spending the night. ⁸

It should be noted that many of the grammatical constructions in the text which feature the “prohibitive *wú*” 毋 do not rhyme, but in this case, these three lines seem to stick out and break the rhythm and flow of the text. As they all deal with time periods and associated ritual behavior, unlike the lines which precede and follow them, perhaps this is a quote from another source. Despite ample grounds for speculation, the exact reason why they do not follow any of the other standard rhyming patterns found in the text must remain an open question.

Further uses for the type of phonetic analysis made possible by the *Digital EDOC* include strip ordering, as when rhyming passages flow from one strip to the next, the following strip can be adduced; this is exactly the case in the arguments put forth by scholars such as Chen Jian 陳劍 to place “*Sān de*” strip 17 after strip 6.⁹ Finally, in cases of clear rhyming patterns where the rhyme-graph is unknown to paleographers or difficult to decipher, use of the rhyme scheme can assist with determining the likely phonetic element in the graph in question, which in turn can lead to greater understanding of the semantic element(s) and thus potential meanings for the graph and the text.

⁸ A note on the transcription: *zhāi sù* 齋宿 is the extended reading the editor Li Ling 李零 gives for the graphs as they appear literally in the text: *qì sù* 齊素 (and they are included in parentheses below the line). One could also reasonably read this line without changing the graphs, to mean “be level and pure”, but as the phonetics are virtually identical, the extension has no real impact on the sounds and overall lack of rhyme in this section.

⁹ See Chen Jian 陳劍, “《三德》竹简编联的一处补正”.

II.4 “Inscription on the Stone Gate-Tower” 〈石闕銘〉 : Use of an Archaic Pronunciation

Anthologized in the “Inscriptions” *míng* 銘 chapter of the 526 C.E. *Wen xuan* 《文選》, the “Inscription on the Stone Gate-Tower” 〈石闕銘〉 is attributed to the poet Lu Chui 陸倕 (470-526). The verse section of the Inscription is comprised of couplets in tetrasyllabic meter, where changes in rhyme indicate breaks in the content of the inscription and also the stanzaic structure (see Chart II.4 below). Phonologically, the poem is quite straightforward, with one notable exception: in the sixth line, the graph *fǒu* 否 is used to rhyme in the middle of a section made up of *zhǐ* rhyme-group (脂部) and *zhī* rhyme-group (之部) words. What is worth noting is not that *fǒu* 否 would here be pronounced with *bù* 不 as its phonetic element (not *kǒu* 口), as throughout their history *fǒu* 否 and *bù* 不 have been used interchangeably (and were likely even considered to be the same word), but as *bù* 不 is listed only under the *yóu* rhyme-group (尤部) in the *Qìeyùn* and under the *yóu* 尤, *yǒu* 有 and *wù* 物 rhyme groups in the *Guǎngyùn* (none of which have any direct phonetic links to the *zhī* rhyme-group), in order to rhyme as a *zhī* rhyme-group (之部) word, the graph would have to be read in its archaic pronunciation: *pə (or *pǰə as reconstructed by Baxter). We can then infer that at least one Chinese poet of the sixth century C.E. used an archaic pronunciation, in at least one case, to widen the available choice of rhyme words. Further studies of this kind may be able to indicate whether this was a more widespread phenomenon or limited to very specific graphs and/or poets.

方	fāng	pjwaŋ	陽	A
梁	liáng	ljaŋ	陽	A
光	guāng	kuaŋ	陽	A
章	zhāng	tśjaŋ	陽	A
指	zhǐ	tśi ^B (*kiʔ)	脂	B
紀	jì	kji ^B (*kəʔ)	之	B
否	fǒu	pjəu ^B (*pə)	幽(之)	B
起	qǐ	k ^h ji ^B (*khəʔ)	之	B
巍	wēi	njwei	微	C
微	wēi	mjwei	微	C
輝	huī	xjwei	微	C
歸	guī	kjwei	微	C
宇	yǔ	ju ^B	幽	D
柱	zhù	ɟju ^B	幽	D
矩	jǔ	kju ^B	幽	D
雨	yǔ	ju ^B	幽	D
房	fáng	bjwaŋ	陽	A
方	fāng	pjwaŋ	陽	A
長	cháng	ɟjaŋ	陽	A
疆	jiāng	kjaŋ	陽	A

Chart II.4 : Rhyme groups and *Qieyùn* Pronunciations of the Rhyme Words and Stanzas as Determined by Rhyme in “Inscription on the Stone Gate-Tower”

II.6 Wang Wei's "Deer Enclosure": Analyzing the Phonetics of Regulated Verse Poetry

Key:	(2-3) Semantic structure	Tone ("level" or "oblique")
	"level tone" (平聲)	Graph (字)
—	"oblique tone" (仄聲: 上, 去, 入)	Modern pronunciation (Mandarin 普通話)
(tone)	Tone is "incorrect" for location	Reconstructed Tang-period pronunciation

〈鹿柴〉“Deer Enclosure” by Wang Wei 王維

(2-3)

—	—			—	()	—	—			。
空	山	不	見	人	，	但	聞	人	語	響
kōng	shān	bù	jiàn	rén	dàn	wén	rén	yǔ	xiǎng	
k ^h uŋ	ʃam	pjəu	kien	ńzjen	dān	mjuən	ńzjen	ŋjwo	xjaŋ	
			—	—			—	—		。
返	景	入	深	林	，	復	照	青	苔	上
fán	jǐng	rù	shēn	lín	fù	zhào	qīng	tái	shàng	
pjwən	kjəŋ	ńzjəp	ʃjəm	ljəm	bjuk	tʃjäu	ts ^h jǎŋ	dái	zjaŋ	

Chart II.5 Tones and *Qieyùn* (Tang-period) Pronunciations for “Deer Enclosure” by Wang Wei

As is clear from Chart II.5 above, “Deer Enclosure” has an extremely irregular tonal structure: the rhymes follow the standard XAXA custom, with no other interlinear rhyming, but the tonal structure [A2--B1--B2--A1] conflicts with the standard for regulated Tang pentasyllabic verse [A1--A2--B1--B2], where the dominant rhyme words would come at the end of A2 and B2; to make “Deer Enclosure” tonally standard, the last five graphs should be put first.¹⁰ The effect this has, however, is to add a wonderful feeling of circularity to the poem, and as the rhymes peculiarly fall on oblique characters (level tone graphs are almost always the rhyme words in Tang poetry), the reader is left hanging at the end and is drawn up¹¹ to reread the first line (“空山

¹⁰ Alternately, one could start at the second line (“但聞人語響”) and then read the first line last, creating a perfect tonal structure, but semantically, the first line is the beginning of the poem.

¹¹ To take it to its extreme, the poem would originally have been written (it was originally painted on the top of a painting) and read top-to-bottom in most likely four (or possibly two) vertical columns (perhaps even

不見人”), which then resolves the tonal structure. However, as “But” (*dàn* 但) begins the next column, the reader might naturally continue reading and the endless cycle continues.

One of the most interesting results of using Tang-period phonetics to analyze this poem is that two examples of alliteration in parallel lines can be found in the poem (and which do not rhyme in modern Mandarin): *jiàn rén* 見人 (*kien n̄zjen*) and *shēn lín* 深林 (*śjəm ljəm*). That *shēn lín* 深林 is a perfect rhyming binome may help explain why it was such a popular trope for poets of the period.

A final note on the tonal structure of the poem: in order to keep tonal harmony, the word “but” (*dàn* 但) would either have to be read incorrectly in a level tone, or would likely serve as a phonetic break: an oblique tone where a level tone would be expected, an eccentricity in the tonality and rhythm of the poem. Whether this is semantically important is left up to the reader, but it is an aspect of Wang Wei’s composition which should not be ignored, and at the very least would likely have had some effect on a reader of the time used to perfect tonal harmony.

III. Final Comments

The above examples have hopefully demonstrated the range and utility of database tools when analyzing early Chinese phonetic systems. In the future, the database can easily be expanded by including William Baxter and Laurent Sagart’s full database of etymological data and proposed reconstructions, as well as those graphs included in Axel Schuessler’s *Minimal Old Chinese and Later Han Chinese* but which were not in his *ABC Etymological Dictionary of Old*

indenting the second and fourth lines), and so the final graph, *shàng* 上, could be a marker to “go back up” (as the basic meaning of *shàng* 上 is “to ascend”, like the sun in the sky) to read from the beginning again.

Chinese. Data from other ancient Chinese dictionaries should likely also be incorporated (such as the *Er ya* 《爾雅》, *Fang yan* 《方言》 and *Shuowen jiezi* 《說文解字》) at some point. Finally, as we continue to refine our understanding of which graphs stood as homophones in which contexts in ancient Chinese, we can further develop the phonetic links and associated reconstructions. As stated above, we are still in a very preliminary stage of this project, and much work is left to be done, but it is my hope that the results so far will be encouraging to those who wish to employ tools such as these in our attempts to understand the full splendor of the literature of ancient China.